

Springer Series on Bio- and Neurosystems 16

Martin Ziegler
Thomas Mussenbrock
Hermann Kohlstedt *Editors*

Bio-Inspired Information Pathways

From Neuroscience to Neurotronics


OPEN ACCESS

 Springer

Springer Series on Bio- and Neurosystems

Volume 16

Series Editor

Nikola Kasabov , Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Penrose, New Zealand

Editorial Board

Shun-ichi Amari, Mathematical Neuroscience, RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

Paolo Avesani, Neuroinformatics Laboratory, University of Trento, Trento, Italy

Lubica Benuskova, Department of Computer Science, University of Otago, Dunedin, New Zealand

Chris M. Brown, Department of Biochemistry, University of Otago, North Dunedin, New Zealand

Richard J. Duro, Grupo Integrado de Ingenieria, Universidade da Coruna, Ferrol, Spain

Petia Georgieva , DETI/IEETA, University of Aveiro, Aveiro, Portugal

Zeng-Guang Hou, Chinese Academy of Sciences, Beijing, China

Giacomo Indiveri, Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zürich, Switzerland

Irwin King , The Chinese University of Hong Kong, Hong Kong, China

Robert Kozma, University of Memphis, Memphis, TN, USA

Andreas König, University of Kaiserslautern, Kaiserslautern, Rheinland-Pfalz, Germany

Danilo Mandic, Department of Electrical and Electronic Engineering, Imperial College London, London, UK

Francesco Masulli, DIBRIS, University of Genova, Genova, Genova, Italy

JeanPhilippe Thivierge, School of Psychology, University of Ottawa, Ottawa, ON, Canada

Allessandro E.P. Villa, Universite de Lausanne, Lausanne, Switzerland

The Springer Series on Bio- and Neurosystems publishes fundamental principles and state-of-the-art research at the intersection of biology, neuroscience, information processing and the engineering sciences. The series covers general informatics methods and techniques, together with their use to answer biological or medical questions. Of interest are both basics and new developments on traditional methods such as machine learning, artificial neural networks, statistical methods, nonlinear dynamics, information processing methods, and image and signal processing. New findings in biology and neuroscience obtained through informatics and engineering methods, topics in systems biology, medicine, neuroscience and ecology, as well as engineering applications such as robotic rehabilitation, health information technologies, and many more, are also examined. The main target group includes informaticians and engineers interested in biology, neuroscience and medicine, as well as biologists and neuroscientists using computational and engineering tools. Volumes published in the series include monographs, edited volumes, and selected conference proceedings. Books purposely devoted to supporting education at the graduate and post-graduate levels in bio- and neuroinformatics, computational biology and neuroscience, systems biology, systems neuroscience and other related areas are of particular interest.

All books published in the series are submitted for consideration in Web of Science.

Martin Ziegler · Thomas Mussenbrock ·
Hermann Kohlstedt
Editors

Bio-Inspired Information Pathways

From Neuroscience to Neurotronics

 Springer

Editors

Martin Ziegler
Institute of Micro- and Nanotechnologies
(IMN) MacroNano@
Technische Universität Ilmenau, Micro-
and Nanoelectronic Systems
Ilmenau, Germany

Thomas Mussenbrock
Chair of Applied Electrodynamics
and Plasma Technology
Ruhr University Bochum
Bochum, Germany

Hermann Kohlstedt
Institute for Electrical Engineering
and Information Engineering
Kiel University
Kiel, Germany



ISSN 2520-8535 ISSN 2520-8543 (electronic)
Springer Series on Bio- and Neurosystems
ISBN 978-3-031-36704-5 ISBN 978-3-031-36705-2 (eBook)
<https://doi.org/10.1007/978-3-031-36705-2>

© The Editor(s) (if applicable) and The Author(s) 2024. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword by The Series Editor

Bio-Inspired Information Pathways: From Neuroscience to Neurotronics

The human brain is the ultimate product of millions of years of evolution of Life. It is also an ultimate inspiration for the development of new technologies and an ultimate target to use these technologies to better understand the brain (reverse engineering) and beyond that—to improve Life on this planet by applying these technologies across disciplines. These technologies, along with the process of creating them, are referred here as *Neurotronics*.

Neurotronics technologies are characterized by low power consumption, ultra-fast performance, and high density integration. Major research in the Neurotronics area is currently devoted to *memristor-based devices* and their applications for neuromorphic computation (computation inspired by the information processing in nervous systems).

For readers not familiar with the concept of *memristor*, but who want to read and understand the book, a memristor is defined as a non-volatile (not depending on power) electronic memory device that is based on a resistor which can manifest memory in terms of retaining its resistance after the current stimuli and changing it based on its previous values. It often is resembled to a synapsis in the human brain. The theoretical concept was first introduced by Leon Chua in 1971 as the fourth fundamental two-terminal circuit element, following the resistor, the capacitor, and the inductor. He called his paper “Memristor—the missing circuit element”. The first memristor circuit was physically realized, though 40 years after the introduction of the concept.

And now neuroscientists and electronic circuit designers need to address the following challenges:

- (1) How a memristor circuit can be physically realized using different and more efficient physical and chemistry nano-properties of materials?

- (2) How useful will these Neurotronics devices be to model the whole human brain at its actual size and functions?
- (3) What problems can be solved with these Neurotronics devices that cannot be efficiently solved with traditional neural networks and AI systems?
- (4) Can memristors, being in one of two non-volatile states (no power supply needed) of low- or high-resistance state, substitute totally the current transistors in the contemporary computers, saving vital power energy for the next generations of societies to come?
- (5) What is the relationship between memristor computation and nano-particle quantum processes?
- (6) How can current brain-inspired spiking neural network models be efficiently implemented on memristor devices?
- (7) How can memristor devices reveal back biological and neurophysiological functions in the human brain and the nervous system that cannot be revealed with the use of the current AI technologies?
- (8) What is beyond memristors, neuromorphic computation, and quantum computing?

The book covers three major areas, leading from neuroscience to the present and future computer devices, and addressing some of the above questions:

- (1) Neuronal Architectures and Information Processing.
- (2) Bio-inspired System Technology.
- (3) Physical Substrates for Neuromorphic Engineering.

Readers of the book can find the state of the art of a wide range of topics that are results of deep interdisciplinary studies involving knowledge from neuroscience, biology, psychology, physics, chemistry, computer science, electrical engineering, material science, network science, mathematics, and nonlinear dynamics.

Neuroscientists can learn more about topics, such as neuronal synchrony; self-organized criticality; plasticity; connectomics; nervous system growth under external stimuli; fundamental biological information pathways that explore the spatiotemporal fundamentals of nervous systems; homeostasis; local versus global aspects of information processing in nervous systems, mimicking neuroplasticity with memristor devices; perceptual decision-making circuit; modeling the biological process of pattern recognition in the box jellyfish nervous system; and many more.

Electronic engineers, computer scientists, and information technologists can learn more about cutting-edge technology, including memristive and memsensor devices; memristive properties, such as $I-V$ linearity, switching time, retention, state numbers, time dependency, device variability; different technologies of bilayer metal oxide memristive devices, floating-gate transistors as memristors, ex situ versus in situ measurements of memristive devices, silver-based filamentary memristive devices, micro-electro-mechanical systems (MEMS), bio-inspired, neuromorphic acoustic sensing, application-specific integrated circuit (ASIC) technologies, chemical oscillators, neuromorphic hardware systems, and many more.

The book is a big step forward toward the development of new neuromorphic computer systems that process not only vector-based and bit-represented information, but create deep information spatiotemporal pathways, similar to the pathways created in the human brain when learning and performing physical and cognitive tasks, as in the example shown in the Fig. 1.

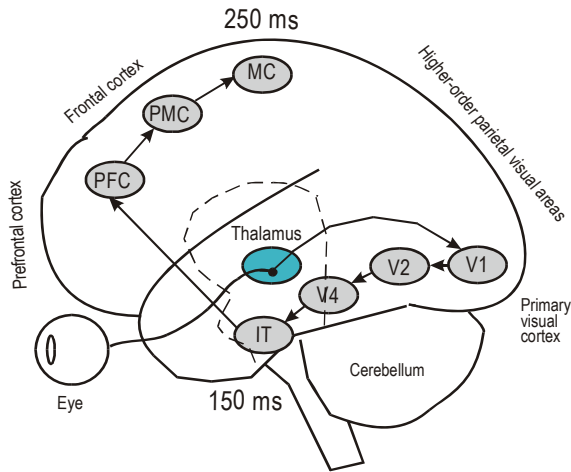


Fig. 1 Example of a deep information pathway for learning and processing visual stimuli in humans for image classification. Location of cortical areas: V1 = primary visual cortex, V2 = secondary visual cortex, V4 = quaternary visual cortex, IT = inferotemporal cortex, PFC = prefrontal cortex, PMC = premotor cortex, MC = motor cortex (from Kasabov, N., *Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence*, Springer Nature (2019) 750p., <https://www.springer.com/gp/book/9783662577134>)

This is a promising step toward the realization of a true model of the human brain and various brain-inspired cognitive and robotic systems.

As the Springer Editor of the series of Bio-and Neuro-systems, I am delighted to welcome this book to our series as it incorporates the state of the art of biology, neuroscience, physics, chemistry, and their impact on the development of novel information technologies.

This book represents the results of many years of collective work of leading European institutions under a large project (CRC1461) titled *Neurotronics: Bio-Inspired Information Pathways* (<https://www.crc1461-neurotronics.de/research/project-overview>) and I would like to encourage other research groups to follow this

example and to share and disseminate their results and knowledge from their projects in our Springer series of Bio-/Neuro-systems.

Prof. Nikola Kasabov
Life Fellow IEEE, Fellow RSNZ
Fellow INNS College of Fellows
Penrose, New Zealand



Preface

Nature is a tinkerer, not an inventor

François Jacob

We are delighted to present this volume which addresses basic and likewise fascinating aspects of information pathways in nervous systems and their undisputed relevance for future computing architectures.

Indeed, creatures are very well adapted to their specific ecological niche, resulting from a billion years of ongoing evolution and the therewith associated interaction of creatures with their environment during their life span. From an engineering point of view, nervous systems process information in such a way that silicon technology, the holy grail of modern digital computing strategies, seems to be outstripped. Electronic devices and circuits, such as transistors, memories, and processors, are optimized for small parameter spreads to run at GHz clock frequencies, applying precise pulse timing. In particular, they exploit nanosecond signal pulses that travel nearly at the speed of light along well-ordered transmission lines interconnecting the system parts in an essential two-dimensional topology. In contrast, information pathways in nervous systems are characterized by highly irregular tissue consisting of neurons, synapses, and axons. Low conduction velocities of the order of several m/s lead to pronounced signal retardation, i.e., delay. Nervous systems can be considered as time-varying networks in which spike dynamics and cellular morphology are intricately linked and reciprocally interwoven. Understanding how living creatures perform pattern recognition and cognitive tasks could be a particularly important ability for future computing architectures.

Understanding how living creatures perform pattern recognition and cognitive tasks is of particular importance for computing architectures: by studying these information pathways refined over eons of evolution, researchers may be able to streamline the process of developing more highly advanced, energy-efficient autonomous systems. With the advent of novel electronic and ionic components along with a deeper understanding of information pathways in living species, a plethora of opportunities to develop completely novel information processing avenues are within

reach. This book encompasses research works on fundamental studies on memristive compartments, such as mesa-type devices and self-assembled nanostructures to neuromorphic circuits. The latter comprises units for pattern recognition, decision-making in order to mimic basal synaptic functionalities up to neuromorphic systems for pattern recognition, decision-making, and bio-inspired sensing. Here for guidelines are basal information pathways in nervous systems, from the local synaptic level to the entire nervous system network. The dual importance of local learning rules is addressed, from spike timing dependent plasticity at the neuron level to the interwoven morphological and dynamical mechanisms of the global network. Biological principles are highlighted, including phylogenies, ontogenesis, and homeostasis, with particular emphasis on network topology and dynamics and their relevance for engineered pulsed-coupled oscillators ensembles.

This volume summarizes research work performed within the Research Unit 2093, entitled: Memristive Devices for Neural Systems between 2014 and 2020, and within the in 2020 established Collaborative Research Center 1461: Neurotronics-Bio-Inspired Information Pathways. Both projects are funded by the German Research Society (DFG).

Ilmenau, Germany
Bochum, Germany
Kiel, Germany
March 2023

Martin Ziegler
Thomas Mussenbrock
Hermann Kohlstedt

Contents

Matter and Mind Matter	1
Tom Birkoben and Hermann Kohlstedt	
Neuromorphic Circuits with Redox-Based Memristive Devices	43
Finn Zahari, Martin Ziegler, Pouya Doerwald, Christian Wenger, and Hermann Kohlstedt	
Redox-Based Bi-Layer Metal Oxide Memristive Devices	87
Finn Zahari, Seongae Park, Mamathamba K. Mahadevaiah, Christian Wenger, Hermann Kohlstedt, and Martin Ziegler	
MemFlash—Floating Gate Transistors as Memristors	115
Henning Winterfeld, Hermann Kohlstedt, and Martin Ziegler	
Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices	129
Ole Gronenberg, Georg Haberfehlner, Finn Zahari, Richard Marquardt, Christian Kübel, Gerald Kothleitner, and Lorenz Kienle	
Modeling and Simulation of Silver-Based Filamentary Memristive Devices	159
Sven Dirkmann, Jan Trieschmann, and Thomas Mussenbrock	
Integration of Memristive Devices into a 130 nm CMOS Baseline Technology	177
Mamathamba Kalishettyhalli Mahadevaiah, Marco Lisker, Mirko Fräschke, Steffen Marschmeyer, Eduardo Perez, Emilio Perez-Bosch Quesada, Christian Wenger, and Andreas Mai	
A Wave Digital Approach Towards Bio-inspired Computing Using Memristive Networks	191
Dennis Michaelis and Karlheinz Ochs	

Memristive Switching: From Individual Nanoparticles Towards Complex Nanoparticle Networks	219
Niko Carstens, Maik-Ivo Terasa, Pia Holtz, Sören Kaps, Thomas Strunskus, Abdou Hassanien, Rainer Adelung, Franz Faupel, and Alexander Vahl	
Photocatalytic Deposition for Metal Line Formation	241
Salih Veziroglu, Moritz Paulsen, Jan Schardt, Blessing Adejube, Cenk Aktas, Alexander Vahl, and Martina Gerken	
Smart Sensor Arrays	265
Maik-Ivo Terasa, Leonard Siebert, Pia Holtz, Sören Kaps, Oleg Lupan, Jürgen Carstensen, Franz Faupel, Alexander Vahl, and Rainer Adelung	
Bio-inspired, Neuromorphic Acoustic Sensing	287
Claudia Lenk, Kalpan Ved, Steve Durstewitz, Tzvetan Ivanov, Martin Ziegler, and Philipp Hövel	
A Bio-inspired Perceptual Decision-Making Circuit Based on the Hassenstein-Reichardt Direction Detector	317
Tom Birkoben, Mirko Hansen, Marina Ignatov, Martin Ziegler, and Hermann Kohlstedt	
Pattern Recognition in the Box Jellyfish Rhopalial Nervous System Mimicked by an Ensemble of Pulsed Coupled Oscillators	335
Jan Bielecki, Tom Birkoben, Maximiliane Noll, Jan-Frederik Freiberg, Peer Wulff, Heinrich Terlau, and Hermann Kohlstedt	
Biologically Inspired and Energy-Efficient Neurons	357
Xiaoyan Cheng, Sebastian Simmich, Finn Zahari, Tom Birkoben, Maximiliane Noll, Tobias Wolfer, Eckhard Hennig, Robert Rieger, Hermann Kohlstedt, and Andreas Bahr	
Synchronization Phenomena in Oscillator Networks: From Kuramoto and Chua to Chemical Oscillators	385
Petro Feketa, Alexander Schaum, and Thomas Meurer	
Emulation of Learning Behavior in the Hippocampus: From Memristive Learning to Behavioral Tests	407
Christian Kaernbach, Thorsten Bartsch, Maximilian Brütt, Annika Hanert, Nick Diederich, and Martin Ziegler	

About the Editors

Martin Ziegler is a Full Professor and the Chair of Micro- and Nanoelectronic Systems at the Faculty of Electrical Engineering and Information Technology of the Technische Universität Ilmenau, Germany. He works in the field of neuromorphic engineering. His main research interests concern the transfer of biological learning and memory into electronic systems. For this purpose, he studies and applies memristive materials and devices, micro-electromechanical systems, and neuromorphic circuits.

Thomas Mussenbrock is a Full Professor and the Chair of Applied Electrodynamics and Plasma Technology in the Faculty of Electrical Engineering and Information Technology at Ruhr University Bochum, Germany. His primary research interests cover modeling and simulation of low-temperature plasmas and plasma processes. Further topics include transport phenomena at the nanoscale and nanoionic devices.

Hermann Kohlstedt is a Full Professor of Nanoelectronics at Kiel University, Germany. His research concerns the development of novel electronics device and computing primitives based on biological fundamentals. His topics span from thin film analytics, device fabrication and their electrical characterization.

Matter and Mind Matter



Tom Birkoben and Hermann Kohlstedt

Abstract As a result of a hundred million years of evolution, living animals have adapted extremely well to their ecological niche. Such adaptation implies species-specific interactions with their immediate environment by processing sensory cues and responding with appropriate behavior. Understanding how living creatures perform pattern recognition and cognitive tasks is of particular importance for computing architectures: by studying these information pathways refined over eons of evolution, researchers may be able to streamline the process of developing more highly advanced, energy efficient autonomous systems. With the advent of novel electronic and ionic components along with a deeper understanding of information pathways in living species, a plethora of opportunities to develop completely novel information processing avenues are within reach. Here, we describe the basal information pathways in nervous systems, from the local neuron level to the entire nervous system network. The dual importance of local learning rules is addressed, from spike timing dependent plasticity at the neuron level to the interwoven morphological and dynamical mechanisms of the global network. Basal biological principles are highlighted, including phylogenies, ontogenesis, and homeostasis, with particular emphasis on network topology and dynamics. While in machine learning system training is performed on virgin networks without any a priori knowledge, the approach proposed here distinguishes itself unambiguously by employing growth mechanisms as a guideline to design novel computing architectures. Including fundamental biological information pathways that explore the spatiotemporal fundamentals of nervous systems has untapped potential for the development of entirely novel information processing systems. Finally, a benchmark for neuromorphic systems is suggested.

Keywords Bio-inspired Computing · Phylogenesis · Ontogenesis · Homeostasis · Artificial spatio-temporal networks

T. Birkoben · H. Kohlstedt (✉)

Chair of Nanoelectronics, Faculty of Engineering, Institute for Electrical Engineering and Information Engineering, Kiel University, Kiel, Germany
e-mail: hko@tf.uni-kiel.de

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_1

1 Introduction

Is it truly possible to implement higher brain functions, such as perception or consciousness, in engineered systems? This question has been frequently raised over the last few decades and has led to distinct views over time, as both neurobiological understanding and available computational capabilities advanced [1–7]. The essence of this question goes back to the fundamental relation between matter and mind, which was addressed as early as ancient Greece, and emerged in the principle of “Dualism” most famously defended by the philosopher René Descartes in the sixteenth century [8]. Descartes postulated that the body (matter) and the mind are distinct and separate units in human beings because he could not imagine that mental phenomena could be explained by natural mechanisms [9]. However, the invention of electroencephalography (EEG) and imaging techniques, such as functional magnetic resonance imaging (fMRI), enabled the study of inner information processing in the human brain and individuals’ states of consciousness [10–13]. As a result, the strict distinction between matter and mind has become blurry [14, 15]. Strong evidence has been found that the inner representation of the human brain (the mind) is related to its neurochemistry (the matter), e.g. the amount and type of neurotransmitters and/or drugs within the nervous system [16, 17]. It is therefore worthwhile to reconsider the relationship between mind and matter when engineering artificial systems to exhibit higher brain functions by considering recent progress in nanoelectronics and neurobiology.

This perspective on future computing is motivated by three key aspects. First by the recent, growing movement to reboot the entire field of computing, i.e. how data are processed. Second by state-of-the-art, fundamental progress in neurosciences, including the fields of complex networks and dynamic brain states. Third by advances in materials science and nanoelectronics that have led to, e.g., memristive devices, nanoparticle/nanowire networks, and fluidic memristors, providing new functionality in electronics, such as synaptic-like plasticity or spatio-temporal networks [18–22]. With the foreseen restrictions on current digital computing, the question “What comes next?” finds its answer in merging novel discoveries made on the nervous system’s information pathways with the development of novel electronic devices, paving the way to an entirely new kind of computing.

In this perspective, we present a concept of an artificial spatio-temporal network which uses temporal and structural mechanisms in nervous systems as guidelines. It addresses the important, interwoven spatiotemporal aspects of information pathways and processing in nervous systems [23–28]. The state of nervous system criticality combined with the blooming and pruning of nervous cells during growth might be an interesting guideline to develop new computing principles [23, 25, 29–34]. Components essential for artificial spatio-temporal networks and a pathway to realize it, are presented, including biological fundamentals such as phylogenies, ontogenesis, and homeostasis [35, 36]. However, these basal biological mechanisms alone might not be sufficient to establish mental functions in artificial systems; we therefore include the temporal binding hypothesis developed in neuroscience as a further essential

guideline [10, 37–40]. The synchronized firing of neural ensembles across different brain regions is treated as a fundamental neural mechanism that defines how a hierarchical network structure, such as the brain, can integrate several sensory inputs to determine the unity of an object (for example linking form, color, size, motion, etc., together) [41, 42]. Opportunities and possible limitations of this approach towards implementing higher brain functions in artificial systems, such as perception and consciousness, will be discussed. The paper is arranged as follows:

In chapter “Neuromorphic Circuits with Redox-Based Memristive Devices”, the current status of computing architectures is summarized. In chapter “Redox-Based Bi-Layer Metal Oxide Memristive Devices”, we present a condensed overview of advanced device components, with a focus on memristive switching devices. We subsequently address spatiotemporal information processing in nervous systems, including network structure, network dynamics, and homeostasis in chapter “MemFlash—Floating Gate Transistors as Memristors”. An artificial spatio-temporal network concept is introduced in chapter “Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices”, where we hypothesize on which information pathways might lead to higher brain functions in engineered systems based on hardware-oriented electrochemical electronics but also discuss current limitations of this approach. In chapter “Modeling and Simulation of Silver-Based Filamentary Memristive Devices” a possible benchmark is discussed for bio-inspired systems. Chapter “Integration of Memristive Devices into a 130 nm CMOS Baseline Technology” provides a discussion on the practical implementation of an artificial spatio-temporal network, to mimic basal biological information pathways.

2 The Current State of Information Technology

The sixties marked the beginning of a glorious time in information technology as the tremendous opportunities of silicon technology merged with the concept of Boolean computing, resulting in the first digital revolution [43].

This development followed the exponential increase over time of electronic components integration on a chip predicted by Gordon E. Moore, combined with a sequential data processing architecture comprising a central processing unit and memory for data storage, for which Alain Turing and John von Neumann laid the foundation years before [44–46]. The tremendous technological and economical success of the digital revolution is still going strong today with seemingly no end in sight. Billions of transistors on a single processor chip, displaying features as small as about 10 nm in size and clock frequencies of a few GHz, are the current standard in CMOS (Complementary Metal Oxide Semiconductor) technology, representing the backbone of today’s semiconductor industry [47–49]. However, during the last couple of years, dark clouds have appeared on the horizon for the semiconductor industry. The envisioned goal of downscaling devices with every new circuit generation to the nm level has created an ongoing need to develop ever more sophisticated and expensive fabrication tools for e.g. lithography, dry-etching, and layer

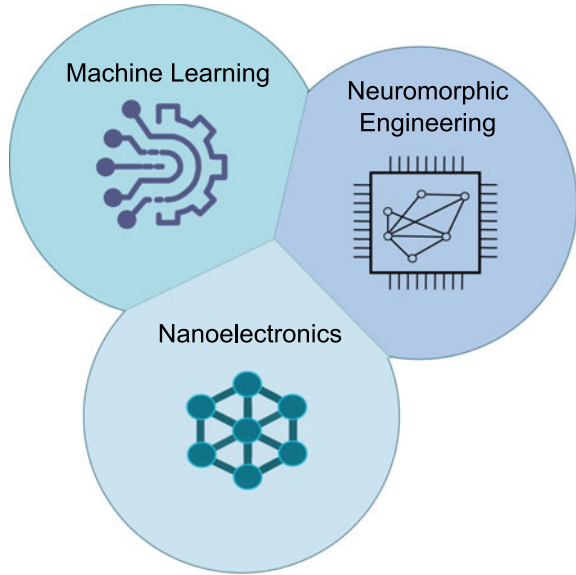
deposition [50–52]. As a result, each new circuit generation entails an increasing economic risk for semiconductor companies. Moreover, over the last few decades, progress in processor core clock rates have overtaken memory access and access times, leading to a cumbersome situation where data transmission between the arithmetic logic unit (ALU) and memories dominates instead of the arithmetic information process itself. This system level-related challenge is called memory latency (or memory gap) and is a consequence of the von Neumann bottleneck, where data is processed sequentially [47, 53, 54]. Two major obstacles restrict the further development of information technology, namely limitations in downscaling at the device level, and memory latency on the architecture. Although society is experiencing a second digital revolution via the resurgence of artificial intelligence (AI) and the Internet of Things (IOT), Moore’s law, which has been driving the computer industry for decades, is becoming outdated as the limits of device integration and/or economical boundaries have now been reached. The incredible advances made by the first digital revolution based on binary “0” and “1” computation combined with the latest achievements in the field of machine learning led to great progress in speech and pattern recognition, while rendering autonomous driving tangible. Yet, additional challenges are growing increasingly problematic behind the scenes. Huge, power consuming hardware systems in the form of cloud servers are now mandatory to support recent advancements in AI and the IOT. This is why global digital players, such as Google, Amazon, and Facebook, as well as bitcoin trading platforms need energy-hungry server farms [55].

On the system level, and in particular since the advent of the internet and the renewed interest in AI, the power consumption of the digital world is growing without limits, in increasing conflict with sustainable and climate-neutral resource management. Moreover, future autonomous electric vehicles require both high recognition capability and low power consumption. It therefore is hardly surprising that the semiconductor world is currently in an era of upheaval, turning a new page on information processing based on novel computing architectures and advanced hardware components.

3 Advanced Computing Architectures and Novel Electronic Devices

The aim of this section is to give a short survey on novel computing architectures and advanced electronic devices. We do not intend to present a comprehensive overview but instead to give a taste of the developments currently being pursued to overcome the limitations of digital computing and to establish new computing primitives. To simplify access to the different research areas for interested readers, we discuss seminal and overview papers and present recently published pioneering research. Nonetheless, we are aware that the given reference list is far from exhaustive. In addition, this section is critical to understanding the similarities, and most importantly the

Fig. 1 A shamrock-like illustration of the three development areas, which characterize the currently expansive development in the field of Artificial Intelligence (AI)



distinctions, between artificial spatio-temporal networks and standard neuromorphic computation presented in Sect. 5. While traditional von Neumann computing continues to dominate the ICT scene, recent groundbreaking innovations in alternative computing architectures and advanced electronic devices have become hard to ignore [22, 56–60].

These developments are threefold. Firstly, somewhat older concepts, such as artificial neural networks (ANN) leading to Deep Learning (DL) systems, have received an impressive performance boost through novel and efficient algorithms paired with more powerful electronics hardware [61]. Secondly, new technologies, such as Quantum Computing and Reservoir Computing (RC), have appeared, leading to remarkable results [62–64]. Thirdly, in the field of nanoelectronics, a plethora of advanced device structures and novel functional components has led to a rethink of traditional computing architecture, paving the way to in-memory computing that circumvents the von-Neumann bottleneck [58, 60, 65, 66]. In Fig. 1 a shamrock-like illustration highlights these three research areas.

The first leaf representing Machine Learning encompasses Artificial Neural Networks (ANN), Spiking Neural Networks (SNN), Reservoir Computing (RC), Long Short Term Memory (LSTM), and Deep Learning (DL) systems [67]. The foundations of Neural Networks were laid by McCulloch and Pitts [68], Rosenblatt’s Perceptron [69] for ANNs, and von Neumann’s postulate of SNNs in 1956. More recent inventions from Jäger (Reservoir Computing) [70], Hochreiter and Schmidhuber (Long-/Short-Term Memory) [71]) and Hinton (DL) [67]) have advanced the field one huge step forward and comprise the backbone of today’s AI.

In the second leaf, the field of neuromorphic engineering, initiated by Carver Mead and Mohawa and Rodney Douglas, seeks to mimic the basal mechanisms of information processing in nervous systems via an essentially hardware-oriented approach [72–74]. In recent years, great progress has been made in the development of bio-inspired processors. Here, event-based spiking neural networks (SNNs) in the form of either mixed (analog and digital) or strictly digital signal processing provides novel opportunities for low-power data processing [59, 60, 75–77]. Interestingly, some of the spiking neuromorphic circuits work at biologically relevant frequencies, exhibiting low energy consumption. One point of merit for neuromorphic engineering is their energy per synaptic operation (SOP), which is in the pJ to nJ range for neuroprocessors [57, 60, 77].

Hence, the incorporation of relatively few basal mechanisms of biological information processing, such as leaky-integrated firing, axon delays, and local learning rules, can lead to significant improvements in resource management.

Recent advances in the field of nanoelectronics devices, such as memristive devices, nanoparticle networks, nanowire networks, or memristive fluids, compose the third leaf of advanced computer architecture. Research in silicon nanoelectronics is dominated by the development of new field effect transistors (FET) [52, 78] for the next generation of CMOS circuits, as well as entirely novel devices and materials exhibiting advanced functionalities [66, 79–84]. In particular, the memristor (originating from memory and resistor, also called memristive device) is a two terminal device that exhibits attractive features for various applications in the post-Moore area, generating considerable interest. Memristive devices were intensively studied in the sixties and seventies [85–87]. The field was further propelled forward by the establishment of the theoretical background of memristors by Leon Chua (1971), with the corresponding experimental realization and interpretation by Hewlett–Packard (HP)-Labs (2008) [88, 89]. Over the years, numerous books and reviews have covered fundamental and practical properties of memristive devices and their related circuits [18, 84, 90–92].

So far we have described nanoelectronic devices fabricated using top-down methods, where the layers are deposited on an entire wafer and the devices are patterned by lithography and dry-etching [93–95]. In bottom-up approaches, functional materials are deposited or synthesized to obtain networks, such as irregular nanowires and/or 3D textures. Often the self-assembly capabilities of materials are exploited to create complex structures. Top-down and bottom-up approaches are habitually combined to create the electrical connections necessary to characterize the structures [96]. In the context of bio-inspired computing, we would like to highlight here the work done on nanowire networks [20, 97–103]. The structure of such networks, and in particular their dynamic properties, reflect basal functionalities as observed in nervous systems, such as small-world connectivity and self-organized criticality (SOC) [23, 104]. We would like to emphasize that the three ICT research areas shown in Fig. 1 are not independent from one another: there is considerable overlap between them, which has proven to be mutually beneficial.

3.1 Advanced Computing Architectures

Here we present a few concepts of novel and reconsidered computing architectures. We would like to emphasize that the following four examples were chosen to demonstrate the diversity of the field but are not intended to give a comprehensive overview. The icons in Fig. 2 represent different computing principles.

For concepts other than those shown in Fig. 2, such as quantum computing, cellular automata, and probabilistic computing, we refer to the literature [62, 80, 105–107]. We focus on comparing today’s digital computing to in-memory computing, vector matrix multiplication, reservoir computing, oscillatory computing, and bio-inspired computing (see icons in Fig. 2).

In order to overcome the von Neumann bottleneck of digital computing (Fig. 2a), near-memory computing was developed in 1990 [108]. Here, the strict separation of an arithmetic logic unit communicating with several distinct memories was eliminated. Part of the computational tasks was performed within the memory itself, leading to more efficient computing. This development has recently shifted to a higher gear, leading to in-memory computing (Fig. 2b) following the invention of memristive crossbar-arrays [58, 109]. Vector matrix multiplication (Fig. 2c) is considered a key hardware booster in Deep Learning. The time and energy consuming task of vector matrix multiplication is performed in a memristive crossbar-array in which the input and output layer are interconnected by an array of weighted, check-

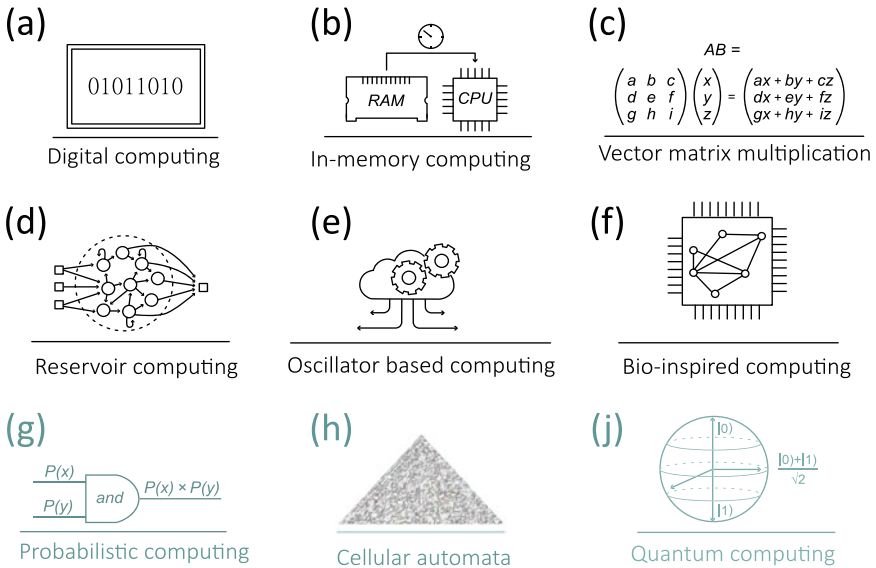


Fig. 2 The illustration shows snapshots of different computing strategies. **a** Digital computing, **b** in-memory computing, **c** matrix multiplication, **d** reservoir computing, **e** oscillatory computing, **f** bio-inspired computing, **g** probabilistic computing, **h** cellular automata, **i** quantum computing

board arranged memristive devices [58, 92, 110]. Vector matrix multiplication is an example of how Deep Learning may benefit from the development of new electronic devices, e.g. memristors. Reservoir Computing (Fig. 2b) was independently invented by Herbert Jäger and Wolfgang Maass and belongs to the general framework of Recurrent Neural Networks (RNN) [70, 111]. In RNNs, a backpropagation through-time procedure is typically applied to adjust (train) the weights of the network to desired target functions. Here, a significant amount of time is required, with no certainty that the optimal weights will be set after learning. Thus in RC the reservoir consists of an ensemble of nonlinear elements coupled to one another. The reservoir projects incoming data and time series to a higher dimension that can be easily readout by conventional classifiers, in which the training is executed by means of a linear regression, for example. This reservoir can be either virtual or physical. These aforementioned reservoirs are designed like neural networks in which the connections are randomized but remain fixed during computation. Physical reservoirs are those which rely on natural systems exhibiting nonlinearity [64, 112, 113].

The goal of analog computing is to mimic complex technical systems by means of electronic circuits which represent key system parameters as a set of voltage levels at nodes. Oscillatory computing (Fig. 2e) refers to a subset of analog computing in which the oscillator frequencies and phases enrich the representation of information. Oscillatory systems are omnipresent in nature and engineering [114–116]. Technically, oscillators can be realized in numerous ways, such as in discrete or integrated semiconductor electronics, spin-torque devices, Josephson junctions, optical devices, or micro electro-mechanical systems (Schneider et al. 2018; Lequeux et al. 2016; Chen et al. 2020; Ignatov et al. 2016; X. Cheng et al. 2021; Feldmann et al. 2019a; C. Lenk, L. Seeber, and M. Ziegler 2020) [82, 117–122]. In general, dynamical systems and their coupled oscillators may offer elegant solutions to compute HP-hard problems. Coupled oscillator networks have been successfully exploited in the field of pattern recognition [123, 124]. However, larger systems have not yet been successfully developed due to noise-stability problems and device constraints in the new class of compact oscillators based on VO_2 or NbO_x , for example [125, 126].

The term bio-inspired computing (Fig. 2f) is only loosely defined. To a large extent, the computing primitives described above (see Figs. 1 and 2b–e) are more or less biologically motivated. The Perceptron is a crude blue print of a neuron and is still today at the heart of Deep Learning systems [69]. In-memory computing is a strategy to abrogate the strict separation of the ALU and memory in digital computing, and is derived from biological information processing where logic and memory are blended. Neuromorphic processors contain circuits that can execute the Leaky Integrate-and-Fire dynamics of neurons, including the biologically motivated winner-take-all (WTA) principle, and introduces axon delays [57, 127]. Coupled oscillators imitate the orchestra of neural ensembles, i.e. the communication of separate brain regions which is considered to be the fundamental mechanism that explains perception [10, 37, 41]. Cellular automata, for example, were introduced by John von Neumann to describe self-reproduction in biology [128]. Probabilistic computing is based on Bayesian inference, which is closely related to the way humans make decisions [129, 130]. Therefore, it is essential to declare precisely to

what extent an artificially built system is bio-inspired and which biological pathway have been applied as design principles [131].

3.2 Novel Electronic Devices

There is an ongoing effort to shrink silicon FETs to feature sizes below 5 nm. The FinFET structure has dominated CMOS technology since its invention in 1989 [132]. Novel designs, such as GAAFET (Gate-All-Around) and MBCFET (Multi-Bridge-Channel), are serious candidates for next generation CMOS chips (see Fig. 3a) [52]. Aside from this ongoing improvement of conventional FETs, devices with novel functionalities and materials have been attracting considerable interest to implement novel computing architecture. Magnetic Josephson Junctions, photonic synapses, and bio-organic memories represent only a fraction of current development strategies [21, 80, 117, 121]. In Fig. 3b–d, unconventional nanoelectronics device structures are illustrated. In Fig. 3b, a memristive device structure is illustrated, comprising two electrodes separated by a memristive layer. In the same Figure, a qualitative I-V curve of a memristive device is shown alongside a sketch of a biological synapse (see also Fig. 4), highlighting that memristive devices are promising artificial synaptic counterparts due to their capability of presenting variable resistive weights in engineered neural networks [19].

One universal property of the memristive device concept is that the memristive state depends on previously induced charge flows, applied currents, or applied electric fields, thus storing a historically-determined resistance state. For details concerning resistive switching and the underlying physical-chemical mechanisms, we refer the reader to the references given in the figure caption (Fig. 3) and the overwhelming literature on the subject [19, 91, 134, 135].

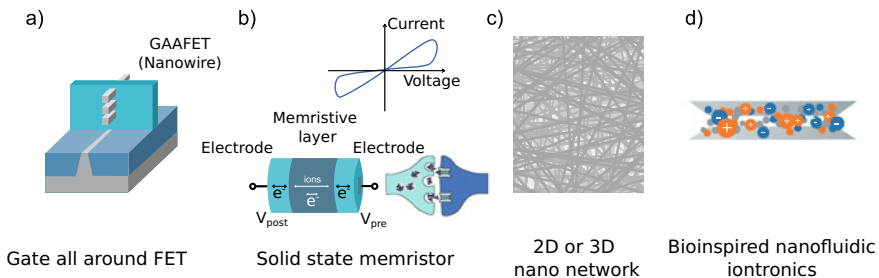


Fig. 3 Schematics of four advanced device components. **a** 3D view graph of a GAAFET as applied in today's latest digital processors [52], **b** sketch of a memristive device including a qualitative I-V curve and illustration of a synapse [18, 91, 133], **c** cartoon of a 3D nanowire network [20, 22, 81, 97, 99–102], **d** 3D cross-sectional graph of a fluidic memristive device adapted from [21] with permission

It is this concurrently complex and simple device concept, together with the tremendous predicted potential for breakthrough technologies in areas such as universal memories and novel non-Boolean computing schemes for cognitive electronic systems, that propels the research and development of memristors and memristor-based circuits worldwide. It is important to mention that, in contrast to the theoretically simplistic memristor concept, in practice the realization of memristive devices by modern thin film technology is a task littered with obstacles. Up until now, a huge number of experimental findings on memristor devices consisting of a broad variety of metal/insulator material combinations have been published, all of which show memristive I-V curves [91, 133, 135]. At first glance, it seems that the toolbox of resistive switching devices is ready for nearly any circuit application: simply pick a device concept and follow the extensive materials and methods laid out in the literature. However, a closer look at the fine details casts a dark shadow on this bright research field, leading to a harsh awakening based on hard facts. These “hard facts” are the requirements and boundary conditions set by the envisaged circuit applications, in which memristors must fit technologically, electronically, and economically. Currently, two main development avenues can be explored for memristive devices. The first focuses on resistive random access memories (RRAMs). It is believed that the zoo of today’s existing memory diversity can be replaced by a single (universal) memory concept. RRAMs are considered attractive candidates for universal memories because they: (i) show non-volatile data storage, (ii) can be densely integrated, (iii) are fast, and (iv) are cheap to produce. In particular, such a universal memory might attenuate the problem known as memory latency in modern digital computers [54, 136]. Besides the RRAM goal which may be categorized under the label “More Than Moore”, novel and very appealing computer architectures have been proposed in which memristors might play a vital role. Another main focus of possible memristive device applications may be associated with such catchphrases as: non-Boolean computing, bio-inspired information processing, neuromorphic engineering, or cognitive electronics [66, 137–140]. On the local, synaptic level, learning in nervous systems is explained by Hebb’s learning rule and Spike-timing dependent plasticity (STDP), amongst others [141]. STDP and other memory-related mechanisms observed in nervous systems, such as Long-term Potentiation (LTP) and Long-term Depression (LTD) [142], have been successfully mimicked by memristive devices [143, 144]. Moreover, traditional studies known from behaviorism, such as classical conditioning (e.g. Pavlov’s dog), anticipation, and optical illusions, were successfully realized experimentally by both single and pairs of memristive devices [119, 145–148]. The extent to which larger networks of memristive devices are able to mimic higher brain functions is still unknown.

In Fig. 3c, a sketch of a nanowire network (NWN) is shown. NWNs have been successfully synthesized for various materials, such as metals, oxides, and semiconductors [113, 139]. Nanowires show appealing features with respect to bio-inspired computing from the point of structure, topology, and inherent dynamics [99–102, 149]. In recent comprehensive reviews by Zhu et al. and Kuncic and Nakayama, hallmarks known from biological systems as small-world connectivity (topology) and self-organized criticality (dynamic) were addressed [22, 101]. Interestingly, brain-

like avalanche effects have been observed in NWNs that exhibit dynamic features found in nervous systems [23, 25, 100, 103]. Finally, we would like to emphasize that emergent neuromorphic materials and devices are not restricted to the solid state phase. In Fig. 3d, a sketch related to a nanofluidic device is shown. Bocquet and co-workers demonstrated by analysis and molecular dynamic simulations that ion transport across quasi-two-dimensional slits under an electric field displays memristive I-V curves, as well as spiking voltage patterns in accordance with the Hodgkin–Huxley model of biological neurons [21, 150]. We would like to emphasize that while these examples of NWNs and nanofluidics clearly demonstrate that the material “tool box” offers novel opportunities to implement higher brain function, its full potential has yet to be fully explored.

4 Information Processing in Nervous Systems

This perspective explores the role of information processing observed in nervous systems as a basis for the development of energy-efficient technological computing systems, and even the possibility of implementing higher brain functions in engineered systems. Nervous systems offer paradigms to improve energy-efficient artificial information processing units. The exploration of signal pathways in nervous systems shows us how evolution led to extremely energy-efficient signal processing units (nervous systems). For example, the human brain dissipates a power of only roughly 20–25 W. This, in addition to the amazing capabilities of humans’ vision and hearing, reveals fascinating opportunities for autonomous vehicles or speech recognition. Hence, processing sensitive data in server clouds may lead to severe security concerns. The data of millions of cars in motion, including their controllability, falling into the wrong hands could lead to fatal attacks; Local data processing in an autonomous car with low power consumption is preferable.

Creatures are very well adapted to their specific ecological niche, a result of a hundred million years of ongoing evolution and the associated interaction between creatures and their environment throughout their life span [151–153]. In particular, information pathways in nervous systems are prototypes for engineers to perform cognitive tasks in quasi-real time with extremely low power consumption [154]. These features alone, and the information processing behind them, represent attractive models for entirely new computing architectures. In Sects. 4.1 and 4.2, local and global aspects of information processing mechanisms are presented, respectively. Distinct differences between digital computing and information pathways in biological systems are highlighted in the framework of topology and dynamics to motivate the concept of artificial spatio-temporal networks, as subset of the field of bio-inspired information processing. In Sects. 4.3 (Phylogenies and Ontogenesis) and 4.4 (Homeostasis), we underline important hallmarks of information processing in biological systems which have so far only been partly considered for artificial systems. Note that in chapter “MemFlash—Floating Gate Transistors as Memristors”, we do not address how such mechanisms can be established in electronics: This is the

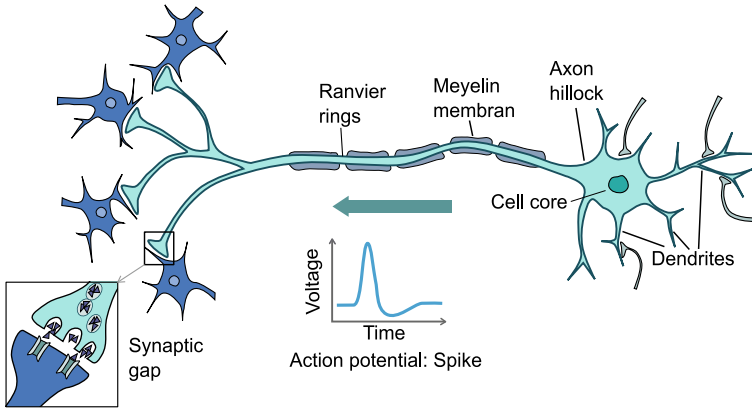


Fig. 4 Blueprint of a neuron including an enlarged sketch of a synapse and the illustration of a single action potential, a spike

subject of chapter “Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices”, where several approaches are proposed to implement an artificial spatio-temporal network. It is not our goal to develop another pattern recognition system but to address the fundamental question: “To what extent can higher brain functions be reproduced in artificial systems?”. We believe that essential information pathways in biology have been to a large extent overlooked, as detailed in this perspective. One important difference between artificial spatio-temporal networks and contemporary AI and neuromorphic engineering is that essential growth mechanisms observed in nervous systems are exploited as a guideline in the former.

4.1 Local Aspects of Information Processing in Nervous Systems

In contrast to current clock-driven Boolean Turing machines, information processing in biological nervous systems is characterized by highly parallel, energy-efficient, and adaptive architecture [53, 155, 156]. When it comes to pattern recognition, failure tolerance, and cognitive tasks, even simple creatures outperform supercomputers, in particular regarding power dissipation. Fundamental building blocks leading to such remarkable properties exploit neurons as central processing units, which are interconnected by synapses to form a complex dynamical three dimensional network, the connectome [157]. In Fig. 4, the structure of a neuron is sketched, including the soma, dendrites, the axon, and connections to other neurons by synapses.

An action potential (spike) is defined as a sudden transitory and propagating change in the resting potential across a membrane. Action potentials sent from presynaptic neurons are received via the dendrites and synapses of the postsynaptic neu-

rons. Those signals are integrated within the cell body of the postsynaptic neuron. When a threshold potential is reached, the neuron generates a new spike or a sequence of new spikes at the axon hillock that are transmitted via the axon to a postsynaptic neuron. This entire process is called Leaky Integrate-and-Fire (LIF). The term leaky reflects the fact that the cell membrane is not a perfect electrical insulator. Numerous LIF models, such as the FitzHugh–Nagumo, Morris–Lecar, or Hindmarsh–Rose models, have been developed to address different aspects of the biological substrate [158–161]. Depending on the electrical activity of two connected neurons, the connection strengths (the weights) can become weaker or stronger. This is at the heart of Donald E. Hebb’s learning rule, who first recognized that “Neurons which fire together wire together” [162]. On the biochemical level the variable strength is explained by the amount of neuro transmitters (vesicles) which are released into the synaptic cleft.

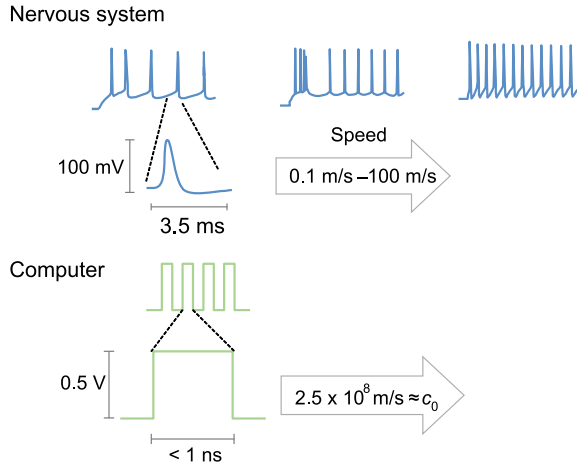
From an engineering point of view, nervous systems process information in such a way that silicon technology, the holy grail of modern digital computing strategies, seems to be outmatched. For example, electronic components and circuits, such as transistors, memories and processors, are optimized for small parameter spreads to run at GHz clock frequencies under a precise pulse timing [47, 54]. In particular, they exploit nanosecond signal pulses that travel at nearly the speed of light along well-ordered transmission lines that connect different system parts in an essentially two-dimensional topology. In contrast, information pathways in nervous systems are characterized by highly irregular tissue consisting of neurons, synapses, and axons. Low conduction velocities on the order of several m/s lead to pronounced signal retardation, i.e. delays. In Fig. 5, characteristic timescales of CMOS processors and nervous systems are compared. In digital computing, the pulse duration is below a ns, and the signal transmission velocity is at nearly the speed of light. In nervous systems, the corresponding values are 3.5 ms for the pulse duration of an action potential or spike, and a few tenths of a ms for the transmission of a spike along myelinated axons [163]. Whereas the clock frequency of a modern Si processor is about 5 GHz, human EEG brain waves range from below 1 Hz to a few 100 Hz [41, 164, 165]. This represents a six orders of magnitude discrepancy between technical and biological parameters. These facts alone point towards fundamental differences between information processing in digital computing and those in natural nervous systems.

4.2 Global Aspects of Information Processing in Nervous Systems

Nervous systems are considered to be time-varying networks in which spike-dynamics and cellular morphology are intricately linked and reciprocally interwoven [27, 151, 166, 167].

Information processing in nervous systems applies a broad range of structurally and temporally related phenomena [163, 168, 169]. At the level of individual

Fig. 5 Comparison between pulse transmission speed, pulse duration, and voltage amplitude in nervous systems and digital computing. The sequence of action potentials were adapted with permission from Fig. 1 of Ref. [160]



synapses, neurons, and axons, the formation and transmission of action potentials (“spikes”) are reasonably well understood. However, a look at the mesoscopic and macroscopic level of the three-dimensional neuronal network leads to an entirely different assessment. Although groundbreaking progress has been reported on *in vivo* and *in vitro* techniques over the last decades, the nervous system’s spatiotemporal information processing is still not well understood [12, 170–172]. The biochemical mechanisms that explain higher brain functions at the cellular level, such as awareness, perception, and in particular consciousness, remain elusive [15, 24, 173–176]. Nonetheless, neuroscientists were able to identify basal mechanisms that define the fundamental platform of the unique and marvelous nervous system’s information processing. Characteristic features, such as STDP [141, 159, 177], stochastic firing and bursting of neurons in the hundred Hz range, recurrent network structures, and aspects of oscillatory synchrony in larger neuronal ensembles [39, 41, 104, 114, 178–183] are essential ingredients in biologically-based information processing. Moreover, factors related to the close interaction of a nervous system with its environment, i.e. external stimuli, are of crucial importance [184]. Therefore, neuronal design principles provide a model for bio-inspired computing systems, which are diametric to development strategies in present binary IT, including GHz clock frequencies, near-light-speed signal transmission, and clearly separated from logic and memory [47, 51].

Beyond that, we would like to emphasize that information-related aspects of nervous systems during evolution (phylogenies), along with their individual development throughout their lifetime (ontogenesis), provide a promising model from which novel electronic architectures may be designed. In the animal kingdom, the intricacy of nervous systems varies tremendously between single- and multi-cellular organisms, and the human brain with its billions of interconnected neurons [185–193]. For the sake of completeness, we would like to specify that the existence of cognitive functionalities in entities without a nervous system, such as plants or the acellular

slime mold *Physarum polycephalum*, is currently heavily debated. For interested readers, more detailed information can be found in the following Refs. [194–197].

Despite their different cognitive capabilities, neurons and nervous systems present many common features in all creature, such as synapses, signal transmission lines (axons), and action potentials (spikes), that act as basic information building blocks. While the term morphology defines the real structure of a nerve net, the topology of a net is more abstract, related to important theoretical graphical parameters that define the connectome of a nervous system [24, 27, 198, 199]. The connectome is considered to be the canonical state describing the cellular wiring diagram of a nerve net. Edges, nodes, cluster coefficients, characteristic path lengths, hubs, and motifs determine the topological quality of a net, for example. An unraveling of the micro- and macro-connectome and nervous system dynamics offer a suitable model for the next generation of bioinspired hardware electronics [169].

The network cube (Fig. 6a) classifies a number of different nets according to theoretical attributes, including randomness, modularity, and heterogeneity [26, 200]. Interestingly, in this framework, cortical maps (lower right corner of the cube) extracted from the structural properties of nervous systems are somewhat isolated from all other nets, which are located in the upper left corner of the cube. In Fig. 6b, dynamical complexity (y-axis) is described as a state between complete asynchrony, with independent, random firing of the individual oscillators, and complete synchrony, with all oscillators firing in phase [201]. Between these two extremes, system dynamics can be characterized by a complex and time-varying interaction of the oscillatory ensemble. This regime exhibits features of self-organized criticality (SOC) typically observed at and near phase transitions and might be identified by avalanches of firing neuron ensembles [20, 23, 202–205]. Avalanche behavior is common in many physical phenomena, such as magnetic systems, earthquakes, and brain dynamics at the critical region of phase transitions, and were first described by Bak et al. [202]. The common feature of all these systems is slow external driving, causing an intermittent, widely distributed response. Avalanches appear in very different sizes, often distributed in the form of power laws. As known from statistical physics, power laws imply the absence of a characteristic scale, a property observed close to a critical point. When describing the dynamics in a nervous system using the SOC and brain-like avalanches models, the type of phase transition associated to each term must be clearly defined. For example, SOC and brain-like avalanches in NWNs (see Sect. 3.2 and Fig. 3c) are related to non-activity - activity phase transitions. In the context of firing neuron ensembles in a brain, SOC and avalanches may describe a temporal phase transition between the asynchronous and synchronous states [203]. In other words, a system could be in the supercritical state (above the critical point) in an inactive -active phase transition, while remaining subcritical (below the critical point) with respect to the asynchronous-synchronous phase transition. However, such phase transitions are not necessarily exclusive and might appear simultaneously in the brain, or the mechanisms could even be interwoven. So far, while nanoparticle networks and NWNs have been studied in context with their activity pattern, neuron-like oscillatory components have yet to be considered. The orchestra of firing neuron ensembles is considered a key underlining mechanism in understanding the

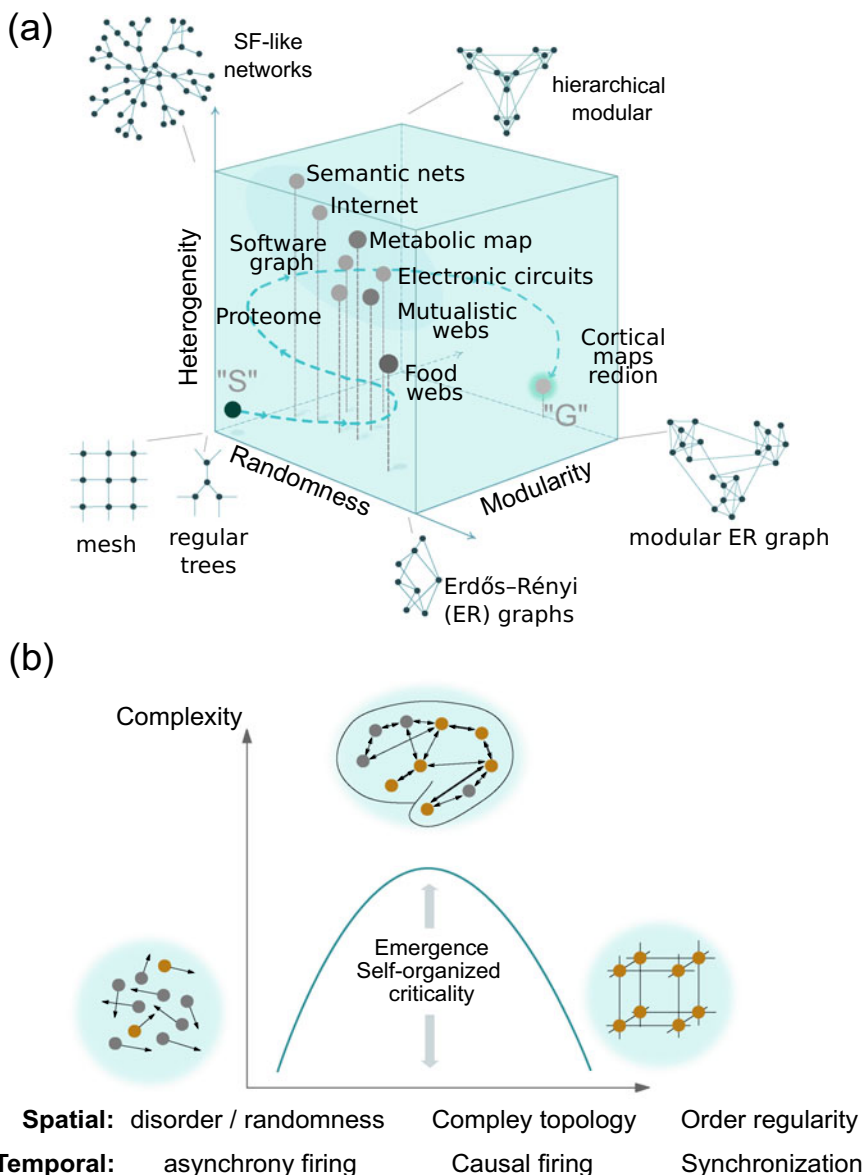


Fig. 6 Network cube and complexity: **a** classification of various networks. The dashed blue line illustrates a fictive, guided “walk” through the cube, starting from “S” and ending at the goal “G”. In this way, it will be possible to push the network properties of neuromorphic circuits towards those of cortical maps (see lower right-hand corner in the cube). This approach is part of an artificial spatio-temporal networks. **b** Qualitative illustration of the complexity term. From a spatial (topological) and temporal (dynamical) point-of-view, a complex system is neither completely random nor entirely ordered, but exhibits a state in between. **a** and **b** are adapted with permission from Solé and Valverde [200] and Huberman and Hogg [201], respectively

binding problem, i.e. the capability of the brain to integrate (bind) different sensory inputs. For example, such a process can occur in the visual system when forming a unified perception of the environment [37–40, 42, 206, 207]. Suggestions on ways to include relaxation-type oscillators to mimic the LIF features of biological neurons and the state of SOC are presented in chapter “Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices”. Finally, the topological and temporal dynamics of the regime are extremely sensitive to external distortions (stimuli) at the critical point, allowing the system to respond in numerous ways to external stimuli [23, 25, 30, 208, 209]. In biological terms, this means that the system can easily adapt to risky environmentally-driven situations. The manifold brain states available near the point of criticality offer a wide repertoire of means to react in a reasonable way to external tasks imposed by the environment. In extreme situations, this improves the chances of survival and is of evolutionary importance.

4.3 *Phylogenies and Ontogenesis*

The origin of bio-inspired computing can be best drawn from the two following neuroscience quotations:

(1) Gilles Laurent pointed out the common evolutionary heritage of living organisms. His contribution “Shall We Even Understand the Fly’s Brain? (see: 23 Problems in Systems Neuroscience edited by J. L. van Hemmen and T. J. Sejnowski, Chap. 1, p. 3, [210]) states: “When it comes to computation, integrative principles, or “cognitive” issues such as perception, however, most neuroscientists act as if King Cortex appeared one bright morning out of nowhere, leaving in the mud a zoo of robotic critters, prisoners of their flawed designs and obviously incapable of perception, feeling, pain, sleep, or emotions, to name but a few of their deficiencies.” [211].

(2) Martijn P. van den Heuvel et al., made in “The Neonatal Connectome During Preterm Brain Development” the following statement: “The adult cerebral brain network is the result of a complex developmental trajectory. From the prenatal formation of the first neurons, throughout the first years of life and all the way into late adolescents, the brain undergoes an elaborate developmental trajectory.” [212].

How are these sayings so important for the design of novel bio-inspired computing primitives? The general idea behind these two quotations is the concept of development. Quotation (1) by Gilles Laurent highlights evolutionary development, the phylogenies of species, and their relevance to the emergence of the human cortex. This bottom-up approach favors the study of less complex creatures that appeared early during evolution, laying the foundation for much more complex nervous systems in vertebrates [190]. In particular, information processing strategies throughout evolution and in completely different species are astonishingly similar, if not the exact same. For example, the basic ingredients of information processing (neurons, synapses, and action potentials, as described in chapter “Redox-Based Bi-Layer Metal Oxide Memristive Devices”) in the nervous systems of squids and macaques are hardly distinguishable from one another. Although François Jacob addressed

the random and playful character of evolution by the phrase “Nature is a tinkerer, not an inventor” [153], evolution can be somewhat conservative in the sense that similar structural and dynamical features appear in very different species throughout phylogenies. This justifies the investigation of information pathways in simpler, easier-to-understand organisms in order to comprehend higher brain functions in more complex vertebrates. A famous example is the research of Eric Kandel on the snail *Aplysia*, relating physiological signaling with behavior [185]. Studying the neural design of biological species with only a few hundred or thousands neurons is a fruitful ansatz to develop novel computing primitives [169, 191–193, 213–215]. We will come back to this issue in chapter “Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices”.

While phylogenetics addresses the development and evolution of groups of similar species, ontogenesis is the study of how an individual member of a species develops as it ages. In quotation (2), Martijn P. van den Heuvel and coworkers underline the intriguing mechanisms of nervous systems development in humans, from conception to late adolescence. We propose that ontogenesis and their functional ingredients could serve as an essential guideline for novel computing primitives. To support this argument, we describe here the fundamentals of ontogenesis in the human nervous system, including the importance of external stimuli during development. Physiology, neurobiology, and behavioral science provide overwhelming experimental evidence showing that the conditions during the growth and regeneration of neuronal nervous systems under external stimuli are of central importance [64, 216–221]. Both the formation and elimination of nerve cells, synapses, and axonal connections occur frequently during the first stages of brain development, belonging to a very creative process that shapes the nervous system to be well-adapted for future environment-related tasks. In addition to the creation of neurons and axons, pruning (programmed cell death or apoptosis) and axonal rewiring are both essential and expedient mechanisms. Finally, myelination of axons is an essential step to improve the nervous system’s performance by shaping and optimizing the signal transfer time between neurons and distributed brain areas. Gerald Edelman coined the expression “Neuronal Darwinism” to highlight the striking parallels between evolution and brain development (Edelman 1987; Tononi, Sporns, and Edelman 1994; Edelman and Tononi 2001; Van Ooyen and Butz-Ostendorf 2017) [222–225]. Neurons, synapses, and axonal connections grow lavishly at first, a growth that is controlled by the genome, epi-genome, and stochastic factors. Subsequent structural shaping and elimination, often called blooming and pruning, are largely influenced by the interaction of the entire nervous system with environmental stimuli, and the nervous system’s subsequent reaction [184, 226, 227]. There have been attempts in the past to design materials and systems that mimic biological information processing, dubbed “evolvable hardware” and “evolution-in-materio” [97, 196, 197, 228, 229]. This work has been recently extended to novel, transistor-based devices by Baek et al. [230]. Although the findings are very promising, basal spatiotemporal and topologically-relevant mechanisms have not been reproducible in electronics hardware so far. In both biological and artificial systems, the connection between these mechanisms should be worked out with regard to the required complexity and functionality (see Fig. 6).

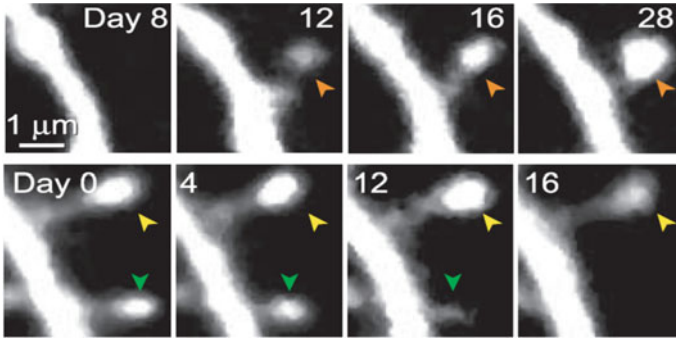


Fig. 7 Structural plasticity. In vivo time snapshots of the appearance and disappearance of dendritic spines in the mice barrel cortex. The top row shows the growth of a persistent spine between days 12 and 28 (orange arrows). The bottom row shows examples of spine retraction (yellow arrows between days 0 and 16, and green arrows between days 0 and 12). Figure from [241] with permission. These structural changes were correlated to external stimuli applied by whisker trimming in mice

Neural network growth in nervous systems has been studied in-depth both theoretically and experimentally [29, 31, 32, 212, 219, 231–234]. In particular, the early stages of nervous system growth under external stimuli is of critical importance for the healthy development of mature creatures [157, 220, 235]. It is known that both external stimuli and genetic factors have tremendous impact on the emergence of functional neural circuits that determine behavior during critical periods of cortical region growth [226, 227, 234, 236]. Cell overproduction and subsequent attrition are likewise important for nervous system development [226, 237, 238]. Morphological aspects, connectivity, growth, regeneration, and the impact of neuronal activity-related spike-based synchronization mechanisms in neuronal network ensembles serve as models for novel electronics [39, 41, 114, 165, 175, 239, 240]. Clear evidence of structural dendritic spine plasticity is shown in a series of photographs taken over a few days in Fig. 7, demonstrating that spines grow and shrink depending on external, touch-related stimuli in mice [241].

A look at a few growth parameters underlines the importance of understanding biological networks during development. A two-year-old human toddler exhibits the maximum number of neurons and synapses of our species, roughly a factor of two more than a fully grown adult. If we estimate 170 billion neurons [187, 226] with 10^3 synapses per neuron, a two-year-old human carries 170×10^{12} synapses. We assume the total axon length of a toddler to be about 850,000 km (<https://aiimpacts.org/transmitting-fibers-in-the-brain-total-length-and-distribution-of-lengths/>). The time between egg fertilization to the age of two is 1000 d or 8.64×10^7 s. This leads to an average net growth of roughly 2000 neuron/s, 2 million synaptic interconnections/s and an axon growth rate of about 10 m/s! These measures alone unambiguously demonstrate the overwhelming significance of network growth in humans, particular during childhood [221, 242]. Moreover, we believe that such a tremendous development is an interesting template for novel computing architectures. It might

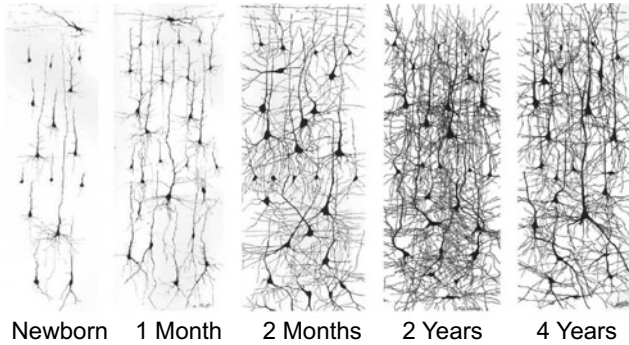


Fig. 8 Blooming and pruning of nerve cells in young humans [157] and J. Conel, “The Post-Natal Development of the Human Cerebral Cortex,” Harvard University Press, Cambridge, 1939–1967. Adapted from [157]

be an essential building block to achieve higher brain functionalities in artificial systems, and constitutes a key aspect artificial spatio-temporal networks.

Figure 8 shows several snapshots taken during human development, where the excessive growth of neurons between the ages of one month to two years is clearly visible. Interestingly, between the ages of two to four years, neuron pruning leads to reduced neuron density. While the net neuron density during adulthood is rather constant, blooming and pruning still continue to occur, albeit at a much lower rate [225].

From the postnatal phase up to the age of around two years, our central nervous system is characterized by enormous development and permanent remodeling, while being simultaneously subject to an exuberant amount of external stimuli via our senses [184, 212, 227, 232]. Genetics, stochastics, and external stimuli (in other words nature and nurture) define who we are and strongly influence higher brain function during adulthood, including perception, awareness, and consciousness.

In Fig. 9, windows of plasticity in human brain development are sketched [227, 243]. Even in much simpler creatures (e.g. the worm *C.-elegans*), external stimuli play an essential role in the healthy development of the nervous system [235].

These windows for sensing, motor skills, and higher cognition are also called critical periods. They reflect the tremendous rearrangement of the human brain during early childhood, accompanied with enormous learning capabilities. It is interesting to assign the above estimated growth parameters and the appearance of critical periods to human altriciality. Altriciality refers to the way creatures are born completely incapable of caring for themselves (Dunsworth et al. 2012). Hence, at the moment of birth (eye opening), a sudden rush of external stimuli, in particular vision, meets a premature nervous system still under heavy construction, reconstruction, and growth, in the case of humans. The concomitant occurrence of environmental input, nervous system growth, and close interaction between the nervous system and its environment may explain the huge plasticity and learning capabilities during these first

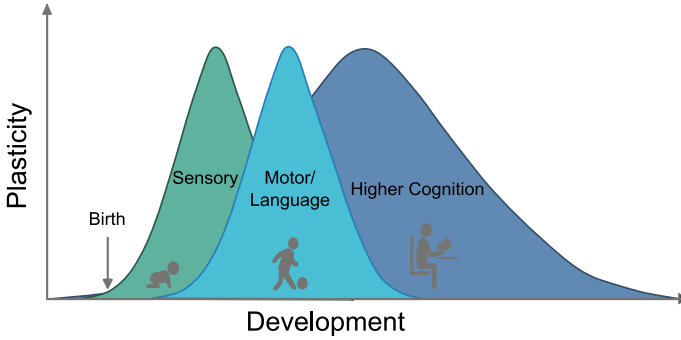


Fig. 9 Illustration of critical or sensitive periods during the first years after birth for humans. The three periods present (from left to right) the development of sensory pathways, motor skills, and higher cognitive functions. Adapted from [227, 243] with permission

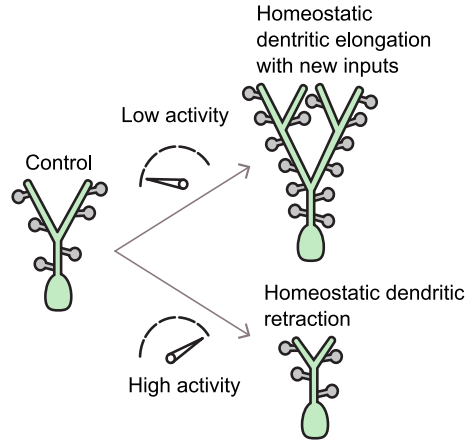
years. This development seems to be essential to form higher brain functions [212, 218]. Although it might be incredibly difficult to mimic such basal neurobiological mechanisms in engineered systems, nervous system development and growth cannot be neglected in establishing higher brain functions in artificial systems. Attempts to achieve this goal are proposed in chapter “Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices”.

4.4 Homeostasis

As in Sect. 4.3, we begin with the following sequentially-labelled neuroscience quotation: (3) Arjen van Ooyen and Markus Butz-Ostendorf emphasized the role of homeostasis on p.133 of their contribution (see: *The Functional Role of Critical Dynamics* edited by Nergis Tomen, J. Michael Herrmann, and Udo Ernst [244]: “In conclusion, during development, homeostatic structural plasticity can guide the formation of synaptic connections to create a critical network that has optimal functional properties for information processing in adulthood.” [245].

Roughly speaking, is homeostasis a kind of counteracting mechanism to network plasticity, and thus an important factor to ensure network robustness and stability. As will be discussed below in more detail, homeostasis comprises dynamical and morphological components, and is thought to explain how a nervous system stabilizes (itself) near the point of criticality [246]. In other words, homeostasis addresses the term “self” in SOC. The role of homeostasis as a stabilizing factor in neural networks is amply described in a huge number of publications, with only a few mentioned here (Abbott 2003; Turrigiano 2012; C. Tetzlaff et al. 2010; Stepp, Plenz, and Srinivasa 2015; Fauth, Wörgötter, and Tetzlaff 2017; van Ooyen 2017; Ma et al. 2019) [225, 239, 247–251]. In homeostatic structural plasticity, all incoming synapses of a cell are modified to stabilize the neuronal activity around a particular level (set point),

Fig. 10 Illustration of homeostasis in a nervous system at the neuron level. Adapted from [253, 254] with permission



and reflect a negative feedback loop between neuronal activity and connectivity [225, 248, 252, 253]. The fundamental principle of homeostasis is sketched in Fig. 10.

Higher firing (dynamic component) of a neuron results in spine deletion (morphological component), whereas reduced firing supports spine formation, keeping the average electrical activity at a set-point, potentially stabilizing the global activity of the entire neural ensemble near the desired critical state, i.e. the state with the largest dynamic range for information processing [30, 208, 209, 254]. While this model appears attractive at first glance, it raises a fundamental question in neuroscience: “How can an individual, local neuron in a huge nervous system access the global network state in order to orientate its own activity accordingly?” [27, 203], or in other words, what defines the activity set-point? This is an example of the poorly understood relation between local, mesoscopic, and global mechanisms in nervous systems.

In chapters “Redox-Based Bi-Layer Metal Oxide Memristive Devices” and “MemFlash—Floating Gate Transistors as Memristors”, we presented various basal local and global information pathways in nervous systems. In the following chapter, we will suggest a number of strategies with the goal of implementing higher brain functions in artificial systems [229].

5 Artificial Spatio-temporal Networks

At this point, an obvious and understandable question might be: Is the goal to achieve higher brain functions in artificial systems possible at all or, more precisely, to what extent can the intriguing and complex biological mechanisms described in chapters “Redox-Based Bi-Layer Metal Oxide Memristive Devices” and “MemFlash—Floating Gate Transistors as Memristors” be merged into a novel computing prim-

itive? How close is neuroscience to understanding higher brain functions and to what extent can the plethora of phenomena set by materials and engineering designs strategies enable mental functions in artificial systems?

Here we discuss possible ways and limitations of using artificial systems to mimic biological fundamentals, including topological and dynamical aspects, such as phylogenies, ontogenies, homeostasis, SOC, memory, oscillatory orchestra (synchrony), and so on. Nonetheless, we are aware that fundamental limits which may impede consciousness in engineered systems. It would be interesting, however, to identify and define those limits.

In Fig. 11, considerations set by materials science and design strategies are illustrated.

Before describing the interplay between the components sketched in Fig. 11, we should first consider a few aspects of biological information pathways which are obviously implementable by materials science and electronics, and might simplify the execution of the proposed artificial spatio-temporal network. In a human brain, the ability to access, and thus measure, the structural, topological, and dynamical states is hindered by both technological and ethical constraints [255, 256]. In contrast, artificial systems should theoretically permit access to all local and global parameters in any conceivable experimental setup. This offers a high degree of freedom in designing artificial systems. In particular, for a system growing in complexity, a designer might decide which segments should be externally controlled and which should develop via self-assembly and self-organization.

Furthermore, the time scales involved in biological information processing may actually facilitate their artificial engineering. In phylogenetic and ontogenetic development, low time scales dominate the scene. Species vary from one generation to the next, with networks growing from days to years. Additionally, nervous system dynamics are in the 100 Hz range, with low transmission velocities on the order of m/s, i.e. the speed of spikes along axons are common. As such, there is no need to build ultrafast artificial systems in order to imitate basal biological information pathway. Indeed, the deposition or synthesis of any material, e.g. nanoparticles or NWNs, is a growing materials network (Fig. 11a), and can be adjusted to low time scales. In addition, low time scales adapt well in many ways to materials transport parameters, including ionic drift, diffusion currents, and mass transport in general. Biological time scales are easily accessible by electronics, facilitating circuit design, and permitting real-time observation of spatio-temporal system development (Fig. 11b, c, d). For example, leaky-integrated-firing of a biological neuron can be technically realized by van der Pol (vdP) oscillators [115, 119, 257], compact devices based on VO_2 or NbO_x , which exhibit a negative differential resistance (NDR) I-V curve [126, 258–260], or integrated, mixed-signal circuits [120]. In general, low time scales known form biological information pathways, including external stimuli that affect them, offer an exploration space attainable by materials-related phenomena, electronics, and parameter monitoring.

How can an artificial spatio-temporal computing system, as sketched in Fig. 11, be practically realized? The goal in a bio-inspired artificial spatiotemporal network is to reach the desired topological and dynamical states simultaneously, in order to

(b) Oscillatory ensemble and external stimuli

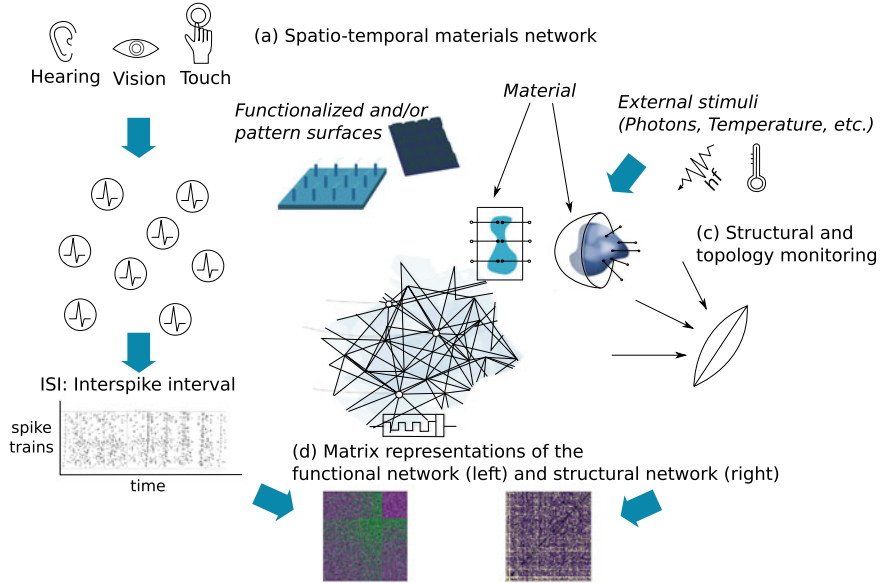


Fig. 11 Artificial spatio-temporal networks: Materials science considerations and design strategies to generate higher brain function in artificial systems. The proposed system takes basal functionalities of bio-inspired information pathways into account discussed in chapters “Redox-Based Bi-Layer Metal Oxide Memristive Devices” and “MemFlash—Floating Gate Transistors as Memristors”. **a** 2D or 3D spatio-temporal materials network. Wires within the network are connected via memristive components. The memristive functionality at cross-sections of the network implements memory and local plasticity in the network. The faded area represents a growing network. In the case of a 2D network structure formed on a planar substrate, network growth might be modified, for example, by pre-pattern substrates, a functionalized surface, additional electrical potentials, optical stimuli, and deposition-related growth parameters (materials, deposition rate, reactive gases, substrate temperature, and so on). A 3D network allows further freedom of design and allows for a nervous system-like connectivity. The 3D network could be in a solid phase, or even multiphase, network, the latter combining materials in the solid, liquid, and gas phase. **b** Representation of a pulse-oscillator ensemble in order to mimic neural spiking activity. The individual oscillators of the ensemble are electrically connected to the network, leading to modifications of the network connectivity by oscillator pulses. Conversely, the network weights in turn influence the dynamic state of the oscillator ensemble via pulse coupling. The oscillatory ensemble allows an input of external stimuli (e.g., touch, vision, and hearing) via fire rate coding. In addition, analyzing the interspike interval (ISI) distributions of the ensemble in quasi-real-time enables permanent monitoring of the dynamic state of individual oscillators, as well as the entire ensemble. **c** Stage to monitor the structure and extract the topology of the network in real time by, for example, optical microscopy, electron microscopy, thermal imaging, or magnetic field distribution detection (similar to MEG (magnetoencephalography)). **d** By monitoring the oscillatory ensemble dynamics (see **b**) and the structural connectivity (see **c**), the spatio-temporal state and its evolution can be analyzed in real-time

mimic the previously discussed characteristic hallmarks of the nervous system. This is handily illustrated in both Fig. 6a, where the topological cortical map region is labeled “G” (Goal) in the network cube, and in Fig. 6b, where the state of SOC is highlighted as the envisaged dynamical state. The main challenge here is to define the appropriate material network properties and dynamical setting for the entire system that will enable a similar spatiotemporal state to that of a nervous system. This global system state is often said to be structurally complex while being temporally close to the edge of chaos [24, 25, 171, 173, 183, 261, 262]. To achieve this goal, we describe the components presented schematically in Fig. 11 and their interactions in accordance to the biological information pathways described in chapters “Redox-Based Bi-Layer Metal Oxide Memristive Devices” and “MemFlash—Floating Gate Transistors as Memristors”. The material network template offers manifold opportunities on either a 2D or 3D platform (Fig. 11a). A network growth mimicking ontogenesis can be realized by continuous film deposition, or ongoing material synthesis of, for example, nanoparticle or nanowire networks. Network growth can be influenced in at least in three ways, the first of which being the oscillatory ensemble that is electrically connected to the network. Here, external stimuli, e.g., hearing, vision, and touch, are imprinted into the material network growth process via fire rate coding (Fig. 11b). Network formation and structure evolution are modified by the additional potential differences between the oscillator contacts within the network. Second, by integrating additional conductive pads (islands) on a 2D substrate platform, the formation of filaments between the oscillator’s electrodes can be controlled by the islands’ shape, number, size, and/or additional applied bias potential (Fig. 11a). The formation of conductive filaments during network growth could also be manipulated via structurally modulated or functionalized surfaces. In this way, not all network pathways are allowed, while others are assisted [263]. Biologically, this corresponds to axon growth and guidance [233]. This approach can also apply to 3D structures, which provides an increased degree of freedom and in principle allows nervous system-like connectivity. The materials network, whether 2D or 3D, does not necessarily have to be in the solid state: Electrolytes may be an appropriate fluid which satisfies the aforementioned requirements, including the state of criticality [21, 99, 264–269]. Third, additional stimuli (Fig. 11a) in the form of, e.g. light or temperature, can also modify the spatio-temporal evolution of the functional material network. An imprint of information during network growth is common to all three methods. This distinguishes the artificial spatio-temporal network approach from common AI systems. In the latter, the training or learning sequence is applied after system manufacture. By applying the three methods described above, it might be possible to imprint information in a similar way to that of a human nervous system during ontogenesis (method 1), as well a kind of a-priori knowledge (methods 2 and 3), i.e. phylogenetic factors.

For the entire system, a simultaneous, in-depth monitoring of network structure during its development and temporal evolution is intended, in accordance with a neuroscience approach to extract the structure and dynamics of complex brain networks [24, 270]. To this end, the spatio-temporal development of time-varying connectivity within the functional materials network (see Fig. 11a) will be monitored, for example, by means by optical microscopy, electron microscopy, or the magnetic

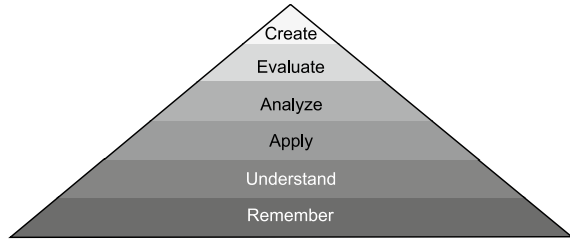
field distribution in accordance with magnetoencephalography (MEG) (see Fig. 11c). This will allow visualization of the time-evolving correlation matrix of the oscillatory ensemble, and the extraction of more theoretical metrics, such as cluster coefficients, characteristic path lengths, motifs, modularity, and hubs [24, 26, 27, 171]. In Fig. 6a, the pale blue dashed line in the cube represents a fictional pathway through the network cube. At first, we assume that the materials network is a topology state labeled “S” (Start). The position “S” within the cube is chosen as an example, but could just as well be any other topological position within the network cube. By constantly monitoring the topology of the system during network growth and intervention via a set of parameters (e.g. added materials, extra potentials, and external stimuli to the oscillatory ensemble), it might be possible to adjust the system to arrive at “G”, defined by a set of characteristic theoretical parameter (hubs, motifs, modularity, cluster coefficient, path length, etc.) [27].

Simultaneously, the ISI distribution and time series of the oscillatory ensemble will be recorded (see Fig. 11b) [271]. Spike train distances provide a means of quantifying neuronal variability and the degree of synchrony in and between oscillatory ensembles, and may indicate the rise of oscillatory avalanche firing as one indicator of the SOC [23, 272–277]. SOC is described as a state located somewhere between the random, independent firing of individual oscillators, and complete synchrony, where all oscillators fire in phase with the same frequency [265, 278]. Between these two extremes, a system’s dynamics can be characterized by a complex and time-varying interaction of the oscillatory ensemble (see Fig. 6b). This regime exhibits features of criticality typically observed close to phase transitions [20, 25, 103, 209, 279–281]. In particular, the topology and temporal dynamics of a system in such a state are extremely sensitive to external distortions (stimuli) and may respond to them in numerous ways.

Practically speaking, we will begin by analyzing coupled nonlinear oscillator network raster plots, phase portraits, phase response curves, bifurcation diagrams, spike distance measurements, and cross-correlation type time-series analysis. Information from these analyses will be subsequently applied to quantify the phase and frequency relationships between network oscillators and their development over time [282, 283]. Finally, we would like to discuss obvious obstacles and challenges. In Sect. 4.4, the rule of homeostasis was highlighted. The concept of homeostasis is of essential importance to stabilize the nervous system dynamics and morphology to a set-point. For the system presented in Fig. 11, homeostasis is not illustrated. It might be possible to reconstruct a feedback parameter from the structural and functional matrices to reduce or enhance, if necessary, the oscillatory activity, or to modify the material growth process. Another challenge might be the implementation of appropriate delay lines to mimic the important signal retardation known from nervous systems [173, 182]. Ionic conductors with slow ionic motion in the form of drift or diffusion currents could be a possible solution.

One important issue remains: Picture a fabricated artificial spatio-temporal system as depicted in Fig. 11, that presents all previously discussed biological information pathways. How can we benchmark the system, and determine how it solves tasks set by external stimuli? Certainly, the functional and structural network states reflect

Fig. 12 Suggested benchmark for bio-inspired systems based on Bloom's taxonomy. The pyramid represents increasing cognitive human skills from bottom to top. Figure 2 from Ref. [287]



the overall system state. As such, one viable approach is to read out the system state and to activate a set of artificial motor neurons to react to an input task. However, this does not accurately represent the process in the human brain, where there is no internal, global system observer to decide on the next step [9]. At this point, we are confronted with a difficult challenge: how can we lead matter to imitate the mind? While the authors can suggest an example system as shown in Fig. 11, this question remains open.

6 Benchmarking for Bio-inspired Computing

Benchmarking in AI is an important approach to measure its performance, and subsequently enable comparisons between different systems. In pattern recognition, for example, MNIS data sets are used, with the recognition rate defining a clear benchmark. While contemporary AI systems show extraordinary capability in performing a single, specific task, their success at task variability is highly limited compared to the nervous system. Nonetheless, a new generation of AI has demonstrated extraordinary capabilities in the field of gaming (Chess and Go), including an aptitude for self-learning [284]. Yet, it remains unclear how to define a fair and comprehensible benchmarking for neuromorphic systems and bio-inspired computing [285]. Computational tasks must be carefully designed in order to assess the overall system's performance in comparison with human mental capabilities, as previously proposed by Alan Turing in his seminal work on Machinery and Intelligence [155]. Bloom's learning taxonomy, which was developed to hierarchically categorize learning in the classroom, can be helpful in assessing how successfully artificial systems mimic higher brain functions [286]. This taxonomy contains six categories of cognitive skills and presents a hierarchy with increasing cognitive functionality from bottom (factual knowledge) to top (creation) (see Fig. 12), or in other words, from lower-order skills that require less cognitive processing to higher-order skills that require deeper learning and a greater degree of cognitive processing [287].

This strategy may serve as a basis for benchmarking in bio-inspired computing systems. However, due to the nervous system's task variability for each of the six cognitive categories, transparent benchmarks must be developed. This goal is extremely important for future comparisons of bio-inspired systems, which are currently devel-

oped on different platforms. In addition, resource-related parameters, such as energy consumption, system weight, and failure tolerance, need to be included.

7 Discussion

This perspective introduces the concept of artificial spatio-temporal networks, which proposes basal hallmarks, such as morphological and dynamical characteristics of nervous systems, to reproduce higher brain functions in artificial systems. In particular, the basal mechanisms known from the growth of nervous systems might play a significant role in their function. This concept will undoubtedly be a way to include biologically-relevant features in future artificial systems. Yet, only the tip of the iceberg has thus far been addressed: to fully realize an artificial spatio-temporal network, several challenges remain unresolved.

In more general terms, artificial spatio-temporal networks again raise the fundamental question: “To what extent can higher brain functions be reproduced in artificial systems?” Seminal books and papers by [1, 3, 5, 7, 176, 234, 288], and many more address this topic in one way or another. According to the authors, higher brain function can be described on the basis of natural sciences and mathematics, permitting us to view this challenge in another light. On an atomistic level, we find in living nature, and therefore in any nervous system, old and well known friends from the periodic table of the elements, including but not limited to Carbon (C), Sodium (Na), Potassium (K), Chlorine (Cl), Oxygen (O), and Hydrogen (H). Any effort to establish higher brain function in an artificial system, whether in silico (software oriented) or in a material-based substrate, as in the case of artificial spatio-temporal networks, should apply another tool box of elements to establish awareness, perception, or consciousness, e.g. Silicon (Si), Gold (Au), Silver (Ag), Tungsten (W), O, and so on. There is no obvious reason why this strategy should not work, but if it cannot, what are the fundamental limits, and how are they defined? A look at biochemical substrates in living species highlights the weaknesses of the simplistic, atomistic view point. There is still unknown genetic information that strongly controls nervous system behavior and function, especially during development, which therefore cannot currently be considered in any artificially-constructed systems. Whether there are shortcuts to bypass the role genes play in neural behavior and development is completely unknown, and might act as a show stopper [234]. On the one hand, it is truly challenging to introduce basal biological functionalities, such as homeostasis, signal delay, growth, and the appropriate states of criticality and topology in an artificial system. On the other hand, the materials tool box may offer plethora of phenomena which have not yet been explored for novel computing architectures [66, 266–269]. Hence, these simple questions and views point towards an even more fundamental aspect: in living systems, the separation between matter and information becomes blurred, making it risky to apply these terms without investigating living and artificial systems equally, or precisely clarifying the respective context [289].

8 Conclusion

In this perspective, we addressed fundamental limits of current ICT and briefly summarized the state-of-the-art. Today's digital electronics work with clock rates in the GHz range, utilizing ns pulses and signal transmissions at nearly light speed in a vacuum. Meanwhile, nervous systems exhibit numerous remarkable and fascinating features, including anticipation, awareness, perception, and consciousness. The associated action potential spikes are 6 orders of magnitude longer, and travel with a velocity 6 orders of magnitude lower, than their electronic analogs, while dissipating only a couple of Watts of power. We touched on the fundamentals of information processing in biological (nervous) and engineered systems. Specifically, we highlighted the dynamical and morphological properties exhibited by nervous systems using the human brain as an example. The exceptional topology of the human cortex in comparison to other biological and technical networks, in addition to the state of SOC, served as guidelines to develop artificial spatio-temporal systems. A pathway to realize artificial spatio-temporal systems in a hardware-orientated system was presented, aiming to emulate higher brain functions in an artificial system. The role of ontogenesis was discussed, revealing that the mechanism of neural network growth provides crucial information useful in designing novel artificial computing systems, which have yet to be addressed in great detail.

Neural network growth illustrates how important the ongoing interaction between the internal and external world is when artificially creating the basic structures that provide the ability to learn specific functions. In our opinion, this emphasizes the importance of basal properties which, while beginning to be applied systems in artificial, have yet to be fully implemented. These properties include individual autonomous dynamic units, time-variable coupling between them, and both positive and negative connection growth. With respect to time-variability, the research field has shown enormous progress in recent years with the development of memristive systems. Although memristive devices can already replicate the phenomena associated with learning to a certain degree, the question remains whether these devices can suitably reproduce both the necessary processes in their entirety, and global dynamics which are shaped by an overwhelmingly complex network. The last point in particular presents immense challenges for a conservative implementation of memristive devices in large-scale systems. Finally, we discussed possible limitations in implementing higher brain functions in artificial systems. We concluded that genetic information plays a key role in the development of neural nervous systems, knowledge that we are still lacking if we want to fully implement this behavior in artificial systems, specifically with regards to awareness, perception, and consciousness. The exploration space for implementation is certainly extraordinary large for artificial spatio-temporal systems. This huge parameter space is both curse and blessing: while such a large number of variables must be monitored and controlled, it also allows for greater flexibility and opportunities. One thing is certain in this context: no matter which engineered solution ultimately prevails, humanity will be confronted with a multitude of ambivalent questions and challenges, in which certainly "Matter and Mind Matter".

Acknowledgements We thank Nora Kohlstedt for preparing part of the figures. “Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 434434223—SFB 1461”. The project is entitled SFB 1461 “Neurotronics: Bio-Inspired Information Pathways” (see for more details <https://www.crc1461-neurotronics.de/index.php/en/>). Moreover, the financial support by the DFG via the Research Unit 2093: “Memristive Devices for neural Systems” is acknowledged. We thank Gitanjali Kolhatkar and Shane Scott for carefully reading the manuscript. We thank Nora Kohlstedt for preparing Fig. 4., i.e. the sketch of a neuron.

References

1. Churchland, P.M., Churchland, P.: Could a machine think? (1990)
2. Churchland, P.M.: Densmore and dennett on virtul machines and consciousness. *Philos. Phenomenol. Res.* **59**(3), 763–767 (1999)
3. Aleksander, I.: How to build a mind: toward machines with imagination. *Maps of the Mind*. Columbia University Press, New York (2001)
4. Hawkins, J., Blakeslee, S.: *On Intelligence*. Times Books, New York (2004)
5. Koch, C., Tononi, G.: Can maschines be consciosus? *IEEE Spectr.* **45**, 55–59 (2008)
6. Dehaene, S.: *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking, New York (2014)
7. Dehaene, S., Lau, H., Kouider, S.: What is consciousness, and could machines have it? p. 8 (2017)
8. Ostenfeld, E.N.: *Ancient Greek Psychology and the Modern Mind-Body Debate*. Number Bd. 63 in *Academia Philosophical Studies*, 2nd edn. Academia Verlag, Baden-Baden (2018)
9. Damasio, A.R.: *Descartes’ Error: Emotion, Reason and the Human Brain*, 18, druck Quill, New York (2004)
10. Varela, F., Lachaux, J.-P., Rodriguez, E., Martinerie, J.: The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* **2**(4), 229–239 (2001)
11. Noirhomme, Q., Laureys, S.: Consciousness and unconsciousness: an EEG perspective. *Clin. EEG Neurosci.* **45**(1), 4–5 (2014)
12. Kriegeskorte, N., Douglas, P.K.: Cognitive computational neuroscience. *Nat. Neurosci.* **21**(9), 1148–1160 (2018)
13. Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., Martial, C., Fernández-Espejo, D., Rohaut, B., Voss, H.U., Schiff, N.D., Owen, A.M., Laureys, S., Naccache, L., Sitt, J.D.: Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv.* **5**(2), eaat7603 (2019)
14. Storm, J.F., Boly, M., Casali, A.G., Massimini, M., Olcese, U., Pennartz, C.M.A., Wilke, M.: Consciousness regained: disentangling mechanisms, brain systems, and behavioral responses. *J. Neurosci.* **37**(45), 10882–10893 (2017)
15. Dehaene, S. (ed.): *The Cognitive Neuroscience of Consciousness*. *Cognition Special Issues*. MIT Press, Cambridge (2001)
16. Tagliazucchi, E., Chialvo, D.R., Siniatchkin, M., Amico, E., Brichant, J.F., Bonhomme, V., Noirhomme, Q., Laufs, H., Laureys, S.: Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. *J. R. Soc. Interface* **13**(114), 20151027 (2016)
17. Perry, E., Walker, M., Grace, J., Perry, R.: Acetylcholine in mind: a neurotransmitter correlate of consciousness? *Trends Neurosci.* **22**(6), 273–280 (1999)
18. Xia, Q., Yang, J.J.: Memristive crossbar arrays for brain-inspired computing. *Nat. Mater.* **18**(4), 309–323 (2019)

19. Bian, H., Goh, Y.Y., Liu, Y., Ling, H., Xie, L., Liu, X.: Stimuli-responsive memristive materials for artificial synapses and neuromorphic computing. *Adv. Mater.* **33**(46), 2006469 (2021)
20. Mallinson, J.B., Shirai, S., Acharya, S.K., Bose, S.K., Galli, E., Brown, S.A.: Avalanches and criticality in self-organized nanoscale networks. *Sci. Adv.* **5**(11), eaaw8438 (2019)
21. Robin, P., Kavokine, N., Bocquet, L.: Modeling of emergent memory and voltage spiking in ionic transport through angstrom-scale slits. *Science* **373**(6555), 687–691 (2021)
22. Kuncic, Z., Nakayama, T.: Neuromorphic nanowire networks: principles, progress and future prospects for neuro-inspired information processing. *Adv. Phys.: X* **6**(1), 1894234 (2021)
23. Beggs, J.M., Plenz, D.: Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**(35), 11167–11177 (2003)
24. Bullmore, Ed, Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**(3), 186–198 (2009)
25. Chialvo, D.R.: Emergent complex neural dynamics. *Nat. Phys.* **6**(10), 744–750 (2010)
26. Sporns, O.: *Networks of the Brain*. MIT Press, Cambridge (2011)
27. Fornito, A., Zalesky, A., Bullmore, E.T.: *Fundamentals of Brain Network Analysis*. Elsevier/Academic Press, Amsterdam; Boston (2016)
28. Bassett, D.S., Sporns, O.: Network neuroscience. *Nat. Neurosci.* **20**(3), 353–364 (2017)
29. Kaiser, M., Hilgetag, C.C.: Spatial growth of real-world networks. *Phys. Rev. E* **69**(3), 036103 (2004)
30. Kinouchi, O., Copelli, M.: Optimal dynamical range of excitable networks at criticality. *Nat. Phys.* **2**(5), 348–351 (2006)
31. Kaiser, M.: Hierarchy and dynamics of neural networks. *Front. Neuroinf.* **4** (2010)
32. Hütt, M.-T., Kaiser, M., Hilgetag, C.C.: Perspective: network-guided pattern formation of neural dynamics. *Philos. Trans. R. Soc. B: Biol. Sci.* **369**(1653), 20130522 (2014)
33. Kaiser, M.: Mechanisms of connectome development. *Trends Cogn. Sci.* **21**(9), 703–717 (2017)
34. Agi, E., Kulkarni, A., Hiesinger, P.R.: Neuronal strategies for meeting the right partner during brain wiring. *Cell. Neurosci.* **63**, 1–8 (2020)
35. Lvrtrup, S.: Phylogenesis, ontogenesis and evolution. *Bolletino di zoologia* **54**(3), 199–208 (1987)
36. Torday, J.: Homeostasis as the mechanism of evolution. *Biology* **4**(3), 573–590 (2015)
37. Singer, W., Gray, C.M.: Visual feature integration and the temporal correlation hypothesis. *Ann. Rev. Neurosci.* **18**(1), 555–586 (1995)
38. Engel, A.K., Fries, P., Singer, W.: Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**(10), 704–716 (2001)
39. Uhlhaas, P.J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolić, D., Singer, W.: Neural synchrony in cortical networks: history, concept and current status. *Front. Integr. Neurosci.* **3**, 17–17 (2009)
40. Sheffield, M.E.J., Dombeck, D.A.: The binding solution? *Nat. Neurosci.* **18**(8), 1060–1062 (2015)
41. Buzsáki, G.: *Rhythms of the Brain*. Oxford University Press, Oxford; New York (2006)
42. Schechter, B.: How the brain gets rhythm: distinctive neural oscillations may link separate brain regions that are responding to the same object. Researchers are now identifying the sources of these vibrations. *Science* **274**(5286), 339–339 (1996)
43. Berlin, L.: *The man behind the microchip: Robert Noyce and the invention of Silicon Valley*. Oxford University Press, Oxford; New York (2005)
44. Moore, G.E.: *Cramming More Components onto Integrated Circuits*. McGraw-Hill, New York, NY, USA (1965)
45. Turing, A.M.: On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lon. Math. Soc.* **s2-42**(1), 230–265 (1937)
46. von Neumann, J.: First draft of a report on the EDVAC. *IEEE Ann. Hist. Comput.* **15**(4), 27–75 (1993)
47. Veendrick, H.J.M.: *Nanometer CMOS ICs: From Basics to ASICs*, 2nd edn. Springer International Publishing, Imprint, Springer, Cham (2017)

48. Masuhara, T.: The future of low-power electronics. In: Höfflinger, B. (ed.) CHIPS 2020, vol. 2, pp. 21–50. Springer International Publishing, Cham (2016)
49. Kurinec, S.K., Walia, S.: Energy Efficient Computing & Electronics: Devices to Systems. Circuits, and Systems. CRC Press, Devices (2019)
50. Dennard, R.H., Gaensslen, F.H., Yu, H.-N., Rideout, V.L., Bassous, E., LeBlanc, A.R.: Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE J. Solid-State Circ.* **9**(5), 256–268 (1974)
51. Höfflinger, B.: (Ed.) New Vistas in Nanoelectronics. In: Hoefflinger, N.B. (ed.) CHIPS 2020, vol. 2. Springer, Cham, Heidelberg, New York, Dordrecht, London (2016)
52. Radamson, H.H., Zhu, H., Zhenhua, W., He, X., Lin, H., Liu, J., Xiang, J., Kong, Z., Xiong, W., Li, J., Cui, H., Gao, J., Yang, H., Yong, D., Buqing, X., Li, B., Zhao, X., Jiahan, Yu., Dong, Y., Wang, G.: State of the art and future perspectives in advanced CMOS technology. *Nanomaterials* **10**(8), 1555 (2020)
53. Backus J (1978) Can Programming Be Liberated from the von Neumann Style? A Functional Style and Its Algebra of Programs
54. Iniewski, K. (ed.): CMOS processors and memories. Analog Circuits and Signal Processing - ACSP. Springer, Dordrecht (2010)
55. Jones, N.: How to stop data centres from gobbling up the world's electricity. *Nature* **561**(7722), 163–166 (2018)
56. Yang, J.J., Strukov, D.B., Stewart, D.R.: Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13 (2012)
57. Schuman, C., Potok, T., Patton, R., Birdwell, J., Dean, M., Rose, G., Plank, J.: A Survey of Neuromorphic Computing and Neural Networks in Hardware (2017)
58. Burr, G.W., Shelby, R.M., Sebastian, A., Kim, S., Kim, S., Sidler, S., Virwani, K., Ishii, M., Narayanan, P., Fumarola, A., Sanches, L.L., Boybat, I., Le Gallo, M., Moon, K., Woo, J., Hwang, H., Leblebici, Y.: Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2**(1), 89–124 (2017)
59. Merolla, P.A., Arthur, J.V., Alvarez-Icaza, R., Cassidy, A.S., Sawada, J., Akopyan, F., Jackson, B.L., Imam, N., Guo, C., Nakamura, Y., Brezzo, B., Vo, I., Esser, S.K., Appuswamy, R., Taba, B., Amir, A., Flickner, M.D., Risk, W.P., Manohar, R., Modha, D.S.: A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**(6197), 668 (2014)
60. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7**(1), 011305 (2020)
61. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
62. Mermin, N.D.: Quantum Computer Science: An Introduction. Cambridge University Press, Cambridge (2007)
63. Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J.C., Barends, R., Biswas, R., Boixo, S., Brandao, F.G.S.L., Buell, D.A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., Fowler, A., Gidney, C., Giustina, M., Graff, R., Guerin, K., Habegger, S., Harrigan, M.P., Hartmann, M.J., Ho, A., Hoffmann, M., Huang, T., Humble, T.S., Isakov, S.V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Kelly, J., Klimov, P.V., Knysh, S., Korotkov, A., Kostritsa, F., Landhuis, D., Lindmark, M., Lucero, E., Lyakh, D., Mandrà, S., McClean, J.R., McEwen, M., Megrant, A., Mi, X., Michielsen, K., Mohseni, M., Mutus, J., Naaman, O., Neeley, M., Neill, C., Niu, M.Y., Ostby, E., Petukhov, A., Platt, J.C., Quintana, C., Rieffel, E.G., Roushan, P., Rubin, N.C., Sank, D., Satzinger, K.J., Smelyanskiy, V., Sung, K.J., Trevithick, M.D., Vainsencher, A., Villalonga, B., White, T., Yao, Z.H., Yeh, P., Zalcman, A., Neven, H., Martinis, J.M.: Quantum supremacy using a programmable superconducting processor. *Nature* **574**(7779), 505–510 (2019)
64. Gauthier, D.J., Bollt, E., Griffith, A., Barbosa, W.A.S.: Next generation reservoir computing. *Nat. Commun.* **12**(1), 5564 (2021)
65. Ielmini, D., Wong, H.-S.P.: In-memory computing with resistive switching devices. *Nat. Electron.* **1**(6), 333–343 (2018)

66. Kaspar, C., Ravoo, B.J., van der Wiel, W.G., Wegner, S.V., Pernice, W.H.P.: The rise of intelligent matter. *Nature* **594**(7863), 345–355 (2021)
67. Hinton, G., Deng, L., Dong, Yu., Dahl, G., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T., Kingsbury, B.: Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process. Mag.* **29**(6), 82–97 (2012)
68. McCulloch, W.S., Pitts, W.: A logical calculus of the ideas immanent in nervous activity, p. 17
69. Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**(6), 386–408 (1958)
70. Jäger, H.: The “echo state” approach to analysing and training recurrent neural networks. Report Corpus ID: 15467150, Bonn (2001)
71. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
72. Mead, C.: Analog VLSI and neural systems. *Computation and Neural Systems Series*. Addison-Wesley, Reading (1989)
73. Mahowald, M., Douglas, R.: A silicon neuron. *Nature* **354**, 515 (1991)
74. Mead, C.: How we created neuromorphic engineering. *Nat. Electron.* **3**(7), 434–435 (2020)
75. Indiveri, G., Linares-Barranco, B., Hamilton, T., van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.-C., Dudek, P., Häfziger, P., Renaud, S., Schemmel, J., Cauwenberghs, G., Arthur, J., Hynna, K., Folorosele, F., SAA–GHI, S., Serrano-Gotarredona, T., Wijekoon, J., Wang, Y., Boahen, K.: Neuromorphic silicon neuron circuits. *Front. Neurosci.* **5**, 73 (2011)
76. Pei, J., Deng, L., Song, S., Zhao, M., Zhang, Y., Shuang, W., Wang, G., Zou, Z., Zhenzhi, W., He, W., Chen, F., Deng, N., Si, W., Wang, Yu., Yujie, W., Yang, Z., Ma, C., Li, G., Han, W., Li, H., Huaqiang, W., Zhao, R., Xie, Y., Shi, L.: Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature* **572**(7767), 106–111 (2019)
77. Frenkel, C., Bol, D., Indiveri, G.: Bottom-up and top-down neural processing systems design: neuromorphic intelligence as the convergence of natural and artificial intelligence. [arXiv:2106.01288](https://arxiv.org/abs/2106.01288) [cs] (2021)
78. Karbalaee, M., Dideban, D., Heidari, H.: A sectorial scheme of gate-all-around field effect transistor with improved electrical characteristics. *Ain Shams Eng. J.* **12**(1), 755–760 (2021)
79. Sengupta, A., Panda, P., Wijesinghe, P., Kim, Y., Roy, K.: Magnetic tunnel junction mimics stochastic cortical spiking neurons. *Sci. Rep.* **6**(1) (2016)
80. Zhang, H.-T., Panda, P., Lin, J., Kalcheim, Y., Wang, K., Freeland, J.W., Fong, D.D., Priya, S., Schuller, I.K., Sankaranarayanan, S.K.R.S., Roy, K., Ramanathan, S.: Organismic materials for beyond von Neumann machines. *Appl. Phys. Rev.* **7**(1), 011309 (2020)
81. Minnai, C., Bellacicca, A., Brown, S.A., Milani, P.: Facile fabrication of complex networks of memristive devices. *Sci. Rep.* **7**(1), 7955 (2017)
82. Lequeux, S., Sampaio, J., Cros, V., Yakushiji, K., Fukushima, A., Matsumoto, R., Kubota, H., Yuasa, S., Grollier, J.: A magnetic synapse: multilevel spin-torque memristor with perpendicular anisotropy. *Sci. Rep.* **6**(1), 31510 (2016)
83. Sangwan, V.K., Hersam, M.C.: Neuromorphic nanoelectronic materials. *Nat. Nanotechnol.* **15**(7), 517–528 (2020)
84. Sung, C., Hwang, H., Yoo, I.K.: Perspective: a review on memristive hardware for neuromorphic computation. *J. Appl. Phys.* **124**(15), 151903 (2018)
85. Hickmott, T.W.: Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **33**(9), 2669–2682 (1962)
86. Argall, F.: Switching phenomena in titanium oxide thin films. *Solid-State Electron.* **11**(5), 535–541 (1968)
87. Dearnaley, G., Stoneham, A.M., Morgan, D.V.: Electrical phenomena in amorphous oxide films. *Rep. Prog. Phys.* **33**(3), 1129–1191 (1970)
88. Chua, L.: Memristor-the missing circuit element. *IEEE Trans. Circuit Theory* **18**(5), 507–519 (1971)

89. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. *Nature* **453**(7191), 80–83 (2008)
90. Tetzlaff, R. (ed.): *Memristors and Memristive Systems*. Springer, New York (2014)
91. Ielmini, D., Waser, R. (eds.): *Resistive switching: from fundamentals of nanoionic redox processes to memristive device applications*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim (2016)
92. Li, H., Wang, S., Zhang, X., Wang, W., Yang, R., Sun, Z., Feng, W., Lin, P., Wang, Z., Sun, L., Yao, Y.: Memristive crossbar arrays for storage and computing applications. *Adv. Intell. Syst.* **3**(9), 2100017 (2021)
93. Pease, R.F.: To charge or not to charge: 50 years of lithographic choices. *J. Vac. Sci. & Technol. B Nanotechnol. Microelectron.: Mater. Proc. Meas. Phenom.* **28**(6), C6A1–C6A6 (2010)
94. Donnelly, V.M., Kornblit, A.: Plasma etching: yesterday, today, and tomorrow. *J. Vacuum Sci. & Technol. A: Vacuum Surf. Films* **31**(5), 050825 (2013)
95. Oluwatosin Abegunde, O., Titilayo Akinlabi, E., Philip Oladipo, O., Akinlabi, S., Ude, U.: Overview of thin film deposition techniques. *AIMS Mater. Sci.* **6**(2), 174–199 (2019)
96. Kronholz, S., Rathgeber, S., Karthäuser, S., Kohlstedt, H., Clemens, S., Schneller, T.: Self-assembly of diblock-copolymer micelles for template-based preparation of PbTiO₃ nanograins. *Adv. Func. Mater.* **16**(18), 2346–2354 (2006)
97. Stieg, A.Z., Avizienis, A.V., Sillin, H.O., Martin-Olmos, C., Aono, M., Gimzewski, J.K.: Emergent criticality in complex turing B-type atomic switch networks. *Adv. Mater.* **24**(2), 286–293 (2012)
98. Asayesh-Ardakani, H., Nie, A., Marley, P.M., Stabile, A., Sarkar, K., Banerjee, S., Ganapathy, S., Yang, Z., Klie, R.F., Shahbazian-Yassar, R.: Atomic resolution studies of metal-insulator transition in VO₂ nanowires. *Microsc. Microanal.* **19**(S2), 492–493 (2013)
99. Pantone, R.D., Kendall, J.D., Nino, J.C.: Memristive nanowires exhibit small-world connectivity. *Neural Netw.* **106**, 144–151 (2018)
100. Hochstetter, J., Zhu, R., Loeffler, A., Diaz-Alvarez, A., Nakayama, T., Kuncic, Z.: Avalanches and edge-of-chaos learning in neuromorphic nanowire networks. *Nat. Commun.* **12**(1), 4008 (2021)
101. Zhu, R., Hochstetter, J., Loeffler, A., Diaz-Alvarez, A., Nakayama, T., Lizier, J.T., Kuncic, Z.: Information dynamics in neuromorphic nanowire networks. *Sci. Rep.* **11**(1), 13047 (2021)
102. Loeffler, A., Zhu, R., Hochstetter, J., Li, M., Kaiwei, F., Diaz-Alvarez, A., Nakayama, T., Shine, J.M., Kuncic, Z.: Topological properties of neuromorphic nanowire networks. *Front. Neurosci.* **14**, 184 (2020)
103. Pike, M.D., Bose, S.K., Mallinson, J.B., Acharya, S.K., Shirai, S., Galli, E., Weddell, S.J., Bones, P.J., Arnold, M.D., Brown, S.A.: Atomic scale dynamics drive brain-like avalanches in percolating nanostructured networks. *Nano Lett.* **20**(5), 3935–3942 (2020)
104. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature* **393**(6684), 440–442 (1998)
105. Serb, A., Bill, J., Khat, A., Berdan, R., Legenstein, R., Prodromakis, T.: Unsupervised learning in probabilistic neural networks with multi-state metal-oxide memristive synapses. *Nat. Commun.* **7**(1), 12611 (2016)
106. Baatar, C., Porod, W., Roska, T.: *Cellular Nanoscale Ensory Wave Computing*. Springer, New York (2010)
107. Kari, S.R.: *Principles of Stochastic Computing: Fundamental Concepts and Applications*, p. 11
108. Patterson, D., Anderson, T., Cardwell, N., Fromm, R., Keeton, K., Kozyrakis, C., Thomas, R., Yelick, K.: A case for intelligent RAM. *IEEE Micro* **17**(2), 34–44 (1997)
109. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R., Eleftheriou, E.: Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**(7), 529–544 (2020)
110. Wang, X., Zidan, M.A., Wei, D.L.: A crossbar-based in-memory computing architecture. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**(12), 4224–4232 (2020)
111. Maass, W., Natschläger, T., Markram, H.: Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14**(11), 2531–2560 (2002)

112. Nakajima, K.: Physical reservoir computing-an introductory perspective. *Jpn. J. Appl. Phys.* **59**(6), 060501 (2020)
113. Milano, G., Pedretti, G., Montano, K., Ricci, S., Hashemkhani, S., Boarino, L., Ielmini, D., Ricciardi, C.: In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. *Nat. Mater.* **21**(2), 195–202 (2022)
114. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., Zhou, C.: Synchronization in complex networks. *Phys. Rep.* **469**(3), 93–153 (2008)
115. Pikovskij, A., Rosenblum, M., Kurths, J.: Synchronization: A Universal Concept in Nonlinear Sciences. Number 12 in Cambridge Nonlinear Science Series. Cambridge University Press, Cambridge, 1st paperback ed., repr edition (2003)
116. Strogatz, S.H.: Exploring complex networks. *Nature* **410**(6825), 268–276 (2001)
117. Schneider, M.L., Donnelly, C.A., Russek, S.E., Baek, B., Pufall, M.R., Hopkins, P.F., Dresselhaus, P.D., Benz, S.P., Rippard, W.H.: Ultralow power artificial synapses using nanotextured magnetic Josephson junctions. *Sci. Adv.* **4**(1) (2018)
118. Chen, J.-R., Smith, A., Montoya, E.A., Lu, J.G., Krivorotov, I.N.: Spin-orbit torque nanoscillator with giant magnetoresistance readout. *Commun. Phys.* **3**(1), 187 (2020)
119. Ignatov, M., Hansen, M., Ziegler, M., Kohlstedt, H.: Synchronization of two memristively coupled van der Pol oscillators. *Appl. Phys. Lett.* **108**(8), 084105 (2016)
120. Cheng, X., Birkoben, T., Kohlstedt, H., Bahr, A.: A CMOS integrated low-power, ultra-low-frequency relaxation oscillator for neuromorphic applications. In: 2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 170–174 (2021)
121. Feldmann, J., Youngblood, N., Wright, C.D., Bhaskaran, H., Pernice, W.H.P.: All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**(7755), 208–214 (2019)
122. Lenk, C., Seeber, L., Ziegler, M.: Tuning acoustic sensing properties of mems cantilever by nonlinear operation. In: Mikro-Nano-Integration; 8th GMM-Workshop, pp. 1–3. VDE (2020)
123. Kantner, M., Schöll, E., Yanchuk, S.: Delay-induced patterns in a two-dimensional lattice of coupled oscillators. *Sci. Rep.* **5**, 8522 (2015)
124. Hölzel, R.W., Krischer, K.: Pattern recognition with simple oscillating circuits. *New J. Phys.* **13**(7), 073031 (2011)
125. Shamsi, J., Avedillo, M.J., Linares-Barranco, B., Serrano-Gotarredona, T.: Hardware implementation of differential oscillatory neural networks using VO 2-based oscillators and memristor-bridge circuits. *Front. Neurosci.* **15**, 674567 (2021)
126. Lee, D., Cha, E., Park, J., Sung, C., Moon, K., Chekol, S.A., Hwang, H.: NbO₂-based frequency storable coupled oscillators for associative memory application. *IEEE J. Electron Devices Soc.* **6**, 250–253 (2018)
127. Hasler, J., Marr, H.: Finding a roadmap to achieve large neuromorphic hardware systems. *Front. Neurosci.* **7**, 118 (2013)
128. Mange, D., Stauffer, A., Petraglio, E., Tempesti, G.: Self-replicating loop with universal construction. *Phys. D* **191**(1–2), 178–192 (2004)
129. Alaghi, A., Hayes, J.P.: Survey of stochastic computing. *ACM Trans. Embed. Comput. Syst.* **12**(2s), 1–19 (2013)
130. Parr, T., Rees, G., Friston, K.J.: Computational neuropsychology and Bayesian inference. *Front. Hum. Neurosci.* **12**, 61 (2018)
131. Venkatesan, T., Williams, S.: Brain inspired electronics. *Appl. Phys. Rev.* **9**(1), 010401 (2022)
132. Colinge, J.-P. (ed.): FinFETs and Other Multi-gate Transistors. Series on Integrated Circuits and Systems. Springer, New York (2008)
133. Sun, K., Chen, J., Yan, X.: The future of memristors: materials engineering and neural networks. *Adv. Func. Mater.* **31**(8), 2006773 (2021)
134. Lanza, M., Waser, R., Ielmini, D., Yang, J.J., Goux, L., Suñe, J., Kenyon, A.J., Mehonic, A., Spiga, S., Rana, V., Wiefels, S., Menzel, S., Valov, I., Villena, M.A., Miranda, E., Jing, X., Campabadal, F., Gonzalez, M.B., Aguirre, F., Palumbo, F., Zhu, K., Roldan, J.B., Puglisi, F.M., Larcher, L., Hou, T.-H., Prodromakis, T., Yang, Y., Huang, P., Wan, T., Chai, Y., Pey, K.L., Raghavan, N., Dueñas, S., Wang, T., Xia, Q., Pazos, S.: Standards for the characterization of endurance in resistive switching devices. *ACS Nano* **15**(11), 17214–17231 (2021)

135. Wang, R., Yang, J.-Q., Mao, J.-Y., Wang, Z.-P., Shuang, W., Zhou, M., Chen, T., Zhou, Y., Han, S.-T.: Recent advances of volatile memristors: devices, mechanisms, and applications. *Adv. Intell. Syst.* **2**(9), 2000055 (2020)
136. Banerjee, W.: Challenges and applications of emerging nonvolatile memory devices. *Electronics* **9**(6) (2020)
137. Zamarreño-Ramos, C., Camuñas-Mesa, L.A., Pérez-Carrasco, J.A., Masquelier, T., Serrano-Gotarredona, T., Linares-Barranco, B.: On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci.* **5** (2011)
138. Ranjan, R., Ponce, P.M., Hellweg, W.L., Kyrmanidis, A., Saleh, L.A., Schroeder, D., Krautschneider, W.H.: Integrated circuit with memristor emulator array and neuron circuits for biologically inspired neuromorphic pattern recognition. *J. Circuits Syst. Comput.* **26**(11), 1750183 (2017)
139. Huang, Q., Zhu, Y.: Patterning of metal nanowire networks: methods and applications. *ACS Appl. & Interf.* **13**(51), 60736–60762 (2021)
140. Wan, Q., Sharbati, M.T., Erickson, J.R., Yanhao, D., Xiong, F.: Emerging artificial synaptic devices for neuromorphic computing. *Adv. Mater. Technol.* **4**(4), 1900037 (2019)
141. Bi, G., Poo, M.: Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**(24), 10464 (1998)
142. Bliss, T.V.P., Lømo, T.: Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* **232**(2), 331–356 (1973)
143. Ohno, T., Hasegawa, T., Tsuruoka, T., Terabe, K., Gimzewski, J.K., Aono, M.: Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater.* **10**, 591 (2011)
144. Winterfeld, H., Ziegler, M., Hanssen, H., Friedrich, D., Benecke, W., Kohlstedt, H.: Technology and electrical characterization of MemFlash cells for neuromorphic applications. *J. Phys. D Appl. Phys.* **51**(32), 324003 (2018)
145. Pershin, Y.V., Di Ventra, M.: Experimental demonstration of associative memory with memristive neural networks. *Neural Netw.* **23**(7), 881–886 (2010)
146. Ziegler, M., Soni, R., Patelczyk, T., Ignatov, M., Bartsch, T., Meuffels, P., Kohlstedt, H.: An electronic version of Pavlov’s dog. *Adv. Func. Mater.* **22**(13), 2744–2749 (2012)
147. Bichler, O., Zhao, W., Alibart, F., Pleutin, S., Lenfant, S., Vuillaume, D., Gamrat, C.: Pavlov’s dog associative learning demonstrated on synaptic-like organic transistors. *Neural Comput.* **25**(2), 549–566 (2013)
148. Ziegler, M., Ochs, K., Hansen, M., Kohlstedt, H.: An electronic implementation of amoeba anticipation. *Appl. Phys. A* **114**(2), 565–570 (2014)
149. Diaz-Alvarez, A., Higuchi, R., Sanz-Leon, P., Marcus, I., Shingaya, Y., Stieg, A.Z., Gimzewski, J.K., Kuncic, Z., Nakayama, T.: Emergent dynamics of neuromorphic nanowire networks. *Sci. Rep.* **9**(1), 14920 (2019)
150. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**(4), 500–544 (1952)
151. Martinez, P., Sprecher, S.G.: Of circuits and brains: the origin and diversification of neural architectures. *Front. Ecol. Evolut.* **8** (2020)
152. Dobzhansky, T.: Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teach.* **35**(3), 125 (1973)
153. Jacob, F.: Evolution and tinkering. **196**(4295), 7 (1977)
154. Poon, C.-S., Zhou, K.: Neuromorphic silicon neurons and large-scale neural networks: challenges and opportunities. *Front. Neurosci.* **5** (2011)
155. Turing, A.M.: I-Computing machinery and intelligence. *Mind* **LIX**(236), 433–460 (1950)
156. Rueckert, U.: Brain-inspired architectures brain-inspired architectures for nanoelectronics. In: Höfflinger, B. (ed.) *CHIPS 2020*, vol. 2, pp. 249–274. Springer International Publishing, Cham (2016)
157. Seung, S.: *Connectome: How the Brain’s Wiring Makes us Who We Are*. Houghton Mifflin Harcourt, Boston (2012)

158. Dayan, P., Abbott, L.F.: *Theoretical Neurosciences: Computational and Mathematical Modeling of Neuronal Systems*. The MIT Press (2001)
159. Gerstner, W., Kistler, W.M.: *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, Cambridge; New York (2002)
160. Izhikevich, E.M.: Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**(6), 1569–1572 (2003)
161. Izhikevich, E.M.: Hybrid spiking models. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **368**(1930), 5061–5070 (2010)
162. Hebb, D.O.: *The Organization of Behavior*, 0 edn. Psychology Press (2005)
163. Kandel, E.R. (ed.): *Principles of Neural Science*, 5th edn. McGraw-Hill, New York (2013)
164. Biyu, J.H.: Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**(9), 480–487 (2014)
165. Buzsáki, G., Draguhn, A.: Neuronal oscillations in cortical networks. *Science* **304**(5679), 1926 (2004)
166. Winfree, A.T.: *The Geometry of Biological Time*. Number v. 12 in *Interdisciplinary Applied Mathematics*, 2nd edn. Springer, New York (2001)
167. Thompson, D.W.: *On Growth and Form*. Dover, New York (1992)
168. Nassim, C.: *Lessons from the Lobster: Eve Marder's Work in Neuroscience*. The MIT Press, Cambridge (2018)
169. Sterling, P., Laughlin, S.: *Principles of Neural Design*. The MIT Press, Cambridge (2015)
170. Kleinfeld, D., Kahler, K.H., Hockberger, P.E.: Controlled outgrowth of dissociated neurons on patterned substrates. *J. Neurosci.* **8**(11), 4098–4120 (1988)
171. Schroeter, M.S., Charlesworth, P., Kitzbichler, M.G., Paulsen, O., Bullmore, E.T.: Emergence of rich-club topology and coordinated dynamics in development of hippocampal functional networks *In Vitro*. *J. Neurosci.* **35**(14), 5459 (2015)
172. van den Heuvel, M.P., Hulshoff Pol, H.E.: Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* **20**(8), 519–534 (2010)
173. Mackey, M.C., Glass, L.: Oscillation and chaos in physiological control systems. *Science* **197**(4300), 287 (1977)
174. Engel, A.K., Fries, P., Konig, P.K., Brecht, M., Singer, W.: Temporal binding, binocular rivalry, and consciousness. *Conscious. Cognit.* **8**(2) (1999)
175. Bassett, D.S., Gazzaniga, M.S.: Understanding complexity in the human brain. *Trends Cognit. Sci.* **15**(5), 200–209 (2011)
176. Melloni, L., Mudrik, L., Pitts, M., Koch, C.: Making the hard problem of consciousness easier. *Science* **372**(6545), 911–912 (2021)
177. Markram, H., Gerstner, W., Sjöström, P.J.: Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synaptic Neurosci.* **4** (2012)
178. Ernst, U., Pawelzik, K., Geisel, T.: Synchronization induced by temporal delays in pulse-coupled oscillators. *Phys. Rev. Lett.* **74**(9), 1570–1573 (1995)
179. Hoppensteadt, F.C., Izhikevich, E.M.: *Weakly Connected Neural Networks*. Applied Mathematical Sciences, vol. 126. Springer, New York (1997)
180. Galizia, C.G., Lledo, P.-M. (eds.): *Neurosciences: From Molecule to Behavior: a University Textbook*. Springer Spektrum, Heidelberg; New York (2013)
181. Gerstner, W., Kistler, W.M., Naud, R., Paninski, L.: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge (2014)
182. Amil, P., Cabeza, C., Masoller, C., Martí, A.C.: Organization and identification of solutions in the time-delayed Mackey-Glass model. *Chaos: Interdiscip. J. Nonlinear Sci.* **25**(4), 043112 (2015)
183. Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd edn. Westview Press, A member of the Perseus Books Group, Boulder, CO (2015)
184. Beer, R.D.: Dynamical approaches to cognitive science. *Trends Cogn. Sci.* **4**(3), 91–99 (2000)
185. Kandel, E.R.: *Cellular Basis of Behavior: An Introduction to Behavioral Neurobiology*. A Series of Books in Psychology. Freeman, San Francisco (1976)

186. Nakagaki, T., Yamada, H., Tóth, Á.: Maze-solving by an amoeboid organism. *Nature* **407**(6803), 470–470 (2000)
187. Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Filho, W.J., Lent, R., Herculano-Houzel, S.: Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Compar. Neurol.* **513**(5), 532–541 (2009)
188. Bosch, T.C.G., Anton-Erxleben, F., Hemmrich, G., Khalturin, K.: The hydra polyp: nothing but an active stem cell community. *Devel. Growth & Differ.* **52**(1), 15–25 (2010)
189. Bielecki, J., Garm, A.: Swim pacemaker response to bath applied neurotransmitters in the box jellyfish *Tripedalia cystophora rhopalium* (2012)
190. Naumann, R., Ondracek, J., Reiter, S., Shein-Idelson, M., Tosches, M., Yamawaki, T.M., Laurent, G.: *The Reptilian Brain*, vol. 25 (2015)
191. Bosch, T.C.G., Klimovich, A., Domazet-Lošo, T., Gründer, S., Holstein, T.W., Jékely, G., Miller, D.J., Murillo-Rincon, A.P., Rentzsch, F., Richards, G.S., Schröder, K., Technau, U., Yuste, R.: Back to the basics: cnidarians start to fire. *Trends Neurosci.* **40**(2), 92–105 (2017)
192. Dupre, C., Yuste, R.: Non-overlapping neural networks in *hydra vulgaris*. *Curr. Biol.* **27**(8), 1085–1097 (2017)
193. Giez, C., Klimovich, A., Bosch, T.C.G.: Neurons interact with the microbiome: an evolutionary-informed perspective. *Neuroforum* **0**(0), 000010151520210003 (2021)
194. Vallverdú, J., Castro, O., Mayne, R., Talanov, M., Levin, M., Baluška, F., Gunji, Y., Dussutour, A., Zenil, H., Adamatzky, A.: Slime mould: the fundamental mechanisms of biological cognition. *Biosystems* **165**, 57–70 (2018)
195. Gagliano, M., Renton, M., Depczynski, M., Mancuso, S.: Experience teaches plants to learn faster and forget slower in environments where it matters. *Oecologia* **175**(1), 63–72 (2014)
196. Adamatzky, A. (ed.): *Advances in Unconventional Computing. Number Volume 22-23 in Emergence, Complexity and Computation*. Springer, Switzerland (2017)
197. Stepney, S., Adamatzky, A.: *Inspired by Nature: Essays Presented to Julian F. Miller on the Occasion of His 60th Birthday*, vol. 28. Springer (2017)
198. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. *Science* **286**(5439), 509 (1999)
199. Barabási, A.-L.: Network science. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **371**(1987), 20120375 (2013)
200. Solé, R.V., Valverde, S.: Information theory of complex networks: on evolution and architectural constraints. In: Ben-Naim, E., Frauenfelder, H., Toroczkai, Z. (eds.) *Complex Networks*, vol. 650, pp. 189–207. Springer, Berlin, Heidelberg (2004)
201. Huberman, B.A., Hogg, T.: Complexity and adaptation. *Phys. D* **22**(1–3), 376–384 (1986)
202. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality. *Phys. Rev. A* **38**(1), 364–374 (1988)
203. Hesse, J., Gross, T.: Self-organized criticality as a fundamental property of neural systems. *Front. Syst. Neurosci.* **8** (2014)
204. Shew, W.L., Clawson, W.P., Pobst, J., Karimippanah, Y., Wright, N.C., Wessel, R.: Adaptation to sensory input tunes visual cortex to criticality. *Nat. Phys.* **11**(8), 659–663 (2015)
205. Cramer, B., Stöckel, D., Kreft, M., Wibrall, M., Schemmel, J., Meier, K., Priesemann, V.: Control of criticality and computation in spiking neuromorphic networks with plasticity. *Nat. Commun.* **11**(1), 2853 (2020)
206. von der Malsburg, C.: The what and why of binding: the modeler’s perspective. *Philos. Explor* (1999)
207. Engel, A.K., König, P., Kreiter, A.K., Singer, W.: Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science* **252**(5009), 1177–1179 (1991)
208. Chialvo, D.R.: Are our senses critical? *Nat. Phys.* **2**(5), 301–302 (2006)
209. Shew, W.L., Yang, H., Yu, S., Roy, R., Plenz, D.: Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J. Neurosci.* **31**(1), 55–63 (2011)

210. van Hemmen, J.L., Sejnowski, T.J. (eds.): 23 Problems in Systems Neuroscience. Oxford University Press, Oxford; New York (2006)
211. Laurent, G.: Shall We Even Understand the Fly's Brain? In: van Hemmen, J.L., Sejnowski, T.J. (eds.) 23 Problems in Systems Neuroscience, pp. 3–21. Oxford University Press (2006)
212. van den Heuvel, M.P., Kersbergen, K.J., de Reus, M.A., Keunen, K., Kahn, R.S., Groenendaal, F., de Vries, L.S., Benders, M.J.N.L.: The neonatal connectome during preterm brain development. *Cereb. Cortex* **25**(9), 3000–3013 (2015)
213. Brenner, S.: The genetics of *caenorhabditis elegans*. *Genetics* **77**(1), 71–94 (1974)
214. Kaiser, M., Varier, S.: Evolution and development of brain networks: from *caenorhabditis elegans* to *Homo sapiens*. *Netw.: Comput. Neur. Syst.* **22**(1–4), 143–147 (2011)
215. Lovas, J.R., Yuste, R.: Ensemble synchronization in the reassembly of Hydra's nervous system. *Curr. Biol.* **31**(17), 3784–3796.e3 (2021)
216. Held, R., Hein, A.: Movement-produced stimulation in the development of visually guided behavior. *J. Comp. Physiol. Psychol.* **56**(5), 872–876 (1963)
217. Huttenlocher, P.R.: Morphometric study of human cerebral cortex development. *Neuropsychologia* **28**(6), 517–527 (1990)
218. Dehaene-Lambertz, G., Spelke, E.S.: The infancy of the human brain. *Neuron* **88**(1), 93–109 (2015)
219. Kaiser, M.: Changing Connectomes: Evolution, Development, and Dynamics in Network Neuroscience. The MIT Press, Cambridge (2020)
220. Wiesel, T.N., Hubel, D.H.: Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.* **26**(6), 1003–1017 (1963)
221. Paredes, M.F., James, D., Gil-Perotin, S., Kim, H., Cotter, J.A., Ng, C., Sandoval, K., Rowitch, D.H., Xu, D., McQuillen, P.S., Garcia-Verdugo, J.-M., Huang, E.J., Alvarez-Buylla, A.: Extensive migration of young neurons into the infant human frontal lobe. *Science* **354**(6308), aaf7073 (2016)
222. Edelman, G.M.: Neural Darwinism: the theory of neuronal group selection. Basic Books, New York (1987)
223. Tononi, G., Sporns, O., Edelman, G.M.: A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. U.S.A.* **91**(11), 5033–5037 (1994)
224. Edelman, G.M., Tononi, G.: A Universe of Consciousness: How Matter Becomes Imagination. Basic Books, New York, 1. paperback ed., [nachdr.] edition (2001)
225. Van Ooyen, A., Butz-Ostendorf, M. (eds.) The Rewiring Brain: A Computational Approach to Structural Plasticity in the Adult Brain. Elsevier, London; Academic, San Diego (2017)
226. Huttenlocher, P.R.: Neural plasticity: the effects of environment on the development of the cerebral cortex. *Neural Plasticity: The Effects of Environment on the Development of the Cerebral Cortex*. Harvard University Press, Cambridge (2002)
227. Hensch, T.K.: Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* **6**(11), 877–888 (2005)
228. Broersma, H., Gomez, F., Miller, J., Petty, M., Tufte, G.: Nascence project: nanoscale engineering for novel computation using evolution. *Int. J. Unconv. Comput.* **8**, 313–317 (2012)
229. Miller, J.F., Harding, S.L., Tufte, G.: Evolution-in-materio: evolving computation in materials. *Evol. Intel.* **7**(1), 49–67 (2014)
230. Baek, E., Das, N.R., Cannistraci, C.V., Rim, T., Bermúdez, G.S.C., Nych, K., Cho, H., Kim, K., Baek, C.-K., Makarov, D., Tetzlaff, R., Chua, L., Baraban, L., Cuniberti, G.: Intrinsic plasticity of silicon nanowire neurotransistors for dynamic memory and learning functions. *Nat. Electron* (2020)
231. Engert, F., Bonhoeffer, T.: Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature* **399**(6731), 66–70 (1999)
232. Kouider, S., Stahhut, C., Gelskov, S.V., Barbosa, L.S., Dutat, M., de Gardelle, V., Christophe, A., Dehaene, S., Dehaene-Lambertz, G.: A Neural Marker of Perceptual Consciousness in Infants. *Science* **340**(6130), 376 (2013)
233. Hiesinger, P.R.: Brain wiring with composite instructions. *BioEssays* **43**(1), 2000166 (2021)

234. Hiesinger, P.R.: *The Self-assembling Brain: How Neural Networks Grow Smarter*. Princeton University Press, Princeton (2021)
235. Ardiel, E.L., Rankin, C.H.: The importance of touch in development. *Paediatr. & Child Health* **15**(3), 53–156 (2010)
236. Lohmann, C., Kessels, H.W.: The developmental stages of synaptic plasticity: the developmental stages of synaptic plasticity. *J. Physiol.* **592**(1), 13–31 (2014)
237. Sanes, D.H., Reh, T.A., Harris, W.A.: *Development of the Nervous System*. Elsevier, Amsterdam; 2nd edn, Boston (2006)
238. Rumpel, S., Triesch, J.: Das dynamische Konnektom. *Neuroforum* **22**(3), 73–79 (2016)
239. Fauth, M., Wörgötter, F., Tetzlaff, C.: Long-term information storage by the interaction of synaptic and structural plasticity (Chap.16). In: van Ooyen, A., Butz-Ostendorf, M. (eds.) *The Rewiring Brain*, pp. 343–360. Academic, San Diego (2017)
240. Singer, W.: Consciousness and the structure of neuronal representations. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **353**(1377), 1829 (1998)
241. Holtmaat, A., Wilbrecht, L., Knott, G.W., Welker, E., Svoboda, K.: Experience-dependent and cell-type-specific spine growth in the neocortex. *Nature* **441**(7096), 979–983 (2006)
242. Kolb, B., Fantie, B.D.: Development of the child’s brain and behavior. In: Reynolds, C.R., Fletcher-Janzen, E. (eds.) *Handbook of Clinical Child Neuropsychology*, pp. 19–46. Springer, US, Boston, MA (2009)
243. Hensch, T.K., Bilimoria, P.M.: Re-opening windows: manipulating critical periods for brain development. *Cerebrum: Dana Forum Brain Sci.* **2012**, 11–11 (2012)
244. Tomen, N., Herrmann, J.M., Ernst, U. (eds.) *The Functional Role of Critical Dynamics in Neural Systems*. Springer Series on Bio- and Neurosystems, vol. 11. Springer International Publishing, Cham (2019)
245. van Ooyen, A., Butz-Ostendorf, M.: Homeostatic structural plasticity can build critical networks. In: *The Functional Role of Critical Dynamics in Neural Systems*, pp. 117–137. Springer (2019)
246. Brütt, M., Kaernbach, C.: On the role of the excitation/inhibition balance of homeostatic artificial neural networks. *Entropy* **23**(12), 1681 (2021)
247. Abbott, L.F.: Balancing homeostasis and learning in neural circuits. *Zoology* **106**(4), 365–371 (2003)
248. Turrigiano, G.: Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. *Cold Spring Harb. Perspect. Biol.* **4**(1), a005736–a005736 (2012)
249. Tetzlaff, C., Okujeni, S., Egert, U., Wörgötter, F., Butz, M.: Self-organized criticality in developing neuronal networks. *PLoS Comput. Biol.* **6**(12), e1001013 (2010)
250. Stepp, N., Plenz, D., Srinivasa, N.: Synaptic plasticity enables adaptive self-tuning critical networks. *PLoS Comput. Biol.* **11**(1), e1004043 (2015)
251. Ma, Z., Turrigiano, G.G., Wessel, R., Hengen, K.B.: Cortical circuit dynamics are homeostatically tuned to criticality In Vivo. *Neuron* **104**(4), 655–664.e4 (2019)
252. Kehayas, V., Holtmaat, A.: Structural plasticity and cortical connectivity. In: *The Rewiring Brain*, pp. 3–26. Elsevier (2017)
253. Tien, N.-W., Kerschensteiner, D.: Homeostatic plasticity in neural development. *Neural Dev.* **13**(1), 1–7 (2018)
254. Butz-Ostendorf, M., van Ooyen, A.: Is lesion-induced synaptic rewiring driven by activity homeostasis? (Chap. 4). In: van Ooyen, A., Butz-Ostendorf, M. (eds.) *The Rewiring Brain*, pp. 71–92. Academic, San Diego (2017)
255. Fukushi, T., Sakura, O.: Ethical challenges and clinical implications of molecular imaging of human consciousness. *Amer. J. Bioeth.* **8**(9), 23–24 (2008)
256. Opitz, A., Falchier, A., Linn, G.S., Milham, M.P., Schroeder, C.E.: Limitations of ex vivo measurements for in vivo neuroscience. *Proc. Natl. Acad. Sci.* **114**(20), 5243 (2017)
257. van der Pol, B.: LXXXVIII. On “relaxation-oscillations”. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **2**(11), 978–992 (1926)
258. Maffezzoni, P., Daniel, L., Shukla, N., Datta, S., Raychowdhury, A.: Modeling and simulation of vanadium dioxide relaxation oscillators. *IEEE Trans. Circuits Syst. I Regul. Pap.* **62**(9), 2207–2215 (2015)

259. Driscoll, T., Quinn, J., Di Ventra, M., Basov, D.N., Seo, G., Lee, Y.-W., Kim, H.-T., Smith, D.R.: Current oscillations in vanadium dioxide: evidence for electrically triggered percolation avalanches. *Phys. Rev. B* **86**(9) (2012)
260. Luo, Z., Bo, Y., Sadaf, S.M., Liu, X.: Van der Pol oscillator based on NbO₂ volatile memristor: a simulation analysis. *J. Appl. Phys.* **131**(5), 054501 (2022)
261. Skarda, C.A., Freeman, W.J.: How brains make chaos in order to make sense of the world. *Behav. Brain Sci.* **10**(2) (1987)
262. King, C.C.: Fractal and chaotic dynamics in nervous systems. *Prog. Neurobiol.* **36**(4), 279–308 (1991)
263. Michaelis, D., Jenderny, S., Ochs, K.: A self-organizing gait pattern generator exploiting an electrical circuit for axon growth. In: 2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 166–169 (2021)
264. Fisher, M.E.: The story of coulombic criticality. *J. Stat. Phys.* **75**(1–2), 1–36 (1994)
265. Aoki, T., Yawata, K., Aoyagi, T.: Self-organization of complex networks as a dynamical system. *Phys. Rev. E* **91**(1), 012908 (2015)
266. Haugland, S.W., Schmidt, L., Krischer, K.: Self-organized alternating chimera states in oscillatory media. *Sci. Rep.* **5**(1), 9883 (2015)
267. Patzauer, M., Krischer, K.: Self-Organized multifrequency clusters in an oscillating electrochemical system with strong nonlinear coupling. *Phys. Rev. Lett.* **126**(19), 194101 (2021)
268. Orlik, M., Orlik, M.: General principles of self-organization: temporal Instabilities. In: Orlik, M. (ed.) Number I in Self-organization in Electrochemical Systems. Springer, Heidelberg, New York, Dordrecht, London, Softcover reprint of the hardcover 1st edition 2012 edition (2012)
269. Orlik, M., Orlik, M.: Spatiotemporal patterns and control of chaos. In: Orlik, M. (ed.) Number II in Self-organization in Electrochemical Systems. Springer, Heidelberg, New York, Dordrecht, London, Softcover reprint of the hardcover 1st edition 2012 edition (2012)
270. Sporns, O.: Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* **15**(3), 247 (2013)
271. Kreuz, T., Mormann, F., Andrzejak, R.G., Kraskov, A., Lehnertz, K., Grassberger, P.: Measuring synchronization in coupled model systems: a comparison of different approaches. *Phys. D* **225**(1), 29–42 (2007)
272. Abbott, L.F., Rohrkemper, R.: A simple growth model constructs critical avalanche networks. In: Cisek, P., Drew, T., Kalaska, J.F. (eds.) *Progress in Brain Research*, vol. 165, pp. 13–19. Elsevier (2007)
273. Scarpetta, S., de Candia, A.: Neural avalanches at the critical point between replay and non-replay of spatiotemporal patterns. *PLoS ONE* **8**(6), e64162 (2013)
274. Priesemann, V.: Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Front. Syst. Neurosci.* **8** (2014)
275. Timme, N.M., Marshall, N.J., Bennett, N., Ripp, M., Lautzenhiser, E., Beggs, J.M.: Criticality maximizes complexity in neural tissue. *Front. Physiol.* **7** (2016)
276. Milton, J.G.: Neuronal avalanches, epileptic quakes and other transient forms of neurodynamics: transient neurodynamics. *Eur. J. Neurosci.* **36**(2), 2156–2163 (2012)
277. di Santo, S., Villegas, P., Burioni, R., Muñoz, M.A.: Landau-Ginzburg theory of cortex dynamics: scale-free avalanches emerge at the edge of synchronization. *Proc. Natl. Acad. Sci.* **115**(7), E1356–E1365 (2018)
278. Bottani, S.: Pulse-coupled relaxation oscillators: from biological synchronization to self-organized criticality. *Phys. Rev. Lett.* **74**(21), 4189 (1995)
279. Beggs, J.M., Timme, N.: Being critical of criticality in the brain. *Front. Physiol.* **3**, 163 (2012)
280. Srinivasa, N., Stepp, N.D., Cruz-Albrecht, J.: Criticality as a set-point for adaptive behavior in neuromorphic hardware. *Front. Neurosci.* **9**, 449 (2015)
281. Chialvo, D.R., Cannas, S.A., Grigera, T.S., Martin, D.A., Plenz, D.: Controlling a complex system near its critical point via temporal correlations. *Sci. Rep.* **10**(1), 1–7 (2020)
282. Kreuz, T., Chicharro, D., Andrzejak, R.G., Haas, J.S., Abarbanel, H.D.: Measuring multiple spike train synchrony. *J. Neurosci. Methods* **183**(2), 287–299 (2009)

283. Hoppensteadt, F.C., Peskin, C.S.: Modeling and Simulation in Medicine and the Life Sciences, vol. 10. Springer Science & Business Media (2012)
284. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., Hassabis, D.: A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**(6419), 1140–1144 (2018)
285. Davies, M.: Benchmarks for progress in neuromorphic computing. *Nat. Mach. Intell.* **1**(9), 386–388 (2019)
286. Adams, N.E.: Bloom’s taxonomy of cognitive learning objectives. *J. Med. Libr. Assoc.: JMLA* **103**(3), 152–153 (2015)
287. Compeau, P.: Establishing a computational biology flipped classroom. *PLoS Comput. Biol.* **15**(5), e1006764 (2019)
288. Churchland, P.M.: *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. MIT Press, Cambridge (1995)
289. Johannsen, W.: *Information und ihre Bedeutung in der Natur: das Leben erfindet die Welt*. Springer, Berlin (2016)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Neuromorphic Circuits with Redox-Based Memristive Devices



Finn Zahari, Martin Ziegler, Pouya Doerwald, Christian Wenger,
and Hermann Kohlstedt

Abstract This chapter addresses opportunities for memristive devices in the framework of neuromorphic computing. Memristive devices are two-terminal circuit elements, comprising resistance and memory functionality. This simple and likewise ingeniously concept allows beneficial applications in numerous neuromorphic circuits. However, the electrical characteristics as well as the materials and technological framework of memristive devices need an optimization for each specific application. The chapter starts with a short overview of basic principles of biological data processing followed by a taxonomy of different bio-inspired computing architectures, divided into time-dependent and time-independent concepts. Furthermore, the requirements on particular memristive device properties, such as I - V linearity, switching time, retention, number of states, time-dependency, and device variability, are discussed. The results of tangible examples of digital and analog memristive switching devices used in a deep neural network based on CMOS-integrated resistive random access memory devices (RRAMs) for chronic obstructive pulmonary disease (COPD) detection, in stochastic learning, in bio-inspired analog learning, and, finally, in oscillatory computing are presented and discussed.

F. Zahari (✉) · H. Kohlstedt
Nanoelectronics, Faculty of Engineering, Kiel University, 24143 Kiel, Germany
e-mail: fnz@tf.uni-kiel.de

H. Kohlstedt
e-mail: hko@tf.uni-kiel.de

M. Ziegler
Department of Electrical Engineering and Information Technology, Technische Universität
Ilmenau, 98693 Ilmenau, Germany
e-mail: martin.ziegler@tu-ilmenau.de

Institute of Micro and Nanotechnologies MacroNano, Technische Universität Ilmenau, 98693
Ilmenau, Germany

P. Doerwald · C. Wenger
IHP Leibniz-Institut für innovative Mikroelektronik, Frankfurt/Oder, Germany
e-mail: wenger@ihp-microelectronics.com

1 Introduction

While traditional von Neumann computing (binary and serial) continues to dominate the information and communication technology (ICT), recent groundbreaking innovations in alternative computing architectures and advanced electronic devices have become hard to ignore [1–6]. The mature silicon (Si)-based, complementary metal oxide semiconductor (CMOS) technology behind von Neumann computing led to tremendous progress in digital computing power over the last six decades. Moore’s law, the prediction that the device integration density on a chip doubles every two to three years, and accompanied by scaling laws, served as secure guidelines for CMOS engineers [7, 8]. Today’s Si-CMOS technology is characterized by impressive technical specifications. To mention but a view, the smallest feature size in advanced Si-CMOS transistor is in the order of 10 nm, several billion functional transistors are integrated into a single arithmetic logic unit (ALU), and a von Neumann computer is running at clock-frequencies of about 5 GHz [5]. Nonetheless, this success should not disguise the fact that Si-based von Neumann computing reaches its limits in the near future, and system performance progress has slowed down for a couple of years. For example, although a further shrinking of transistor dimension is still possible, limitations, such as quantum mechanical tunneling, will set fundamental constraints. The so-called memory gap appeared as another architecture-related show stopper. While the clock frequency increased with every new chip generation, the memory access time did not follow this trend. Therefore, the system performance is limited because the data handling between the ALU and memory presents a data bottleneck [9]. Aware of this stagnation, worldwide efforts by universities, research institutions, and industry focus on numerous novel computing architectures and advanced functional devices. Besides quantum computing, bio-inspired computing attracted considerable attention [10–15]. The term “bio-inspired” embraces various classes of computing architectures and hardware systems that take, to a more or less extent, biological information pathways as guidelines into account. In order to categorize such bio-inspired computing systems, basic principles of information pathways in nervous systems are summarized here. We want to emphasize, that the following introduction to the biological background describes the information processing in nervous systems in a very simplified way and represents only the tip of the iceberg. For more details on the intriguing biochemical and spatio-temporal mechanisms in this context, we refer the reader to the literature [16, 17]. A common way to explain information pathways in nervous systems is to distinguish the processing on the local neuron level from that of the entire system. Neurons are the information processing units in nervous systems. Neurons receive and process information in the form of action potentials. In Fig. 1a, the structure of a neuron, including the soma, dendrites, the axon, and connections to other neurons by synapses, is sketched [16]. An action potential (spike) is an abrupt transitory and transmitted change of the resting potential across the cell membrane. The amplitude of a spike is 100 mV, and its pulse duration is about 3.5 ms. Post-synaptic neurons receive action potentials (signals) from pre-synaptic neurons via dendrites and synapses. Those action potentials are

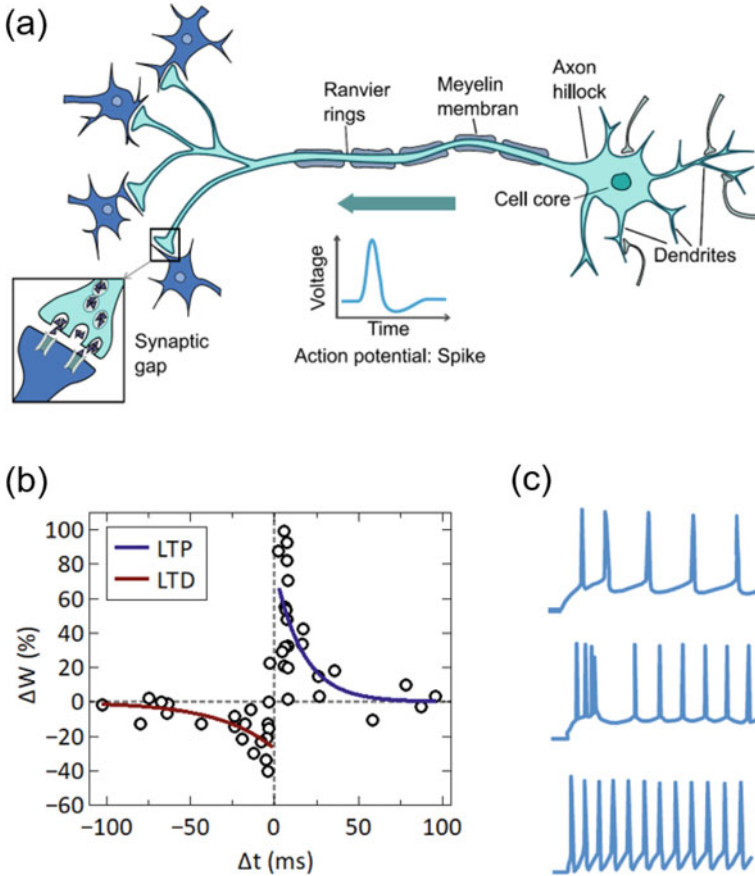


Fig. 1 **a** Blueprint of a neuron, including an enlarged sketch of a synapse and the illustration of a single action potential (i.e., a spike). Reprinted with permission of the corresponding author from [18]. **b** Change of synaptic weight (ΔW) in dependency on the timing ($\Delta t = t_{post} - t_{pre}$) of pre-synaptic spikes at time t_{pre} and post-synaptic spikes at time t_{post} illustrating spike-timing dependent plasticity (STDP). ΔW is measured 20-30 min after inducing the synaptic change with 60 correlated pre- and postsynaptic spikes with a frequency of 1 Hz at synapses between hippocampal glutamatergic neurons in culture. The data is fitted with exponential functions of the form $\Delta W = e^{(-\Delta t/\tau)}$ for $\Delta t > 0$ (LTP) and $\Delta t < 0$ (LTD). Adapted from [19] [G.-q. Bi & M.-m. Poo, Annu. Rev. Neurosci. 24, 139–166 (2001)], with the permission of Annual Reviews. **c** Different irregular signal patterns, which occur in nervous systems [20]. ©2003 IEEE. Adapted, with permission, from [E. Izhikevich, IEEE Transactions on Neural Networks 14, 1569–1572 (2003)]

integrated within the post neuron's cell body. If the potential exceeds a threshold at the axon hillock, a new or a series of new spikes are generated at the axon hillock. Such a spike or spike trains are transmitted via the axon to subsequent neurons. The process is called leaky-integrate-and-fire (LIF), reflecting that the cell membrane is not a perfect insulator. The conduction velocity along the axon is between a few cm/s up to a few tenth m/s. We like to underline that the spike duration (3.5 ms) and the conduction velocity (e.g., 100 m/s) are six orders of magnitudes smaller than the pulse duration (1 ns) and signal transmission speed (roughly the speed of light in vacuum, $3 \cdot 10^8$ m/s) in modern processors, respectively. Facts, which indicate a different information processing of von Neumann computers and nervous systems. The mentioned interconnects are called synapses. Interestingly, the synaptic efficacies are not fixed and can change, which refers to the information coupling strength between the pre- and post-neuron. The efficacy can be increased (i.e., synaptic potentiation) or decreased (i.e., synaptic depression). This plastic behavior can last from milliseconds to minutes [called short-term potentiation (STP) or short-term depression (STD)] or from hours to years and up to the whole lifetime of animals [called long-term potentiation (LTP) or long-term depression (LTD)] [17]. The change of synaptic coupling strength depends on the common activity of the pre- and post-synaptic cells, e.g., in accordance to signal timing between the pre- and post-neuron. This mechanism is called spike-timing-dependent plasticity (STDP) [21]. STDP is closely related to Hebb's learning rule, which says that neurons that fire together wire together [22]. This rule describes the learning process in neural networks on the local cellular level by adjusting the synaptic efficacy dependent on the common activity of pre- and post-neurons. This further contributes to associative learning on the network level [23, 24]. Accordingly, STDP represents a local learning rule and is an essential process for establishing learning and memory in nervous systems [see Fig. 1b][19, 25]. In stark contrast to von Neumann computers, the spike generation in nervous systems is far from being regular. In dependency on the signal input of a neuron, a plethora of different firing rates (ranging from Poisson-like to bursting) are observed [26–28]. In Fig. 1c, examples of such irregular signal patterns are sketched [20]. Any signal transmission in nervous systems is accompanied by errors and noise [29]. For example, random potential fluctuations and the granular structure of neurotransmitters lead to a pronounced stochastic component in biological information pathways. Nevertheless, neural networks can store and retrieve information reliably. So noise is not a bug in biological systems, it is a feature [30, 31]. Nonetheless, although processes at the local neural level are highly stochastic, the entire nervous system exhibits rhythmic brain waves. In the human brain, they appear in the form of alpha, beta, and gamma waves [32]. Characteristic features, such as STDP [21, 25, 33], stochastic firing and bursting of neurons in the hundred Hz range, recurrent network structures, and aspects of oscillatory synchrony in larger neuronal ensembles [34–43] are essential ingredients in biological-based information processing. Moreover, factors related to the close interaction of a nervous system with its environment, i.e., external stimuli, are of crucial importance [44]. Therefore, neuronal design principles provide a model for neuromorphic systems, which are diametric to development strategies in present binary ICT, including precise GHz clock

frequencies, near-light-speed signal transmission, and clearly separated logic and memory [45, 46]. In contrast to clock-driven von Neumann machines [9, 47], information processing in biological nervous systems is characterized by highly parallel, energy efficient, and adaptive architectures [10, 11, 14, 48].

Now we turn back to the field of neuromorphic circuits [10–15]. Bio-inspired computing aims to realize biological plausible information pathways (a few are mentioned above) in engineered systems. However, this goal immediately leads to numerous questions and challenges: Which of the manifold and intriguing information pathways observed in biology need to be mimicked by neuromorphic circuits to establish novel computing architecture with superior properties to conventional von Neumann systems? Could any biological process be simulated by a von Neumann computer to achieve similar performances as the biological model? Should novel neuromorphic circuits be made on today’s mature Si-CMOS platform or might it be wise to open the material “tools box” apart from Si technology in order to facilitate the integration of novel devices and/or to incorporate self-assembly strategies similar to those observed in nature?

Might it be beneficial to design circuits working at biologically plausible time scales, i.e., with pulse durations of about a few ms and signal conduction velocities of about m/s, which lead to strong signal retardations? Finally, yet importantly, the consideration of stochasticity could be an additional design goal in accordance with its fundamental role in biology. These questions display only a small selection of possible approaches to mimic biological information pathways in engineered systems. This chapter focuses on neuromorphic circuits that take up those hallmarks of biological information processing that have been discussed above, i.e., STDP, stochasticity, oscillatory computing, and so on. The circuits comprise redox-based memristive devices as key components. Memristive devices consist of capacitor-like layer sequence, i.e., metal-memristive material-metal stacks. A universal property of the memristive devices is that the memristive state depends on previously induced charge flows, applied currents, or applied electric fields, thus storing any resulting resistance state. Memristive devices can be engineered to exhibit either binary switching or an analog resistance variation. Both device categories are subjects of this chapter and show beneficial features in pattern recognition and oscillatory computing. For details concerning resistive switching and the underlying physical-chemical mechanisms, we refer the reader to Chap. 3 and the overwhelming literature on the subject [49–52].

2 Requirements for Memristive Devices for Neuromorphic Computing

Memristive devices are being explored for many different types of neuromorphic computation schemes, where their non-volatility allows computation to be performed in memory [53]. In this respect, memristive devices allow overcoming the von Neumann bottleneck, where memory and computation are separated from each other

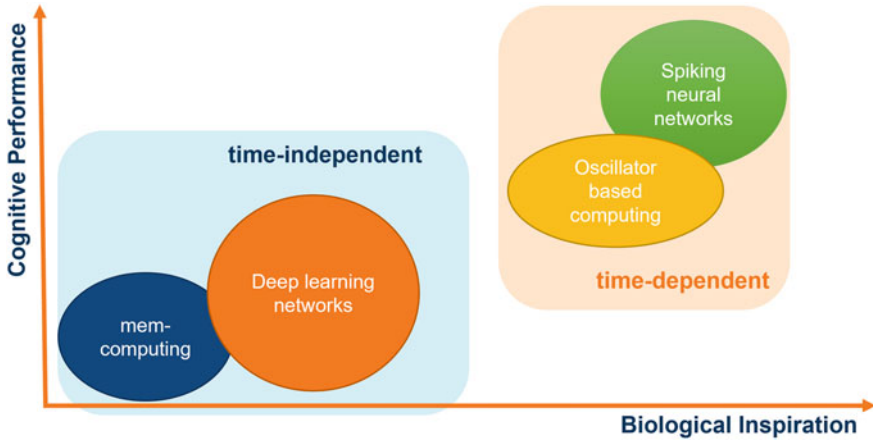


Fig. 2 Classification of different biologically inspired computing architectures within time-dependent and time-independent computational schemes: the more biological the computational scheme, the higher the expected cognitive performance

[9]. However, there are many different neuromorphic computing architectures, which place different requirements on memristive devices [54, 55]. Figure 2 shows a classification that relates different network architectures to their biological inspiration and cognitive performance. While time-independent computing schemes are most widely used, time-dependent computing schemes are more biologically realistic. However, their technical implementation is much more challenging. While oscillatory computing and spiking neural networks (SNNs) taking the temporal dimension of the signals into account [14, 56–58], mem-computing, and deep neural networks (DNNs) working in a time-independent way using synchronized signals (i.e., they are based on clock signals) [57, 58]. This section will show the differences in the requirements for the memristive devices. For this purpose, a short overview of the architectures will be given, and the requirements they impose on memristive devices will be elaborated. Concrete examples of memristive networks that use time-independent computational schemes are discussed in Sect. 3, while examples of time-dependent neuromorphic architectures are presented in Sect. 4.

Mem-Computing

Mem-computing was originally invented for non-bio-inspired applications which aim to implement new computing architectures tasks like stateful [59, 60] and non-stateful logic [61, 62], in-memory arithmetic operations [63], solving linear and partial differential equations [64–66], optimization [67–69], and signal processing [70, 71] (for a review, the reader is referred to [3, 55, 58]). In all of those applications,

memristive devices are required which own a fast and low power switching, high cycling endurance (e.g., $> 10^{12}$ cycles for stateful logic) as well as a low device-to-device and cycle-to-cycle variability [55, 58]. However, the inherent randomness of memristive switching mechanisms are challenging for high-precision computing tasks and different techniques are needed to increase the overall precision of the devices [64, 65, 69].

Deep Neural Networks (DNNs)

DNNs [72] (including convolutional neural networks (CNN) [73]) are bio-inspired computing schemes that benefit from the in-memory computing architectures incorporating memristive devices. In these networks, a large number of artificial neurons arranged in layers connected in a feed-forward structure by adjustable weights. These analog weights are trained with the backpropagation algorithm, which implements the delta rule between each neural layer [72]. These networks are specialized in pattern recognition tasks and build the backbone of today's machine learning applications [3, 58]. However, several drawbacks come along with DNNs. The networks are usually set up in software running on traditional von Neuman architecture, i.e., mostly on general-purpose graphics processing units (GPUs) [74] or tensor processing units (TPUs) [75]. Since a huge amount of data is needed for training and many learning cycles are required for real applications, these networks consume a large amount of energy and space. In addition, the training is very time-consuming [3, 57]. In this context, memristive devices can provide a solution as their in-memory computing properties enable parallelization of processes, reducing power consumption and training time by orders of magnitude [76–79]. The significantly increased efficiency lies in implementing matrix-vector multiplications (MVMs) in hardware utilizing Ohm's law and Kirchoff's current law [54]. For this, however, strict requirements must be met by the memristive devices to be used as artificial synapses in DNNs, which are summarized in Fig. 3. A linear, gradual, and symmetric change in resistance is required for training [77, 80, 81]. For example, it has been shown that a 2% deviation from perfect symmetry increases the required number of analog states to train a DNN from 100 to 1000 [82]. In addition, endurance is essential since a lot of data is needed for training, combined with high energy efficiency and low latency to enable training directly on edge devices [83]. On the other hand, the training algorithm can compensate for device variability and yield to a certain degree [82]. Transferring pre-trained weights to memristive devices, furthermore, leads to less strict requirements the devices have to fulfill. In this respect, it has been shown that a resolution of four to eight bits is sufficient to compete with floating-point precision weights in the inference process [75, 84]. For this purpose, multiple binary devices can be combined to mimic the weights of a synapse [85, 86]. However, the device variability and yield are more critical for pre-trained networks, and endurance becomes less crucial [82]. Retention is also of particular relevance. While for training, short-term

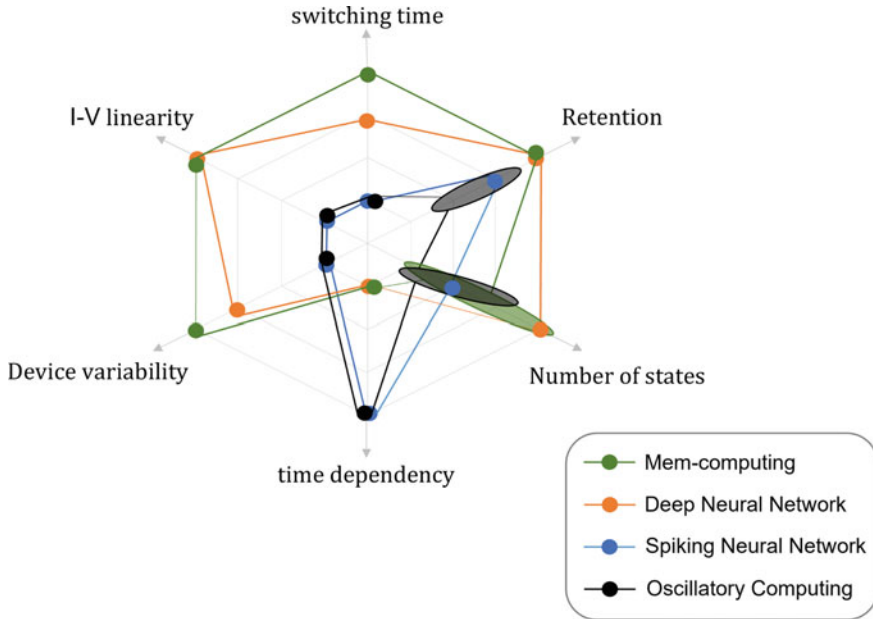


Fig. 3 Comparison of the requirements memristive devices have to fulfill to be suitable for different computational architectures

retention and stability are sufficient [78, 82, 87], for inference, long-term retention and stability must be given [78, 82, 87].

Spiking Neural Networks (SNNs)

In SNNs, the computation is based on synaptic connectivity and asynchronous, event-driven, and temporally precise signals [14, 15, 56–58, 88]. This enables such networks to adapt to changing environmental conditions and react accordingly. However, adequate local learning algorithms are required to exploit those advantages of SNNs over DNNs. These algorithms must satisfy the special needs of memristive devices and neuromorphic network structures in equal measure [89–97]. At the memristive device level, several criteria, and in some cases different requirements from DNNs, must be met in order to satisfy the bio-inspired learning rules designed for SNNs. Figure 3 summarizes these and compares them to the device requirements of DNNs. Since nonlinear time-dependent conductivity modulations are the basis of learning in SNNs [57, 58], a gradual but nonlinear resistance change in memristive devices can be beneficial [57, 97]. Furthermore, both the endurance and energy-efficient switching and read-out are crucial properties for memristive devices in SNNs [57, 83]. However, device-to-device and cycle-to-cycle variability [57] and yield are

less critical since the effect of defective devices can be mitigated during learning [93]. Furthermore, the switching speed is expected to be less important [57, 88], especially if the networks run on a biologically relevant time scale with a spike duration of milliseconds and frequencies of a few Hz [33]. The performance of SNNs is often investigated with pattern recognition tasks with frame-based datasets [3, 98]. In this respect, typical benchmark datasets containing static images are MNIST [99], CIFAR [100], and ImageNet [101]. These datasets do not contain temporal information. Thus, they do not allow to show the full potential of the time-dependent SNNs [102] and an outperformance in comparison to DNNs has not been reported so far [58]. More suitable benchmarks for SNNs should contain real-world spatio-temporal data, e.g., collected with event-based sensors [98, 103]. Thus, instead of executing pattern recognition on static data, SNNs are expected to be superior in interacting with the real world in a dynamically changing environment by processing continuous but sparse input streams on an energy-efficient way [57, 98, 102]. More suitable benchmarks for those tasks are dealing, e.g., with hand and arm gesture detection (DvsGesture) [104], automated driving [105, 106], or robotics [107]. Moreover, generally applicable learning algorithms and network structures, which can cope with several different tasks, are within the focus of research [3, 15].

Oscillatory Computing

Nature uses time-coherent dynamics for information processing based on the formation of context-dependent, self-organized, and transient network structures. These enable us to react adequately to changing environmental conditions. Furthermore, these self-organized network structures are an important property for sensory integration [36, 108]. Even if the underlying mechanisms are only partially understood, the interaction between dynamics and topology has been identified as one of the essential building blocks of information processing in the brain in recent years [109]. In the current understanding, it is assumed that information is encoded into coherent states by temporally correlated neural activity patterns [110]. This concept offers, particularly, an elegant explanation for the binding problem - the question of the mechanism of sensory integration, which allows our brain to construct uniform perceptions from the multitude of sensory information. First evidence of these concepts could have been gathered from experiments with sensorimotor networks [111]. More recent studies have shown the universality of these concepts for the entire brain [110]. In this respect, memristive devices allow a new degree of freedom for the concept of neural synchrony: a local memory that supports a transient connectivity pattern [112–114]. The requirements for memristive devices needed for this have been less studied and are the content of Sect. 4.2. However, the inherent stochasticity of memristive devices has been shown to be helpful for this application [113].

3 Time-Independent Neural Networks

This section deals with two memristive networks that use time-independent computational schemes, both relying on a fully CMOS-integrated 4 kbit resistive random-access memory (RRAM) array [115–117]. The first network discussed in Sect. 3.1 is a mixed-signal-circuit implementation of a DNN for the detection of chronic obstructive pulmonary disease (COPD) [118]. Here, the devices are used to store the pre-trained weights of the DNN while the neurons are implemented in software. In that way, the possibility for on-chip recognition of saliva samples using in-memory techniques to detect COPD in a Point-of-Care application is shown. The second network introduced in Sect. 3.2 exploits the RRAM cells' inherent stochasticity to solve a pattern recognition task [119, 120]. The stochastic artificial neural network (StochANN) is able to learn a subset of the MNIST benchmark through the adaptation of synaptic weights directly in hardware (in a mixed-signal realization) through a supervised local stochastic learning rule. Additional simulations of StochANNs with a larger number of devices show the performance limits of such a network for the whole MNIST benchmark. The results are compared to state-of-the-art approaches using time-independent and time-dependent networks.

3.1 Deep Neural Network Implemented in CMOS-Integrated RRAM Arrays Used for Chronic Obstructive Pulmonary Disease Detection

In this section, a time-independent DNN trained for disease detection is introduced [118]. The three-layer network is composed of a binary, fully CMOS-integrated 4 kbit RRAM array [115–117] emulating the synaptic weights, while neurons are implemented in software. The network was trained entirely in software, and the weights were subsequently transferred to the RRAM array. In that way, memristive devices can be used for inference, while they do not have to fulfill the same requirements as needed for training (see Sect. 2). In the following, first, the disease to be detected and the relevant input parameters for the machine learning (ML) method are introduced. Afterward, the network implementation and its performance are described.

COPD, one of the most prevalent lung diseases worldwide, runs a perfidious course with an often long-lasting undiagnosed initial phase. Clinical treatment approaches for COPD result in repeated clinical visits and extended hospitalization for patients. This fact, apart from being an economic burden for healthcare infrastructures, drastically impacts patients' life quality. To address this issue, today's healthcare systems have encouraged the development of personalized solutions through which patients can receive appropriate medical assistance in an outpatient clinic or a home-care environment [121]. Recent advances in point-of-care medical devices have facilitated the early detection, prevention, and treatment of various diseases [122]. However, without analytical insight, collected data from medical sensors are merely raw data

with low clinical value. For instance, in our previous work, a portable biosensor for managing COPD in home-care environments was presented [123]. The developed biosensor was capable of characterizing the viscosity of saliva samples for diagnostic purposes. However, saliva samples' viscosity properties are one parameter of various parameters required for COPD detection. As a result, upon viscosity measurements by the developed biosensor, a sophisticated diagnostic algorithm is required to detect COPD by concurrent consideration of all essential parameters related to a patient's personal and medical background. These demographic parameters include, but are not limited to, age, gender, weight, cytokine level, pathogen load, and the smoking background of subjects. Therefore, machine learning tools, or more specifically pattern recognition methods, could make the diagnostic procedure more efficient by converting collected data from medical sensors into meaningful clinical information. Moreover, machine learning can be used for identifying diagnostic links between symptoms and diseases that have been previously unknown and providing treatment plans and recommendations to healthcare specialists.

As a result, implementing ML tools is crucial for converting collected raw data from subjects into meaningful clinical-diagnostic information. Furthermore, advanced ML analytics could make the management of COPD in Point-of-Care applications more efficient. Nevertheless, drawbacks of cloud-based ML techniques for medical applications such as data safety, immense energy consumption, and enormous computation requirements need to be addressed for this application. To address these challenges, CMOS-integrated RRAM arrays can be used for the hardware-based implementation of ML methods. Therefore, a memristive neuromorphic platform is presented in this work for on-chip recognition of saliva samples of COPD patients and healthy controls. Two groups of saliva samples, 160 for Healthy Controls (HC) and 79 for COPD patients, were collected in the frame of a joint research project at the Research Center Borstel, BioMaterialBank Nord (Borstel, Germany) [124, 125]. Patient materials were collected and anonymized prior to accessibility. The sampling procedure of the saliva samples was approved by the local ethics committee of the University of Luebeck under the approval number AZ-16-167. Figure 4 demonstrates a hierarchy chart, categorizing the collected saliva samples into extended subgroups with respect to their diagnosis, gender, and smoking status [124]. As shown in Fig. 5, analog values of these four attributes were converted into 23 binary bits [gender (1), smoking status (3), age (9), dielectric permittivity (10)]. Dielectric sensors could be used to characterize sputum samples collected from patients for early diagnosis of COPD. The CMOS-based dielectric sensor system used for the real-time monitoring of sputum samples is described in [126].

The neuromorphic hardware implementation [118] of the developed ANN model was performed with a 4-kbit array of CMOS-integrated RRAM devices based on amorphous HfO₂ developed by IHP [115–117]. The array consists of 64 × 64 memristive cells in a 1-Transistor-1-Resistor (1T-1R) configuration. The two distinct states, low resistance state (LRS) and high resistance state (HRS), were used for the implementation. The mean read-out currents are 30.8 mA and 3.2 mA at 0.2 V for LRS and HRS, respectively. For the deployment of the 10-level model, a mixed-signal

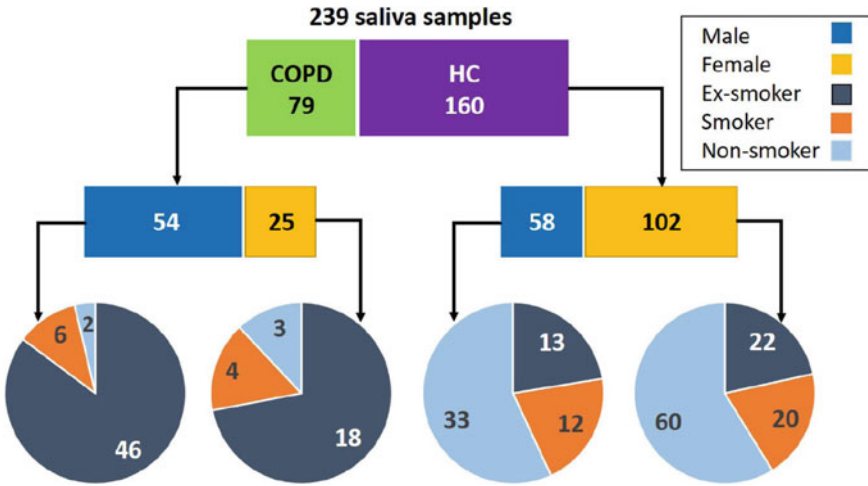


Fig. 4 Hierarchical categorization of collected saliva samples into extended subgroups with respect to their diagnosis, gender, and smoking status. Reprinted from [124] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

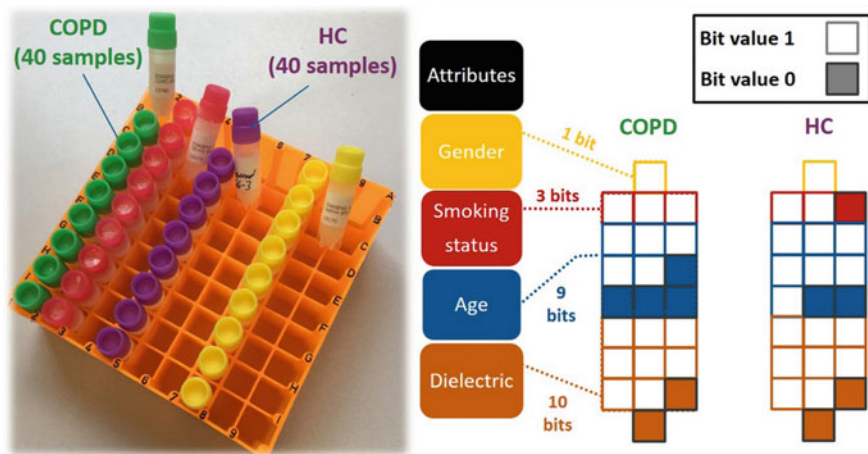
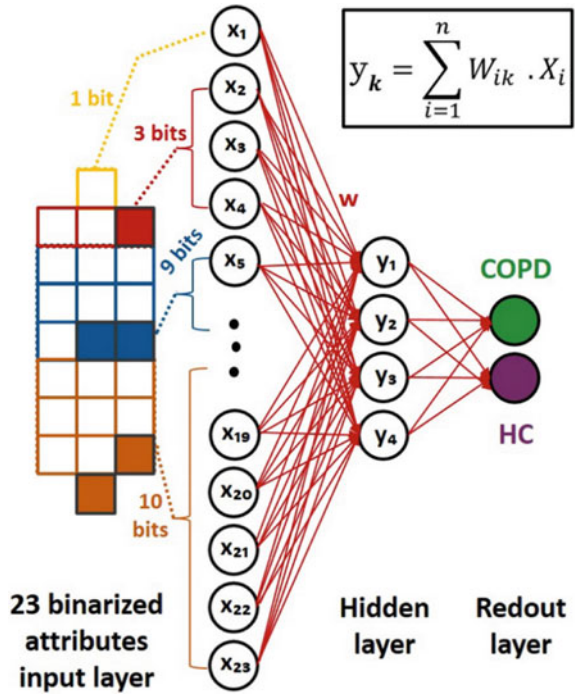


Fig. 5 Conversion of analog attributes of the dataset (gender, smoking status, age, and dielectric properties) into 23 binary bits. Reprinted from [118] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

Fig. 6 ANN topology with one hidden layer for the classification of saliva samples of COPD patients and HC. Reprinted from [118] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)



neuromorphic circuit with software-based neurons and hardware synapses was used [119, 120]. Considering the topology of the ANN model (see Fig. 6) with one hidden layer and one read-out layer with four and two neurons per layer, respectively, 106 parameters (i.e., synaptic weights and biases) were required for connecting the network layers. The resistance states of 1060 memristive devices on a single chip were set to the HRS or LRS, respective to the pre-trained weights. Every network parameter is represented by the combination of ten devices where five devices represent positive values and five devices represent negative values, respectively. The sum of ten read-out currents at 0.2 V represents the total value of one synaptic weight. After successfully implementing pre-trained weights on the hardware, the test subset of data was used to evaluate the performance of the neuromorphic model for the recognition of COPD and HC samples. In order to recognize the COPD samples with the mixed-signal approach, the 23 input bits of the test-subset data were applied to the simulated neurons within the input layer. The output neurons of every subarray are perceptrons with a sigmoidal activation function, which receive the sum of current values passing through the connected devices together with a specific bias value. These current values are normalized to the maximum value of the pre-trained analog network to guarantee that the sigmoid function is activated with a reasonable range of values. The output values of the third layer (read-out layer) perceptrons denote whether a test sample belongs to COPD or HC categories. This hardware realization

agrees with the theory of neural networks that the weighted sum of inputs determines the value of a perceptron in the subsequent layer, as illustrated in Fig. 6.

In summary, the concept of on-chip recognition of saliva samples of COPD patients using a memristive neuromorphic platform was studied. A hardware-friendly artificial neural network model was developed and trained for classifying COPD and HC samples using real clinical data. Subsequently, a 10-level conversion of the trained classification model was transferred onto a memristive neuromorphic platform for the on-chip recognition. The memristive chip provided a remarkable accuracy of 89%, offering an alternative approach to cloud-based methods required for diagnosing COPD in Point-of-Care applications.

3.2 *Stochastic Learning with Binary CMOS-Integrated RRAM Devices*

The inherent stochastic nature of the filament formation and dissolution in RRAM devices is challenging for many applications, especially if a high numerical precision is needed (see Sect. 2). On the other hand, different approaches benefit from the randomness of resistive switching and exploit it explicitly for the technical emulation of biological information processing. Such networks include noise tolerant stochastic computing technologies [127], synchronization of oscillatory neurons to emulate neuronal coherence [113, 128] as described in Sect. 4.2, stochastic switching neurons [129, 130] and stochastic learning rules realized with single binary synapses [119, 120, 131–133], as well as compounds of several binary devices as one synapse [85, 129, 131, 133].

The stochastic learning algorithm [119, 120] described in this section utilizes the stochastic nature of binary fully CMOS-integrated 4 kbit RRAM arrays [115–117] in a 1-transistor-1-resistor (1T-1R) configuration, the same technology as used in Sect. 3.1, in a stochastic artificial neural network (StochANN) to learn the MNIST benchmark [99]. In that way, it is shown that the proposed StochANN is able to process analog information with binary memory cells. The devices are composed of a $\text{HfO}_{2-x}/\text{TiO}_{2-y}$ bi-layer sandwiched between TiN electrodes. They can be switched between two distinct resistance levels, i.e., HRS and LRS. As an initial step, an electro-forming process is required. This is reliably done by the incremental step pulse with verify algorithm (ISPVA) [134]. The electrical device properties depend on the crystalline phase of the HfO_{2-x} [120, 135]. The switching probability for polycrystalline and amorphous HfO_{2-x} films is shown in Fig. 7a, b and c, d, respectively. The device-to-device (D2D) variability of 128 1T-1R devices is therefore determined by applying single voltage pulses in the set [Fig. 7a, c] and reset [Fig. 7b, d] regime, i.e., for the transition from HRS to LRS and vice versa. To obtain the measured data shown as dots in Fig. 7, resistance states were measured with a read-out voltage of 0.2 V after applying a positive or negative voltage pulse for set and reset transition, respectively. A current of 20 mA has to be exceeded to count

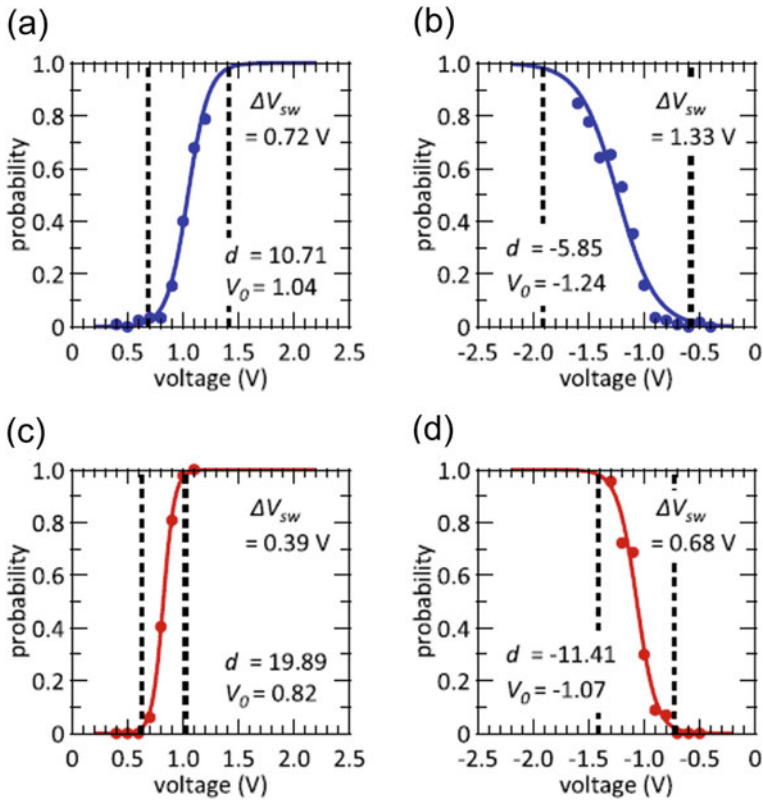


Fig. 7 Switching probability of the used RRAM devices dependent on the applied voltage amplitude of 10 ms pulses. Dots represent measured data, while the solid lines are fits with Eq. 1. The fit parameters d and V_0 , as well as the switching window ΔV_{sw} are also depicted. In **a**, **b**, the set and reset behavior of the polycrystalline devices are shown, respectively, while in **c**, **d**, the set and reset behavior of the amorphous devices are depicted, respectively. For each technology, 128 devices were measured. Reprinted from [120] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

as an effective set operation, while the current has to be lower than 5 mA to count as a successful reset process. All pulses had a length of 10 ms. The cycle-to-cycle (C2C) variability shows no significant deviation from the D2D variability in similar devices [136]. Furthermore, the switching voltages determined here do not depend on the devices' position within the 4 kbit array. Thus, taking the D2D variability into account for designing the learning rule is reasonable. The switching probability dependence on an applied voltage pulse can be described by a Poisson distribution taking voltage amplitude and pulse width into account [113, 119]. The distribution function for N voltage pulses (neural activity level) with a voltage amplitude V can be expressed as [113, 120]

$$f_N = \frac{1}{1 + e^{-d(V-V_0)}} \quad (1)$$

Here, V_0 denotes the voltage at which the probability f_N is equal to 0.5, and d is a measure of the distribution functions slope and, therefore, of the switching variability. The larger the absolute value of d , the smaller the switching window ΔV_{sw} in which a stochastic encoding of analog data is possible. The switching window is defined as the voltage interval in which the switching probability f_N is between 2 and 98 %. Fitting the measured D2D variability with Eq. 1 leads to the solid line in Fig. 7 as well as to the depicted d and ΔV_{sw} values. In summary, ΔV_{sw} is smaller for amorphous HfO_{2-x} than for polycrystalline HfO_{2-x} devices due to the grain boundaries' impact on the D2D variability and a more homogeneous defect distribution in amorphous hafnia films [135, 137]. Furthermore, ΔV_{sw} is smaller in the set transition compared to the reset transition for both technologies.

To emulate synaptic plasticity, the activity A of a neuron is encoded in voltage pulses with amplitudes V within the switching windows:

$$V = V_1 + A \cdot \Delta V, \quad (2)$$

with

$$A = \frac{N}{\Delta t}. \quad (3)$$

Here, N is the number of action potentials arriving at a neuron in the time interval Δt , and V_1 is the lower bound of the switching window. Exploiting the whole switching window to map analog activity, i.e., analog data, to the stochastic nature of the switching event is possible by a proper choice of ΔV . In the following, the StochANN utilizing this local learning rule is described, and the influence of the switching window size on the learning performance is shown.

The MNIST benchmark [99] of static visual patterns is used within this work. The learning data set contains 60,000 images of handwritten digits from 250 different writers. Each image consists of $28 \cdot 28$ greyscale pixels with 256-levels. Some representations of the ten included patterns (i.e., digits from zero to nine) are shown in Fig. 8. A test data set contains additional 10,000 images, which can be used to determine classification accuracy. For the StochANN [119, 120] described in this section, averaged images are used, as shown in Fig. 9. These are obtained by calculating the average greyscale values of 100 randomly chosen representations of each pattern. For learning, the pixel intensities of every image are mapped to the interval $[0,1]$ by dividing the values of every pixel by the maximum pixel value of the respective image. Learning takes place in a supervised manner in a time-independent StochANN. The network topology is illustrated in Fig. 9. Each pixel has one corresponding input neuron connected to every of the output neurons (one for each pattern sketched here) in a two-layer feed-forward configuration similar to earlier approaches [89–93]. Each connection is made by a binary stochastic synapse. The input neurons map the pixel

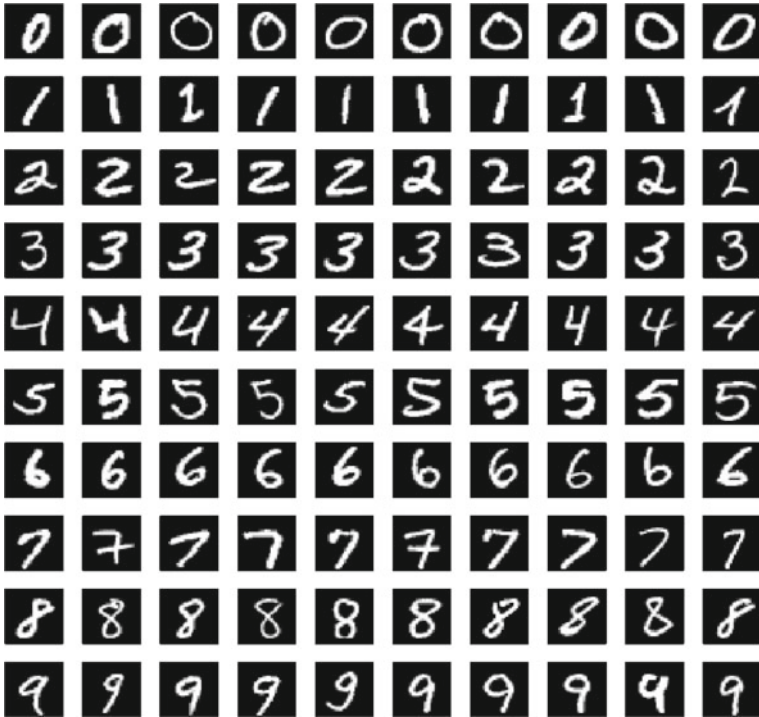


Fig. 8 Examples of MNIST images. Each row shows different representations of one pattern

intensities into voltage pulses, inducing a switching event with respective probability for the connections to the dedicated output neuron. In that way, the trained synaptic connections, i.e., memristive devices, form receptive fields of the output neurons. Either set or reset transitions of both technologies are used for learning. If the set transition is used for learning, reset pulses are applied to each synaptic device prior to execute the learning rule corresponding to a low probability p_{sat} . Accordingly, a low probability set pulse is used if the reset transition is exploited for learning. Thus, saturation effects are avoided. In that way, all training images can be used several times to train the network. It should be noted that the number of output neurons can vary for two reasons. First, if only a subset of the patterns is learned, fewer output neurons are necessary (one for each pattern). Second, each pattern can be learned by several output neurons to increase the network performance. The StochANN performance is evaluated experimentally with a mixed-signal circuit emulating the synapses in hardware using the fully CMOS-integrated RRAM arrays and neurons in software. Details about the circuit design can be found in [120]. Moreover, the network is simulated without taking device variabilities and imperfections into account. Here, the stochastic learning rule is simulated by generating a random number $r_{i,j}$ uniformly distributed over the interval (0, 1) for every pixel j of every learning image i and

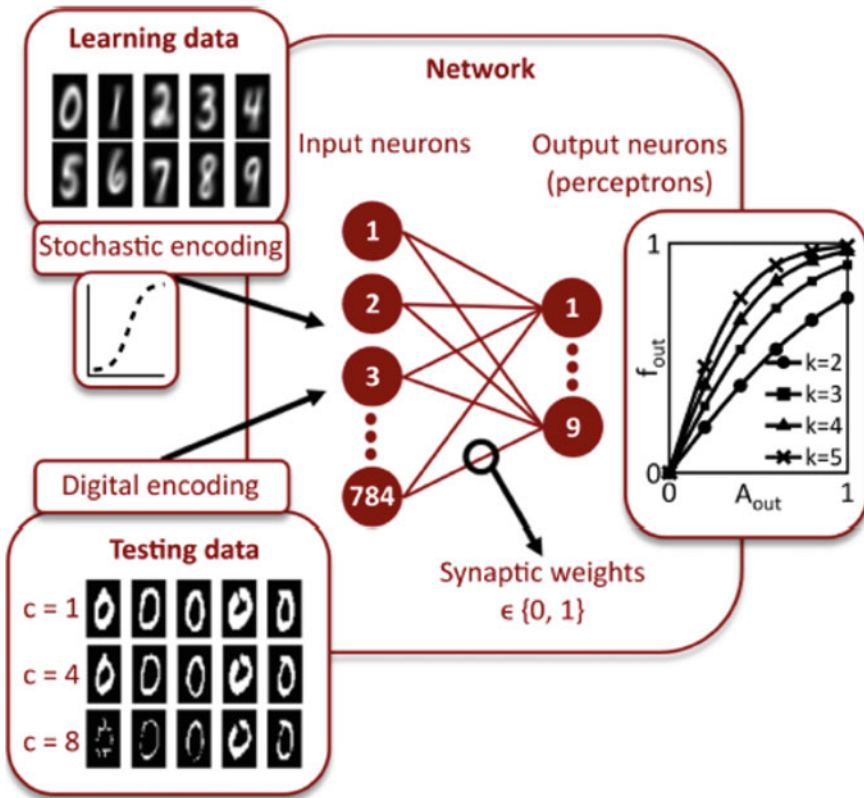


Fig. 9 Schematic visualization of the network structure (center) with 784 input neurons and 10 output neurons connected with memristive devices as binary synaptic weights. The learning data (left top) and the test data (left bottom) are sketched as well as the output neurons activation function (right). Reprinted from [120] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

the respective synaptic weight $w_{i,j}$ is set to 1 if the pixel intensity is larger than $r_{i,j}$. The simulations serve to determine the maximum possible network performance by omitting device imperfections. Furthermore, the number of available devices limits the experimentally realized network size. Thus, larger networks can be simulated to compare the stochastic learning algorithm to state-of-art networks. After learning, the classification accuracy can be evaluated by applying test images to the network. In experiments, 50 randomly chosen images of each pattern are used, while all 10,000 test images are used in simulations. Therefore, the pixel intensities are binarized. For each test image, a threshold value proportional to a global constant c and the mean pixel intensity of that image is determined. Pixel intensities, which do not exceed the threshold, are set to 0, and the others are set to 1. The larger c , the fewer pixels are active, as shown in Fig. 9. Every test image is shown once to the network by the

input neurons, which induce a 0.2 V read-out pulse for pixel values of 1. The output neurons are modeled as perceptrons with an activation function

$$f_{out} = \frac{1 - e^{-k \cdot A_{out,i,m}}}{1 + e^{-k \cdot A_{out,i,m}}}. \quad (4)$$

Here, k is a positive constant that defines the slope, as shown in Fig. 9 and $A_{out,i,m}$ is the normalized activity of the input neurons for the test image i weighted by the synaptic connections $w_{j,m}$ of input neurons j to the output neuron m according to

$$A_{out,i,m} = \frac{1}{784} \cdot \sum_{j=1}^{784} p_{i,j,bin} \cdot w_{j,m}, \quad (5)$$

where $p_{i,j,bin}$ is the binarized intensity value of pixel j being part of image i . The weights $w_{j,m}$ are determined as logical 1 if the read current exceeds 10 mA and the set transition is used for learning. If the reset transition is used, $w_{j,m}$ is assigned a logical 0 if the current is larger than 10 mA. Thus, Eq. 5 is valid for both cases. The output neuron, which receptive field re-samples the test image best, has the highest activation $A_{out,i,m}$, and associates the test image to the pattern it learned. If several output neurons are used to learn the same pattern, the sums of all activation functions belonging together are evaluated. A classification accuracy, named recognition rate in the following, is determined by calculating the percentage of correctly assigned test images.

Two MNIST subsets were used first to compare experimental results with simulations. One subset contains the digits “0”, “1” and “9”, while the second subset consists of “0”, “3” and “8”. Thus, the patterns differ more from one another in the first set compared to the second set, where the patterns have more pixels in common. In total, $3 \cdot 784 = 2,352$ individual synaptic connections are needed. This number of functional devices is selected from each type (polycrystalline or amorphous devices) of the 4 kbit chips, and they are randomly assigned to the output neurons. The parameters $c = 4$ (binarization of test images), $k = 5$ (slope of output neurons activation function in Eq. 4) and $p_{sat} = 35\%$ (for to avoid saturation of weights) were optimized in simulations. Here, recognition rates (mean value and standard deviation) of 84.5% ($\pm 4.6\%$) for the subset {0, 3, 8}, and 87.0% ($\pm 4.8\%$) for the subset {0, 1, 9} were determined with 100 simulation runs for each subset with five learning epochs each. It should be emphasized here that the algorithm converges within these five training epochs. Combining the results of both subsets, a recognition rate of 85.7% ($\pm 4.9\%$) is obtained. This is plotted in Fig. 10a as dashed line (mean value) and grey area (standard deviation). Receptive fields trained experimentally are shown in Fig. 11 for both device technologies, both subsets, and both state transitions used for learning. It is obvious that both types of devices and both transitions can be used within the network to learn the respective patterns in hardware. Furthermore, the switching windows size ΔV_{sw} affects the learned patterns. In particular, the smallest switching window corresponding to the set transition of the amorphous

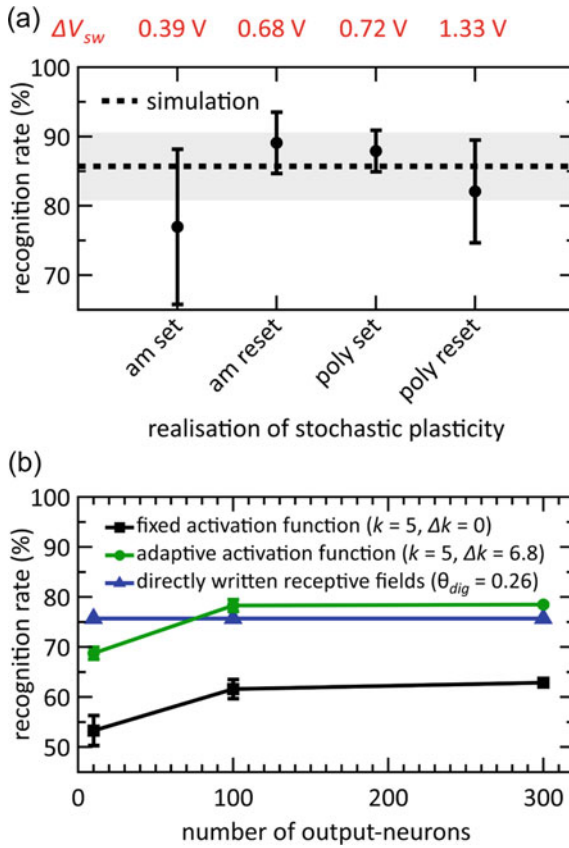


Fig. 10 Recognition rates of the StochANN. **a** Combined experimental results for MNIST subsets $\{0, 1, 9\}$ and $\{0, 3, 8\}$. Mean values and standard deviations of five experimental runs for each pattern (i.e., ten runs in total for each data point) are shown as black dots while standard deviations are given as error bars. Simulation results (100 runs for each subset) are given as dashed line (mean value) and gray area (standard deviation). The abbreviations “poly” and “am” denote the polycrystalline and the amorphous HfO_{2-x} -based devices, and “set” and “reset” denote the transition used to emulate stochastic plasticity. **b** simulation results for the whole MNIST dataset are shown (mean values and standard deviation of five runs with five learning epochs each). These results are achieved with a fixed activation function (black squares) and an adaptive activation function (green circles) of the output neurons. The recognition rates with directly written receptive fields are also shown for comparison (blue triangles). Reprinted from [120] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

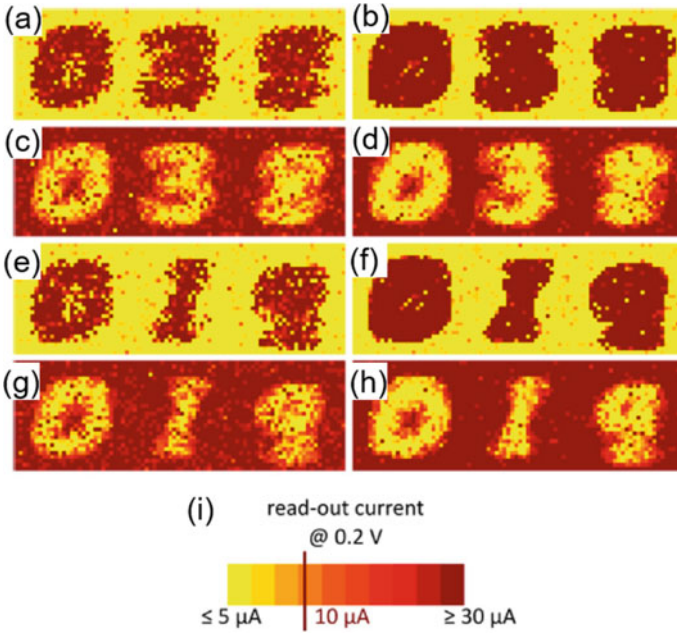


Fig. 11 Receptive fields of learned patterns in hardware. The read-out currents of the RRAM devices measured with 0.2 V are shown. In **a–d**, the used patterns are {0, 3, 8} learned with the set transition of polycrystalline devices **a**, the set transition of amorphous devices **b**, the reset transition of polycrystalline devices **c**, and the reset transition of amorphous devices **d**. In **e–h**, the used patterns are {0, 1, 9} learned with the set transition of polycrystalline devices **e**, the set transition of amorphous devices **f**, the reset transition of polycrystalline devices **g**, and the reset transition of amorphous devices **h**. Read-out currents are encoded in the pixel color as specified in **i**. Reprinted from [120] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

devices leads to receptive fields, which are more challenging to differentiate visually [Fig. 11b, f] compared to the largest switching window belonging to the reset transition of the polycrystalline devices [Fig. 11c, g]. However, the determined recognition rates show deviations from the simplified assumption that an increased switching window leads to better classification accuracy. In Fig. 10a, the recognition rates for both technologies and both state transitions are shown. Each data point denotes combined mean values and standard deviations for five experimental runs of each subset (i.e., ten experimental runs in total for each data point). On the x-axis, “am” denotes amorphous and “poly” denotes polycrystalline HfO_{2-x} devices. Furthermore, ΔV_{sw} is given in Fig. 10a. In summary, the medium-sized switching windows corresponding to the reset transition of amorphous HfO_{2-x} devices and the set transition of polycrystalline HfO_{2-x} devices show performances within the error margin of the simulations. Thus, the simulations accurately reproduce the experimental results. A larger switching window for the reset transition of the polycrystalline HfO_{2-x} devices and a narrower switching window for the set transition of the amorphous HfO_{2-x}

devices, however, result in worse recognition performance. Thus, evidence is shown that device variability has an impact on network performance. In particular, the size of the switching window must not be too small to optimize the stochastic synapses in the proposed StochANN. The reason for the worse accuracy obtained with the largest ΔV_{sw} has to be evaluated in the future.

The network performance is tested with all ten patterns in simulations as well. All results were obtained with five learning epochs in five simulation runs. The recognition rates are summarized in Fig. 10b. Black dots show the StochANN performance for a fixed slope $k = 5$ of the output neurons' activation function. Recognition rates were determined as 53.3% ($\pm 3.0\%$) for ten output neurons (i.e., one for each pattern) 61.6% ($\pm 1.9\%$) for 100 output neurons and 62.9% ($\pm 0.7\%$) for 300 output neurons. An increase in learning epochs or the number of output neurons does not lead to any improvement. Furthermore, these results are compared to patterns written directly into the synaptic states without any learning algorithm involved. For this purpose, the input patterns were binarized using a fixed threshold Θ_{bin} . The synaptic weights were set to 1 if the corresponding pixel intensity of the input images were larger than the threshold. As shown by the blue triangles in Fig. 10b, a 75.7% recognition rate was obtained for a fixed Θ_{bin} of 0.26, which was optimized in simulations. No standard deviation can be denoted since no stochasticity is involved but only one deterministic prototype of each pattern is stored. Thus, binarizing the input images and writing them directly into the receptive fields improves the performance compared to the stochastic learning rule. However, a thorough optimization of Θ_{bin} for the specific dataset has to be performed, which becomes more tedious as the number of different input patterns increases. Moreover, an adaptive slope of the output neurons' activation function in combination with the stochastic learning rule leads to even higher recognition rates [green squares in Fig. 10b]. Here, the slope k is adapted with

$$k_m = k_0 - \Delta k \cdot \frac{\sum_{j=1}^{784} w_{j,m}}{784}, \quad (6)$$

where k_0 is the base value, and Δk is a positive constant weighted by the total strength of the synaptic connections. Thus, the slope is steeper for neurons that have learned patterns with less active pixels, leading to a stronger activation of those neurons for less input strength, as can be seen in Fig. 9. The slope adaptation only depends on the final weights stored in the synaptic connection. No adaptation during learning is necessary. The adaptive slope has similar functionality to variable threshold values for the output neurons reported for other pattern recognition networks [89, 90, 92, 93, 138, 139]. Here, the adaptive thresholds are essential to obtain a high recognition performance by emulating homeostasis. With $k_0 = 5$ and $\Delta k = 6.8$, a classification accuracy of 68.8% ($\pm 1.2\%$), 78.3% ($\pm 1.2\%$) and 78.5% ($\pm 0.2\%$) were achieved for 10, 100, and 300 output neurons, respectively. This shows that the stochastic learning rule slightly outperforms the directly written receptive fields for more than 100 neurons in the output layer. This can be explained by the fact that only one

prototype of each pattern exists for directly written patterns, while variations of the prototypes exist in the learned receptive fields.

Using more complex time-independent networks performing supervised learning leads to recognition rates $> 98.5\%$ for SNNs [140, 141] and 99.87% for a thoroughly optimized CNN [142] and thus comparable to human performance estimated to be approx. 99.8% [143]. Using unsupervised learning in time-dependent neural networks is reported to achieve 93.5% with 300 output neurons [90] as well as 91.9% and 95.0% with 1600 and 6400 output neurons and the same amount of inhibitory neurons [138]. The latter approach was extended in Ref. [95] to a so-called lattice map (LM)-SNN leading to an accuracy of 94.07% for 1600 excitatory and inhibitory neurons, respectively. For a broad overview of different time-dependent and time-independent networks using supervised or unsupervised learning methods, the reader is referred to the overwhelming literature [98, 138, 142, 144–146]. All results named so far were obtained in simulations where only in [90] memristive devices were modeled to be used as synaptic connections. A fully hardware-implemented CNN based on multilevel RRAM devices can achieve recognition rates of 96.2% [147]. Here, a five-layer network is trained in software in a supervised manner, and the weights are subsequently transferred to eight 128×16 1T-1R arrays using two devices as one synapse to obtain positive and negative weights. Moreover, re-training of the last feature extraction layer was done in hardware. Another approach, in which two analog RRAM devices are used as one hardware synapse together with software neurons, reaches an accuracy of 91.7% for a re-scaled MNIST dataset of $8 \cdot 8$ pixel size. Here, a three-layer network using one array of 128×64 1T-1R devices can be utilized for learning directly in hardware using a supervised learning scheme [148]. Simulations of an extended network show a recognition rate of 97.3% taking device variability into account. A neuromorphic processor implementing a multi-layer SNN with static random access memory (SRAM) allowing on-line supervised learning reaches a recognition accuracy of 97.83% [149]. Furthermore, hardware acceleration of DNN inference with pre-trained weights transferred to PCM devices is reported to lead to a recognition rate of 98.3% [150]. Another integrated circuit utilizes memristive devices as synaptic connections with the possibility of on-line learning has also been published [151].

The StochANN performance shown here is promising for such a simple network structure but has to be improved to compete with other reported networks. One big drawback of the proposed concept is that only the averaged pattern can be learned. Transfer the supervised learning rule into an unsupervised learning approach can potentially help to extract more prototypes of each pattern [89, 90, 93] without the need for supervised learning by computing and gradually improving an error function, as done in classical backpropagation algorithms [77, 80, 81]. Furthermore, using several binary devices as one synapse can help to improve the network performance [85, 129, 131, 133]. The advantage of the proposed network is that learning can be done directly in hardware with a mature technology using fully CMOS-integrated RRAM devices as synapses.

4 Time-Dependent Neural Networks

In this section, examples of time-dependent memristive networks are covered. In Sect. 4.1, an SNN based on analog memristive devices performing bio-inspired learning and pattern recognition is presented [93]. Simulations reproducing real device behavior on learning the MNIST benchmark are provided while the impact of device variability and yield is investigated. A mixed-signal circuit implementation using real crossbar-integrated double barrier memristive devices (DBMDs) [152] to learn basal patterns experimentally is furthermore shown [94]. In Sect. 4.2, examples for oscillator computing with memristive devices are provided. According to Fig. 2, these networks show the highest degree of biological inspiration and, therefore, the highest amount of cognitive performance is expected. The oscillator-based computing scheme shown below emulates perception by transient synchronization of memristively coupled oscillators. Thereby, it establishes a certain analogy to biology to solve the binding problem [113]. The influence of the switching dynamics of two types of memristive devices on the synchronization of oscillators is, furthermore, investigated, and device requirements for oscillatory computing are deduced.

4.1 Bio-Inspired Learning with Analog Memristive Devices

In this section, a time-dependent neural network utilizing analog memristive devices to emulate bio-inspired learning for a pattern recognition task is presented. The network performance is investigated by simulations incorporating real device behavior [93] and by the realization of a mixed-signal circuit using real crossbar-integrated devices [94]. Here, LTP and LTD are induced by replicating the Hebbian learning rule described in Sect. 1 for unsupervised bio-inspired learning. Hebbian learning was already realized a decade ago with single memristive devices by emulating STDP [153, 154] as well as LTP and LTD [155].

The used devices are so-called double barrier memristive devices (DBMDs) with the layer sequence $\text{Au}/\text{Nb}_x\text{O}_y(2.5\text{nm})/\text{Al}_2\text{O}_3(1.3\text{ nm})/\text{Nb}$ [152] which are explained in detail in Chap. 3. Here, memristive switching is reported to take place by field-driven oxygen ion movement within Nb_xO_y , modulating the effective Schottky barrier height and the effective tunneling width of the Al_2O_3 [152, 156]. Thus, a homogeneous interface-based switching leading to a gradual resistance change is performed. The amount of resistance change depends on the applied voltage amplitude and time. A mathematical description of experimentally determined switching data is given by the memristive plasticity model of Ziegler et al. [24]. This model is compatible with advanced biophysical plasticity models that can fit experimental data on STDP, while it is also suitable to describe plasticity emulation with memristive devices. In that way, a behavioral model is obtained, which can be used for network-level simulations to explore how the modeled devices can be utilized to emulate Hebbian plasticity in trainable neuromorphic networks. Therefore, the degree of conductance change, i.e.,

the change of synaptic weight ω , is linked to the applied voltage pulses. The weight change is given by [24]

$$\frac{d\omega}{dt} = \beta(\omega) \omega(t) \left(1 - \frac{1}{\omega_{max}} \omega(t) \right). \quad (7)$$

Here, β is the weight-dependent learning rate, and ω_{max} is the maximum achievable weight. The switching dynamics of memristive devices are expressed in β , which depends not only on the electrical stimuli but also on the present conductance state. Thus β depends on the switching mechanism of the memristive device and can lead to various learning behaviors [24, 157]. The learning rate can be different for potentiation β_p and depression β_d . Furthermore, the synaptic weight ω represents the conductance G of a memristive device. Since the conductance change usually depends on the voltage pulse amplitude ΔV as well as the width Δt and the number of pulses n , the learning rates β_p and β_d are also modeled to be dependent on these parameters [24]:

$$\beta_p(G, n, \Delta t, \Delta V) = k_p \alpha(\Delta V) \lambda(\Delta t) (1 - \gamma G(n-1)) \quad (8)$$

$$\beta_d(G, n, \Delta t, \Delta V) = -k_d \alpha(\Delta V) \lambda(\Delta t) \gamma G(n-1), \quad (9)$$

where k_p , k_d , and γ are positive constants, while α and λ account for the non-linearity of the memristive devices' switching process.

Figure 12 shows the plasticity measurements (conductance vs. pulse number) of DBMDs investigated with voltage pulses of different amplitude and widths on 86 single devices. Black dots denote the average data from 86 individual devices, while error bars denote the standard deviation. Red solid lines show a replication with the introduced plasticity model. Model parameters are given in the respective original paper [93]. The gradual conductance modulation is clearly visible in Fig. 12a. Here, 1000 equivalent positive voltage pulses inducing potentiation and subsequent 1000 equivalent negative voltage pulses inducing depression were applied to the devices, as illustrated in the inset. A pulse duration of $\Delta t = 1$ ms was chosen together with $\Delta V = 3.9$ V and $\Delta V = -2.5$ V for potentiation and depression, respectively. Device conductance was read out by applying a voltage of 0.48 V, i.e., well below the threshold to change the device state [152], after every 100 potentiation or depression pulses. The data are depicted relative to the average maximum conductance $G_{max} = 100$ nS of all devices after 1000 potentiation pulses. To determine the variations of the device conductance in dependency on the pulse amplitude and width, potentiation pulses with ΔV between 2.4 V and 3.7 V and a fixed width of 1 ms [Fig. 12b] and Δt ranging from 1 ms to 30 ms and a fixed amplitude of 3.9 V [Fig. 12c] were used. The data points in both figures show device conductance after 1000 pulses. The impact of depression pulses is shown in Fig. 12d for 1000 voltage pulses of 30 ms length, and ΔV between -1.4 V and -2.6 V applied to previously fully potentiated devices. An asymmetry between positive and negative voltages is obvious. Moreover, the device conductance is nearly unaffected for positive voltages of 2.4 V and below

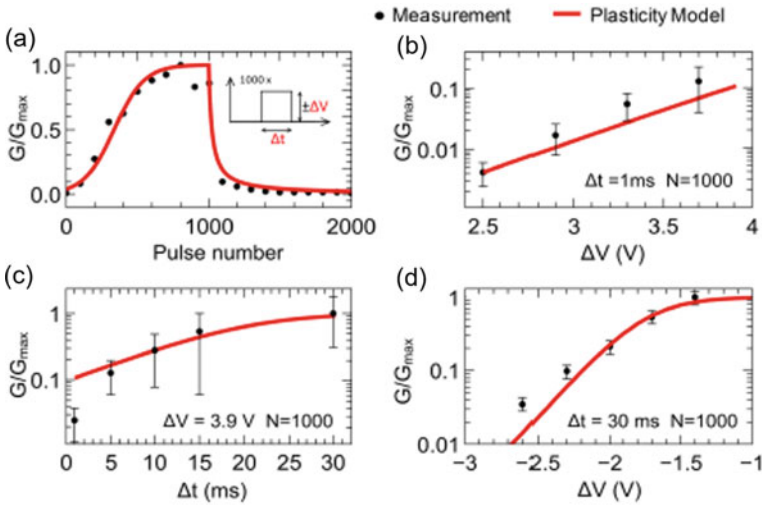


Fig. 12 Plasticity measurements of DBMDs: Mean values from 86 individual devices are given as black dots, while error bars denote the standard deviation. Conductances are normalized with the average maximum conductance G_{max} achieved with 1000 potentiation pulses of $\Delta V = 3.9$ V amplitude and $\Delta t = 1$ ms width. Fits with the plasticity model are given as red solid lines. **a** Typical conductance modulation with 1000 potentiation pulses ($\Delta V = 3.9$ V, $\Delta t = 1$ ms) and 1000 depression pulses ($\Delta V = -2.5$ V, $\Delta t = 1$ ms) is shown. Each data point depicts the normalized conductance measured after 100 pulses. The conductance change after 1000 potentiation pulses with varying pulse amplitudes and widths is shown in **b** and **c**, respectively. The conductance change with 1000 depression pulses of different amplitude is shown in **d**. Reprinted from [93] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

as well as for negative voltages with an absolute value of 1.4 V or below. By using potentiation pulses with $\Delta V = 3.9$ V, however, the conductance can be increased by two orders of magnitude, which can be fully turned back with $\Delta V = -2.6$ V. A thorough analysis of the data is given in [93].

The measured data incorporated in the plasticity model can now be used to simulate a neuromorphic network capable of learning visual patterns by adjusting the synaptic weights emulated by DBMDs. The MNIST dataset [99] of handwritten digits, which is introduced in Sect. 3.2, shall be learned. The network operates similarly as networks reported in other works [89–92]. The two-layer feedforward network is schematically shown in Fig. 13 [93]. Here, each input neuron (blue circles) is connected to every output neuron (red circles) by DBMDs (symbols of memristive devices) arranged in a crossbar array. Every input neuron stochastically encodes the intensity of one pixel into voltages pulses [91–93]. Therefore, the pixel intensities are normalized to the interval $(-1, 1)$. The absolute values of the normalized intensities denote the probability of an input voltage pulse generation, while the sign stands for the voltage polarity. In that way, every input neuron either generates no spike or voltages pulses of $+0.6$ V or -0.6 V amplitude, both not affecting the device conductance by themselves. The currents flowing to the output neurons depend not only on

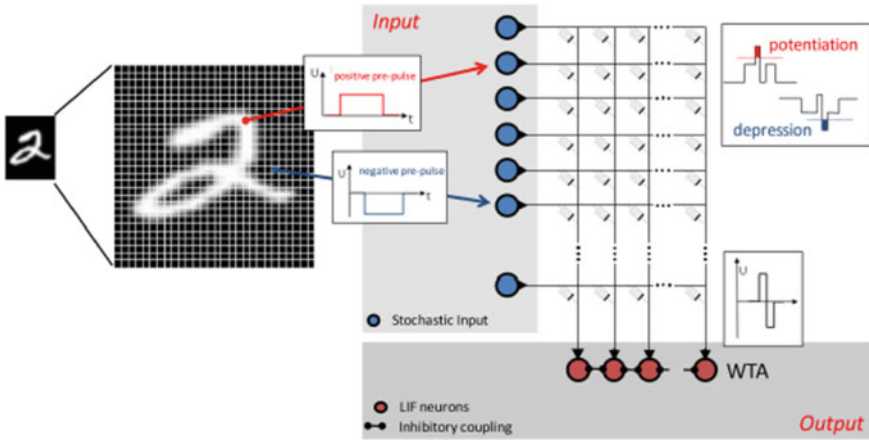


Fig. 13 Schematic of the simulated neural network. Reprinted from [93] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

the voltage pulse amplitudes but also on the conductance of the devices. The leaky integrate-and-fire (LIF) output neurons [33, 158] integrate the incoming stimuli until a certain threshold is reached [89]. When an output neuron reaches its threshold, a voltage pulse consisting of a positive part with $\Delta V = 2.9$ V and a negative part with $\Delta V = -2.3$ V. This post-synaptic pulse overlaps with the pre-synaptic pulses and changes the conductance of memristive devices connected to the respective output neuron. If a positive pre-synaptic pulse superimposes with the post-synaptic pulse, a potentiation takes place for the positive part ($V_{sum} = 3.5$ V) while the negative part does not affect the memristive state significantly ($V_{sum} = -1.7$ V). Vice versa, a net depression takes place when a negative input pulse overlaps with an output pulse ($V_{sum} = 2.3$ V and $V_{sum} = -2.9$ V, respectively). If no input pulse occurs, the post-synaptic pulse's voltage amplitudes alone do not significantly impact device conductance. Thus, input pixels with a strong intensity lead to positive input pulses, which, superimposed with induced output pulses, increase the device conductance. Pixels with low intensity induce negative input pulses, which lead to decreased conductance if an output spike simultaneously occurs. Thus, prototypes of the patterns are stored in the resistance states of the memristive devices. All devices connected to an output neuron are building the receptive field of this specific neuron. In that way, unsupervised associative learning based on local Hebbian plasticity is realized. Essential for network performance is, furthermore, an inhibitory coupling network for the output neurons implementing a winner-takes-it-all (WTA) mechanism, in which the first spiking neuron resets the integration of all other neurons [89]. Moreover, an adapting individual threshold is implemented for the output neurons to allow that all output neurons participate equivalently in learning. This mimics homeostasis in biological systems [89].

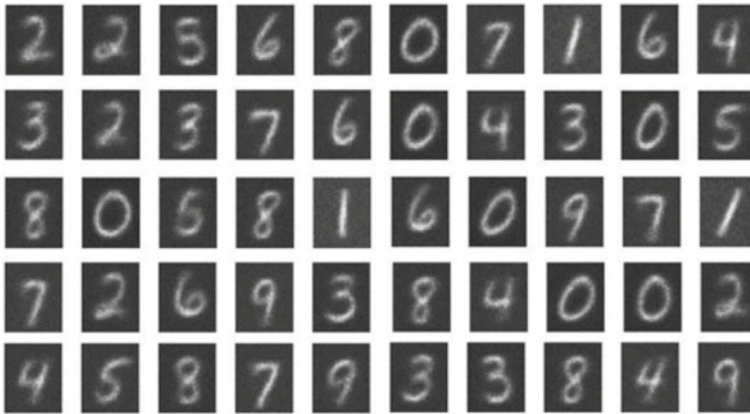


Fig. 14 Obtained receptive fields after unsupervised learning with 50 output neurons. In the gray-scale used to visualize device conductance white indicates maximum device conductance (strong synaptic weight), while black represents minimum conductance values (weak synaptic weight) of the memristive devices. Reprinted from [93] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

After learning, the network can be used to classify unknown images. Therefore, the output neuron, which receptive field matches the input image best, creates an output pulse. To assign every output neuron to its learned pattern, a small amount of pre-classified images is applied to the network and it is evaluated for which pattern the output neurons get activated. Afterward, the network performance can be evaluated by applying all 10,000 images from the MNIST test dataset to the network and calculating recognition accuracy by determining the percentage of correctly assigned patterns. Therefore, only the pre-synaptic pulses to encode the images are used, and the spiking events of the output neurons are tracked while post-synaptic pulse generation is suppressed to stop changing the device states. For 10, 20, 50, and 100 output neurons, recognition rates of 65%, 70%, 77%, and 82%, respectively, were determined. These rates are in good agreement with similar networks [89–92]. A typical set of learned receptive fields obtained in a simulation with 50 output neurons is shown in Fig. 14. It can be seen that the implemented network using the Hebbian learning scheme is able to learn different prototypes of all ten patterns (digits from zero to nine). The obtained performance is significantly lower than those from other spiking networks, as described in Sect. 3.2. However, network requirements for using memristive devices in SNNs are examined in this work, while improving pattern recognition computing schemes was not intended.

The results presented so far were generated without taking device variability into account. Now, the impact of different degrees of D2D and C2C variability is investigated [93]. Therefore, networks with ten output neurons were trained with three iterations of the whole MNIST learning dataset. Results are shown in Fig. 15(a,b), in which every data point was determined as the average of three simulation runs. The findings are then compared in Fig. 15c to the measured variability of real

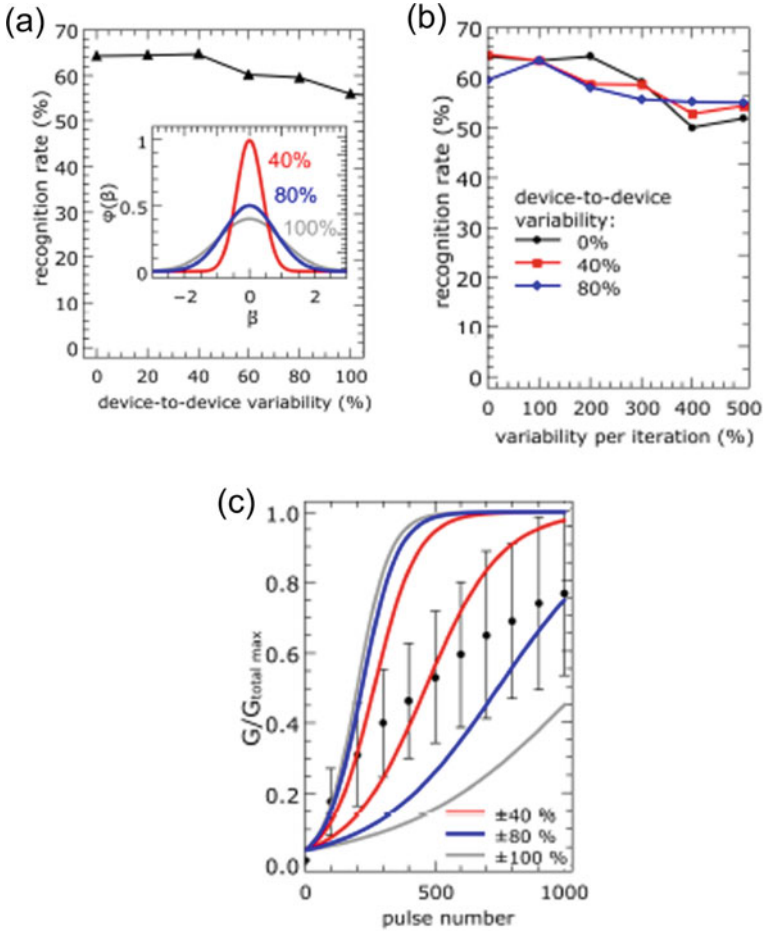


Fig. 15 Impact of device variability, i.e., variability of local learning rate of individual devices, on network performance. Device-to-device (D2D) and cycle-to-cycle (C2C) variability are depicted as standard deviation of Gaussian distributions. Each data point represents the mean value of three simulation runs. **a** D2D variability **b** C2C variability for a D2D variability of 0% (black dots), 40% (red squares) and 80% (blue diamonds) **c** Experimentally determined D2D variation for potentiation with $\Delta V = 3.9$ V and $\Delta t = 1$ ms. Each data point shows the normalized mean conductance of 86 devices after 100 potentiation pulses. The red, blue, and gray lines indicate the range of learning rates with, 40%, 80%, and 100% D2D variability, respectively. Reprinted from [93] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

DBMDs. To model D2D variability, the learning rates of all devices were varied with Gaussian distributions [see inset of Fig. 15a] initially in every simulation run. As depicted in Fig. 15a, a standard deviation of up to 50% does not affect the recognition rate, while a further increase in variability slightly influences the performance. The C2C variability was modeled by varying the individual devices' learning rate for every applied image with a Gaussian distribution. Figure 15b (black dots) shows that a C2C variability of up to 200% does not affect the network performance. A combination of D2D and C2C variability, the most realistic scenario, is given in Fig. 15b with red squares and blue diamonds. Robust performance is achieved for a C2C variability of 100% combined with a D2D variability of 80%. Furthermore, the numerical investigation provides evidence that the most crucial performance losses result from a constant D2D variability since this effect does not average out in many learning iterations like it is the case for C2C variability. To estimate if the D2D variability of DBMDs does allow to use them as artificial synapses in the investigated network, the measured D2D variability is compared to the theoretically obtained boundaries. Figure 15c shows experimentally recorded device conductance of 86 individual devices after every 100th potentiation pulse with amplitudes of 3.9 V and 1 ms normalized by the highest recorded conductance $G_{max,total}$. Furthermore, the solid lines in Fig. 15c indicate the variation range of learning rates with 40%, 80%, and 100% D2D variability. Thus, the experimentally obtained D2D variability lies within the required variation interval. Moreover, in the original paper [93], the experimentally obtained yield, i.e., the percentage of functional devices, is shown to be approx. 98%, which does not influence the network performance significantly. In conclusion, evidence is provided that DBMDs are attractive candidates to be used as artificial synapses in neuromorphic circuits. In particular, the gradual conductance change under voltage pulsing, as well as the variability and yield of such real devices are believed to be suitable for the investigated network and learning rule.

The possibility of using DBMDs as artificial synapses has also been shown experimentally [94]. Here, a two-layer network, like described above, has been implemented in a mixed-signal-circuit. The synaptic connections were emulated with a real crossbar array containing $16 \cdot 16 = 256$ devices. Due to the high I - V non-linearity and the diode-like character of the DBMDs, no additional selector devices are needed to avoid the sneak path problem (see Sect. 2 and Chap. 3). The I - V characteristics of crossbar integrated devices were proven to be similar to single devices [94], as it is also shown in Chap. 3. The neurons were emulated in software. During learning, the conductances measured with a pre-synaptic pulse of 0.9 V amplitude were used to compute the current flowing to the output neurons. If a post-synaptic spike is triggered, potentiation and depression pulses with $\Delta V = 3.6$ V and $\Delta t = 100$ ms as well as $\Delta V = -1.1$ V and $\Delta t = 300$ ms, respectively, were used. In that way, the overlapping of pre- and post-synaptic pulses was not realized, but the potentiation and depression took place according to the Hebbian learning scheme explained above. This deviation from simulations was needed because the simple circuit design did not allow for real parallel data processing [94]. However, the work aimed to provide a proof of principle that DBMDs can be used in a crossbar configuration without the need for additional selector devices as artificial synapses in neuromorphic networks.

Fig. 16 Experimental results **a** Used training data **b** Obtained receptive fields during unsupervised learning with five output neurons. The pixel color indicates the memristive devices' resistance values (synaptic weights). Reprinted from [94] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

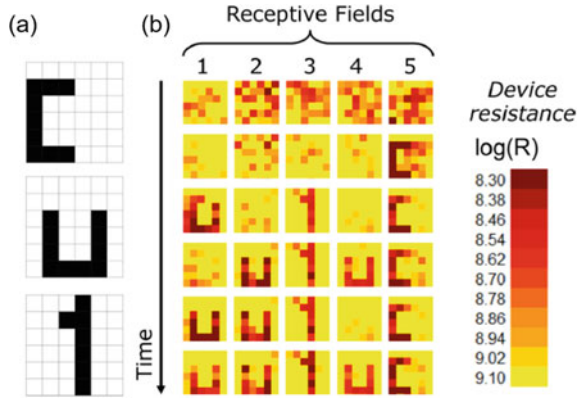


Figure 16a shows the simple 6 · 6 pixel pattern that the network shall learn. Figure 16b shows the developing receptive fields for five output neurons. Every pixel represents the resistance of one memristive device emulating the synaptic connection between one input and one output neuron. Initially, the resistances encoded in the pixel color are randomly distributed. After using 22,000 learning images in total, all three patterns are learned. Thus, the realization of unsupervised bio-inspired learning was possible with real DBMDs arranged in a crossbar structure thanks to the high $I-V$ non-linearity and the diode-like character, as well as due to no required initial electro-forming step. As shown in the simulations above, the presented system can, in principle, cope with more complex tasks. However, a much larger amount of memristive cells is necessary for that. Due to the high resistances of DBMDs (even in LRS) and the several orders of magnitude lower wiring resistance ($\approx 100 \Omega$ for an individual wire with the size $1100 \cdot 40 \cdot 0.5 \text{ mm}^3$) in the present crossbar array, larger arrays are believed to work as well.

4.2 Oscillatory Computing

Biological information processing relies heavily on nonlinear dynamics [36, 109]. This enables the integration of the multitude of information in an enormous and massively parallel network of neurons divided into functionally specialized regions such as the visual cortex, auditory cortex, or dorsolateral prefrontal cortex. Each of these regions participates as a context-dependent, self-organized, and transient subnetwork [36, 108]. Even if the underlying mechanisms are only partially understood, the interaction between dynamics and topology has been identified as one of the essential building blocks of information processing in the brain in recent years [109]. In the current understanding, it is assumed that information is encoded into coherent states by temporally correlated neural activity patterns [110]. This concept offers, particularly, an elegant explanation for the binding problem - the question

of the mechanism of sensory integration, which allows our brain to construct uniform perceptions from the multitude of sensory information. First, evidence of these concepts could have been gathered from experiments with sensorimotor networks [111]. More recent studies have shown the universality of these concepts for the entire brain [110]. Recently, oscillators coupled by memristive devices have been shown to emulate this kind of information processing to some extent [112, 113].

In this section, we report on oscillator computing with memristive devices. We show how the dynamics of oscillator networks, coupled by memristive devices, is affected by the resistance of these devices. Therefore, important requirements for memristive devices are discussed as well as applications with the possibility to open up new pathways towards the construction of cognitive electronics.

4.2.1 Oscillator Computing with Memristive Connectivity

From studies of the thalamocortical system, Hoppensteadt and Izhikevich proposed a computational scheme based on oscillators with different frequencies that are weakly coupled to an externally changed medium, causing dynamic connectivity [159]. In their model, information is encoded in the oscillators' phase and/or frequency synchrony. The weak coupling, thereby, allows a dynamic change of their connectivity patterns depending on an external signal. A similar approach follows the idea of memristive coupled oscillator structures [112, 113], which will be explained in the following.

Figure 17 shows the model of two memristively coupled oscillators. Both oscillators are initially oscillating in their own frequencies f_i and f_j [Fig. 17a]. As long as their coupling is weak, they are not affecting each other. However, if the coupling strength between the oscillators increases, they start to interfere and synchronize in frequency and phase for sufficient high coupling strengths [Fig. 17b]. However, if the coupling strength is decreased thereafter, the oscillators will desynchronize again due to their different frequencies. This model can be realized with two self-sustained van der Pol oscillators with resistive coupling, as shown in Fig. 17c [112], i.e., via the conductance g_m of memristive devices. While in the following the model of memristive coupled oscillators will be discussed in the framework of van der Pol oscillators, any other type of oscillator may also be suitable [160]. The oscillator system shown in Fig. 17 can be described by the following set of second-order dimensionless nonlinear equations:

$$\frac{d^2 y_1}{dt^2} = -\alpha (1 - y_1^2) \frac{dy_1}{dt} - \beta \frac{(y_1 + \gamma_1)^2}{\gamma_1} = g_m(x, t) \left(\frac{dy_2}{dt} - \frac{dy_1}{dt} \right) \quad (10)$$

$$\frac{d^2 y_2}{dt^2} = -\alpha (1 - y_2^2) \frac{dy_2}{dt} - \beta \frac{(y_2 + \gamma_1)^2}{\gamma_1} = g_m(x, t) \left(\frac{dy_1}{dt} - \frac{dy_2}{dt} \right) \quad (11)$$

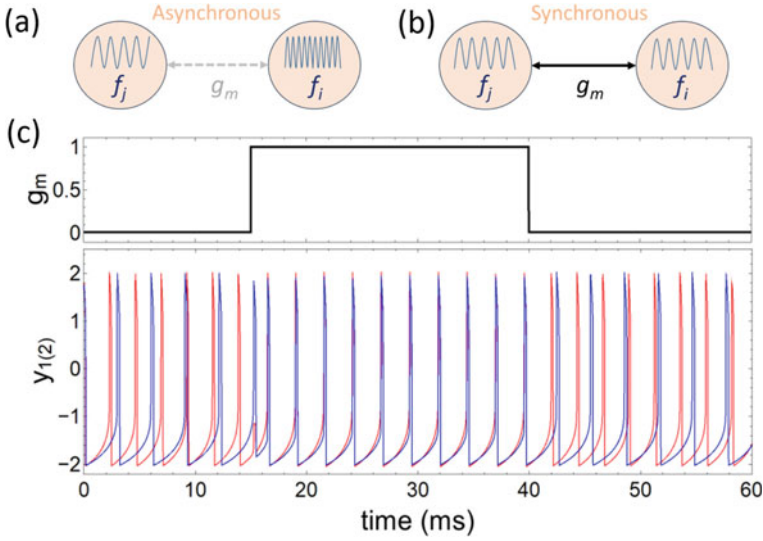


Fig. 17 **a, b** Concept of memristively coupled oscillators. **c** Depending on the coupling strength g_m (conductance of the memristive device) a frequency and phase synchronization of the oscillators occurs, as seen from the time course of the oscillator voltage y_1 (y_2) in the lower graph

Here, β , γ , and α are positive constants that define the uncoupled oscillators' damping, non-linearity, and frequency behaviors, respectively. Furthermore, g_m is the mutual coupling, representing the conductance of the memristive device. The memristive device can be modeled via

$$I = g_m(x, t) \cdot V \quad \text{with} \quad \frac{dx}{dt} = f(x, V, t). \quad (12)$$

Here, x is the memristive state variable, V the voltage across the memristive device (for example, $V = y_1 - y_2$), and f a dynamic function describing the voltage-driven atomic reconfiguration in the particular material system [160]. The device conductance influences the coupling strength via $g_m(x, t) = x \cdot G_{on} + (1 - x) \cdot G_{off}$, where G_{on} and G_{off} are the maximum and minimum conductance of the device, respectively. The obtained result is shown in Fig. 17c. Here, the change of the coupling strength [upper graph in Fig. 17c] synchronizes the initially asynchronous oscillators. Thus, memristively coupled oscillators allow emulating the initially described essential principle of biological information processing: synchronization (information encoding into coherent states by correlated neural activity) and memory (change of connectivity).

4.2.2 Memristive Devices for Oscillator Computing

Using suitable memristive devices is an important point for realizing the previously presented model in hardware. The typical characteristics of memristive devices can be roughly divided into two classes. The first class of devices shows an abrupt resistance jump at a specific voltage (type 1), while the second class exhibits a gradual resistance change under an applied voltage (type 2). In Fig. 18, representatives of the two classes of memristive devices are compared. The first device (type 1) has the layer sequence Al/TiO_x/Ag and is an electro-chemical metallization cell (ECM), while the second device (type 2) is composed of the layer sequence TiN/TiO_x/HfO_x/Au, and its resistance change is functionally based on a valence-change mechanism (VCM). For the latter, a bi-layer oxide structure has been used, which is known to stabilize the resistance switching mechanism [161, 162]. Figure 18b shows measurements for the two types of devices. While for type 1 cell an abrupt jump in the resistance is observed, and the investigated type 2 device shows a more gradual transition of the conductance under voltage cycling. To analyze the consequences of the device type on the dynamics of coupled oscillators, two scenarios were simulated: (i) an abrupt and fast change of the coupling strength, when a threshold voltage is reached (behavior type 1) and (ii) a gradual change of the coupling strength, when the threshold value $V_{th,p}$ is exceeded (grey line in Fig. 18). Furthermore, the bi-directional switching behavior was taken into account such that g_m is reduced again when the voltage falls below a negative threshold value $V_{th,n}$. The obtained results are shown in Fig. 18c. It can be seen that the binary switching devices show a much faster synchronization.

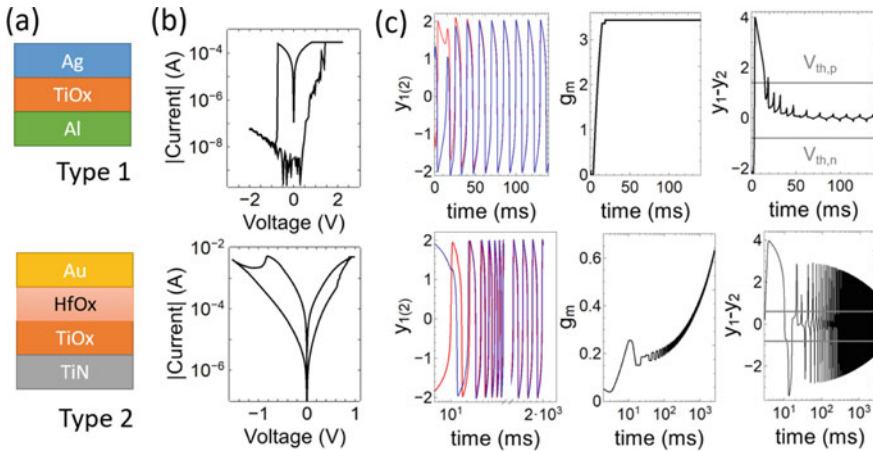


Fig. 18 a Typical memristive devices with archetypal switching characteristics. b I - V characteristics of an Al/TiO_x/Ag cell (type1) and TiN/TiO_x/HfO_x/Au (type2). c Simulation of the dynamics of the two coupled oscillators (y_1 and y_2) via a memristive device (conductance g_m) with an abrupt change in resistance (red) and a gradual change (blue). Here $y_1 - y_2$ represents the voltage across the memristive device. The threshold values for set and reset of the devices are shown in gray

4.2.3 Experimental Realization and Networks

A hardware implementation of memristively coupled oscillators is shown in Fig. 19. Here, van der Pol oscillators are realized via a programmable unijunction transistor (PUT) based circuit, as described in detail in Ref. [112]. In this circuit, the two oscillators are coupled via a memristive device so that their voltage oscillations (at points u_j and u_i) are smaller than the thresholds of the devices. Thus, the voltage oscillations cannot change the resistance value of the memristive device [see Fig. 19a]. The ECM cell shown in Fig. 18a was used as the memristive device, which changes its conductance g_m depending on an external voltage signal u_{learn} . As shown in Fig. 19b, the oscillators can be synchronized with this circuit depending on g_m . For the example shown in Fig. 19b, a voltage sequence of seven voltage pulses of 2.2 V amplitude was applied to the memristive device. This yields a conductance change of the device.

The transition to a simple artificial neural network is shown in Fig. 19c and d. The four oscillators of the first layer are connected to the two oscillators a and b of the second layer. Initially, all six oscillators of the network oscillate in their own frequency, as indicated in the contour plot by the different colors [Fig. 19c]. By applying a voltage sequence u_{learn} , the connectivity matrix g_m between the oscillators changes, which strengthens the coupling between oscillators 1, 2, and a as well as oscillators 3, 4, and b . As a result, the system oscillates with only two frequencies, as shown in Fig. 19d (different colors in contour plot). This computing scheme

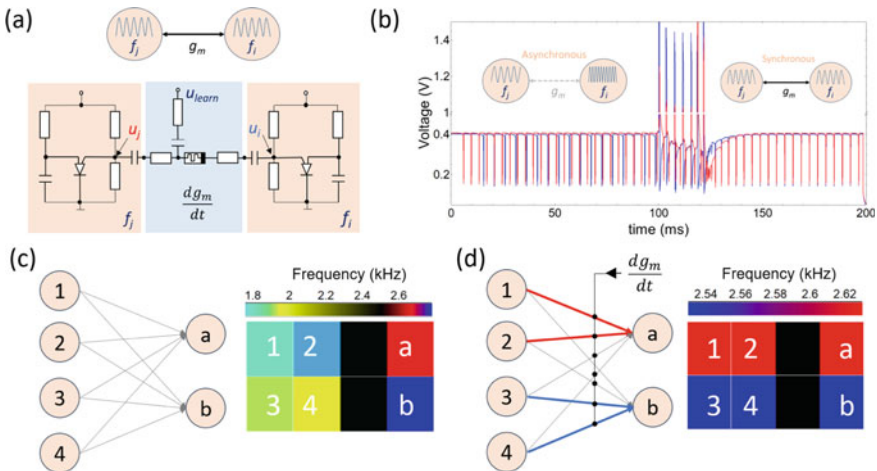


Fig. 19 **a** Experimental realization of memristively coupled van der Pol oscillators. **b** Voltage courses u_j and u_i are labeled in the circuit diagram in **a**. The voltage train u_{learn} (starting at 100 ms) decreases the resistance of the memristive device and the previously asynchronous oscillators synchronize. The ECM cell shown in Fig. 18a was used as memristive device. **c** Network of 6 van der Pol oscillators. By changing the coupling matrix g_m as a consequence of altered external conditions, the oscillators are synchronized in two frequencies (red and blue) **d**

was used in Ref. [113] to emulate perception. Thereby, u_{learn} can be identified as a level of perception that strengthens or weakens different connections depending on a temporal change of perception. For that, the inherent stochasticity of the memristive devices was used. In this respect, the number of voltage pulses defines the probability that the memristive device changes its resistance state. Thus, the number of voltage pulses of u_{learn} simulates the level of perception. Therefore, a certain analogy to biology can be established in the solution of the binding problem [113].

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 434434223 - SFB 1461. Moreover, the financial support by the DFG via the Research Unit 2093 is acknowledged. We thank Nora Kohlstedt for preparing Fig. 1(a), i.e. the sketch of a neuron.

References

1. Yang, J.J., Strukov, D.B., Stewart, D.R.: Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13–24 (2013)
2. Schuman, C.D., Potok, T.E., Patton, R.M., et al.: A survey of neuromorphic computing and neural networks in hardware (2017). [arXiv:1705.06963](https://arxiv.org/abs/1705.06963) [cs]. [arXiv: 1705.06963](https://arxiv.org/abs/1705.06963) [cs]
3. Burr, G.W., Shelby, R.M., Sebastian, A., et al.: Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2**, 89–124 (2017)
4. Merolla, P.A., Arthur, J.V., Alvarez-Icaza, R., et al.: A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**, 668–673 (2014)
5. Radamson, H.H., Zhu, H., Wu, Z., et al.: State of the art and future perspectives in advanced CMOS technology. *Nanomaterials* **10**, 1555 (2020)
6. Frenkel, C., Bol, D., Indiveri, G. Bottom-up and top-down neural processing systems design: neuromorphic intelligence as the convergence of natural and artificial intelligence (2021). [arXiv: 2106.01288](https://arxiv.org/abs/2106.01288) [cs]
7. Moore, G.E.: Cramming more components onto integrated circuits. *Electronics* **38**, 4 (1965)
8. Dennard, R., Gaensslen, F., Yu, H.-N., et al.: Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE J. Solid-State Circuits* **9**, 256–268 (1974)
9. Backus, J.: Can Programming be liberated from the von neumann style? A functional style and its algebra of programs. *Commun. ACM* **21**, 613–641 (1978)
10. Mead, C., Ismail, M.: *Analog VLSI Implementation of Neural Systems First*. Springer, Boston (1989)978-1-4613-1639-8
11. Mead, C.: Neuromorphic electronic systems. *Proc. IEEE* **78**, 1629–1636 (1990)
12. Chicca, E., Stefanini, F., Bartolozzi, C., et al.: Neuromorphic electronic circuits for building autonomous cognitive systems. *Proc. IEEE* **102**, 1367–1388 (2014)
13. Indiveri, G., Liu, S.-C.: Memory and information processing in neuromorphic systems. *Proc. IEEE* **103**, 1379–1397 (2015)
14. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7**, 011305 (2020)
15. Christensen, D.V., Dittmann, R., Linares-Barranco, B., et al.: Roadmap on neuromorphic computing and engineering (2022). [arXiv:2105.05956](https://arxiv.org/abs/2105.05956) [cond-mat]. [arXiv: 2105.05956](https://arxiv.org/abs/2105.05956) [cond-mat] (2021)
16. Kandel, E.R.: *Principles of Neural Science*, 5th edn. McGraw-Hill, New York (2013)978-0-07-139011-8
17. Luo, L.: *Principles of Neurobiology*. Garland Science, New York (2015)978-0-8153-4492-6

18. Birkoben, T., Kohlstedt, H., Matter mind matter (2022). [arXiv:2204.12774](https://arxiv.org/abs/2204.12774) [cond-mat, q-bio]. [arXiv: 2204.12774](https://arxiv.org/abs/2204.12774) [cond-mat, q-bio]
19. Bi, G.-Q., Poo, M.-M.: Synaptic modification by correlated activity: Hebb's Postulate Revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001)
20. Izhikevich, E.: Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**, 1569–1572 (2003)
21. Markram, H., Gerstner, W., Sjöström, P.J.: Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synapt. Neurosci.* **4**, 2 (2012)
22. Hebb, D.O.: *The Organization of Behavior: A Neuropsychological Theory*. Psychology Press, New York (2002)978-1-4106-1240-3
23. Gerstner, W.: To Appear in: *From Neuron to Cognition via Computational Neuroscience*, M. Arbib, J. Bonaiuto (eds.), vol. 21. MIT Press, Cambridge (2011)
24. Ziegler, M., Riggert, C., Hansen, M., et al.: Memristive Hebbian plasticity model: device requirements for the emulation of Hebbian plasticity based on memristive devices. *IEEE Trans. Biomed. Circuits Syst.* **9**, 197–206 (2015)
25. Bi, G.-Q., Poo, M.-M.: Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10464–10472 (1998)
26. He, B.J.: Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**, 480–487 (2014)
27. Izhikevich, E.M.: *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. First MIT Press paperback edition. MIT Press, Cambridge (2010). ISBN:978-0-262-51420-0
28. Schroeter, M.S., Charlesworth, P., Kitzbichler, M.G., et al.: Emergence of rich-club topology and coordinated dynamics in development of hippocampal functional networks in vitro. *J. Neurosci.* **35**, 5459–5470 (2015)
29. Stein, R.B., Gossen, E.R., Jones, K.E.: Neuronal variability: noise or part of the signal? *Nat. Rev. Neurosci.* **6**, 389–397 (2005)
30. Smetters, D.K., Zador, A.: Synaptic transmission: noisy synapses and noisy neurons. *Curr. Biol.* **6**, 1217–1218 (1996)
31. Zhang, C., Zhang, D. Stepanyants, A.: Noise in neurons and synapses enables reliable associative memory storage in local cortical circuits. *eNeuro* **8** (2021)
32. Buzsáki, G., Draguhn, A.: Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004)
33. Gerstner, W., Kistler, W.M.: *Spiking Neuron Models: Single Neurons, Populations. Plasticity*. Cambridge University Press, Cambridge (2002)978-0-521-89079-3
34. Ernst, U., Pawelzik, K., Geisel, T.: Synchronization induced by temporal delays in pulse-coupled oscillators. *Phys. Rev. Lett.* **74**, 1570–1573 (1995)
35. Hoppensteadt, F.C., Izhikevich, E.M.: *Weakly Connected Neural Networks Applied Mathematical Sciences*, vol. 126. Springer, New York (1997)978-0-387-94948-2
36. Buzsáki, G.: *Rhythms of the Brain*. Oxford University Press, New York (2006)978-0-19-530106-9
37. Arenas, A., Díaz-Guilera, A., Kurths, J., et al.: Synchronization in complex networks. *Phys. Rep.* **469**, 93–153 (2008)
38. Galizia, C.G., Lledo, P.-M. (eds.): *Neurosciences: From Molecule to Behavior: A University Textbook*. Springer Spektrum, Heidelberg (2013)978-3-642-10768-9
39. Gerstner, W., Kistler, W.M., Naud, R., et al.: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge (2014)978-1-107-06083-8
40. Amil, P., Cabeza, C., Masoller, C., et al.: Organization and identification of solutions in the time-delayed mackey-glass model. *Chaos* **25**, 043112 (2015)
41. Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd edn. Westview Press, a Member of the Perseus Books Group, Boulder (2015). ISBN: 978-0-8133-4910-7

42. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘Small-World’ networks. *Nature* **393**, 440–442 (1998)
43. Uhlhaas, P., Pipa, G., Lima, B., et al.: Neural synchrony in cortical networks: history, concept and current status. *Front. Integrat. Neurosc.* **3** (2009)
44. Beer, R.D.: Dynamical approaches to cognitive science. *Trends Cogn. Sci.* **4**, 91–99 (2000)
45. Höfflinger, B.: CHIPS 2020 VOL. 2: New Vistas in Nanoelectronics. Springer International Publishing, Cham (2016). ISBN: 978-3-319-22093-2
46. Veendrick, H.: Nanometer Cmos Ics. Springer Science+Business Media, LLC, New York (2017)978-3-319-47595-0
47. Turing, A.M.I.: Computing Machinery and Intelligence. *Mind* **LIX**, 433–460 (1950)
48. Rueckert, U.: CHIPS 2020 VOL. 2: New Vistas in Nanoelectronics, Höfflinger, B., (ed.), pp. 249–274. Springer International Publishing, Cham (2016). ISBN: 978-3-319-22093-2
49. Ielmini, D., Waser, R.: Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications. Wiley-VCH Verlag GmbH Co. KGaA, Weinheim (2016)978-3-527-33417-9
50. Lanza, M., Waser, R., Ielmini, D., et al.: Standards for the characterization of endurance in resistive switching devices. *ACS Nano* **15**, 17214–17231 (2021)
51. Bian, H., Goh, Y.Y., Liu, Y., et al.: Stimuli-responsive memristive materials for artificial synapses and neuromorphic computing. *Adv. Mater.* **33**, 2006469 (2021)
52. Wang, R., Yang, J.-Q., Mao, J.-Y., et al.: Recent advances of volatile memristors: devices, mechanisms, and applications. *Adv. Intell. Syst.* **2**, 2000055 (2020)
53. Di Ventra, M., Pershin, Y.V.: The parallel approach. *Nat. Phys.* **9**, 200–202 (2013)
54. Truong, S.N., Min, K.-S.: New memristor-based crossbar array architecture with 50- multiplication of analog neuromorphic computing. *J. Semicond. Technol. Sci.* **14**, 356–363 (2014)
55. Ielmini, D., Wong, H.-S.P.: In-memory computing with resistive switching devices. *Nat. Electron.* **1**, 333–343 (2018)
56. Maass, W.: Networks of spiking neurons: the third generation of neural network models. *Neural Netw.* **10**, 1659–1671 (1997)
57. Ziegler, M., Wenger, C., Chicca, E., et al.: Tutorial: concepts for closely mimicking biological learning with memristive devices: principles to emulate cellular forms of learning. *J. Appl. Phys.* **124**, 152003 (2018)
58. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R., et al.: Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020)
59. Borghetti, J., Snider, G.S., Kuekes, P.J., et al.: Memristive’ switches enable ‘stateful. logic operations via material implication. *Nature* **464**, 873–876 (2010)
60. Xu, N., Fang, L., Kim, K.M., et al.: Time-efficient stateful dual-bit-memristor logic. *physica status solidi (RRL) – Rapid Res. Lett.* **13**, 1900033 (2019)
61. Li, S., Xu, C., Zou, Q., et al.: Pinatubo: A Processing-in-Memory Architecture for Bulk Bitwise Operations in Emerging Non-Volatile Memories in 2016 53rd ACM/EDAC/IEEE Design Automation Conference (DAC), pp. 1–6 (2016)
62. Maan, A.K., Jayadevi, D.A., James, A.P.: A survey of memristive threshold logic circuits. *IEEE Trans. Neural Netw. Learn. Syst.* **28**, 1734–1746 (2017)
63. Wright, C.D., Liu, Y., Kohary, K.I., et al.: Arithmetic and biologically-inspired computing using phase-change materials. *Adv. Mater.* **23**, 3408–3413 (2011)
64. Le Gallo, M., Sebastian, A., Mathis, R., et al.: Mixed-precision in-memory computing. *Nat. Electron* **1**, 246–253 (2018)
65. Zidan, M.A., Jeong, Y., Lee, J., et al.: A general memristor-based partial differential equation solver. *Nat. Electron* **1**, 411–420 (2018)
66. Sun, Z., Pedretti, G., Ambrosi, E., et al.: Solving matrix equations in one step with cross-point resistive arrays. *PNAS* **116**, 4123–4128 (2019)
67. Pershin, Y.V., Di Ventra, M.: Solving mazes with memristors: a massively parallel approach. *Phys. Rev. E* **84**, 046703 (2011)
68. Mostafa, H., Müller, L.K., Indiveri, G.: An event-based architecture for solving constraint satisfaction problems. *Nat. Commun.* **6**, 8941 (2015)

69. Bojnordi, M.N. Ipek, E.: Memristive Boltzmann Machine: A Hardware Accelerator for Combinatorial Optimization and Deep Learning in 2016 IEEE International Symposium on High Performance Computer Architecture (HPCA), pp. 1–13 (2016)
70. Li, C., Hu, M., Li, Y., et al.: Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **1**, 52–59 (2018)
71. Le Gallo, M., Sebastian, A., Cherubini, G., et al.: Compressed sensing with approximate message passing using in-memory computing. *IEEE Trans. Electron Devices* **65**, 4304–4312 (2018)
72. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436–444 (2015)
73. Rawat, W., Wang, Z.: Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput.* **29**, 2352–2449 (2017)
74. LeCun, Y.: 1.1 Deep Learning Hardware: Past, Present, and Future in 2019 IEEE International Solid-State Circuits Conference - (ISSCC), pp. 12–19 (2019)
75. Jouppi, N.P., Young, C., Patil, N., et al.: In-datacenter performance analysis of a tensor processing unit. In: Proceedings of the 44th Annual International Symposium on Computer Architecture, pp. 1–12. Association for Computing Machinery, New York (2017). ISBN: 978-1-4503-4892-8
76. Alibart, F., Zamanidoost, E., Strukov, D.B.: Pattern classification by memristive crossbar circuits using ex situ and in situ training. *Nat. Commun.* **4**, 2072 (2013)
77. Burr, G.W., Narayanan, P., Shelby, R.M., et al.: Large-scale neural networks implemented with non-volatile memory as the synaptic weight element: comparative performance analysis (accuracy, speed, and power). In: 2015 IEEE International Electron Devices Meeting (IEDM), pp. 4.4.1–4.4.4 (2015)
78. Ambrogio, S., Narayanan, P., Tsai, H., et al.: Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature* **558**, 60–67 (2018)
79. Gokmen, T. Vlasov, Y.: Acceleration of deep neural network training with resistive cross-point devices: design considerations. *Front. Neurosci.* **10** (2016)
80. Burr, G.W., Shelby, R.M., Sidler, S., et al.: Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element. *IEEE Trans. Electron Devic.* **62**, 3498–3507 (2015)
81. Yu, S.: Neuro-inspired computing with emerging nonvolatile memories. *Proc. IEEE* **106**, 260–285 (2018)
82. Cartier, E.A., Kim, W., Gong, N., et al.: Reliability Challenges with Materials for Analog Computing in 2019 IEEE International Reliability Physics Symposium (IRPS), pp. 1–10 (2019)
83. Milo, V., Malavena, G., Monzio Compagnoni, C., et al.: Memristive and CMOS devices for neuromorphic computing. *Materials* **13**, 166 (2020)
84. Jia, Z., Maggioni, M., Smith, J., et al. Dissecting the NVidia Turing T4 GPU via microbenchmarking (2019). [arXiv:1903.07486](https://arxiv.org/abs/1903.07486) [cs]. [arXiv: 1903.07486](https://arxiv.org/abs/1903.07486) [cs]
85. Bill, J., Legenstein, R.: A compound memristive synapse model for statistical learning through STDP in spiking neural networks. *Front. Neurosci.* **8**, 412 (2014)
86. Valentian, A., Rummens, F., Vianello, E., et al.: Fully Integrated Spiking Neural Network with Analog Neurons and RRAM Synapses in 2019 IEEE International Electron Devices Meeting (IEDM), pp. 14.3.1–14.3.4 (2019)
87. Agarwal, S., Jacobs Gedrim, R.B., Hsia, A.H., et al.: Achieving ideal accuracies in analog neuromorphic computing using periodic carry. In: 2017 Symposium on VLSI Technology, pp. T174–T175 (2017)
88. Chicca, E., Indiveri, G.: A recipe for creating ideal hybrid memristive-CMOS neuromorphic processing systems. *Appl. Phys. Lett.* **116**, 120501 (2020)
89. Querlioz, D., Bichler, O., Gamrat, C. Simulation of a memristor-based spiking neural network immune to device variations In: The 2011 International Joint Conference on Neural Networks, pp. 1775–1781 (2011)
90. Querlioz, D., Bichler, O., Dollfus, P., et al.: Immunity to device variations in a spiking neural network with memristive nanodevices. *IEEE Trans. Nanotechnol.* **12**, 288–295 (2013)

91. Sheridan, P., Ma, W., Lu, W.: Pattern Recognition with Memristor Networks in 2014 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1078–1081 (2014)
92. Zahari, F., Hansen, M., Mussenbrock, T., et al.: Pattern recognition with TiO_x-based memristive devices. *Materials* **2**, 203–216 (2015)
93. Hansen, M., Zahari, F., Ziegler, M., et al. Double-barrier memristive devices for unsupervised learning and pattern recognition. *Front. Neurosci.* **11** (2017)
94. Hansen, M., Zahari, F., Kohlstedt, H., et al.: Unsupervised Hebbian learning experimentally realized with analogue memristive crossbar arrays. *Sci. Rep.* **8**, 8914 (2018)
95. Hazan, H., Saunders, D., Sanghavi, D. T., et al.: Unsupervised Learning with Self-Organizing Spiking Neural Networks in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (2018)
96. Brivio, S., Conti, D., Nair, M.V., et al.: Extended memory lifetime in spiking neural networks employing memristive synapses with nonlinear conductance dynamics. *Nanotechnology* **30**, 015102 (2018)
97. Brivio, S., Ly, D.R.B., Vianello, E., et al.: Non-linear memristive synaptic dynamics for efficient unsupervised learning in spiking neural networks. *Front. Neurosci.* **15**, 27 (2021)
98. Pfeiffer, M., Pfeil, T.: Deep learning with spiking neurons: opportunities and challenges. *Front. Neurosci.* **12**, 774 (2018)
99. Lecun, Y., Bottou, L., Bengio, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998)
100. Krizhevsky, A.: Learning multiple layers of features from tiny images. Technical report, University of Toronto, p. 60 (2009)
101. Russakovsky, O., Deng, J., Su, H., et al.: ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015)
102. Roy, K., Jaiswal, A., Panda, P.: Towards spike-based machine intelligence with neuromorphic computing. *Nature* **575**, 607–617 (2019)
103. Dora, S., Kasabov, N.: Spiking neural networks for computational intelligence: an overview. *Big Data Cognit. Comput.* **5**, 67 (2021)
104. Amir, A., Taba, B., Berg, D., et al.: A Low Power, Fully Event-Based Gesture Recognition System in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7388–7397 (2017)
105. Binas, J., Neil, D., Liu, S.-C., et al.: DDD17: End-to-end DAVIS driving dataset (2017) . [arXiv:1711.01458](https://arxiv.org/abs/1711.01458) [cs]. [arXiv: 1711.01458](https://arxiv.org/abs/1711.01458) [cs]
106. Sironi, A., Brambilla, M., Bourdis, N., et al.: HATS: histograms of averaged time surfaces for robust event-based object classification (2018). [arXiv:1803.07913](https://arxiv.org/abs/1803.07913) [cs]. [arXiv: 1803.07913](https://arxiv.org/abs/1803.07913) [cs]
107. Mueggler, E., Rebecq, H., Gallego, G., et al.: The event-camera dataset and simulator: event-based data for pose estimation, visual odometry, and SLAM. *Int. J. Robot. Res.* **36**, 142–149 (2017)
108. Chialvo, D.R.: Emergent complex neural dynamics. *Nat. Phys.* **6**, 744–750 (2010)
109. Engel, A.K., Fries, P., Singer, W.: Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001)
110. Fries, P.: A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474–480 (2005)
111. Gray, C.M., König, P., Engel, A.K., et al.: Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**, 334–337 (1989)
112. Ignatov, M., Hansen, M., Ziegler, M., et al.: Synchronization of two memristively coupled van der pol oscillators. *Appl. Phys. Lett.* **108**, 084105 (2016)
113. Ignatov, M., Ziegler, M., Hansen, M., et al.: Memristive stochastic plasticity enables mimicking of neural synchrony: memristive circuit emulates an optical illusion. *Sci. Adv.* **3**, e1700849 (2017)
114. Torrejon, J., Riou, M., Araujo, F.A., et al.: Neuromorphic Computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017)

115. Walczyk, D., Walczyk, C., Schroeder, T., et al.: Resistive switching characteristics of CMOS embedded HfO₂-based 1T1R cells. In: *Microelectronic Engineering. Proceedings of the 17th Biennial International Insulating Films on Semiconductor Conference*, vol. 88, pp. 1133–1135 (2011)
116. Kim, S., Ishii, M., Lewis, S., et al.: NVM Neuromorphic Core with 64k-Cell (256-by-256) Phase Change Memory Synaptic Array with on-Chip Neuron Circuits for Continuous in-Situ Learning in 2015 IEEE International Electron Devices Meeting (IEDM), pp. 17.1.1–17.1.4 (2015)
117. Mahadevaiah, M.K., Perez, E., Wenger, C. et al.: Reliability of CMOS Integrated Memristive HfO₂ Arrays with Respect to Neuromorphic Computing in 2019 IEEE International Reliability Physics Symposium (IRPS), pp. 1–4 (2019)
118. Zarrin, P.S., Zahari, F., Mahadevaiah, M.K., et al.: Neuromorphic on-chip recognition of saliva samples of COPD and healthy controls using memristive devices. *Sci. Rep.* **10**, 19742 (2020)
119. Wenger, C., Zahari, F., Mahadevaiah, M.K., et al.: Inherent stochastic learning in CMOS-integrated HfO₂ arrays for neuromorphic computing. *IEEE Electron Device Lett.* **40**, 639–642 (2019)
120. Zahari, F., Pérez, E., Mahadevaiah, M.K., et al.: Analogue pattern recognition with stochastic switching binary CMOS-integrated memristive devices. *Sci. Rep.* **10**, 14450 (2020)
121. Dong, T., Santos, S., Yang, Z., et al.: Sputum and salivary protein biomarkers and point-of-care biosensors for the management of COPD. *Analyst* **145**, 1583–1604 (2020)
122. Baker, S.B., Xiang, W., Atkinson, I.: Internet of Things for Smart Healthcare: Technologies, Challenges, and Opportunities. *IEEE Access* **5**, 26521–26544 (2017)
123. Guha, S., Jamal, F. I. Wenger, C.: A review on passive and integrated near-field microwave biosensors. *Biosensors (Basel)***7**, E42 (2017)
124. Zarrin, P.S., Roeckendorf, N., Wenger, C.: In-vitro classification of saliva samples of COPD patients and healthy controls using machine learning tools. *IEEE Access* **8**, 168053–168060 (2020)
125. Zarrin, P.S., Roeckendorf, N.: Database containing saliva samples of COPD patients and healthy controls. <https://archive.ics.uci.edu/ml/datasets/Exasens>
126. Zarrin, P., Ibne Jamal, F., Roeckendorf, N., et al.: Development of a portable dielectric biosensor for rapid detection of viscosity variations and its in vitro evaluations using saliva samples of copd patients and healthy control. *Healthcare* **7**, 11 (2019)
127. Yang, Y., Lu, W.: Nanoscale resistive switching devices: mechanisms and modeling. *Nanoscale* **5**, 10076–10092 (2013)
128. Fries, P.: A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474–480 (2005)
129. Hu, M., Wang, Y., Qiu, Q., et al.: The Stochastic Modeling of TiO₂ Memristor and Its Usage in Neuromorphic System Design in 2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC), pp. 831–836 (2014)
130. Tuma, T., Pantazi, A., Gallo, M.L., et al.: Stochastic phase-change neurons. *Nat. Nanotech* **11**, 693–699 (2016)
131. Gaba, S., Sheridan, P., Zhou, J., et al.: Stochastic memristive devices for computing and neuromorphic applications. *Nanoscale* **5**, 5872–5878 (2013)
132. Suri, M., Querlioz, D., Bichler, O., et al.: Bio-inspired stochastic computing using binary CBRAM synapses. *IEEE Trans. Electron Devic.* **60**, 2402–2409 (2013)
133. Payvand, M., Nair, M.V., Müller, L.K., et al.: A neuromorphic systems approach to in-memory computing with non-ideal memristive devices: from mitigation to exploitation. *Faraday Discuss.* **213**, 487–510 (2019)
134. Grossi, A., Zambelli, C., Olivo, P., et al.: Electrical characterization and modeling of pulse-based forming techniques in RRAM arrays. *Solid-State Electron.* **115**, 17–25 (2016)
135. Milo, V., Zambelli, C., Olivo, P., et al.: Multilevel HfO₂-based RRAM devices for low-power neuromorphic networks. *APL Mater.* **7**, 081120 (2019)
136. Fantini, A., Goux, L., Degraeve, R., et al.: Intrinsic Switching Variability in HfO₂ RRAM in 2013 5th IEEE International Memory Workshop, pp. 30– 33 (2013)

137. Grossi, A., Perez, E., Zambelli, C., et al.: Impact of the precursor chemistry and process conditions on the cell-to-cell variability in 1T-1R based HfO₂ RRAM devices. *Sci. Rep.* **8**, 1–11 (2018)
138. Diehl, P. U. Cook, M.: Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Front. Comput. Neurosci.* **9** (2015)
139. Yousefzadeh, A., Stomatias, E., Soto, M., et al.: On practical issues for stochastic STDP hardware with 1-Bit synaptic weights. *Front. Neurosci.* **12** (2018)
140. Lee, J. H., Delbruck, T. Pfeiffer, M.: Training deep spiking neural networks using backpropagation. *Front. Neurosci.* **10** (2016)
141. Wu, Y., Deng, L., Li, G., et al.: Spatio-temporal backpropagation for training high-performance spiking neural networks. *Front. Neurosci.* **12** (2018)
142. Ahlawat, S., Choudhary, A., Nayyar, A., et al.: Improved handwritten digit recognition using convolutional neural networks (CNN). *Sensors* **20**, 3344 (2020)
143. LeCun, Y., Jackel, L.D., Denker, J.S., et al.: Learning algorithms for classification: a comparison on handwritten digit recognition, 16 (1995)
144. Kheradpisheh, S.R., Ganjtabesh, M., Thorpe, S.J., et al.: STDP-based spiking deep convolutional neural networks for object recognition. *Neural Netw.* **99**, 56–67 (2018)
145. Zhang, T., Cheng, X., Jia, S., et al.: Self-backpropagation of synaptic modifications elevates the efficiency of spiking and artificial neural networks. *Sci. Adv.* **7**, eabh0146 (2021)
146. Martin, E., Ernault, M., Laydevant, J., et al.: EqSpike: spike-driven equilibrium propagation for neuromorphic implementations. *iScience* **24**, 102222 (2021)
147. Yao, P., Wu, H., Gao, B., et al.: Fully hardware-implemented memristor convolutional neural network. *Nature* **577**, 641–646 (2020)
148. Li, C., Belkin, D., Li, Y., et al.: Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nat. Commun.* **9**, 1–8 (2018)
149. Park, J., Lee, J., Jeon, D.: A 65-Nm neuromorphic image classification processor with energy-efficient training through direct spike-only feedback. *IEEE J. Solid-State Circuits* **55**, 108–119 (2020)
150. Khaddam-Aljameh, R., Stanisavljevic, M., Mas, J.F., et al.: HERMES Core - A 14nm CMOS and PCM-based In-Memory Compute Core Using an Array of 300ps/LSB Linearized CCO-based ADCs and Local Digital Processing in 2021 Symposium on VLSI Technology, pp. 1–2 (2021)
151. Ishii, M., Kim, S., Lewis, S., et al.: On-Chip Trainable 1.4M 6T2R PCM Synaptic Array with 1.6K Stochastic LIF Neurons for Spiking RBM in 2019 IEEE International Electron Devices Meeting (IEDM)), pp. 14.2.1–14.2.4 (2019)
152. Hansen, M., Ziegler, M., Kolberg, L., et al.: A double barrier memristive device. *Sci. Rep.* **5**, 13753 (2015)
153. Jo, S.H., Chang, T., Ebong, I., et al.: Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **10**, 1297–1301 (2010)
154. Linares-Barranco, B., Serrano-Gotarredona, T., Camuñas-Mesa, L., et al.: On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci.* **5**, 26 (2011)
155. Ohno, T., Hasegawa, T., Tsuruoka, T., et al.: Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater.* **10**, 591–595 (2011)
156. Dirkmann, S., Hansen, M., Ziegler, M., et al.: The role of ion transport phenomena in memristive double barrier devices. *Sci. Rep.* **6**, srep35686 (2016)
157. Querlioz, D., Bichler, O., Vincent, A.F., et al.: Bioinspired programming of memory devices for implementing an inference engine. *Proc. IEEE* **103**, 1398–1416 (2015)
158. Gerstner, W.: Time structure of the activity in neural network models. *Phys. Rev. E* **51**, 738–758 (1995)
159. Hoppensteadt, F.C., Izhikevich, E.M.: Oscillatory neurocomputers with dynamic connectivity. *Phys. Rev. Lett.* **82**, 2983–2986 (1999)
160. Pikovsky, A., Rosenblum, M., Kurths, J.: Synchronization: A Universal Concept in Nonlinear Sciences. ISBN: 978-0-521-53352-2. Cambridge University Press, Cambridge (2001)

161. Stathopoulos, S., Khat, A., Trapatseli, M., et al.: Multibit memory operation of metal-oxide bi-layer memristors. *Sci. Rep.* **7**, 17532 (2017)
162. Park, S., Klett, S., Ivanov, T., et al.: Engineering method for tailoring electrical characteristics in TiN/TiOx/ HfOx/Au bi-layer oxide memristive devices **3**, 16 (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Redox-Based Bi-Layer Metal Oxide Memristive Devices



Finn Zahari, Seongae Park, Mamathamba K. Mahadevaiah,
Christian Wenger, Hermann Kohlstedt, and Martin Ziegler

Abstract In a memristor or a so-called memristive device, the resistance state depends on the previous charge flow through the device. The new resistance state is stored and classifies a memristor as a non-volatile memory device. This likewise unique and simple feature qualifies memristive devices as attractive compartments with regard to the development of a universal memory and beyond von Neumann computing architectures, including in-memory computing and neuromorphic circuits. In this chapter, we present studies on two kinds of bi-layer metal oxide memristive devices with the layer sequences Nb/NbO₂/Al₂O₃/Nb_xO_y/Au and TiN/TiO_x/HfO_x/Au, either prepared by reactive DC-magnetron sputtering, etching and optical lithography. It is shown that the memristive properties of such devices can be engineered, which enables tailoring of the memristive devices for particular applications.

F. Zahari (✉) · H. Kohlstedt
Nanoelectronics, Faculty of Engineering, Kiel University, 24143 Kiel, Germany
e-mail: fnz@tf.uni-kiel.de

H. Kohlstedt
e-mail: hko@tf.uni-kiel.de

S. Park · M. Ziegler
Department of Electrical Engineering and Information Technology, Technische Universität
Ilmenau, 98693 Ilmenau, Germany
e-mail: seongae.park@tu-ilmenau.de

M. Ziegler
e-mail: martin.ziegler@tu-ilmenau.de

Institute of Micro and Nanotechnologies MacroNano, Technische Universität Ilmenau, 98693
Ilmenau, Germany

M. K. Mahadevaiah
IHP Leibniz-Institut für innovative Mikroelektronik, Frankfurt, Germany
e-mail: kalishettyhalli@ihp-microelectronics.com

C. Wenger
IHP Leibniz-Institut für innovative Mikroelektronik, Frankfurt/Oder, Germany
e-mail: wenger@ihp-microelectronics.com

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series
on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_3

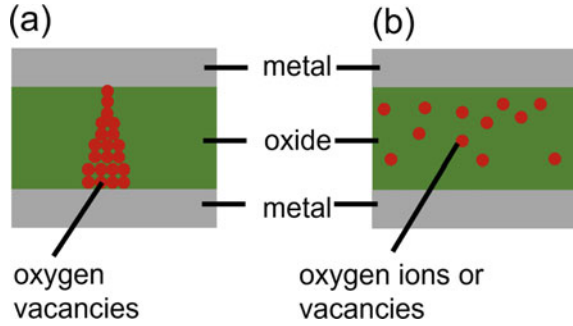
1 Introduction

The term “Memristor” is a portmanteau from the two words “Memory” and “Resistor”. A memristor (also called a memristive device) typically consists of capacitor-like layer sequences, i.e., a metal-insulator-metal structure. A universal property of the memristive device concept is that the memristive state depends on previously induced charge flows, applied currents, or applied electric fields, which result in different stored resistance states. This simultaneously unique and simple device principle, together with the predicted potential for breakthrough technologies in areas such as universal memories and novel non-Boolean computing schemes for cognitive electronic systems, drives the research and development of memristors and memristor-based circuits. Memristive devices were intensively studied in the last century’s sixties and seventies [1–3]. The field was further driven forward by the establishment of the theoretical background of memristors by Leon Chua (1971), with the corresponding experimental realization and interpretation by the Hewlett-Packard (HP)-Labs (2008) [4, 5].

Currently, two main development avenues can be explored for memristive devices. The first focuses on resistive random access memories (RRAMs). It is believed that the zoo of today’s existing memory diversity can be replaced by a single (universal) memory concept. RRAMs are considered attractive candidates for universal memories because they: show non-volatile data storage, can be densely integrated, are fast, and are cheap to produce. In particular, such a universal memory might attenuate the problem known as memory latency in modern digital computers [6, 7]. Besides the RRAM goal, which may be categorized under the label “More Than Moore”, novel and very appealing computer architectures have been proposed in which memristors might play a vital role. Another main spotlight of possible memristive device applications may be assigned to such catchphrases as non-Boolean computing, bio-inspired information processing, neuromorphic engineering, or cognitive electronics [8–16]. On the local synaptic level, learning in nervous systems is explained by the Hebbian learning rule [17] and, amongst others, spike-timing dependent plasticity (STDP) [18]. STDP and other memory-related mechanisms observed in nervous systems, such as long term potentiation (LTP) and long term depression (LTD) [19], were successfully mimicked by memristive devices [8, 20–22]. To what extent larger networks of memristive devices can mimic higher brain functions is still unknown.

Until now, a huge number of experimental findings on memristor devices consisting of a wide variety of metal/insulator material combinations have been published, all of which show memristive I - V curves [23]. For details concerning resistive switching and the underlying physical-chemical mechanisms, we refer the reader to the overwhelming literature on the subject [12, 14, 23–27]. These physical-chemical mechanisms include electronic effects (e.g., charging of electron traps) [25, 26], magnetic effects (e.g., spin-transfer torque - STT) [28] or ferroelectricity (e.g., ferroelectric tunnel junctions - FTJ) [29]. Moreover, nano-ionic mechanisms are widely utilized to build memristive devices. Included are phase change memory (PCM) devices [30] in which the resistance of the active layer is modulated by switching

Fig. 1 Schematic illustrations of VCM devices based on filament formation (a) and homogeneous ion movement (b)



between crystalline and amorphous phase. Moreover, devices based on the formation and dissolution of conductive filaments composed of metal ions, i.e., the electrochemical metalization effect (ECM) [31], or composed of oxygen vacancies, i.e., the valance change mechanism (VCM) [23, 32], belong to this class. Furthermore, VCM can also be realized without conductive filaments. In these interface-based memristive devices, the resistance is changed by modulating interfacial properties like the Schottky-barrier height [24, 32–34] or the tunneling probability [35–37] by the movement of oxygen vacancies or oxygen ions [23, 25]. Thus, the functionality of VCM devices is based on redox reactions. A sketch of a filamentary-based device and an interface-based device are shown in Fig. 1a and b, respectively. This chapter deals with VCM devices composed of bi-layer metal oxides.

Since the requirements the devices have to fulfill depend on the application, both types of devices are investigated for different purposes, as described in more detail in chapter [Neuromorphic Circuits with Redox-Based Memristive Devices](#) and in further literature [23, 38–41]. In this respect, filamentary-based devices are available in a more mature technology compared to interface-based devices [41]. However, the need for an initial electro-forming step, a current compliance for switching, a pronounced intrinsic randomness of the switching process, and a limited amount of achievable resistance states are some drawbacks compared to interface-based devices [24, 38–42]. On the other hand, endurance, retention, switching speed, and scalability are some of the reported benefits of filamentary devices [38–42]. The properties of interface-based devices will be discussed in further detail in Sect. 2 while the properties of filamentary-based devices are described in Sect. 3.

The quality of the solid-state electrolyte, in which ion movement occurs, determines the switching properties. In this regard, it has been theoretically shown that an oxygen-deficient HfO_x with x between 1.5 and 1.75 is required for the most efficient nucleation of vacancies as the starting point for filament growth [43]. Park et al. [44] experimentally verified this by showing that $x = 1.8$ leads to filamentary switching while $x = 1.98$ leads to interface-based switching. However, stoichiometric HfO_2 matrices are also reported to show filamentary switching enabled through local reduction to HfO_{2-x} [45, 46]. Similar observations are reported for Sr_2TiO_4 [47]. However, in these devices, only the filament-type switching is explained by the

movement of vacancies. In contrast, the interface-based switching is explained by the charging and discharging of electron traps. Moreover, it has been shown that the oxygen content in HfO_x [44, 45] and TaO_x [48] can be used to modify important device parameters in filamentary devices, such as forming and switching voltages, memory window, and high resistance state (HRS) and low resistance state (LRS) currents. Furthermore, adjusting the switching window by tailoring the defect density has also been reported for interface-based devices in simulation studies [44, 49]. Engineering device properties will be discussed in more detail in Sect. 4. The second oxide in the bi-layer system can have different functions, like acting as an oxygen reservoir or a diffusion barrier. A reservoir like TiO_x [50, 51], TaO_x [37], or a TiON formed at a TiN electrode interface [52, 53] allows to increase or decrease oxygen vacancy or ion concentration in the solid-state electrolyte while performing memristive switching. On the other hand, a diffusion barrier like Al_2O_3 [54–57] restricts ion movement within the electrolyte and, thus, can improve retention times [55] or switching linearity [56], and can lead to multi-state operation [56, 57]. It should be noted here that bi-layer metal oxide memristive devices are not necessarily switching by ion migration. As a second switching mechanism, the charging and discharging of electron traps are reported to be responsible for memristive switching in metal oxide junctions [58–63].

Both interface-based and filamentary-based VCM devices consisting of bi-layer metal oxides are described in the following, and exemplary realizations are shown in Sects. 2 and 3, respectively. Strategies to engineer the electrical parameters of these memristive devices to meet the requirements for specific applications are provided in Sect. 4.

2 Interface-Based Devices

The current transport is homogeneously distributed across the electrode area in interface-based memristive devices. Here, the resistance is determined by the interfacial properties of the incorporated materials. Thus, the resistance change is based on altering interfacial properties like Schottky-barrier height or width [24, 32–34] or tunneling probability [35–37]. This is achieved by a rearrangement of oxygen vacancies or oxygen ions [23, 25] or by charging and discharging electron traps [58–60, 62]. In that way, interface-based memristive devices show analog, i.e., gradual switching [24, 35, 41]. Moreover, no electro-forming is needed, and devices with a high I - V non-linearity [33, 64] to avoid the need for additional selector elements for passive crossbar integration [65] are reported. Furthermore, the intrinsic randomness is less pronounced than in filamentary devices [38, 41]. Moreover, interface-based VCM devices have already been integrated into 3D stacks in complementary metal-oxide-semiconductor (CMOS) technology with a minimum feature size of 5 nm [64]. However, CMOS integration and scaling to the nanoscale is challenging for most reported devices [41]. CMOS integration is challenging if the incorporated materials or the needed deposition steps are not used in CMOS fabrication lines or if

the used materials do not withstand the CMOS fabrication steps. This is explained in detail in chapter [Integration of Memristive Devices into a 130 nm CMOS Baseline Technology](#). Downscaling is challenging because the resistance of interface-based devices scales with the electrode area, and most reported devices have high resistances even on the micrometer scale. Drawbacks are, furthermore, a worse retention time and slower switching speed [41] compared to filamentary devices. Therefore, interface-based devices are good candidates to emulate synapses in bio-inspired neuromorphic networks [11, 40], as described in detail in chapter [Neuromorphic Circuits with Redox-Based Memristive Devices](#). Two different devices are described in this section. The first device is based on a $\text{Nb}_x\text{O}_y/\text{Al}_2\text{O}_3$ bi-layer in which Nb_xO_y acts as the solid-state electrolyte while Al_2O_3 is a diffusion barrier restricting ion movement in the Nb_xO_y layer [55]. The second device is based on a $\text{HfO}_2/\text{TiO}_x$ bi-layer. Here, TiO_x acts as an oxygen reservoir, enabling an oxygen exchange with HfO_2 [44].

2.1 Devices Based on $\text{Nb}_x\text{O}_y/\text{Al}_2\text{O}_3$

The first device is named double barrier memristive device (DBMD) [55]. The double oxide layer $\text{Nb}_x\text{O}_y/\text{Al}_2\text{O}_3$ is sandwiched between an Au top electrode and an Nb bottom electrode, as it is sketched in Fig. 2a. The Au leads to a Schottky barrier formation, while alumina acts as a tunnel barrier. The movement of oxygen ions within Nb_xO_y under an applied electric field is reported to be responsible for memristive switching [55, 66], as modeled by a kinetic Monte Carlo simulation [67] which is explained in more detail in chapter [Modeling and Simulation of Silver-Based Filamentary Memristive Devices](#). It should be noted that the charging and discharging of electron traps as the switching mechanism cannot be ruled out [55]. However, a variation in the charge distribution within Nb_xO_y leads to both a variation of the Schottky-barrier height and effective tunneling width. By scaling the solid-state electrolyte down to 2.5 nm, i.e., to the length scale of a single electron wave, the impact of the bulk resistance is minimized, and a very strong mutual interdependency of both barriers is achieved. Thus, both interfaces are altered simultaneously when the charge distribution varies.

Devices with active areas between $70 \mu\text{m}^2$ and $2500 \mu\text{m}^2$ are produced on 100 mm Si wafers, which are passivated by a thermal SiO_2 layer. All layers are deposited by DC magnetron sputtering in one and the same vacuum chamber without breaking the vacuum. The devices are structured using standard optical lithography in combination with lift-off and etching steps. Additional passivation with SiO_x enables the deposition of Ti metal contact pads, which allow an automatic measuring of device properties across the whole wafer. This will be discussed in more detail in Sect. 4.1. The fabrication process is explained in detail in [55, 66]. First, Nb and Al are deposited in an Ar atmosphere. The subsequent deposition of Nb_xO_y from a metallic target in an Ar/O₂ atmosphere leads to the fully oxidation of Al to Al_2O_3 , as observed in electron energy loss spectroscopy (EELS) and energy-dispersive X-ray spectroscopy (EDX) in a transmission electron microscope (TEM) [69]. The Nb_xO_y stoichiometry was found to be between the oxidation states Nb_2O_5 and NbO_2 . All

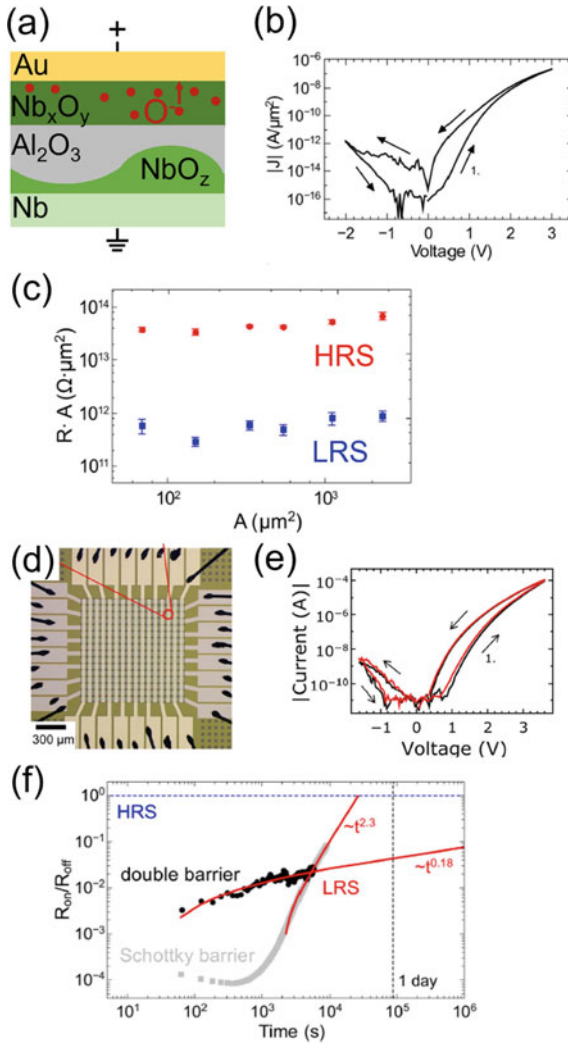


Fig. 2 Double barrier memristive devices (DBMDs). **a** Schematic material stack. **b** Typical $|J| - V$ hysteresis curve. **c** Area-dependent current transport in HRS (red) and LRS (blue). **d** Optical microscope photograph of a $16 \cdot 16$ crossbar array containing DBMDs. **e** Typical $|I| - V$ curves of a crossbar-integrated device (red) and a single device (black). **f** Retention of a DBMD (black) and a device without an Al_2O_3 layer. **b**, **f** are reprinted from [55] (licensed under CC BY 4.0), **c** is adapted from [55] (licensed under CC BY 4.0), and **d**, **e** are reprinted from [68] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

interfaces but the Nb/Al₂O₃ interface appeared to be smooth on the nanoscale. Due to the roughness at the lower interface, the thinnest parts of the alumina were determined to be 1.3 nm, as sketched in Fig. 2a. Furthermore, the interface between the Nb bottom electrode and Al₂O₃ was also slightly oxidized. The TEM measurements, including EELS and EDX, are described in more detail in chapter [Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices](#).

Figure 2b shows a typical current density versus voltage (J - V) hysteresis curve of a DBMD, which was measured with a grounded Nb electrode, while voltage was applied to the Au electrode. Switching from HRS to LRS is taking place gradually, i.e., in an analog way, by applying a positive voltage to the Au top electrode. Switching back to HRS takes place by applying a negative bias to the top electrode. No initial electro-forming is needed, and an intrinsic current compliance is present. Both HRS and LRS show an area-dependent current transport, as shown in Fig. 2c, since the product of device area A and resistance R results in a constant value for both states. The absolute value of the current density is depicted on a log scale in Fig. 2b, showing the high J - V non-linearity and asymmetry. For dense device integration in a crossbar array, a selector device in series to every memristive element or a strong built-in I - V non-linearity is needed to avoid parasitic currents, i.e., the sneak path problem, when addressing a specific device [65, 70]. The built-in diode-like characteristics of DBMDs allows crossbar integration without needing additional selector devices [68]. A photograph of a crossbar array consisting of 256 DBMDs is shown in Fig. 2d. Here, the devices are located at the cross points of the metallization lines. In that way, each top and bottom contact pad connects the top electrodes in one column of the array. Each left and right pad connects the bottom electrodes of one row. Thus, only 32 contact pads are needed to address all 256 memristive devices. The I - V characteristics of crossbar integrated devices are similar to those of reference devices, as shown in Fig. 2e, in which the black line belongs to a reference device and the red line to a crossbar-integrated device, respectively. Furthermore, Fig. 2f shows the retention time of a DBMD. After the device was set from R_{off} (HRS) to R_{on} (LRS), the resistance was recorded every 60 s with a voltage pulse of 0.5 V amplitude and 2 s duration. In that way, a non-destructive read-out of the resistance state was performed, and the relaxation of the state was measured. The resistance increased fast in the first 600 s (black data points), while afterward, the resistance change was significantly less pronounced. An R_{on}/R_{off} ratio of more than one order of magnitude was still present after one day. This behavior can be quantitatively described by fitting the experimental data to a $\sim t^\beta$ power law (red lines) [61]. In the first 600 s, the relaxation can be described by $\sim t^{0.65}$, while afterward, $\sim t^{0.18}$ holds. That two different power laws are dominant for different time intervals suggests that different physical processes may be responsible for memristive switching, e.g., charging of traps and movement of ions on a faster and slower timescale, respectively. It is also shown in Fig. 2f that the Al₂O₃ considerably affects the retention time. The grey squares show the same measurement on devices without alumina. The retention characteristics of those devices show no resistance increase for the first 700 s but a steep increase afterward fitted by a $\sim t^{2.3}$ power law. Thus, the Al₂O₃ does significantly increase the retention time. Furthermore, the fact that alumina acts as a

diffusion barrier for oxygen [54, 56, 57] is also believed to lead to the self-limited switching behavior due to the limited amount of oxygen ions within Nb_xO_y and no oxygen exchange between both oxides.

The retention time of DBMDs does not allow to use these devices for classical memory applications but for neuromorphic circuits [11]. Bio-inspired learning of visual patterns utilizing crossbar integrated DBMDs as hardware synapses have already been demonstrated [68]. This is explained in chapter [Neuromorphic Circuits with Redox-Based Memristive Devices](#) in detail. However, the drawbacks of DBMDs are those named above for interface-based devices [41]. In this respect, the scalability is challenging due to the high resistance, even on the micrometer scale. For quadratic devices with a feed size of 10 μm , the current in HRS for a typical read-out voltage of 0.5 V is in the pA range. Thus, for scaling the devices down to the nanoscale, the resistance has to be tailored by orders of magnitude. On the other hand, high ohmic devices are desired for low-power applications. Thus, the resistance has to be adjusted so that currents are high enough to be suitable for electronic circuits but low enough to account for energy efficiency. Strategies for engineering device performances are given in Sect. 4. Furthermore, compatibility with CMOS technology is crucial for developing integrated circuits. While Al_2O_3 is a well-known dielectric in CMOS technology, Nb_xO_y in the needed stoichiometry might be more challenging to be integrated into state-of-the-art fabrication lines. Moreover, gold is not desired in CMOS fabrication. Also, the impact of other fabrication steps (e.g., tempering in back-end-of-line processes [71]) has to be considered. CMOS integration of memristive devices is explained in chapter [Integration of Memristive Devices into a 130 nm CMOS Baseline Technology](#). However, if CMOS integration is not possible, devices can be fabricated directly onto CMOS processed substrates in other fabrication lines [72, 73].

After describing device performance, including benefits and drawbacks of the introduced DBMDs, the model for current transport and memristive switching is explained in more detail in the following. Figure 3a shows an equivalent circuit emulating the I - V characteristics [55, 66]. Here, the $\text{Au}/\text{Nb}_x\text{O}_y$ interface is modeled as a Schottky diode D_s , the Nb_xO_y is modeled as a parallel circuit of a resistor R_{NbO} and a capacitor C_{NbO} , and the Al_2O_3 layer is modeled as a voltage-controlled current source I_t emulating elastic tunneling in parallel to a capacitor C_{AlO} . The variable describing the memristive state x denotes the average ion position within the Nb_xO_y . More precisely, thermionic emission theory was employed to model the Schottky diode [74, 75]. According to this theory, charge carriers have enough energy due to an applied electric field to cross the potential barrier between a metal and a semiconductor. For an applied forward bias V , this current can be calculated by

$$I_S = I_R \left(e^{\frac{qV}{nk_B T}} - 1 \right), \quad (1)$$

where q is the elemental charge, k_B is the Boltzmann constant, T is the temperature, and n is the ideality factor. If $n > 1$, deviation from ideal thermionic emis-

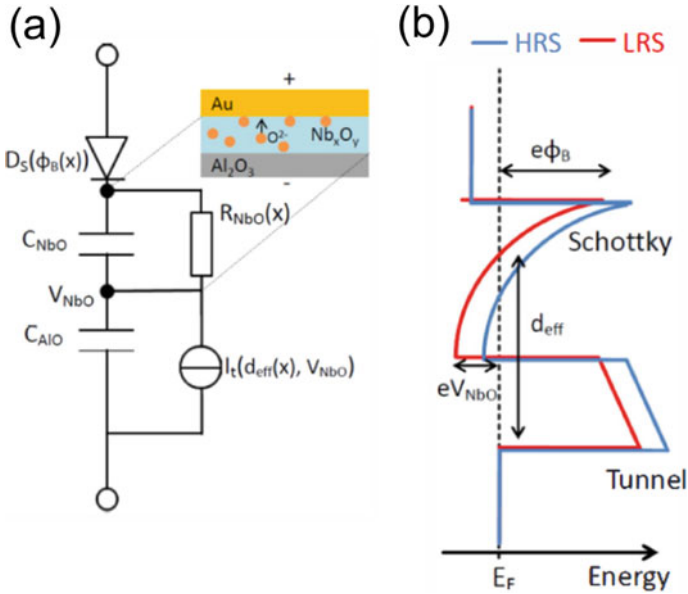


Fig. 3 Model of a DBMD [66] **a** Equivalent circuit **b** Schematic of the band diagram. ©2016 IEEE. Reprinted, with permission, from [M. Hansen et al., IEEE International Conference on Rebooting Computing (ICRC) (2016)]

sion theory is indicated. This deviation can be caused by, e.g., additional tunneling through the barrier (thermionic field emission), additional interface states, or barrier in-homogeneity [74, 75]. The reverse current I_R is given by

$$I_R = A^* AT^2 e^{-\frac{\Phi_B}{k_B T}}. \quad (2)$$

Here, A^* is the effective Richardson constant, A is the device area, and Φ_B is the Schottky barrier height. For reverse bias, however, and if the apparent barrier height Φ_B is reasonably smaller than the band gap of the insulator, the reverse current decreases gradually with the applied negative bias. This can be described by

$$I_{R, V < 0} = -A^* AT^2 e^{-\frac{\Phi_B}{k_B T}} e^{-\frac{\alpha_r \sqrt{|V|}}{k_B T}}. \quad (3)$$

The device-dependent parameter α_r is used to fit the experimentally observed reverse voltage dependency. As it is described in [55], the Schottky diode dominates the current transport for voltages below 0.5 V (LRS) and 1.0 V (HRS). For higher voltages in the forward regime, i.e., higher positive voltages on the Au electrode, the Al_2O_3 tunnel barrier gets relevant and acts as in intrinsic current compliance. The tunneling current can be modeled by the Simmons equation [76]

$$I_{tun} = \frac{KA}{d_{tox}^2} [x_1^2 e^{(-\alpha d_{tox} x_1)} - x_2^2 e^{(-\alpha d_{tox} x_2)}]. \quad (4)$$

In Eq. 4 holds $x_1 = \sqrt{\Phi - eV_I/2}$ and $x_2 = \sqrt{\Phi + eV_I/2}$. Here, Φ is the tunnel barrier height, and V_I is the resulting voltage across the tunnel barrier. Furthermore, A is the device area, K is a constant value, d_{tox} is the barrier thickness, and $\alpha = 2\sqrt{2m}/\hbar$ (m : free electron mass; \hbar : Planck's constant divided by 2π).

As it is indicated in Fig. 3a, the Schottky barrier height Φ_B , the effective tunnel barrier width d_{eff} (corresponds to d_{tox} in Eq. 4) as well as the resistance of Nb_xO_y R_{NbO} are dependent on the state variable x , i.e., on the average ion position. Thus, redistribution of oxygen ions within the Nb_xO_y leads to a change in the interfacial properties of the Schottky barrier and the tunnel barrier, while the resistance of Nb_xO_y is only slightly affected [55]. Details about the evolution of x under applied bias and the impact on the interfacial properties can be found in [55, 66]. This model was adapted by Dirkmann et al. [67]. Here, a lumped element circuit model containing the Schottky diode and the tunnel barrier is consistently coupled with a 3D kinetic Monte Carlo model for the ion transport. The simulation results show that the drift of charged oxygen ions within the Nb_xO_y can account for resistive switching behavior. In that way, a Schottky barrier height lowering from 0.90 eV to 0.83 eV, a decrease of the ideality factor from 4.0 to 3.4, and a decrease of the effective tunnel barrier thickness from 1.3 nm to 1.2 nm were determined to explain the R_{on}/R_{off} ratio of more than one order of magnitude. The model also covers the retention behavior. Furthermore, the simulation indicates that the whole applied voltage drops across the Schottky barrier for reverse bias and low forward bias, while the tunnel barrier is responsible for the built-in current compliance at higher forward biases. The model is also described in more detail in chapter [Modeling and Simulation of Silver-Based Filamentary Memristive Devices](#). A sketch of the band diagram for HRS and LRS is given in Fig. 3b and shows the decrease of the Schottky barrier height as well as the tunnel barrier width for switching from HRS to LRS [66].

2.2 Devices Based on HfO_2/TiO_x

Another device based on the oxide bi-layer HfO_2/TiO_x sandwiched between an Au top electrode and a TiN bottom electrode is discussed now [44]. All used materials but the Au are already integrated into CMOS fabrication lines [77]. A typical I - V characteristic for a device with an area of $100 \mu m^2$ is shown in Fig. 4a with grey dots. The switching mechanism is reported to be similar to that of the DBMDs. A Schottky barrier is formed between Au and HfO_2 . Mobile oxygen ions within HfO_2 lead to resistive switching, as sketched in Fig. 4b, c. However, as for the DBMDs, charging and discharging of electron traps cannot be omitted as a reason for switching [44]. The I - V characteristics are qualitatively similar to the DBMDs. Thus, gradual and electro-forming free switching for voltages exceeding a certain threshold and a high I - V non-linearity leading to a diode-like behavior are present.

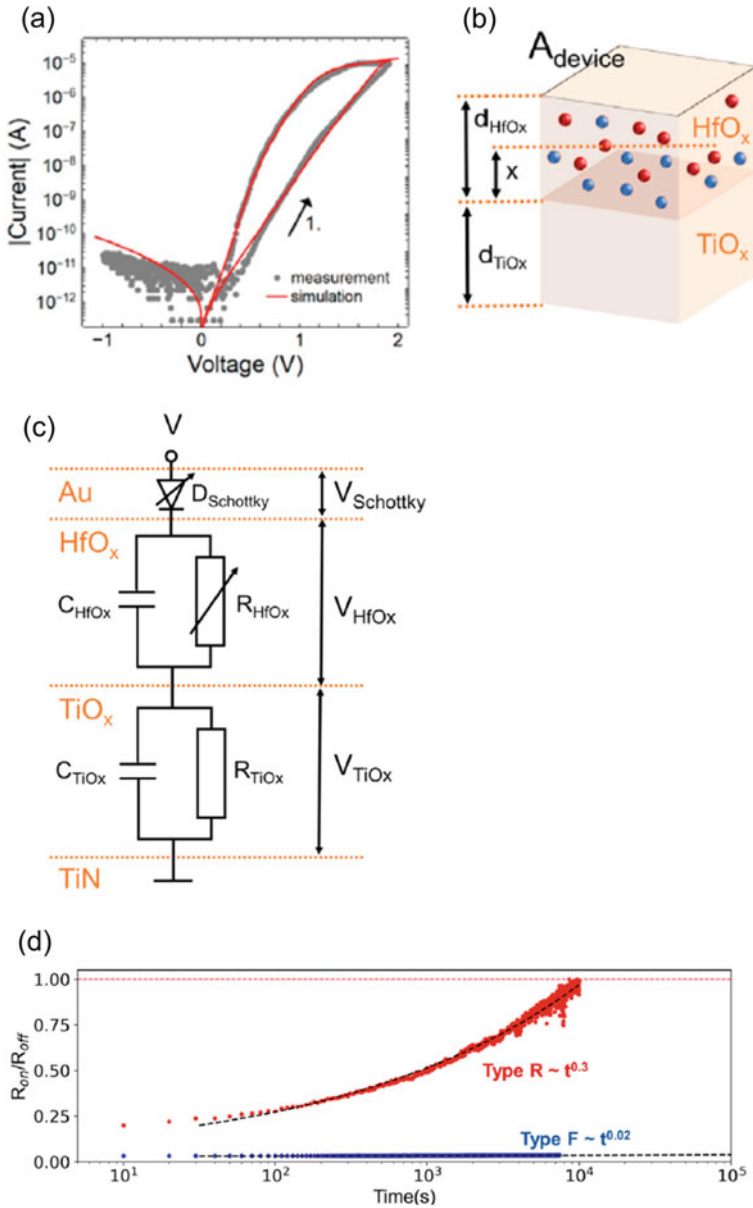


Fig. 4 Interface-based TiN/TiO_x/HfO_x/Au device. **a** Typical $|I| - V$ hysteresis curve. **b** Schematic material stack with indicated switching model. Red and blue spheres indicate mobile oxygen ions and stationary oxygen vacancies, respectively. **c** Equivalent circuit **d** Retention of interface-based devices (red) and filamentary devices (blue). Reprinted from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

However, one difference is that a current compliance of $10\ \mu\text{A}$ is used for switching to prevent dielectric breakdown. Furthermore, the switching window, i.e., the difference between HRS and LRS, decreases with increasing device area. This leads to the fact that the HRS shows an area-dependent current transport while the LRS seems to be not area dependent. However, as it is described in detail in [44], the drift velocity of oxygen ions is modeled to be area dependent, which leads to less switching for increasing device areas and thus to a decreased $R \cdot A$. The retention time can be fitted with a power law function of $\sim t^{0.3}$, depicted in Fig. 4d by the red dots.

The devices are produced with DC magnetron sputtering of all materials without breaking the vacuum and subsequent standard optical lithography and edging steps on 100 mm Si wafers passivated with thermal SiO_2 . Device areas between 100 and $2500\ \mu\text{m}^2$ are fabricated on one and the same wafer. One clear advantage of this material system is that the stoichiometry of HfO_2 can easily be adjusted during sputtering. A $\text{HfO}_{1.98}$, which is named HfO_2 in this chapter for simplicity reasons, leads to the described device performance, while $\text{HfO}_{1.8}$ leads to filamentary-based switching, which is covered by Sects. 3 and 4.2.

The device model shown in Fig. 4c contains a Schottky diode formed between Au and HfO_2 , which is described by thermionic emission theory (see Eqs. 1–3). The oxides are both modeled by parallel RC circuits. The model described in detail in [44] reproduces the I - V characteristics very well, as depicted in Fig. 4a by the red solid line. Here, the resistive switching is dominated by Schottky-barrier height decreasing from 0.71 eV to 0.61 eV and the ideality factor increasing from 3.9 to 4.45 by considering oxygen ion movement as switching mechanism. The model also covers the decreasing $R \cdot A$ with increasing A .

3 Filamentary-Based Devices

Many transition metal oxide-based memristive devices show resistive switching associated with filamentary-type VCM [23, 32, 78]. VCM is triggered by the migration of field-assisted oxygen ion-related defects in transition metal oxides. The oxygen ion-related defects are typically oxygen vacancies that are much more mobile than the transition metal cations [79]. The oxygen vacancy migration and valence change of the cation sub-lattice often appear in the form of the formation and dissolution of a conductive filament in a localized area in the oxide leading to filamentary-type switching.

The resistive switching is typically observed after an initial electro-forming step in filamentary-type devices [79]. Double-positively charged oxygen vacancies (V_0^{++}) are generated by electro-forming, and the V_0^{++} s migrate under the applied electric field. If V_0^{++} s are close to each other, a single charged state (V_0^+) becomes more favorable. The cluster of V_0^+ s are arranged in a filament structure, where the V_0^+ has a significantly higher energy barrier for diffusion than V_0^{++} [80]. The localized conductive filament leads to reversible changes of local resistivity and overall device resistance [81]. The vacancies located in the filament have a high diffusion

energy barrier. Therefore, the filamentary-type devices typically show long retention properties.

The bi-layer metal oxide devices typically consist of an oxygen vacancy reservoir layer and a solid-state electrolyte layer for filamentary-type memristive devices [82]. Under an applied electric field, oxygen vacancies are injected into the solid-state electrolyte from the reservoir layer. The oxygen vacancies form a localized conductive filament in the electrolyte layer, which reduces the overall resistance. If the electrolyte layer builds a Schottky-like barrier at the interface to the contacting metallic electrode, the conductive filament leads to a lowering of the barrier height and, thus, the overall device resistance [50]. The filamentary devices can be recognized by the resistance-area product [83]. The resistance in the ON state (LRS) is independent of the electrode area size because the conductive filament is created in a localized form.

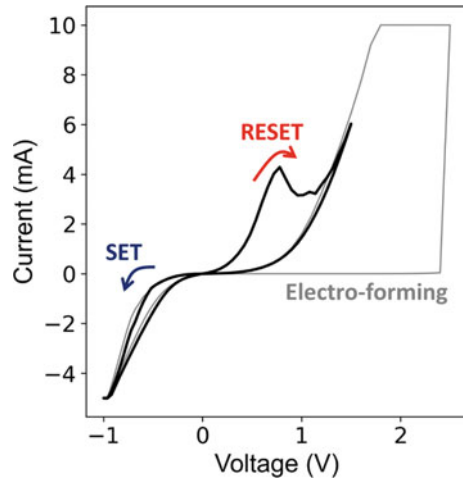
The benefits of filamentary RRAM devices are manifold. These devices can show a high endurance, high retention time, a fast switching speed (up to ns regime), and great scalability in the nm regime [38–42]. Moreover, they have already been integrated into CMOS technology, allowing to combine traditional electronic circuits with RRAM devices on the very same chip [77, 84–87], and even 3D integration has been shown [84, 88, 89]. Drawbacks are the need for an initial electro-forming step and a current compliance for switching, the need for selector devices to integrate them in memory arrays, pronounced intrinsic randomness of the switching process, a limited amount of achievable resistance states, and a rather low resistance [38–42, 65, 70]. However, a bi-layer metal oxide concept has recently attracted attention to tackle the intrinsic randomness issue in filamentary devices [57]. Besides, randomness can be exploited for certain tasks [90, 91]. Moreover, even though most RRAM devices can only be switched between two resistance states (binary devices), multilevel devices have already been reported. For this multi-state operation, a variable current compliance for switching to LRS and a variable switching voltage for switching to HRS are usually needed [45, 52, 92–94], or variable pulse widths can be used [45]. However, devices showing a gradual switching with identical pulses have also been reported [56, 95]. Moreover, more sophisticated programming algorithms are reported to reach specific resistive states [57, 85, 87, 88, 94, 96, 97].

In this section, a $\text{TiO}_x\text{-HfO}_x$ bi-layer oxide device is mainly presented as an example of a filamentary-based VCM memristive device [44]. Using the $\text{TiO}_x\text{-HfO}_x$ bi-layer oxide device, typical electrical characteristics of filamentary-based devices are shown.

The $\text{TiO}_x\text{-HfO}_x$ memristive devices consist of TiN (50 nm) as the bottom electrode, TiO_x (30 nm)/ HfO_x (2-8 nm) metal oxide bi-layers, and Au (50 nm) as the top electrode. The devices were fabricated on a 4-inch wafer, where the devices have 6 different active area sizes with a thickness gradient of the HfO_x layer from 2 nm to 8 nm. This 4-inch wafer fabrication process enables to exclude wafer-to-wafer variability, and study the effect of the area size and the thickness of HfO_x on the electrical properties of the memristive devices.

In Fig. 5, the $\text{TiO}_x\text{-HfO}_x$ bi-layer memristive devices show a typical I - V curve as a filamentary-type device. At first, an electro-forming step was required to initiate the

Fig. 5 I - V curve obtained by DC voltage sweep measurement showing RS and electro-forming in the filamentary $\text{TiO}_x/\text{HfO}_x$ devices



resistive switching in the $\text{TiO}_x/\text{HfO}_x$ memristive device. The electro-forming process was characterized by an abrupt jump in the current at a higher voltage (2.3 V) than the SET/RESET voltages. However, a high voltage through electro-forming could make it difficult for memristive devices to be integrated into CMOS circuits. Strategies to overcome this issue are already reported [45, 98, 99] and are addressed in Sect. 4.2.

After the electro-forming cycle, the resistive switching is observed in the $\text{TiO}_x/\text{HfO}_x$ memristive device. The switching required external current compliance of 5 mA for the SET process, while the RESET exhibited a compliance-free behavior. Bipolar switching properties were observed, and the amplitudes of switching voltages were symmetric, with -0.7 V and 0.7 V for SET and RESET, respectively. The $\text{TiO}_x/\text{HfO}_x$ device featured gradual switching characteristics, while typical filamentary type devices show an abrupt SET process. The origin of the abrupt transition during SET is related to positive feedback between the current increase and Joule heating [100]. In order to achieve a gradual switching behavior, a layer stack modification [100] and a multi-layer stack approach [51] have been proposed. In the multi-layer stack concept, for example, the $\text{TiO}_x/\text{HfO}_x$ bi-layer oxide can be exploited, where TiO_x serves as a reservoir for oxygen vacancies and stabilizes the switching process [57]. A gradual switching can promote uniformity during device operation. Furthermore, a gradual resistance transition during both SET and RESET is favorable for analog computing applications [101].

Gradual switching is often related to analog behavior. The analog behavior can be investigated using pulse measurements, applying a train of voltage pulses to memristive devices. Figure 6a shows the switching behaviors of the $\text{TiO}_x/\text{HfO}_x$ device in a pulse measurement. For the measurement, 20 SET voltage pulses followed by 20 RESET voltage pulses are applied, while a reading pulse is applied after each switching pulse to read the resistance of the device. The $\text{TiO}_x/\text{HfO}_x$ device shows a non-linear transition from the SET to the RESET, as shown in Fig. 6a. However,

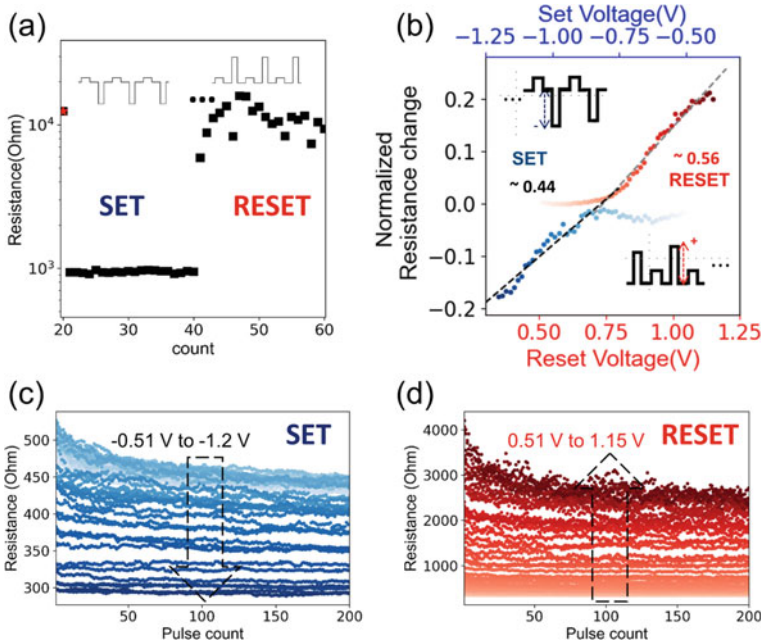


Fig. 6 **a** A train of 20 set voltage pulses and the subsequent 20 reset voltage pulses. A non-linear transition from the SET to the RESET is observed. **b** Linear trend in the resistance change in dependency on the amplitude of the switching pulse voltage. **c** Multi-state in the SET and **d** in the RESET resulting from the adjusted amplitude of switching voltages. **b** is reproduced from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

the magnitude of resistance change is influenced by the amplitude of the applied switching voltage pulses in Fig. 6c, d. The impact of pulse amplitude is observed by applying switching voltage pulses with varied amplitude, each followed by 200 reading pulses. Furthermore, the median values of resistance were extracted from Fig. 6c, d, and the resistance change from the initial state was calculated at each switching voltage amplitude. The resistance change showed symmetry in Fig. 6b between the SET and the RESET operation. Symmetric and analog resistive switching is desirable for neuromorphic computing since this makes the programming of the memristive device more efficient [102].

Another indication of the filamentary-type device can be the retention property. In Fig. 4d, the retention property of the $\text{TiO}_x/\text{HfO}_x$ device is compared to the interface-type device described in Sect. 2. The interface-type device has the same material configuration but fewer oxygen vacancies in the HfO_x layer ($x = 1.98$). The filamentary-type device shows a much longer retention time than the interface-type device due to the high diffusion barrier of the vacancies within the filament. This has been quantitatively shown by fitting the retention measurement to the $\sim t^\beta$ power law [61] introduced in Sect. 2, leading to $\beta = 0.02$ compared to $\beta = 0.3$ for the

interface-based devices. Thus, the relaxation of the resistance is proceeding orders of magnitude slower in the filamentary devices.

4 Engineering *I-V* Characteristics of Memristive Devices

The possibility of adapting electrical parameters of memristive devices allows engineering the device performance to meet the requirements for specific applications [45, 49]. Parameters such as switching voltage, forming voltage, HRS, LRS, the switching dynamics, or the number of resistance states can be engineered. Typical strategies are tailoring the switching oxide itself by adjusting the stoichiometry or the density of oxygen vacancies or ions [44, 45, 47–49], doping with additional cations [94], or using different oxide phases [91, 94]. Furthermore, adding additional oxide layers to affect the oxygen exchange [55–57] or the local temperature distribution [51] can impact the memristive behavior. Geometrical parameters like the switching oxide thickness [44, 71] or the device area [44, 55] can also influence the *I-V* characteristics. In Sect. 4.1, strategies for tailoring the interface-based devices introduced in Sect. 2 are described. This includes a thorough analysis of the process plasma during sputter deposition of DBMDs. Section 4.2 deals with engineering approaches to tailor the filament-based devices shown in Sect. 3.

4.1 Engineering of Interface-Based Devices

The devices described in Sect. 2 are fabricated by magnetron sputtering [103, 104]. The thin film deposition conditions like pressure, gas composition, structure and strength of the magnetic field, amplitude and form of discharge, and substrate position and orientation strongly influence electrical and optical material properties [105–109]. In this respect, gas flow rates and the resulting pressure significantly impact particle energies and, thereby, the structure and density of deposited films [106, 107]. Furthermore, the gas composition affects chemical reactions in the gas phase and at surfaces (target, substrate), influencing, e.g., the stoichiometry. Moreover, the substrate position and orientation relative to the sputter target affect film properties since particle fluxes and electric potentials in the plasma environment can show strong spatial in-homogeneity due to, e.g., structure and strength of the involved magnetic field [105, 107–109]. Another crucial parameter during reactive sputter deposition in an electronegative gas like O^2 is the density of negatively charged ions. If the target surface gets completely poisoned, negative oxygen ions are formed and accelerated by the electric field towards the substrate. This can cause defects or stress in deposited materials due to the high kinetic energy [105, 110]. The radial distribution, the energy, and the impinging angle of these ions depend on the erosion of the target (i.e., the target geometry) and the discharge voltage. It has been shown that the erosion state of a target can be correlated with the resistivity of oxides,

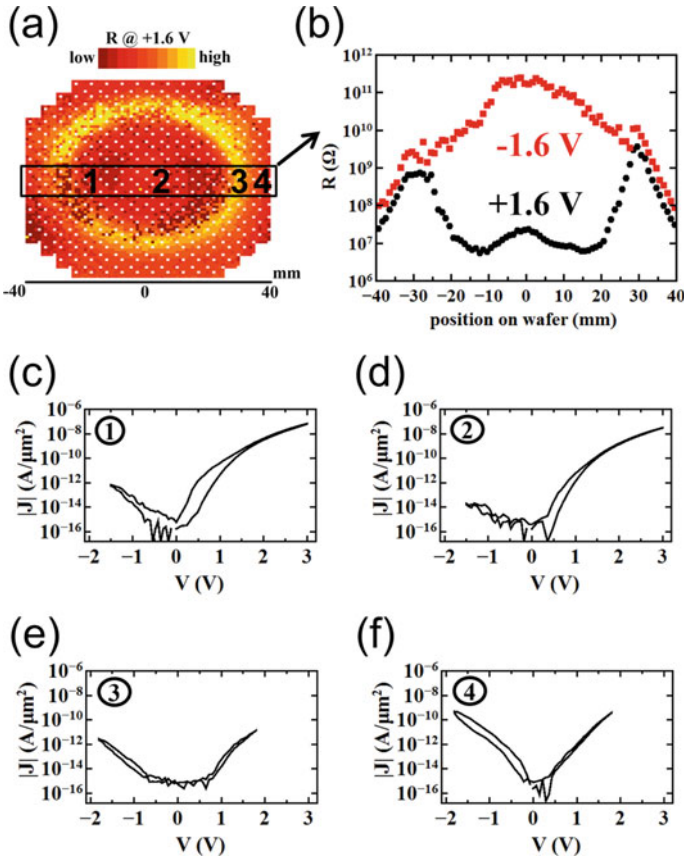


Fig. 7 Electrical characterization of DBMDs at different positions on a 100 mm wafer: **a** Resistance map measured with 1.6 V. **b** Device resistance dependent on radial position for positive bias (black dots) and negative bias (red squares). **c–f** Typical $|J|$ - V curves at different radial positions 1 to 4 (-25 mm, $+5$ mm, $+30$ mm, and $+35$ mm). Reprinted from [49] [F. Zahari et al., *Journal of Vacuum Science & Technology B* 37, 061203 (2019)], with the permission of AIP Publishing

attributed to the impact of negatively charged oxygen ions [105, 111]. The erosion is almost only taking place in a ring-shaped area (called race track) determined by the magnetic field.

To correlate the plasma properties to the electrical characteristics of DBMDs, several plasma parameters of the Nb_xO_y deposition process were measured at different positions underneath the Nb target since the I - V characteristics vary at different positions on the 100 mm wafer [49]. Figure 7 shows the device-to-device variation across the whole wafer. Figure 7a contains a resistance map measured with read-out voltages of 1.6 V on a spatial resolution of 1 mm, showing circular symmetry. Here, 5336 devices with an area of $400 \mu m^2$ were measured. Figure 7b depicts the dependence of the resistance for positive (black dots) and negative (red squares) bias on the radial position. Therefore, the resistances of up to ten devices

were averaged for each position. Defective devices (short-circuited or not properly connected to the wiring) and test structures [gray squares in Fig. 7a] were not used for averaging. In the inner area of the wafer with a radius of about 25 mm, current rectification is observed, and the devices show memristive behavior as described in Sect. 2. However, not only the resistance but also the switching window varies with position, as can be seen by comparing Fig. 7c with d. The yellow ring in Fig. 7a, starting at about 30 mm away from the center, contains devices with relatively high resistance. The I - V characteristics differ significantly from the functional devices in the center [Fig. 7e]. No rectification and no hysteresis can be observed. Here, the race track of the target is located. In the outer red area in Fig. 7a, the resistance is rather low and comparable to the resistance in the inner area. However, no rectification and no hysteresis are apparent [Fig. 7f]. It should be noted that I - V curves with lower voltages are depicted in Fig. 7e, f because the devices suffer dielectric breakdown for higher voltages at these locations.

As described above, devices located at different positions on a 100 mm wafer show significant different I - V characteristics even though they were produced in one and the same process. Thus, the DBMDs are well suited to correlate plasma process parameters with material properties and, finally, with electrical properties [49]. The overall goal is first to understand the impact of several plasma parameters on the electrical properties and second to use these findings to develop a process to fabricate plasma-engineered devices tailored for specific applications. Thus, plasma parameters were recorded with a probe consisting of a sensor copper plate (11 mm in diameter) connected at the backside to a thermocouple for temperature measurement ($40 \mu\text{V/K}$) and a copper bias wire [109]. This was used as a passive thermal probe (PTP) [109, 112] and as a Langmuir probe (LP) [113, 114]. A PTP is used to measure the energy balance at the location of the sample. The energy flux is dominated by impinging particles, surface reactions, incoming radiation, and loss processes such as heat conduction and convection through the surrounding gas and by emitted radiation [106]. Therefore, the energy balance depends on the process parameters named above, and it has a crucial impact on film properties [106, 107, 115]. The energy balance during Nb_xO_y deposition at different positions is shown in Fig. 8a. With an LP, on the other hand, additional process parameters like the electron temperature T_e [Fig. 8b], the floating potential of the probe Φ_{fl} [Fig. 8c], the plasma potential Φ_{pl} [Fig. 8d], and the current density of positively charged ions to the probe j_{ion} [Fig. 8e] are measured by applying a voltage sweep to the bias wire and evaluating the measured current [116, 117]. While the energy flux and j_{ion} show a Gaussian distribution with maxima in the center as expected for an unbalanced magnetron, the other parameters show a more complex dependency on the position. In particular, T_e and Φ_{pl} have local maxima at the location of the race track, while Φ_{fl} has local minima approx. 20 mm away from the center, where the devices show the largest switching window and the lowest resistance for positive bias together with a pronounced rectification [compare Fig. 7a–d]. Thus, it is expected that the effect of the energy balance superimposes with the other plasma parameters and with the distribution of negatively charged oxygen ions to obtain a Nb_xO_y film having properties (e.g., oxidation state and defect concentrations) as needed for memristive switch-

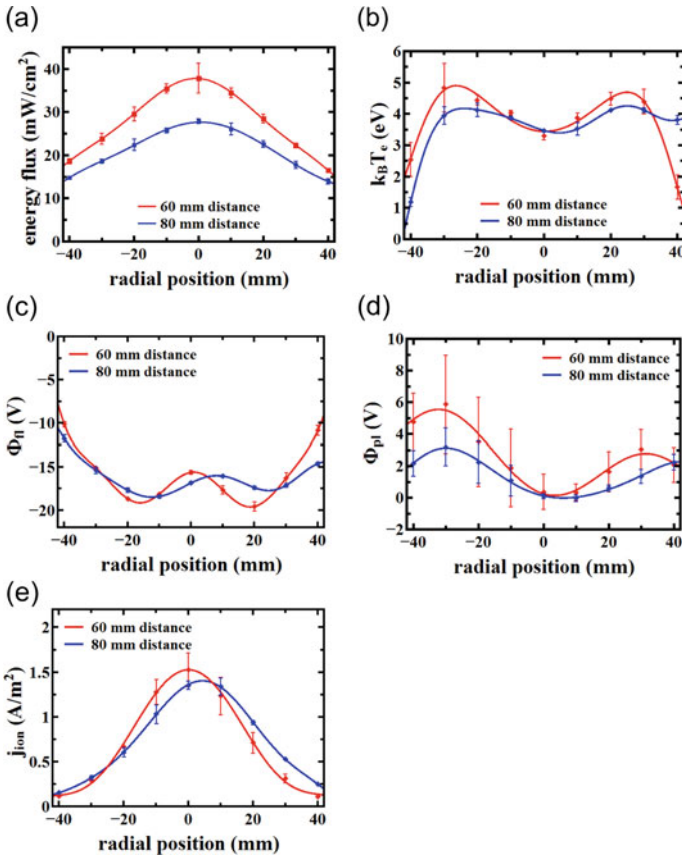


Fig. 8 Plasma parameters in dependency on the radial position. **a** Energy flux **b** Electron temperature, **c** Floating potential of probe **d** Plasma potential **e** Current of positive ions. In **a-e**, the mean values and the standard deviations determined with two measurements for each position (three for the center position) are depicted (squares with error bars). The solid and dashed lines show fits of the mean values. Results obtained with a target to probe distance of 60 mm and 80 mm are shown in red (squares/solid lines) and blue (squares/dashed lines), respectively. Reprinted from [49] [F. Zahari et al., Journal of Vacuum Science & Technology B 37, 061203 (2019)], with the permission of AIP Publishing

ing in DBMDs. This knowledge can potentially be exploited for engineering device performance by using multiple frequency capacitively coupled plasmas (MFCCPs) for sputter deposition. An MFCCP allows to control the ion energy and the ion flux independently of one another during sputtering and thus can help to control different process parameters [118, 119].

To show the influence of the plasma parameters on the stoichiometry of Nb_xO_y , TEM/EELS measurements were performed on two different devices, one located at a radial position of 5 mm (showing memristive behavior) and one at a radial position of -30 mm (showing relatively high resistance and no switching), respectively [49].

While Nb_xO_y of the functional device is in a high oxidation state between Nb_2O_5 and NbO_2 , as already shown before [69], the Nb_xO_y from the non-functional device shows significantly less incorporated oxygen and cannot be assigned to one of the known niobium oxidation states (i.e., Nb_2O_5 , NbO_2 , or NbO). Furthermore, the film thickness does vary only by 0.1 nm between both positions and thus within the standard deviation of the TEM measurement. The deviation in film thickness does not explain the substantial deviation in I - V characteristics since devices with both thicknesses lead to memristive switching devices if they are located within the center of the wafer. Thus, the deviations of the I - V characteristics can be mainly attributed to the chemical and electronic properties of the niobia, while the geometrical properties are less important. This can indeed be correlated to the plasma process parameters. Moreover, the kinetic Monte Carlo simulations [67] introduced in Sect. 2 were extended to show the impact of the concentration of the mobile charged defects on the switching window [49]. The results are depicted in Fig. 9a. In particular, it is shown that an increase in defect concentration n_d from $2 \cdot 10^{20} \text{cm}^{-3}$ to $8 \cdot 10^{20} \text{cm}^{-3}$ increases the switching window at 0.5 V from a factor of approx. 3 to a factor of approx. $4 \cdot 10^5$. Thus, increasing the defect concentration by a factor of 4 does lead to an increased switching window by five orders of magnitude. Increasing ion concentration in the interface-based $\text{HfO}_x/\text{TiO}_x$ device also leads to a larger switching window shown by simulations [44]. Furthermore, as it is shown in Fig. 9b, also the voltage range in which a hysteresis loop can be observed changes with ion concentration, and the reverse current is also more affected as in the model of the DBMDs.

4.2 Engineering of Filamentary-Based Devices

This section presents methodologies to engineer the characteristics of the filamentary-type memristive devices, such as electro-forming, I - V non-linearity, retention, and variability. In particular, engineering methods for the $\text{TiO}_x/\text{HfO}_x$ memristive devices, that are presented in Sect. 3, are further described.

The concentration of oxygen vacancies in the metal oxide is one of the decisive parameters for the filamentary-type switching behaviors [43]. In the $\text{TiO}_x/\text{HfO}_x$ memristive device, the HfO_x showed a substoichiometric property of $x = 1.8$ in X-ray photoelectron spectroscopy (XPS) measurements. The HfO_x layer had an optimal environment for efficient electro-forming and filamentary-type switching [43]. In the $\text{TiO}_x/\text{HfO}_x$ devices, oxygen deficiency in the HfO_x was achieved by deliberately growing a substoichiometric oxide film using a wedge sputtering method. Another approach to modifying the oxygen vacancies in the oxide layer is inserting an oxygen scavenging metal layer [120]. As already stated above, adjusting the vacancy concentration is a common way to adjust electrical switching parameters like forming voltage or switching voltages [45, 48, 98, 99]. Another method for scaling the forming voltage is tuning the thickness of the metal oxide layer. In Fig. 10, the electro-forming voltage in the $\text{TiO}_x/\text{HfO}_x$ device showed a dependency on the thickness of

Fig. 9 Modeled hysteresis curves dependent on the mobile defect concentration. **a** $|J|$ - V curves of DBMDs. Reprinted from [49] [F. Zahari et al., Journal of Vacuum Science & Technology B 37, 061203 (2019)], with the permission of AIP Publishing. **b** $|I|$ - V curves of interface-based TiN/TiO_x/HfO_x/Au device. Reprinted from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)

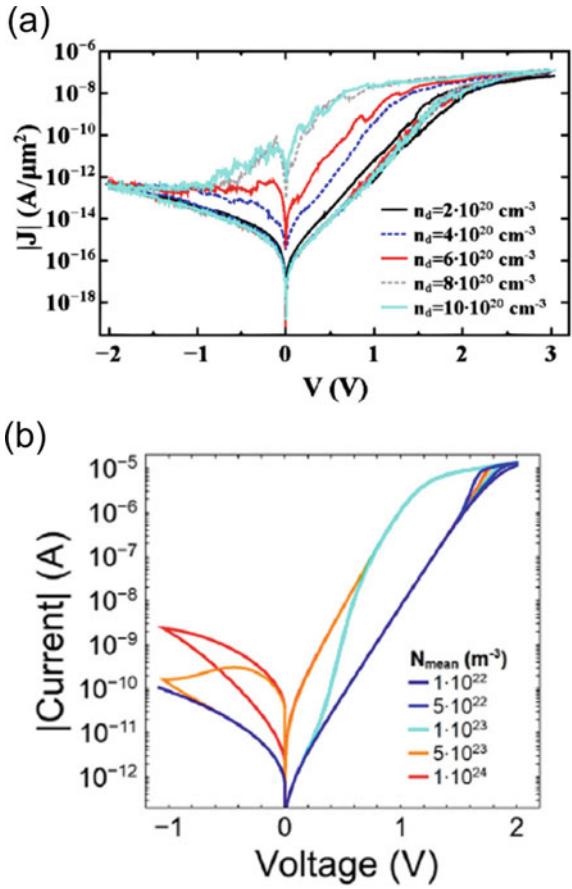
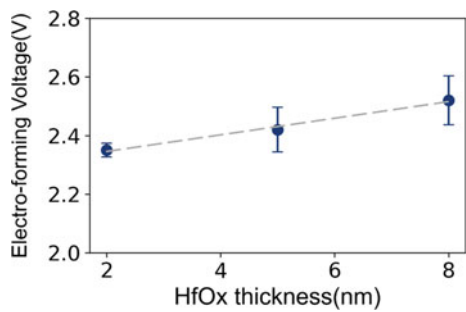


Fig. 10 Tailoring of the electro-forming voltage in the filamentary TiO_x/HfO_x devices by tuning the HfO_x thickness. Reprinted from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)



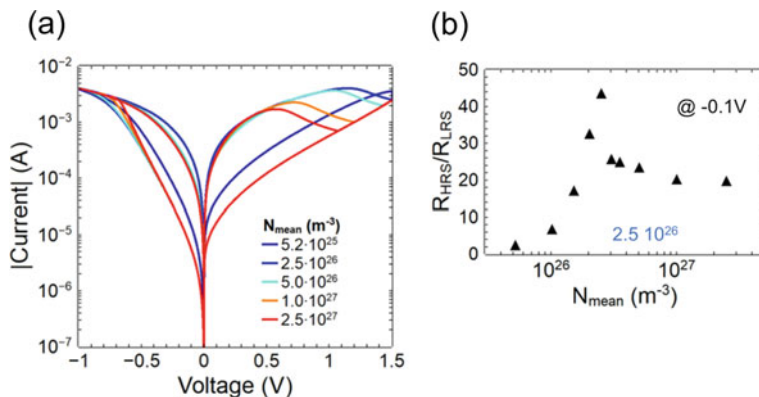
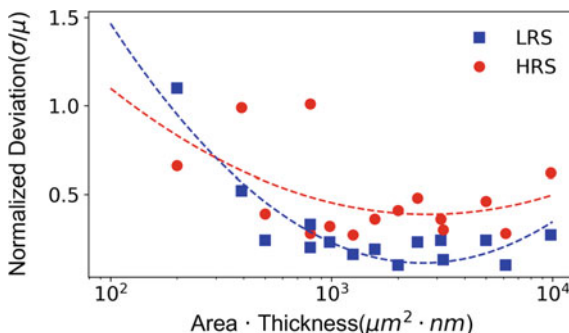


Fig. 11 Impact of the oxygen vacancy concentration on the switching hysteresis determined by simulations. Reprinted from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>). **a** Concentration of oxygen vacancies and resistive switching behavior. **b** Switching window as a function of the concentration of the oxygen vacancies

Fig. 12 TiO_x/HfO_x device variability dependent on the product of the active area and the HfO_x thickness. Reprinted from [44] (licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>)



the HfO_x . The thicker the switching oxide layer, the higher the voltage required to initiate resistive switching.

Besides impacting forming and switching voltages, the vacancy concentration also affects the switching hysteresis. In Fig. 11a, the simulation result shows such an impact on the TiO_x/HfO_x device. There is an optimum concentration of oxygen vacancies at $2.5 \cdot 10^{26} m^{-3}$, where the TiO_x/HfO_x device has the maximum switching window at a read-out voltage of $-0.1 V$, shown in Fig. 11b. The simulation shows that the device suffers from a higher concentration of oxygen vacancies due to the threshold value of oxygen diffusion [44].

Formation and dissolution of filaments are stochastic processes, which means the filamentary-type devices typically show a high variability during the switching. However, scaling the active area size down can reduce the variability [120]. Moreover, the study of the TiO_x/HfO_x filamentary devices shows that the variability can be influenced by the correlation between the thickness of the oxide layer and the active

area size. As shown in Fig. 12, an optimal combination of the oxide thickness and the active area size exists to obtain reliable switching operation.

The comparison of the retention time of interface-based and filamentary-type $\text{TiO}_x/\text{HfO}_x$ devices [see Fig. 4d] implies that the modification of oxygen deficiency in the HfO_x layer can control the retention property. In this respect, many studies show methods to enhance the retention time by introducing retention stabilization layers to modify the oxygen deficiency [54, 121].

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 434434223 - SFB 1461. Moreover, the financial support by the DFG via the Research Unit 2093 is acknowledged.

References

1. Hickmott, T.W.: Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **33**, 2669–2682 (1962)
2. Argall, F.: Switching phenomena in titanium oxide thin films. *Solid-State Electron.* **11**, 535–541 (1968)
3. Dearnaley, G., Stoneham, A.M., Morgan, D.V.: Electrical phenomena in amorphous oxide films. *Rep. Prog. Phys.* **33**, 1129–1191 (1970)
4. Chua, L.: Memristor-the missing circuit element. *IEEE Trans. Circuit Theory* **18**, 507–519 (1971)
5. Strukov, D.B., Snider, G.S., Stewart, D.R., et al.: The missing memristor found. *Nature* **453**, 80–83 (2008)
6. Iniewski, K. (ed.): *CMOS Processors and Memories*. Springer Netherlands (2010). ISBN: 978-90-481-9215-1. <https://doi.org/10.1007/978-90-481-9216-8>
7. Banerjee, W.: Challenges and applications of emerging nonvolatile memory devices. *Electronics* **9**, 1029 (2020)
8. Linares-Barranco, B., Serrano-Gotarredona, T., Camu-as-Mesa, L., et al.: On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci.* **5**, 26 (2011)
9. Ranjan, R., Ponce, P.M., Hellweg, W.L., et al.: Integrated circuit with memristor emulator array and neuron circuits for biologically inspired neuromorphic pattern recognition. *J. Circuit Syst. Comp.* **26**, 1750183 (2017)
10. Hu, M., Graves Catherine, E., Li, C., et al.: Memristor-based analog computation and neural network classification with a dot product engine. *Adv. Mater.* **30**, 1705914 (2018)
11. Ziegler, M., Wenger, C., Chicca, E., et al.: Tutorial: concepts for closely mimicking biological learning with memristive devices: principles to emulate cellular forms of learning. *J. Appl. Phys.* **124**, 152003 (2018)
12. Wang, R., Yang, J.-Q., Mao, J.-Y., et al.: Recent advances of volatile memristors: devices, mechanisms, and applications. *Adv. Intell. Syst.* **2**, 2000055 (2020)
13. Spiga, S., Sebastian, A., Querlioz, D., et al.: *Memristive Devices for Brain-Inspired Computing: From Materials, Devices, and Circuits to Applications - Computational Memory, Deep Learning, and Spiking Neural Networks*. Woodhead Publishing (2020). ISBN: 978-0-08-102787-5
14. Bian, H., Goh, Y.Y., Liu, Y., et al.: Stimuli-responsive memristive materials for artificial synapses and neuromorphic computing. *Adv. Mater.* **33**, 2006469 (2021)
15. Sun, K., Chen, J., Yan, X.: The future of memristors: materials engineering and neural networks. *Adv. Funct. Mater.* **31**, 2006773 (2021)

16. Kaspar, C., Ravoo, B.J., van der Wiel, W.G., et al.: The rise of intelligent matter. *Nature* **594**, 345–355 (2021)
17. Hebb, D.O.: *The Organization of Behavior: A Neuropsychological Theory*. Psychology Press, New York (2002). ISBN: 978-1-4106-1240-3. <https://doi.org/10.4324/9781410612403>
18. Bi, G.-Q., Poo, M.-M.: Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10464–10472 (1998)
19. Bliss, T.V.P., Lømo, T.: Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* **232**, 331–356 (1973)
20. Jo, S.H., Chang, T., Ebong, I., et al.: Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **10**, 1297–1301 (2010)
21. Ohno, T., Hasegawa, T., Tsuruoka, T., et al.: Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater* **10**, 591–595 (2011)
22. Winterfeld, H., Ziegler, M., Hanssen, H., et al.: Technology and electrical characterization of memflash cells for neuromorphic applications. *J. Phys. D: Appl. Phys.* **51**, 324003 (2018)
23. Ielmini, D., Waser, R. (eds.): *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*. Wiley-VCH Verlag GmbH & Co., KGaA, Weinheim (2016). ISBN: 978-3- 527-33417-9
24. Sawa, A.: Resistive switching in transition metal oxides. *Mater. Today* **11**, 28–36 (2008)
25. Waser, R., Dittmann, R., Menzel, S., et al.: Introduction to new memory paradigms: memristive phenomena and neuromorphic applications. *Faraday Discuss.* **213**, 11–27 (2019)
26. Wang, J., Zhuge, X., Zhuge, F.: Hybrid oxide brain-inspired neuromorphic devices for hardware implementation of artificial intelligence. *Sci. Technol. Adv. Mater.* **22**, 326–344 (2021)
27. Lanza, M., Waser, R., Ielmini, D., et al.: Standards for the characterization of endurance in resistive switching devices. *ACS Nano* **15**, 17214–17231 (2021)
28. Kawahara, T., Ito, K., Takemura, R., et al.: Spin-transfer torque RAM technology: review and prospect. *Microelectron. Reliab.* **52**, 613–627 (2012)
29. Garcia, V., Bibes, M.: Ferroelectric tunnel junctions for information storage and processing. *Nat. Commun.* **5**, 4289 (2014)
30. Burr, G.W., Brightsky, M.J., Sebastian, A., et al.: Recent progress in phase-change memory technology. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **6**, 146–162 (2016)
31. Valov, I., Waser, R., Jameson, J.R., et al.: Electrochemical metallization memories-fundamentals, applications, prospects. *Nanotechnology* **22**, 254003 (2011)
32. Waser, R., Dittmann, R., Staikov, G., et al.: Redox-based resistive switching memories - nanoionic mechanisms, prospects, and challenges. *Adv. Mater.* **21**, 2632–2663 (2009)
33. Park, S., Jung, S., Siddik, M., et al.: Self-formed schottky barrier induced selector-less rram for cross-point memory applications. *Physica status solidi (RRL) – Rapid Res. Lett.* **6**, 454–456 (2012)
34. Hsu, C.-W., Wang, Y.-F., Wan, C.-C., et al.: Homogeneous barrier modulation of TaO_x/TiO₂ bilayers for ultra-high endurance three-dimensional storage-class memory. *Nanotechnology* **25**, 165202 (2014)
35. Meyer, R., Schloss, L., Brewer, J., et al.: Oxide dual-layer memory element for scalable non-volatile cross-point memory technology. In: 2008 9th Annual Non-Volatile Memory Technology Symposium (NVMTS), pp. 1–5 (2008). <https://doi.org/10.1109/NVMT.2008.4731194>
36. Govoreanu, B., Redolfi, A., Zhang, L., et al.: Vacancy-modulated conductive oxide resistive ram (vmco-rram): an area-scalable switching current, self-compliant, highly nonlinear and wide on/off-window resistive switching cell. In: 2013 IEEE International Electron Devices Meeting, pp. 10.2.1–10.2.4 (2013). <https://doi.org/10.1109/IEDM.2013.6724599>
37. Baeumer, C., Heisig, T., Arndt, B., et al.: Spectroscopic elucidation of ionic motion processes in tunnel oxide-based memristive devices. *Faraday Discuss.* **213**, 215–230 (2019)
38. Burr, G.W., Shelby, R.M., Sebastian, A., et al.: Neuromorphic computing using non-volatile memory. *Adv. Phys. X* **2**, 89–124 (2017)

39. Ielmini, D., Wong, H.-S.P.: In-memory computing with resistive switching devices. *Nat. Electron* **1**, 333–343 (2018)
40. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R., et al.: Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020)
41. Christensen, D.V., Dittmann, R., Linares-Barranco, B., et al.: 2022 Roadmap on Neuromorphic Computing and Engineering (2021). [arXiv:2105.05956](https://arxiv.org/abs/2105.05956) [cond-mat]
42. Wong, H.S.P., Lee, H.Y., Yu, S., et al.: Metal-oxide RRAM. *Proc. IEEE* **100**, 1951–1970 (2012)
43. McKenna, K.P.: Optimal stoichiometry for nucleation and growth of conductive filaments in HfO_x . *Modell. Simul. Mater. Sci. Eng.* **22**, 025001 (2014)
44. Park, S., Klett, S., Ivanov, T., et al.: Engineering method for tailoring electrical characteristics in $\text{TiN}/\text{TiO}_x/\text{HfO}_x/\text{Au}$ bi-layer oxide memristive devices **3**, 16 (2021)
45. He, W., Sun, H., Zhou, Y., et al.: Customized binary and multi-level HfO_2 -x-based memristors tuned by oxidation conditions. *Sci. Rep.* **7**, 10070 (2017)
46. Sokolov, A.S., Jeon, Y.-R., Kim, S., et al.: Influence of oxygen vacancies in ALD HfO_2 -x thin films on non-volatile resistive switching phenomena with a Ti/HfO_2 -x/ Pt structure. *Appl. Surf. Sci.* **434**, 822–830 (2018)
47. Shibuya, K., Dittmann, R., Mi, S., et al.: Impact of defect distribution on resistive switching characteristics of Sr_2TiO_4 thin films. *Adv. Mater.* **22**, 411–414 (2010)
48. Skaja, K., Andrä, M., Rana, V., et al.: Reduction of the forming voltage through tailored oxygen non-stoichiometry in tantalum oxide ReRAM devices. *Sci. Rep.* **8**, 10861 (2018)
49. Zahari, F., Schlichting, F., Strobel, J., et al.: Correlation between sputter deposition parameters and I-V characteristics in double-barrier memristive devices. *J. Vac. Sci. Technol. B* **37**, 061203 (2019)
50. Asanuma, S., Akoh, H., Yamada, H., et al.: Relationship between resistive switching characteristics and band diagrams of $\text{Ti}/\text{Pr}_{1-x}\text{Ca}_x\text{MnO}_3$ junctions. *Phys. Rev. B* **80**, 235113 (2009)
51. Hardtdegen, A., Torre, C.L., Cüppers, F., et al.: Improved switching stability and the effect of an internal series resistor in $\text{HfO}_2/\text{TiO}_x$ bilayer ReRAM cells. *IEEE Trans. Electron Devices* **65**, 3229–3236 (2018)
52. Yu, S., Wu, Y., Jeyasingh, R., et al.: An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans. Electron Devices* **58**, 2729–2737 (2011)
53. Niu, G., Calka, P., Huang, P., et al.: Operando diagnostic detection of interfacial oxygen ‘breathing’ of resistive random access memory by bulk-sensitive hard x-ray photoelectron spectroscopy. *Mater. Res. Lett.* **7**, 117–123 (2019)
54. Baeumer, C., Schmitz, C., Ramadan, A.H.H., et al.: Spectromicroscopic insights for rational design of redox-based memristive devices. *Nat. Commun.* **6**, 8610 (2015)
55. Hansen, M., Ziegler, M., Kolberg, L., et al.: A double barrier memristive device. *Sci. Rep.* **5**, 13753 (2015)
56. Woo, J., Moon, K., Song, J., et al.: Improved synaptic behavior under identical pulses using $\text{AlO}_x/\text{HfO}_2$ bilayer RRAM array for neuromorphic systems. *IEEE Electron Device Lett.* **37**, 994–997 (2016)
57. Stathopoulos, S., Khiat, A., Trapatseli, M., et al.: Multibit memory operation of metal-oxide bi-layer memristors. *Sci. Rep.* **7**, 17532 (2017)
58. Sawa, A., Fujii, T., Kawasaki, M., et al.: Hysteretic current–voltage characteristics and resistance switching at a rectifying $\text{Ti}/\text{Pr}_0.7\text{Ca}_0.3\text{MnO}_3$ interface. *Appl. Phys. Lett.* **85**, 4073–4075 (2004)
59. Seong, D.-J., Jo, M., Lee, D., et al.: HPHA effect on reversible resistive switching of Pt/Nb -doped SrTiO_3 schottky junction for nonvolatile memory application. *Electrochem. Solid-State Lett.* **10**, H168 (2007)
60. Yoon, J.H., Song, S.J., Yoo, I.-H., et al.: Highly uniform, electroforming-free, and self-rectifying resistive memory in the $\text{Pt}/\text{Ta}_2\text{O}_5/\text{HfO}_2$ -x/ TiN structure. *Adv. Funct. Mater.* **24**, 5086–5095 (2014)

61. Mikheev, E., Hoskins, B.D., Strukov, D.B., et al.: Resistive switching and its suppression in Pt/Nb:SrTiO₃ junctions. *Nat. Commun.* **5**, 3990 (2014)
62. Kim, Y., Kwon, Y.J., Kwon, D.E., et al.: Nociceptive memristor. *Adv. Mater.* **30**, 1704320 (2018)
63. Choi, S., Kim, Y., Nguyen, T.V., et al.: Low-power self-rectifying memristive artificial neural network for near internet-of-things sensor computing. *Adv. Electron. Mater.* **7**, 2100050 (2021)
64. Luo, Q., Xu, X., Gong, T., et al.: 8-layers 3D vertical RRAM with excellent scalability towards storage class memory applications. In: 2017 IEEE International Electron Devices Meeting (IEDM), pp. 2.7.1–2.7.4 (2017). <https://doi.org/10.1109/IEDM.2017.8268315>
65. Burr, G.W., Shenoy, R.S., Virwani, K., et al.: Access devices for 3D crosspoint memory. *J. Vac. Sci. Technol. B* **32**, 040802 (2014)
66. Hansen, M., Ziegler, M., Kohlstedt, H.: Double barrier memristive devices for neuromorphic computing. In: 2016 IEEE International Conference on Rebooting Computing (ICRC), pp. 1–8 (2016). <https://doi.org/10.1109/ICRC.2016.7738713>
67. Dirkmann, S., Hansen, M., Ziegler, M., et al.: The role of ion transport phenomena in memristive double barrier devices. *Sci. Rep.* **6**, srep35686 (2016)
68. Hansen, M., Zahari, F., Kohlstedt, H., et al.: Unsupervised Hebbian learning experimentally realized with analogue memristive crossbar arrays. *Sci. Rep.* **8**, 8914 (2018)
69. Strobel, J., Hansen, M., Dirkmann, S., et al.: In depth nano spectroscopic analysis on homogeneously switching double barrier memristive devices. *J. Appl. Phys.* **121**, 245307 (2017)
70. Zidan, M.A., Fahmy, H.A.H., Hussain, M.M., et al.: Memristor-based memory: the sneak paths problem and solutions. *Microelectron. J.* **44**, 176–183 (2013)
71. Kim, H.-D., Crupi, F., Lukosius, M., et al.: Resistive switching characteristics of integrated polycrystalline hafnium oxide based one transistor and one resistor devices fabricated by atomic vapor deposition methods. *J. Vac. Sci. Technol. B* **33**, 052204 (2015)
72. Li, C., Hu, M., Li, Y., et al.: Analogue signal and image processing with large memristor crossbars. *Nat. Electron* **1**, 52–59 (2018)
73. Kiani, F., Yin, J., Wang, Z., et al.: A fully hardware-based memristive multilayer neural network. *Sci. Adv.* **7**, eabj4801 (2021)
74. Sze, S.M.: *Physics of Semiconductor Devices*, pp. 134–196. Wiley (2006). ISBN: 978-0-470-06832-8
75. Tung, R.T.: The physics and chemistry of the schottky barrier height. *Appl. Phys. Rev.* **1**, 011304 (2014)
76. Simmons, J.G.: Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**, 1793–1803 (1963)
77. Walczyk, D., Walczyk, C., Schroeder, T., et al.: Resistive switching characteristics of CMOS embedded HfO₂-based 1T1R Cells. In: *Microelectronic Engineering. Proceedings of the 17th Biennial International Insulating Films on Semiconductor Conference* **88**, 1133–1135 (2011)
78. Lim, E.W., Ismail, R.: Conduction mechanism of valence change resistive switching memory: a survey. *Electronics* **4**, 586–613 (2015)
79. Yang, J.J., Miao, F., Pickett, M.D., et al.: The mechanism of electroforming of metal oxide memristive switches. *Nanotechnology* **20**, 215201 (2009)
80. Duncan, D., Magyari-Köpe, B., Nishi, Y.: Filament-induced anisotropic oxygen vacancy diffusion and charge trapping effects in hafnium oxide RRAM. *IEEE Electron Device Lett.* **37**, 400–403 (2016)
81. Zidan, M.A., Chen, A., Indiveri, G., et al.: Memristive computing devices and applications. *J. Electroceramics* **39**, 4–20 (2017)
82. Xiong, W., Zhu, L.Q., Ye, C., et al.: Bilayered oxide-based cognitive memristor with brain-inspired learning activities. *Adv. Electron. Mat.* **5**, 1900439 (2019)
83. Bagdzevicius, S., Maas, K., Boudard, M., et al.: Interface-type resistive switching in perovskite materials. *J. Electroceramics* **39**, 157–184 (2017)
84. Chen, H.-Y., Yu, S., Gao, B., et al.: HfO_x based vertical resistive random access memory for cost-effective 3D cross-point architecture without cell selector. In: 2012 International Electron Devices Meeting, pp. 20.7.1–20.7.4 (2012). <https://doi.org/10.1109/IEDM.2012.6479083>

85. Mochida, R., Kouno, K., Hayata, Y., et al.: A 4M synapses integrated analog ReRAM based 66.5 TOPS/W neural-network processor with cell current controlled writing and flexible network architecture. In: 2018 IEEE Symposium on VLSI Technology, pp. 175–176 (2018). <https://doi.org/10.1109/VLSIT.2018.8510676>
86. Valentian, A., Rummens, F., Vianello, E., et al.: Fully integrated spiking neural network with analog neurons and RRAM synapses. In 2019 IEEE International Electron Devices Meeting (IEDM), pp. 14.3.1–14.3.4 (2019). <https://doi.org/10.1109/IEDM19573.2019.8993431>
87. Le, B.-Q., Grossi, A., Vianello, E., et al.: Resistive RAM with multiple bits per cell: array-level demonstration of 3 bits per cell. *IEEE Trans. Electron Devices* **66**, 641–646 (2019)
88. Adam, G.C., Hoskins, B.D., Prezioso, M., et al.: 3-D memristor crossbars for analog and neuromorphic computing applications. *IEEE Trans. Electron Devices* **64**, 312–318 (2017)
89. Lin, P., Li, C., Wang, Z., et al.: Three-dimensional memristor circuits as complex neural networks. *Nat. Electron.* **3**, 225–232 (2020)
90. Suri, M., Bichler, O., Querlioz, D., et al.: CBRAM devices as binary synapses for low-power stochastic neuromorphic systems: auditory (Cochlea) and visual (Retina) cognitive processing applications. In: 2012 International Electron Devices Meeting, pp. 10.3.1–10.3.4 (2012). <https://doi.org/10.1109/IEDM.2012.6479017>
91. Zahari, F., Pérez, E., Mahadevaiah, M.K., et al.: Analogue pattern recognition with stochastic switching binary CMOS-integrated memristive devices. *Sci. Rep.* **10**, 14450 (2020)
92. Wang, S.-Y., Huang, C.-W., Lee, D.-Y., et al.: Multilevel resistive switching in Ti/Cu_xO/Pt memory devices. *J. Appl. Phys.* **108**, 114110 (2010)
93. Wu, Y., Yu, S., Wong, H.-S. P., et al.: AlO_x-based resistive switching device with gradual resistance modulation for neuromorphic device application. In: 2012 4th IEEE International Memory Workshop, pp. 1–4 (2012). <https://doi.org/10.1109/IMW.2012.6213663>
94. Milo, V., Zambelli, C., Olivo, P., et al.: Multilevel HfO₂-based RRAM devices for low-power neuromorphic networks. *APL Mater.* **7**, 081120 (2019)
95. Perez, E., Mahadevaiah, M.K., Perez-Bosch Quesada, E., et al.: In-depth characterization of switching dynamics in amorphous hfo₂ memristive arrays for the implementation of synaptic updating rules. *Jpn. J. Appl. Phys.* (2022). <https://doi.org/10.35848/1347-4065/ac6a3b>
96. Alibart, F., Gao, L., Hoskins, B.D., et al.: High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm. *Nanotechnology* **23**, 075201 (2012)
97. Pérez, E., Pérez-Ávila, A.J., Romero-Zalaz, R., et al.: Optimization of multi-level operation in RRAM arrays for in-memory computing. *Electronics* **10**, 1084 (2021)
98. Bousoulas, P., Michelakaki, I., Skotadis, E., et al.: Low-power forming free TiO_{2-x}/HfO_{2-y}/TiO_{2-x}-trilayer RRAM devices exhibiting synaptic property characteristics. *IEEE Trans. Electron Devices* **64**, 3151–3158 (2017)
99. Milano, G., Raffone, F., Luebben, M., et al.: Water-mediated ionic migration in memristive nanowires with a tunable resistive switching mechanism. *ACS Appl. Mater. Interfaces* **12**, 48773–48780 (2020)
100. Cüppers, F., Menzel, S., Bengel, C., et al.: Exploiting the switching dynamics of HfO₂-based ReRAM devices for reliable analog memristive behavior. *APL Mater.* **7**, 091105 (2019)
101. Tan, Z.J., Somjit, V., Toparli, C., et al.: Electronegative metal dopants improve switching consistency in Al₂O₃ resistive switching devices (2021). [arXiv:2104.13301](https://arxiv.org/abs/2104.13301)
102. Wang, Z., Yin, M., Zhang, T., et al.: Engineering incremental resistive switching in TaO_x based memristors for brain-inspired computing. *Nanoscale* **8**, 14015–14022 (2016)
103. Kelly, P.D., Arnell, R.D.: *Magnetron Sputtering : A Review of Recent Developments and Applications* (2000)
104. Bräuer, G., Szyszka, B., Vergöhl, M., et al.: Magnetron sputtering – milestones of 30 years. In: *Vacuum. Selected Papers from the Proceedings of The Tenth International Symposium on Sputtering and Plasma Processes (ISSP 2009)*, 8th–10th July 2009, Kanazawa, Japan **84**, 1354–1359 (2010)
105. Minami, T., Miyata, T., Yamamoto, T., et al.: Origin of electrical property distribution on the surface of ZnO: Al films prepared by magnetron sputtering. *J. Vac. Sci. Technol. A* **18**, 1584–1589 (2000)

106. Kersten, H., Deutsch, H., Steffen, H., et al.: The energy balance at substrate surfaces during plasma processing. *Vacuum* **63**, 385–431 (2001)
107. Bornholdt, S., Itagaki, N., Kuwahara, K., et al.: Characterization of the energy flux toward the substrate during magnetron sputter deposition of ZnO thin films. *Plasma Sources Sci. Technol.* **22**, 025019 (2013)
108. Haase, F., Lundin, D., Bornholdt, S., et al.: On the impact of electron temperature in magnetron sputtering benchmarked with energy flux measurements. *Contrib. Plasma Phys.* **55**, 701–713 (2018)
109. Gauter, S., Haase, F., Kersten, H.: Experimentally unraveling the energy flux originating from a DC magnetron sputtering source. *Thin Solid Films* **669**, 8–18 (2019)
110. Ellmer, K., Welzel, T.: Reactive magnetron sputtering of transparent conductive oxide thin films: role of energetic particle (Ion) bombardment. *J. Mater. Res.* **27**, 765–779 (2012)
111. Welzel, T., Ellmer, K.: The influence of the target age on laterally resolved ion distributions in reactive planar magnetron sputtering. *Surface and coatings technology. PSE 2010 Special Issue* **205**, S294–S298 (2011)
112. Thornton, J.A.: Substrate heating in cylindrical magnetron sputtering sources. *Thin Solid Films* **54**, 23–31 (1978)
113. Mott-Smith, H.M., Langmuir, I.: The theory of collectors in gaseous discharges. *Phys. Rev.* **28**, 727–763 (1926)
114. Cherrington, B.E.: The use of electrostatic probes for plasma diagnostics—a review. *Plasma Chem. Plasma Process* **2**, 113–140 (1982)
115. Kersten, H., Kroesen, G.M.W., Hippler, R.: On the energy influx to the substrate during sputter deposition of thin aluminium films. *Thin Solid Films* **332**, 282–289 (1998)
116. Piejak, R., Godyak, V., Alexandrovich, B., et al.: Surface temperature and thermal balance of probes immersed in high density plasma. *Plasma Sources Sci. Technol.* **7**, 590–598 (1998)
117. Piel, A.: *Plasma Physics: An Introduction to Laboratory, Space, and Fusion Plasmas* isbn: 978-3-642-43631-4. Springer, Berlin Heidelberg (2010)
118. Heil, B.G., Czarnetzki, U., Brinkmann, R.P., et al.: On the possibility of making a geometrically symmetric RF-CCP discharge electrically asymmetric. *J. Phys. D: Appl. Phys.* **41**, 165202 (2008)
119. Bienholz, S., Bibinov, N., Awakowicz, P.: Multiple frequency capacitively coupled plasmas as a new technology for sputter processes. *J. Phys. D: Appl. Phys.* **46**, 084010 (2013)
120. Lee, J., Shin, J., Lee, D., et al.: Diode-less nano-scale ZrO_x/HfO_x RRAM device with excellent switching uniformity and reliability for high-density cross-point memory applications. In: *2010 International Electron Devices Meeting*, pp. 19–5 (2010)
121. Huang, H.-M., Wang, Z., Wang, T., et al.: Artificial neural networks based on memristive devices: from device to system. *Adv. Intell. Syst.* **2**, 2000149 (2020)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



MemFlash—Floating Gate Transistors as Memristors



Henning Winterfeld, Hermann Kohlstedt, and Martin Ziegler

Abstract The idea of resistive switching devices is originally based on the fact that the application of electric fields changes the atomic structure locally and thus also the electronic structure of the material. This leads globally to a sustained change in the resistance of the material layer, which is generally referred to as resistive switching. In resistive switching devices, these atomic reconfigurations are reversible and allow the state to be maintained for a long time, which is why the devices are referred to as memristive devices (also named Memristor). Memristive devices can be realized as two terminal devices in a metal-insulator-metal structure. In the MemFlash cell, there is no atomic rearrangement in the device and therefore is a purely electronic based switching device. The basic components of the MemFlash cell are floating gate transistors, which are reduced from a three-terminal to a two-terminal device by means of a diode-like wiring scheme and thus exhibit memristive switching behavior. In this book chapter, the MemFlash cell is introduced.

Keywords Neuromorphic computing · Memristive devices · Floating gate transistors

1 Introduction

Memristive devices (also called memristors) are mostly based on the effect of resistive switching. In resistive switching, the application of an electric field causes a local change in the atomic bonding structure of the crystalline atoms of the memristive layer. Resistive switching is caused by different mechanism, ranging from magnetic

H. Winterfeld · H. Kohlstedt
Nanoelectronics, Department of Electrical and Information Engineering, Kiel University,
Kaiserstraße 2, 24143 Kiel, Germany
e-mail: hko@tf.uni-kiel.de

M. Ziegler (✉)
Department of Micro- and Nanoelectronic Systems, Institute of Micro- and Nanotechnologies
(IMN) MacroNano, Technische Universität Ilmenau, 98684 Ilmenau, Germany
e-mail: martin.ziegler@tu-ilmenau.de

effects through electrostatic effects to atomic configurations. A good overview is given in [1]. Based on these effects, a number of device structures for memristive devices have been developed [2] and the research on resistive switching devices goes back to [3, 4].

Even though research on resistive switching materials goes back decades [3, 4], interest in the subject has been renewed in recent years. In this context, memristive switching materials for non-volatile memory architectures [5] and especially for neuromorphic systems [6] are considered as promising candidates. However, both of these applications require technology to manufacture memristive devices on a large scale, as well as integration or interconnection with the prevailing silicon technology [7]. The latter puts research into CMOS-compatible memristive materials and their CMOS integration at the center of research interest. This goes hand in hand with the development of CMOS integrated memristive devices [7, 8].

In this context, a MemFlash cell offer an interesting alternative to conventional memristive devices [9, 10], because it is a modified version of a floating gate transistor (FGT), which are CMOS integrable devices and commercially available. In a MemFlash cell resistive switching is caused exclusively by a pure electronic effect and not based on an atomic reconstruction.

In this chapter we will introduce the MemFlash cell. First, the concept of the MemFlash cell will be explained and it will be shown how to create a memristive device from a FGT. Subsequently, in chapter “Redox-Based Bi-Layer Metal Oxide Memristive Devices”, aspects of the MemFlash cell that are technological important will be discussed including an overview of different realizations and extensions of the basic concept. Thereafter, the range of applications for neuromorphic computing is discussed in this chapter and important properties and conditions for the cell design are presented.

2 The MemFlash Concept

Memristive devices are two-terminal devices, which in their simplest structure can be realized as a metal-insulator-metal sandwich structure. In a memristive device the resistance depends on the charge flow through in the past [11]. It was concluded that this is a fundamental property that classifies memristive devices, along with the resistor, the inductor and the capacitor, to one a fundamental device of electronics [11]. In this context, the reference to the Memristor postulated by Leo Chua already in 1971 could be drawn [12]. Even if this conclusion is critically questioned [13, 14], memristive devices or memristors show fundamental properties which cannot be reproduced with one of the other three circuit elements: common with the capacitor and the inductor is the possibility of an information storage (memory effect). However, common with the resistor, memristors are passive devices in which no energy storage takes place. The memristor is therefore, as the name suggests, a memory-resistor.

Mathematically, memristive devices can be formally represented by the following set of equations:

$$I = G(x, V, T) \cdot V \quad \text{with} \quad \frac{dx}{dt} = f(x, V, t), \quad (1)$$

here x is called a state variable, which is for example a measure of the change in atomic configuration within the resistive switching material. The function $f(x, V, t)$, describes the voltage driven changes in the atomic structure [11, 15] and is depending on the particular memristive system. The conductance of the device $G(x, V, t)$ is named as Memductance and a simple relationship can be made between G and the state variable x via the relationship [11]:

$$G = x \cdot G_{on} + (1 - x) \cdot G_{off}, \quad (2)$$

where G_{on} and G_{off} are the maximum and minimum conductance of the device, respectively. Please note, that the state variable x varies between 0 and 1. However, the MemFlash cell is not based on atomic reconfigurations but is relies on the FGT. An FGT is purely function on electron charge transfer. In this context, it will be shown in the following how this device can be converted into a memristive device.

2.1 Functional Principle of a MemFlash Cell

A schematic of a simple memory transistor is shown in Fig. 1a. Basically, it consists of a MOS field-effect transistor (MOSFET) with a MOS structure (Metal Oxide Semiconductor) in whose gate stack. This MOS stack gives a memory capacitance as a gate which controls the channel of the MOSFET and is called a floating gate (FG) [16]. Depending on the charge on the FG, the vertical charge flow from source to drain through the transistor is influenced. For example, if we consider an n-MOS transistor as shown in Fig. 1a, the drain and source terminals are both heavily n-doped while the channel region is p-doped (die Bulk region könnte in der Abbildung mit p-Si versehen werden). Depending on the voltage difference between the gate electrode and the bulk junction, electrons are accumulated at the interface to the gate oxide, resulting in a conductive connection (channel) between the drain and the source junction. If the floating gate is sufficiently well isolated from its environment and the potential on the external accessible control gate is kept constant, the charge state of the floating gate defines the memory state of the transistor. However, floating gate transistors are not memristors at first glance, since memristors are two-terminal devices while the transistor essentially has three terminals—source, drain and the externally addressable control gate (cf. Fig. 1a). At this respect, the core idea of a MemFlash cell is an external circuit that allows the device to become externally a two-terminal device. A circuit that allows this is shown in Fig. 1a as well. Here, the control gate (CG) and source are combined and connected to a common potential,

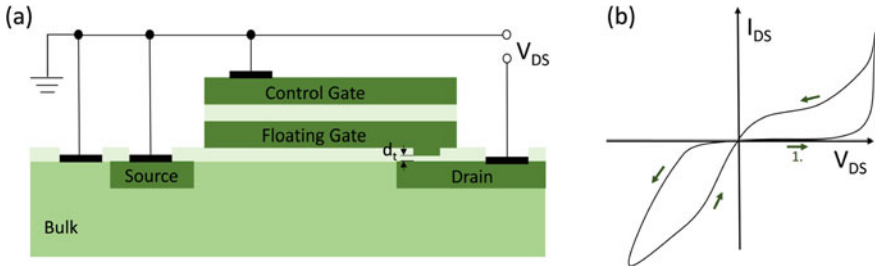


Fig. 1 The concept a MemFlash cell: wiring scheme and I-V curve. **a** Schematic drawing of a MemFlash cell and the diode-like circuitry that turns a FGT into a MemFlash cell (see bold black connection between source and gate). **b** Typical current-voltage curve of a MemFlash cell through the circuit shown in **(a)**. The pinched hysteresis is a fingerprint of a memristor

i.e. the ground potential. Applying a bipolar voltage signal V_{DS} (see Fig. 1) the FG is charged or discharged in dependency of the bias direction via the tunneling window with the thickness d_t and leads to a pinched hysteresis for I_{DS} vs. V_{DS} [9]. In other words, the charging or discharging of FG changes immediately the channel current I_{DS} .

As shown in Fig. 1b, a pinched hysteresis loop can be realized with this connection as a fingerprint of a memristor. The current-voltage characteristic shown in Fig. 1b reflects the curve measured in Ref. [9] and describes the dependence of the channel current (I_{DS}) when a bipolar voltage sweep is applied to the drain contact (V_{DS}). Here fore, the voltage was increased linearly from zero and then decreased linearly to a negative voltage before returning to the origin (see arrows in Fig. 1b). The result is a hysteresis. In order to understand this memristive behavior more precisely, a capacitive model is developed and described in the following section.

2.2 Physical Device Model of a MemFlash Cell

The model used to describe the MemFlash cell is sketched in Fig. 2a and is taken from Refs. [9, 10]. In this capacitive model, the potential of the floating gate is expressed via

$$V_{FG} = \frac{Q_{FG}}{C_T} + k_C V_C + k_D V_D + k_B V_B \quad (3)$$

Here, Q_{FG} is the charge stored on the floating gate, and V_C , V_D , V_S , and V_B are the potentials of the control gate, drain, source, and bulk terminal, respectively. Furthermore, k_C , k_D , k_S , and k_B are the respective coupling constants to the floating gate which are defined by $k_i = C_i/C_T$ ($i = C, D, S, B$). Here, C_C , C_D , C_S , and C_B are the respective capacitances and C_T is the total capacitance given by $C_T = C_C + C_D + C_S + C_B$. In the memristive mode, as already mentioned, a three-terminal

floating gate transistor is reconfigured to a two-terminal cell. Therefore, the control gate and source potentials are grounded. If we also neglect coupling to the bulk, the floating gate potential equation reduces to the following form:

$$V_{FG} = \frac{Q_{FG}}{C_T} + k_D V_D \tag{4}$$

Thus, the floating gate potential is determined by the drain voltage V_D and by the charge on the floating gate Q_{FG} , which in turn is determined by V_D and can be calculated by

$$Q_{FG} = Q_{FG}(t_0) + \int_{t_0}^{t_1} I_t(V_D, V_{FG})dt \tag{5}$$

Here, $I_t(V_D, V_{FG})$ is the current that charges or discharges the F_G through the tunnel window (the tunnel window is marked by a capacitance C_t in circuit model in Fig. 2a). In general, there are different contributions to that current, but the main current contributions come from the Fowler–Nordheim tunnel effect and from the injection of hot charge carriers [10]. These two effects are outlined in Fig. 2b. In the case of Fowler–Nordheim tunneling, the application of a high electric field effectively thins the thickness of the tunnel oxide, leading to the tunneling of electrons. While hot charge carrier injection is a thermally generated charge transport mechanism in which charge carriers can overcome the potential barrier between drain and floating

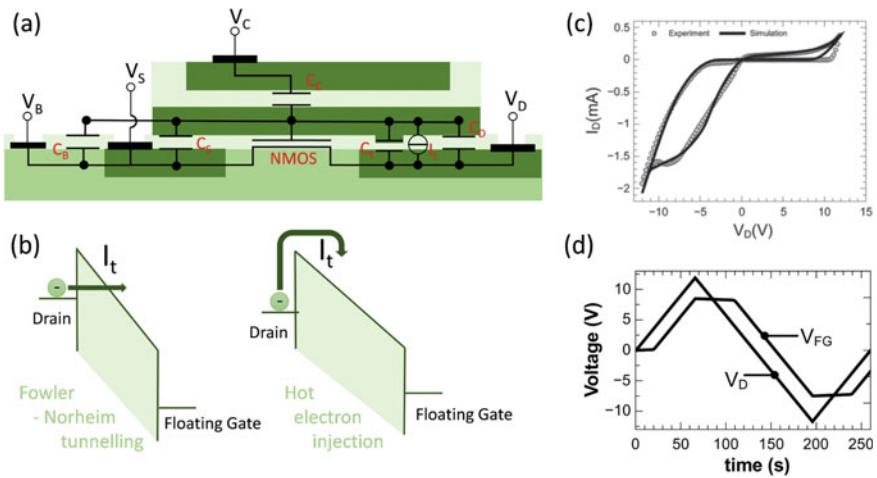


Fig. 2 Physical model of a MemFlash device. **a** Equivalent circuit model of a FGT. **b** Different charge and discharge mechanisms for the floating gate. Current-voltage characteristics of a MemFlash cell (taken from Ref. [10]). **c** The solid line shows the model, while the dots correspond to a real measurement. **d** Floating gate potential curve when a triangular voltage is applied to the drain (taken from Refs. [9, 10])

gate at high voltages. Both mechanisms can be described within analytical equations and are represented in the model shown in Fig. 2 over two current sources [10]. Here, the Fowler–Nordheim current is given by

$$I_{FN} = \pm A_{tox} \cdot A_{FN} \cdot (E_{tox})^2 \exp(-B_{FN}/E_{tox}) \quad (6)$$

where A_{tox} is the tunneling oxide area, and A_{FN} and B_{FN} are the Fowler–Nordheim tunneling constants. E_{tox} is the electrical field during programming and erasing, $E_{tox} = (V_D - V_{FG})/d_{tox}$, and depends on the thickness of the tunneling gate oxide d_t . According to Ref. [10] the equation for the hot carrier injection current can be approximated by

$$I_{inj} = \pm A_{tox} \cdot A_{inj} \exp\left(-\frac{B_{inj}}{(c_{inj} + V_{FG})^2} + D_{inj} \cdot V_D\right) \quad (7)$$

where A_{inj} , B_{inj} , C_{inj} and D_{inj} are positive fit constants, which have to be estimated from experimental data. In order to describe the current-voltage characteristics with the equations, a theoretical description of the MOSFET is still missing. This can best be described by the following set of equations:

$$I_D = \begin{cases} \frac{1}{2}\beta \left[(V_G - V_{th})V_D - \frac{V_D^2}{2} \right] \cdot (1 + \lambda V_D) & \text{for } V_G > V_{th} \text{ and } V_D < V_G - V_{th} \\ \frac{1}{2}\beta (V_G - V_{th})^2 \cdot (1 + \lambda V_D) & \text{for } V_G > V_{th} \text{ and } V_D > V_G - V_{th} \\ 0 & \text{for } V_G < V_{th} \end{cases} \quad (8)$$

Here V_{th} , β and λ are the threshold voltage, transconductance and channel length modulation parameter of the MOS transistor, respectively. V_G is the gate potential, which is equal to $V_D - V_{FG}$ for $V_D < 0$ and $V_G = V_{FG}$ for $V_D > 0$. Figure 2c shows the current-voltage characteristic calculated with the model by using a linear triangular voltage sweep and the following set of parameters: $\lambda = 0.0625 \text{ V}^{-1}$, $\beta = 28.3 \mu\text{SV}^{-1}$ and $V_{th} = 1.052 \text{ V}$. In addition, Fig. 2c compares the model with a typical experimental I-V curve.

Due to the good agreement between modeled and measured I-V characteristics, the physical function mechanism can be well described via the model: for positive drain voltage, the floating gate charge is reduced, while for negative voltages, the QFG is charged. This changes the channel resistance from the high resistance state (HRS) to the low resistance state (LRS) for positive drain voltages and vice versa from LRS to HRS for negative voltages. The associated change in floating gate potential is shown in Fig. 2d and demonstrates that by changing the voltage at the drain terminal, the floating gate potential can be varied. This forms the kernel of the MemFlash cell's memristive behavior and thus the charge on the floating gate can be identified as the memristive state variable x , where the following relationship holds:

$$\frac{dx}{dt} = f(x, V, t) \rightarrow \frac{dQ_{FG}}{dt} = I_t(G_{FG}, V_D, V_{FG}) \quad (9)$$

Furthermore, the dynamic function f in the case of the MemFlash cell can be represented by the charge current of the floating gate. This allows the memristive behavior to be adjusted directly via the charge transport mechanism of the floating gate and thus to tailor the cell for its application. The latter will be described in more detail below.

3 Technology and Scaling Issues of the MemFlash

A major advantage of MemFlash cells over most of the conventional Memristors is their compatibility with CMOS technology. The floating gate transistor described in the previous chapters was manufactured in a NMOS technology and is thus already present in an established technology [17]. This enables the production of large numbers of memristive cells with low variability and high yield. This is crucial for the application of Memristors in neuromorphic computing and remains one of the major challenges of conventional Memristor technologies. However, as shown in the last paragraph, in the MemFlash cell the memristive behavior is mainly described by the tunnel current for charging and discharging the floating gate. This places a special technological emphasis on the tunnel oxide and the associated requirements in terms of storage duration and transport mechanisms. This defines the technological parameters for the current-voltage characteristics of the MemFlash cell and will be discussed in the following section. In addition, it will be shown that the concept of the MemFlash cell is not limited to FGTs, but can be applied to different types of thin film junction-less memory transistors in general. This offers a wide range of technological possibilities to tailor the behavior of the MemFlash cell for particular application.

3.1 Tunneling Oxide Scaling

Memory transistors from the last section, in the form of the EEPROM (electrically erasable programmable read-only memory) cell, form the basis of today's non-volatile storage media such as flash memory. These memory transistors are manufactured using the floating gate tunneling oxide (FLOTIX) technology shown in Fig. 3a [16]. Within this technology, the gate oxide is thinned at the drain connection so that electrons can pass through the tunneling oxide via the transport processes described in the last section. This changes the charge on the floating gate, but depends on a number of factors, including programming time (the longer the time, the greater the change), temperature (the higher the temperature, the fewer hot electrons are available for injection), tunneling oxide quality and imposes stringent conditions on the

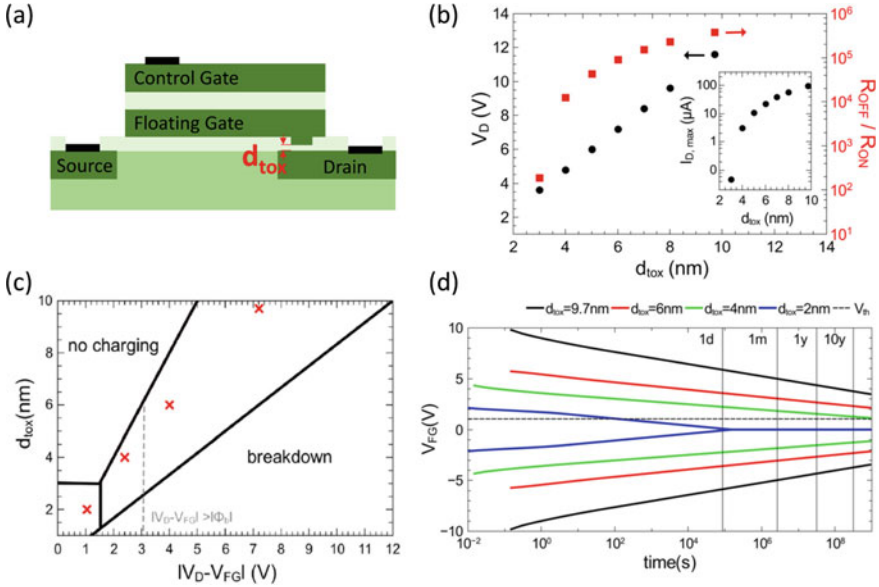


Fig. 3 Scaling of tunneling oxide layer: **a** the thickness of the gate oxide d_t has a significant influence on the voltage required and thus on the power consumption. **b** Dependence of V_D and the maximum change in resistance R_{off}/R_{on} on the thickness of the tunnel oxide. **c** Simplified schematic representation of the effective tunneling regime via the Fowler–Nordheim tunneling process. No charge or discharge was assumed for tunnel currents smaller than 10^{-19} A. An electric field strength of 12 MV cm^{-1} was used for the breakdown condition. **d** Retention of the memory cell. Shown here via the change in the floating gate potential as a function of the layer thickness. (Figures are taken from Ref. [10])

programming voltage required. In particular, a higher voltage amplitude and larger programming times are required compared to Memristors, whose resistor switching mechanism is based on atomic reconfiguration. Thus, the technological factor that plays the leading role here is the tunnel oxide thickness, which is chosen in such a way that the memory cell allows information storage for more than 10 years [10, 17]. However, in a number of applications, such as neuromorphic computing, short storage times are acceptable or even desirable and the required voltage of the MemFlash cell can be reduced accordingly by decreasing the layer thickness [10].

An estimation of the drain voltage for different film thicknesses was made in Ref. [10] and the main results are shown in Fig. 3. Obviously, by reducing the film thickness, the required voltage can be significantly reduced to below 10 V (see black data points in Fig. 3b), while still maintaining relatively strong resistance changes as depicted as R_{off}/R_{on} in Fig. 3b (see red data points). Furthermore, the maximum current is also significantly reduced (see inset in Fig. 3b), which leads to a significant reduction in the power consumption of the cell [10].

In addition, the effective intervals for a charge transport between drain and FG were estimated in Ref. [10]. In this case, the voltage interval is limited by the break-

down field strength of SiO₂ to large voltages, while the use of a too low voltage leads to the fact that no change can be achieved at the floating gate. As shown in Fig. 3c, this interval becomes smaller for quite thin oxide thicknesses. It has been found that oxide thicknesses in the range of 4–5 nm lead to a required voltage at the drain between 4 and 5 V for FGTs [10]. However, these values lead to quite short storage times, as shown in Fig. 3d, and thus rule out a broad application of the MemFlash cell in this form. Nonetheless, the concept of the MemFlash cell can be applied to different types of memory transistors. In the following section, this issue will be discussed in more detail.

3.2 Different Types of MemFlash

Figure 4 gives an overview of different MemFlash cell designs. Besides the MemFlash cell which is based on a floating gate transistor (labeled as FLOTOX in Fig. 4), SONOS (silicon oxide nitride oxide polysilicon) transistors [18], quantum dot devices [19] and also memory transistors with a gate which splits into a y geometry [20] have been realized so far. In the following section, these other types of MemFlash cells will be briefly discussed.

In Ref. [18], the MemFlash concept was applied to industrially manufactured SONOS field-effect transistors. To account for the special characteristics of depletion-type SONOS transistors, the two-terminal wiring of the MemFlash cell was extended to include a resistance-bridged pn diode. Common SONOS devices have very thin layers of the ONO gate stack, with an approximate thickness of 2 nm for the tunnel oxide (O), 5–10 nm for the nitride layer (N), and 5 nm for the barrier oxide (O) toward

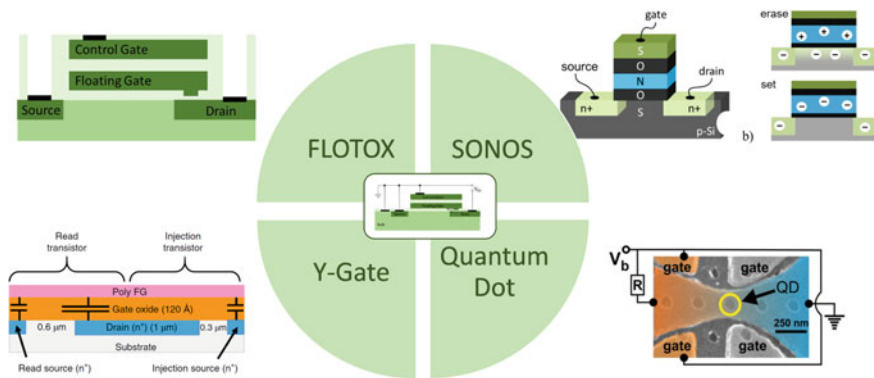


Fig. 4 Different memory transistors used as MemFlash device: these include floating-gate oxide (FLOTOX) transistors [9], silicon oxide nitride oxide polysilicon (SONOS) field-effect transistors [18], transistors with a Y-gate [19], and quantum dot (QD) floating gate transistors [20]

the gate electrode. Thus, this type of MemFlash cells promise reliable operation, low parameter spread and high integration density [18].

A completely different approach was taken in Ref. [19]. Here a MemFlash cell was realized with a quantum dot (QD) floating gate transistor by shorting the source contact of the quantum wire (QW) and lateral gates. Here, the memristive behavior arises from the Coulomb interaction of localized QD charges with the nearby QW. This leads to a change in conductance that depends on the voltage polarity.

Finally, an energy-efficient memristive floating-gate device was presented in Ref. [20]. This device was fabricated in a standard 180 nm CMOS process and connected to a readout transistor. This concept, known as a Y-flash cell, allows the device to operate in an energy-efficient subthreshold memristive mode. This allows small signal changes to be linearized, so that a dynamic resistance range of two orders of magnitude can be achieved. This allows 65 discrete resistance levels to maintain long data retention, high endurance and a low noise margin.

An extension of a memory transistor by another transistor was proposed in Ref. [21]. The idea of this approach is to combine a MemFlash based on a complementary n-MOS transistor/p-MOS transistor arrangement, similar to conventional CMOS inverters (not shown in Fig. 4). This concept, called C-MemFlash, allows the voltage polarity and the direction of the resistor switching to be adjusted by means of an appropriate wiring scheme.

4 MemFlash Cells for Neuromorphic Computing

One of the increasingly important areas of application for Memristors is neuromorphic computing [22]. Here, Memristors are used to emulate synaptic plasticity, i.e. the changes in coupling strengths between neurons. In neuronal networks, synapses play a special role, since these variable changes enable local and decentralized information processing. Thus, synaptic plasticity is a basic building block for learning and memory processes in neuronal networks and a detailed modeling is required to achieve functional neuronal networks [23]. A large number of different models that mimics neuronal learning are known, whereby the theory of Hebbian learning models allows a quite extensive description of these [24]. In the following, the possible applications of MemFlash cells in the field of neuromorphic computing will be presented in the context of this model.

4.1 Hebbian Learning Models

Hebbian learning models goes back to a rule postulated by Donald Hebb in 1949 [25] which states that “neurons that fire together wire together.” This means that neurons that are simultaneously active change their connection strength also named as synaptic weight. This learning rule can be easily translated into a mathematical

framework by denoting the change in the coupling strength is defined by a function which describes the activities A_j or A_i of the adjacent neurons, i.e. the pre- and post-synaptic neurons, and the actual weight of the synapse ω_{ij} [26]:

$$\frac{d\omega_{ij}}{dt} = F(\omega_{ij}, A_j, A_i) \quad (10)$$

In the simplest case the function $F(\omega_{ij}, A_j, A_i)$ can be written as:

$$F(\omega_{ij}, A_j, A_i) = \alpha A_j \cdot A_i \quad (11)$$

Here, the coefficient α is called learning rate and is usually positive: $\alpha > 0$. A learning rule with $\alpha < 0$ is usually called anti-Hebbian [26]. However, the exciting question is how this learning theory can be applied to the MemFlash cell. Here, a structural similarity between the equation for the change of the synaptic weight ω_{ij} and the change of the memristive state x can directly be recognized, which allows us to derive the following relation [27]:

$$\frac{d\omega}{dt} \rightarrow \frac{dx}{dt} \rightarrow \frac{dQ_{FG}}{dt} = I_t(Q_{FG}, V_D, V_{FG}) \quad (12)$$

Thus, the change in synaptic coupling strength can be directly modeled by the change in floating-gate charge [24, 27].

Furthermore, the learning function $F(\omega_{ij}, A_j, A_i)$ is simulated via the tunnel current and the pre- and postsynaptic activities can be identified as the potentials at the drain and floating gate. For this purpose, voltage pulse trains at the drain terminal can be used, as sketched in Fig. 5a, which must be adjusted in their voltage level, pulse length and frequency to emulate the desired learning behavior. An example that emulates the change of synaptic weight via the conductance of the MemFlash cell is shown in Fig. 5b. Here, constant frequency pulse trains of 175 10 V set pulses (also called potentiation pulses) and 175 -7.5 V reset pulses (also called depression pulses) were used, which were varied in their pulse width. A read pulse of $+2.5$ V was consecutively applied after each potentiation/depression pulse to non-destructively readout the channel conductance. Figure 5c shows the normalized synaptic weight changes $\omega(t)$ (floating-gate charges) for the different pulse widths. Here, a strong nonlinearity of the weight change in the respective first potentiation and depression pulses can be observed. An exception is the 5 ms curve, which shows no unlearning behavior. This result can also be seen from the curves shown in Fig. 5d. Good agreement with a variety of biological plasticity processes has been discussed in Ref. [24]. The application of cellular learning paradigms in the area of associative learning was also demonstrated in [24]. In addition, a broad range of applications of Y-gate MemFlash cells for neuromorphic architectures was presented in Ref. [20].

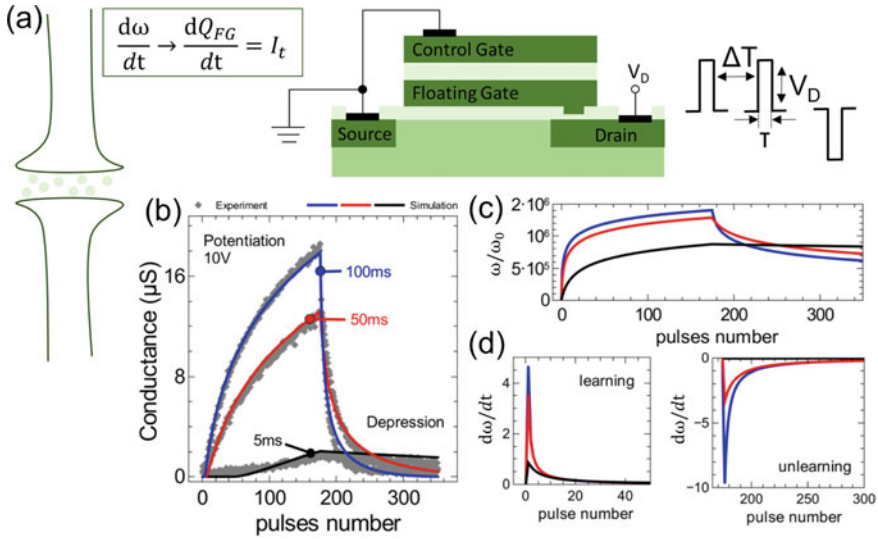


Fig. 5 Application of MemFlash cell in neuromorphic computing: **a** replication of synaptic plasticity as a change in FG charge. For this purpose, the floating gate charge can be adjusted by suitable pulse trains at the drain potential and thus the conductance of the MemFlash cell. **b** an example of an emulated plasticity curve from Ref. [24]. This shows a strong non-linear behavior in the change of conductance (c). **d** The learning rates are thus initially very high and become lower with the number of pulses. (Figure **b–d** are taken from [24])

5 Conclusion

In summary, it can be stated that by means of defined charge changes of the floating gate, controlled learning models can be realized in hardware. In addition, the existing technology integration of MemFlash cells with the associated good reproducibility and high production capability of a large number of cells offers a good possibility for the realization of neuromorphic computing architectures in hardware.

Acknowledgements The authors acknowledge financial support via the Deutsche Forschungsgemeinschaft (DFG) by the Research Unit 2093 entitled: Memristive Devices for Neuronal Systems. We thank Georg Schöneweger for carefully reading the manuscript.

References

1. Ielmini, D., Waser, R. (eds.): Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications. Wiley, New York (2015)
2. Waser, R., Dittmann, R., Menzel, S., Noll, T.: Introduction to new memory paradigms: memristive phenomena and neuromorphic applications. *Faraday Discuss.* **213**, 11–27 (2019)

3. Hickmott, T.W.: Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **33**(9), 2669–2682 (1962)
4. Dearnaley, G., Stoneham, A.M., Morgan, D.V.: Electrical phenomena in amorphous oxide films. *Rep. Prog. Phys.* **33**(3), 1129 (1970)
5. Wang, Z., Wu, H., Burr, G.W., Hwang, C.S., Wang, K.L., Xia, Q., Yang, J.J.: Resistive switching materials for information processing. *Nat. Rev. Mater.* **5**(3), 173–195 (2020)
6. Jeong, D.S., Hwang, C.S.: Nonvolatile memory materials for neuromorphic intelligent machines. *Adv. Mater.* **30**(42), 1704729 (2018)
7. Chen, W.H., Dou, C., Li, K.X., Lin, W.Y., Li, P.Y., Huang, J.H., Chang, M.F.: CMOS-integrated memristive non-volatile computing-in-memory for AI edge processors. *Nat. Electr.* **2**(9), 420–428 (2019)
8. Mahadevaiah, M.K., Pérez, E., Wenger, C., Grossi, A., Zambelli, C., Olivo, P., ... , Ziegler, M.: Reliability of cmos integrated memristive hfo₂ arrays with respect to neuromorphic computing. In: 2019 IEEE International Reliability Physics Symposium (IRPS), pp. 1–4. IEEE (2019)
9. Ziegler, M., Oberländer, M., Schroeder, D., Krautschneider, W.H., Kohlstedt, H.: Memristive operation mode of floating gate transistors: a two-terminal MemFlash-cell. *Appl. Phys. Lett.* **101**(26), 263504 (2012)
10. Riggert, C., Ziegler, M., Schroeder, D., Krautschneider, W.H., Kohlstedt, H.: MemFlash device: floating gate transistors as memristive devices for neuromorphic computing. *Semicond. Sci. Technol.* **29**(10), 104011 (2014)
11. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. *Nature* **453**(7191), 80–83 (2008)
12. Chua, L.: Memristor-the missing circuit element. *IEEE Trans. Circuit Theory* **18**(5), 507–519 (1971)
13. Vongehr, S., Meng, X.: The missing memristor has not been found. *Sci. Rep.* **5**(1), 1–7 (2015)
14. Meuffels, P., Soni, R.: Fundamental issues and problems in the realization of memristors (2012). [arXiv:1207.7319](https://arxiv.org/abs/1207.7319)
15. Chua, L.O., Kang, S.M.: Memristive devices and systems. *Proc. IEEE* **64**(2), 209–223 (1976)
16. Sze, S.M., Li, Y., Ng, K.K.: *Physics of Semiconductor Devices*. Wiley, New York (2021)
17. Winterfeld, H., Ziegler, M., Hanssen, H., Friedrich, D., Benecke, W., Kohlstedt, H.: Technology and electrical characterization of MemFlash cells for neuromorphic applications. *J. Phys. D Appl. Phys.* **51**(32), 324003 (2018)
18. Himmel, N., Ziegler, M., Mähne, H., Thiem, S., Winterfeld, H., Kohlstedt, H.: Memristive device based on a depletion-type SONOS field effect transistor. *Semicond. Sci. Technol.* **32**(6), 06LT01 (2017)
19. Maier, P., Hartmann, F., Mauder, T., Emmerling, M., Schneider, C., Kamp, M., Worschech, L.: Memristive operation mode of a site-controlled quantum dot floating gate transistor. *Appl. Phys. Lett.* **106**(20), 203501 (2015)
20. Danial, L., Pikhay, E., Herbelin, E., Wainstein, N., Gupta, V., Wald, N., Kvatinisky, S.: Two-terminal floating-gate transistors with a low-power memristive operation mode for analogue neuromorphic computing. *Nat. Electron.* **2**(12), 596–605 (2019)
21. Ziegler, M., Günther, R., Kohlstedt, H.: Complementary floating gate transistors with memristive operation mode. *IEEE Electron Device Lett.* **37**(2), 186–189 (2016)
22. Burr, G.W., Shelby, R.M., Sebastian, A., Kim, S., Kim, S., Sidler, S., Virwani, K. et al.: Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2**(1), 89–124 (2017). <https://doi.org/10.1080/23746149.2016.1259585>
23. Ziegler, M., Wenger, C., Chicca, E., Kohlstedt, H.: Tutorial: concepts for closely mimicking biological learning with memristive devices: principles to emulate cellular forms of learning. *J. Appl. Phys.* **124**(15), 152003 (2018)
24. Ziegler, M., Riggert, C., Hansen, M., Bartsch, T., Kohlstedt, H.: Memristive Hebbian plasticity model: device requirements for the emulation of Hebbian plasticity based on memristive devices. *IEEE Trans. Biomed. Circuits Syst.* **9**(2), 197–206 (2015)

25. Hebb, D.O. : The Organization of Behavior: A Neuropsychological Theory, 11. [print]. Wiley, New York (1974)
26. Gerstner, W., Kistler, W.M., Naud, R., Paninski, L.: Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition. Cambridge University Press, Cambridge (2014)
27. Ziegler, M., Kohlstedt, H.: Mimic synaptic behavior with a single floating gate transistor: a MemFlash synapse. J. Appl. Phys. **114**(19), 194506 (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Critical Discussion of Ex situ and In situ TEM Measurements on Memristive Devices



Ole Gronenberg, Georg Haberfehlner, Finn Zahari, Richard Marquardt, Christian Kübel, Gerald Kothleitner, and Lorenz Kienle

Abstract Memristors are promising candidates for new memory technologies and are capable to mimic synapses in artificial neural networks. The switching in memristive devices occurs typically in few nanometer thin dielectric layers. The direct observation of the switching mechanism is crucial for better comprehension and improvements of memristors. Therefore, in situ experiments are conducted in a transmission electron microscope (TEM). However, sample preparation processes and electron beam irradiation can lead to a chemical and structural modification of the active layers. Moreover, devices may show significant device-to-device variability due to the details of processing parameters. Thus, it is essential to characterize the identical device electrically before microstructural analysis.

O. Gronenberg · L. Kienle (✉)

Synthesis and Real Structure, Department of Materials Science, Kiel University, Kaiserstr. 2, 24143 Kiel, Germany
e-mail: lk@tf.uni-kiel.de

O. Gronenberg

e-mail: og@tf.uni-kiel.de

G. Haberfehlner · G. Kothleitner

Institute for Electron Microscopy and Nanoanalytics, Technical University Graz, Steyrergasse 17, 8010 Graz, Austria
e-mail: georg.haberfehlner@felmi-zfe.at

G. Kothleitner

e-mail: gerald.kothleitner@felmi-zfe.at

F. Zahari · R. Marquardt

Nanoelectronics, Department of Electrical Engineering and Information Engineering, Kiel University, Kaiserstr. 2, 24143 Kiel, Germany
e-mail: fnz@tf.uni-kiel.de

R. Marquardt

e-mail: rvm@tf.uni-kiel.de

C. Kübel

Institute of Nanotechnology, Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany
e-mail: christian.kuebel@kit.edu

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_5

1 Introduction

Is Moore's Law at an end? Memristive devices can help to extend the functionality of today's state-of-the-art digital computers while simultaneously decrease energy consumption for data intensive and associative tasks in novel in-memory and bio-inspired computing architectures [1, 2]. The usual two-terminal device structure can further help to decrease the size of integrated circuits and memory arrays.

Simple two-terminal memristive devices can help to further decrease the feature size in integrated circuits, while simultaneously decreasing energy consumption. However, with decreasing feature size, every single structural and electronic defect becomes more important and analytical techniques become the limiting factor for knowledge-driven engineering of memristive devices. In this chapter, the limits of TEM are critically discussed. TEM, as one of the suitable analytical methods, provides sufficient resolution as well as a plethora of different techniques to obtain information from real space, reciprocal space and spectroscopy. Recent progress in TEM analysis of memristive devices is reviewed in publications of Strobel et al. [3], Yang et al. [4], Sun et al. [5], Xu et al. [6] and Yang et al. [7].

The final goal of most TEM studies on memristive devices is to identify the switching mechanism by directly observing structural or chemical changes or inhomogeneities inside the memristive material. However, the identification of fingerprints for memristive switching is not straightforward as discussed below.

The most studied switching mechanism is the growth and rupture of a conducting filament which is believed to be essential for the electrochemical metalization mechanism (ECM) and in valence change memories (VCMs) [8]. In memristive devices based on the ECM, a metallic filament forms due to the partial dissolution of an active electrode (typically Ag or Cu) into a dielectric layer in high electric fields. Such metallic filaments were often observed in situ in TEM experiments, as summarized in the review from Yang et al. [4]. However, radiolysis and developing electric fields in the sample due to the electron beam complicate such experiments and can lead to unambiguous interpretations. Especially for the thin TEM samples, surface-migration and surface oxidation should be prevented to reduce effects from radiolysis and electric fields.

One of the main defects discussed for resistive switching in the field of VCMs are oxygen vacancies, which are reported to be responsible for the formation of one or multiple conducting filaments inside the dielectric matrix [9]. These vacancies, however, are generally difficult to detect. In the case of TiO_2 memristors, such filaments could be identified by small crystallized regions with decreased oxygen content in the so called Magnéli phases [10]. These can be directly identified by diffraction and imaging techniques in the TEM. In other oxides, however, such direct fingerprints for the filaments are under debate and thus indications are used to identify the filaments, e.g., by spectroscopic means. Li et al. [11] and Jang et al. [12] observed such an indication for HfO_2 in the low loss electron energy loss spectrum (EELS): A plasmon peak change under in situ biasing was interpreted as the change of oxy-

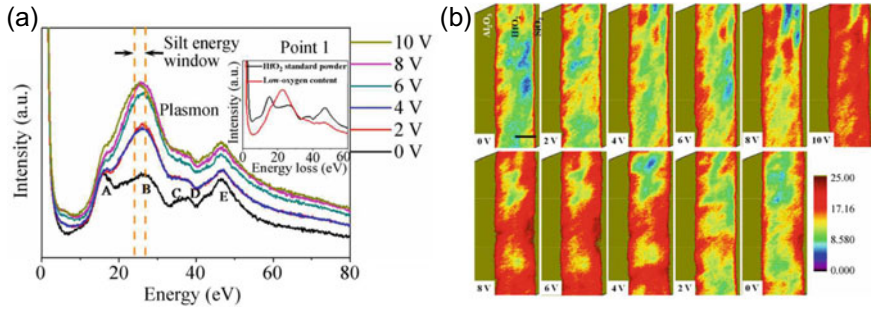


Fig. 1 a Low-loss EEL spectra with bias voltages from 0 to 10 V. The slit energy window for the energy filtered (EF)TEM images in b is indicated by orange dashed lines. The color code indicates higher oxygen vacancy concentrations. Adapted by permission from [Springer Nature Customer Service Centre GmbH]; [Springer] [Nano Research] [11]

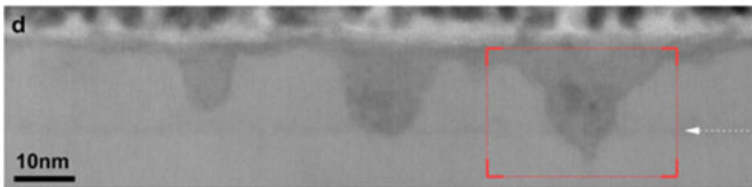


Fig. 2 Annular bright field STEM image of a Au/Fe:SrTiO₃/Nb:SrTiO₃ memory device after electroforming and delamination of the Au top electrode. Adapted with permission from [ACS Publications][Chemistry of Materials] [9]

gen vacancy concentration and distribution, which could then be mapped in energy filtered (EF)TEM images (see Fig. 1b).

An additional difficulty is presented by the fact, that conducting filaments can be formed anywhere in the memristive device, while TEM can only analyze a limited region of interest. Yang et al. circumvent this problem by electrically probing a HfO₂ thin film with conductive-AFM and fabricating FIB lamellas from probed regions [4]. Such a filament was subsequently observed by means of scanning TEM (STEM) energy dispersive X-ray spectroscopy (EDX) elemental mapping, as it exhibited a reduced oxygen signal. Du et al. delaminated the top electrode of a device after cycling and prepared FIB-lamellas from regions where they detected defects [9]. The observed Ti³⁺ ions indicated oxygen vacancies in SrTiO₃ which could be identified by X-ray absorption spectroscopy and X-ray photoemission electron microscopy. Consequently, the researchers were able to show that multiple filaments can be formed simultaneously, but with different velocities (see Fig. 2) [9]. Accordingly to this, an observed filament can be in any resistive state, since a different filament might dominate the resistive switching. Nevertheless, observing and characterizing the filaments present great opportunities to gain a fundamental understanding of switching mechanisms and towards knowledge-based improvements of memristive devices.

However, such experimental results impose a high margin of error since artefacts could be introduced by TEM sample preparation and by beam damage during measurements. Accordingly, great effort should be made to rule out or minimize artifacts. Therefore, the following chapter presents means to achieve suitably pure samples, as well as aspects of best practice. These include TEM sample preparation and the possible induction of artifacts, performing TEM measurements without beam damage and ultimately dedicated *in operando* methods.

2 TEM Sample Preparation

TEM measurements require a specimen thickness below 100 nm and for reliable (electron energy loss) spectroscopic information the thickness should be below 50 nm. This requirement makes the preparation equally important as the experiment and the analysis itself, because unwanted alterations of the material or the device are to be avoided. When analyzing memristive devices, electrical characterization prior to preparation and structural characterization via TEM is crucial and should be carried out on the same device that is to be measured. Memristive devices generally exhibit a high variability from device to device, which complicates the subsequent analysis. Even devices on the same wafer can differ strongly, depending on their position on the wafer. These differences could be attributed to local variances in plasma conditions during sputter-deposition, which result in different oxidation states of the metal oxide components on the wafer, which was proven by EELS fingerprinting of the O-K edge in this case of a niobium oxide thin film [13]. Challenges of preparations of specific devices are summarized below.

2.1 FIB Preparation of Specific Devices

The preparation of cross-sections of single devices on a lithographically structured wafer relies on Focused Ion Beam (FIB) milling with Ga ions combined with scanning electron microscopy to select a specific region of interest. The FIB method enables the preparation of a cross-sectional lamella, a few μm in size, with a homogeneous thickness for high quality measurements. Prior to the Ga ion milling, a protective Pt-layer is deposited with a gas injection source (GIS) that reduces both charging and Ga-implantation in the device. In case of insulating specimens, this Pt-layer might be insufficient as strong charging hinders the Pt-deposition itself. In case of sapphire substrates, electron and Ga ion bombardment can lead to an amorphization of the surface (see. Fig. 3). An alternative preparation method would be using the precision ion polishing system (PIPS), which is described further below. However, due to the method requiring steps of mechanical grinding and thinning steps, preparation proved

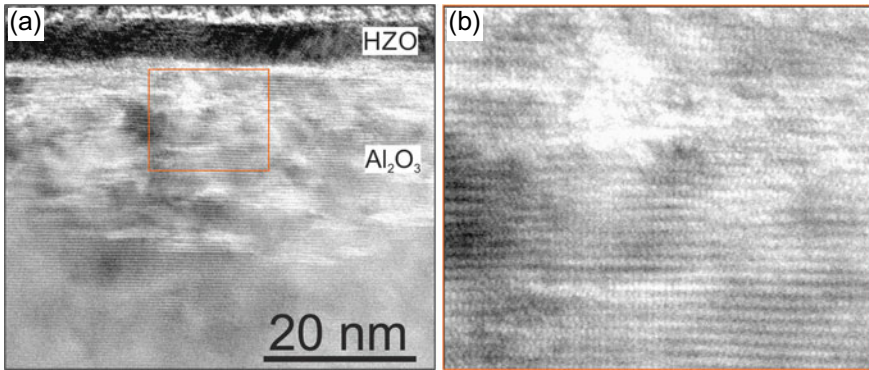


Fig. 3 **a** HRTEM micrograph of hafnium zirconium oxide (HZO) grown on sapphire substrate, prepared as a FIB cross-section. A high defect density is visible at the surface of the substrate. The orange rectangle indicates the magnified region in **b**

to be impossible, as sapphire is one of the hardest materials. In the case of memristive devices, the top electrode protects the underlying layers and reduces charging effects during the FIB preparation.

Independent from the charging effects in the FIB, (e.g., by a sputtered top electrode) dedicated TEM measurements require a final thinning step of the FIB lamella. This can be done in the FIB with Ga ions accelerated with lower voltage from the top of the cross-section. However, this procedure is prone to Ga implantation and surface amorphization.

An alternative approach is to perform the final thinning with Ar ions with low voltages to gently sputter material from the FIB lamella. However, for sapphire substrates, low-energy milling is required, as it has been observed that sapphire is sensitive to Ar ion milling at energies above 3 keV [14]. Sapphire becomes susceptible to electron beam damage in the TEM, which could be prevented when milling specimens with Ar ions with acceleration voltages of 2.5 keV or less [14]. This can be explained by the incorporated Ar point defects that act as nucleation centers for the reduction of Al_2O_3 in the electron beam [15]. Accordingly, oxides, especially with high defect concentrations, should be Ar-milled only if necessary and with low accelerating voltages (e.g., 500 eV) to prevent the creation of defects in the preparation.

Different milling rates along the lamella in the FIB might occur in cases, where the top most layer is inhomogeneous or structured with different materials (e.g., with top electrodes or dielectrics). In these cases it might be helpful to flip the lamella by 180° and to perform the FIB thinning from the backside, i.e., the side of the substrate [16]. Particularly for the most common substrate in many fields, a crystalline Si wafer, this procedure enhances the homogeneity of the FIB lamella. Still, a slightly misoriented Si substrate can result in strong anisotropic etching from different etching rates of different crystal orientations, as can be seen in Fig. 4a, while in Fig. 4b homogeneous

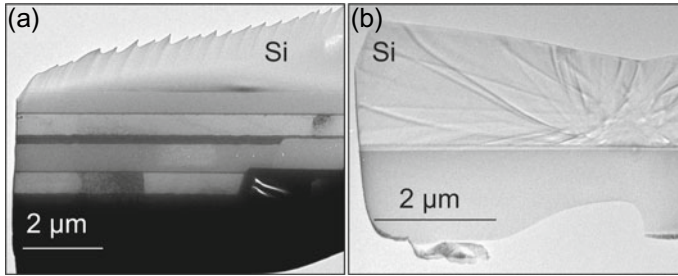


Fig. 4 Two FIB lamellas prepared with a backside thinning procedure, developed in-house, using a rotation tip. In **a**, the lamella was slightly tilted due to stress during transfer to the rotation holder for the 180° rotation. As a result, a saw tooth structure developed due to anisotropic etching in the crystalline Si substrate. In contrast, the lamella in **b** was perfectly aligned for backside thinning resulting in a smooth Si surface

milling was achieved. In Fig. 4a, the FIB lamella was lifted out and attached to a rotation tip for the 180° rotation. During this transfer, the lamella was slightly tilted by stress, leading to a sawtooth structure due to anisotropic etching of the crystalline Si substrate.

In cases where it is suspected that the preparation has produced artifacts, a reference sample must be prepared with a well-established and more sensitive preparation method. Such a method could be using the PIPS at low voltages, which mills a hole inside the sample with Ar ion impact at an oblique angle. With the PIPS, no specific devices can be selected as TEM specimens, though high-quality specimens can be produced. A comparison between FIB and PIPS preparation, acting as check of reliability, has been carried out by the analysis of low-temperature superconducting Josephson junctions as a reference system [17]. Figure 5a exemplifies the slightly different outcome of these procedures: the sample prepared with FIB shows lower contrast and resolution than the sample prepared via PIPS because of higher thickness and some FIB induced amorphization of the surface layers [18]. Importantly, though, the STEM-EDX scan (Fig. 5b) across the layers yielded qualitatively the same result for both cases, showcasing the partial oxidation of the Al layer on the side facing the surface. Since the results show such congruence, the preparation by FIB - which is necessary for the actual devices - is deemed suitable [19].

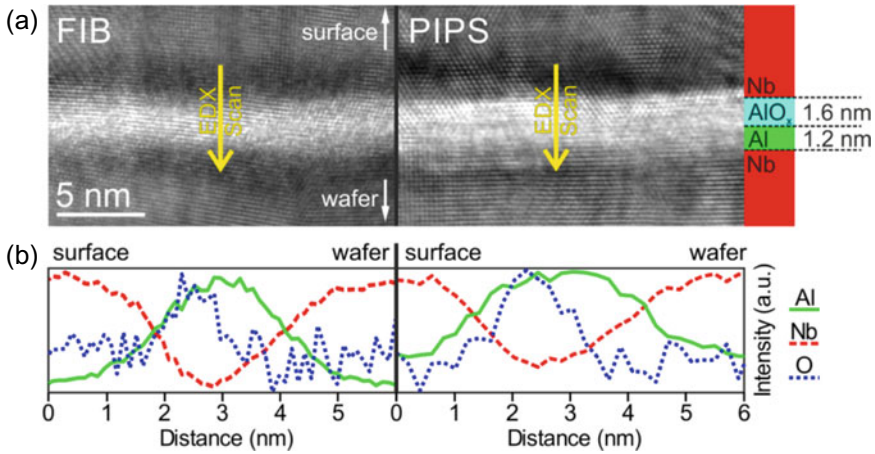


Fig. 5 Comparison of a Josephson junction prepared via FIB (left) and PIPS (right). In **a**, HRTEM shows the polycrystalline nature of the Nb top and bottom electrodes as well as the partially amorphous Al/AIO_x layer in between. The EDX scans in **b**, performed along the yellow arrows in **a**, show both that Al is partially oxidized on the side facing the surface. Adapted with permission of AIP Publishing [19]

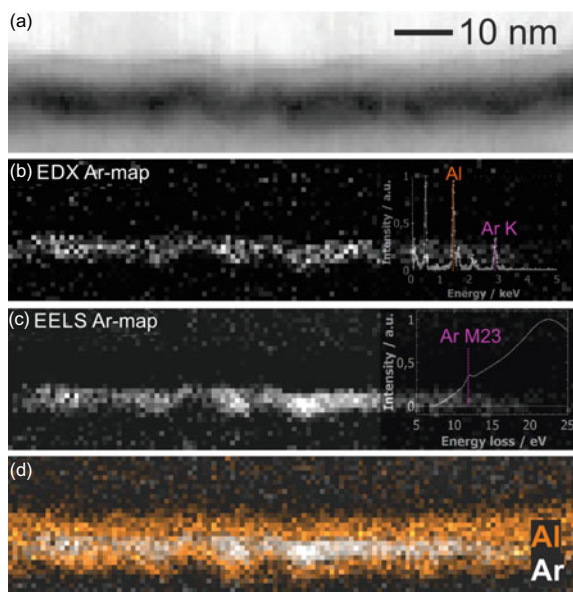
2.2 Conclusion for the Preparation

- Protect the topmost layer before preparation to prevent Ga implantation and related artefacts
- Use sensitive preparation methods with low acceleration voltages
- Check for preparation artefacts in the beginning, before starting sophisticated and time-consuming measurements
- If preparation artefacts are likely, compare the same or a well-comparable sample prepared with a more sensitive method (e.g., PIPS)
- Best practice: if possible fabricate a (prototype) device on a TEM-compatible grid to check functionality before, after, or even during measurements.

3 Spectroscopic Methods

Modern TEMs are equipped with EDX spectrometers as standard, since characteristic X-rays are readily formed when the electron beam passes through the sample. These X-rays can be used to determine and even quantify the elements in respect to their concentrations in the sample. Not only the primary beam generates X-rays, but also secondary electrons that interact inside the microscope column and generate spurious X-ray signals. These spurious X-rays, as well as further artifacts and pitfalls, will be discussed in the following chapter.

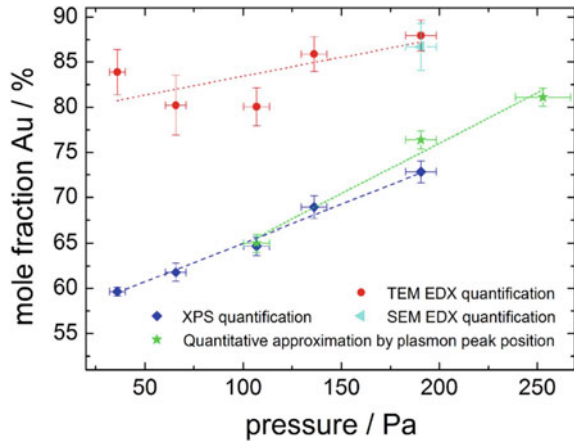
Fig. 6 **a** STEM high angle annular dark field (HAADF) image showing the thin film stack of a DBMD build up from bottom to top of Nb, Al_2O_3 , HfO_2 and Au. The spectrum image in **b** is created from the Ar EDX signal and in **c** from the Ar M_{23} EELS signal. The inset in **c** shows the EEL spectrum with the Al_2O_3 plasmon and the sharp M_{23} threshold peak of Ar at 12 eV. The overlaid image in **d** shows the Al EDX signal in orange and the Ar EELS map of **c**



Spurious X-rays are characteristics of the respective microscope and TEM grid in use. Reference spectra next to the sample give a clear indication of which signal is spurious or intrinsic. This procedure fails when the TEM grid contains elements which are also present in the actual specimen. Accordingly, the right TEM grid needs to be chosen beforehand. Further limiting this choice are overlapping EDX peaks between TEM grid and elements that are expected to be in the sample. For instance, peaks of Si and Mo or Ti and Ba overlap and cannot be distinguished with the energy resolution of common EDX detectors.

Artificial peaks can also occur and are produced as sum-peaks when two characteristic X-rays are detected within the processing time of a signal. These sum-peaks can usually be ignored but become critical when an element at the same energy is falsely detected. One prominent example is the sum-peak of Al, which lies at the same position as Ar [20] and could lead to the wrong exclusion of Ar in Al. In a study of a double barrier memristive device (DBMD) including a roughly 7 nm thin Al_2O_3 layer (compare Sect. 4 dealing with this device), an Ar EDX signal was measured exactly in the middle of this layer as shown in Fig. 6b. In this case, an Al sum-peak could be ruled out by the complementary technique EELS. The Ar M_{23} edge at 12 eV (see inset in Fig. 6c) and the L_{23} edge at 320 eV prove the presence of Ar in the Al_2O_3 thin film, which is also indicated by the EDX signal. The EEL spectral image exhibits a significantly better signal-to-noise ratio at the same acquisition time in comparison to the EDX map. The compound map in Fig. 6d of Al (EDX signal) and Ar (EELS signal) shows a clear anti-correlation, suggesting Ar segregation in less dense regions inside the Al_2O_3 . Dark contrast in the STEM-HAADF image in Fig. 6c supports this hypothesis.

Fig. 7 Mole fraction of Au in AgAu bimetallic nanoparticles deposited with a Haberland type GAS at different Ar pressures. The mole fractions were determined by XPS (blue squares), TEM-EDX (red circles), SEM-EDX (cyan triangle) and approximated from UV-vis spectra (green stars). Adapted with permission from IOP Publishing [21]



The usage of complementary techniques can not only support the measurements but also point the way to further conclusions. As an example, a combination of EDX, X-ray photoelectron spectroscopy (XPS) as well as UV-vis spectroscopy was used to investigate a system of bimetallic nanoparticles. These nanoparticles were fabricated via a custom-made multi-component (e.g., AgAu) target in a Haberland-type gas aggregation cluster source (GAS) and subsequently studied. Since XPS and the optical behavior of the sample rely strongly on the surface properties, the comparison with EDX can reveal the difference of bulk and surface composition, even though the particles are 10 nm in diameter. XPS quantifications showed a systematically higher Ag content in comparison to EDX, which led to the assumption that an Ag enrichment had occurred on the surface of the AgAu nanoparticles, even though the Ag-Au system is perfectly miscible. This hypothesis is supported by quantitative approximations of plasmon peak positions in UV-vis spectra shifting with the composition (see Fig. 7), which can be controlled with the pressure inside the GAS [21].

4 Example Characterization of Double Barrier Memristive Devices

In this section, an extensive investigation of double barrier memristive devices (DBMDs) is presented, from which knowledge-based improvements could be derived. These DBMDs were designed with an active layer (i.e., NbO_x or HfO_x) that is deposited on the bottom part of a Josephson junction ($\text{Nb}/\text{Al}/\text{Al}_2\text{O}_3$) [17] and a Schottky-type contact to an Au top electrode. This device is believed to show interfacial resistive switching which is reported to result from oxygen ion movement affecting the interfacial properties [22]. Therefore, no fingerprints for resistive

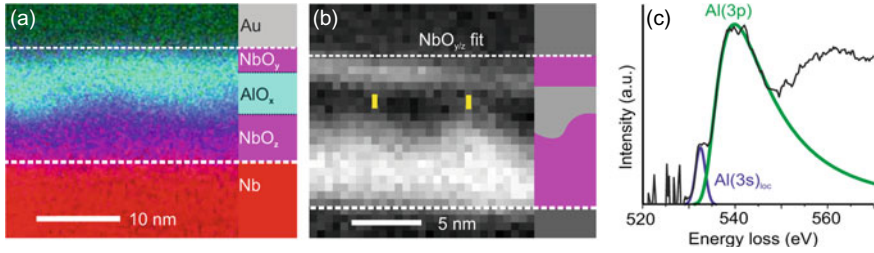


Fig. 8 STEM and spectroscopic investigations of the first generation DBMD. The compound map from EDX signals in **a** illustrates the layout of the device. The goodness-of-fit map in **b** of the O-K edge of the NbO_{y/z} shows the surface-oxidation of the Nb bottom electrode and its roughness. The O-K edge of AlO_x in **c** contains a pre-peak feature which was fitted with a Gaussian function, while the main edge was fitted with a Pearson I(V). Adapted with permission from AIP Publishing [19]

switching, such as filaments that could be directly observed with TEM, are known to characterize these devices.

These devices were analyzed by spatially-resolved STEM-EELS and -EDX to simultaneously map light and heavy elements. ELNES (energy loss near edge structure) fingerprinting was utilized to determine the oxidation states of the Nb atoms. In this approach, the shape of the edges is compared to known reference spectra. In the first-generation DBMD, the ELNES of the alumina resembles that of amorphous Al₂O₃. Unlike the targeted (Nb/Al/Al₂O₃)-Josephson junction, the entire Al layer and even the surface of the Nb bottom electrode are oxidized. This niobia (NbO₂) could be assigned to Nb(II) oxide, thus providing sufficient conductivity for device functionality. The Nb bottom electrode and Al layer were sputtered as metallic thin films, which subsequently oxidize during the reactive sputtering of the memristive layer on top (i.e., NbO_x). However, the ELNES of this NbO_x could not be clearly assigned to a certain Nb oxide. A mixture of Nb(IV) and Nb(V) appears probable. Further, the electronic structure of the AlO_x layer measured with EELS contains an additional pre-peak feature at the O-K edge which is located roughly 6 eV before the main edge (see Fig. 8c). A similar feature, which evolves due to beam damage in oxides, is reported in literature - its origin will be discussed in more detail later on [23, 24]. In this case, however, beam damage appears unlikely, as no evolution of this pre-peak over time and no morphological changes even after extended electron irradiation could be observed. Another explanation of this pre-peak feature could be oxygen vacancies in AlO_x as suggested by Nigo et al. [25]. This feature might be caused by localized Al(3s) states in the band gap of AlO_x. By comparing the area ratio of the pre-peak and of the main edge, the concentration of these vacancies could be estimated to be in the range of 1.7 to 3.2% with respect to the entire oxygen content [19]. A third possible explanation for this feature is a superposition of the O-K edges of alumina and the neighboring niobia, which was observed by Bruley et al. [26]. The onset of the O-K edge from niobia lies few electron volts lower than that of alumina. This explanation seems quite convincing, as the Nb-M₃ and Nb-M₂ edges are present at the same position and the shape and the position of the pre-peak

resembles the onset of the two neighboring niobia O-K edges. Both interdiffusion as well as an imperfectly aligned FIB lamella with respect to the electron beam could explain this scenario.

In addition to spectroscopy, simultaneous morphological analysis can yield further valuable information. The roughness of the bottom electrode leads to a minimal distance between NbO_x and NbO_2 of around 1 nm across the AlO_x (see Fig. 8, the yellow lines), which might enable tunneling through this oxide. In contrast to the targeted $\text{Nb/Al/Al}_2\text{O}_3$, no metallic Al smooths the interface to the dielectric AlO_x [27] which should lead to an inhomogeneous electric field between the electrodes.

According to [19], these devices demonstrated a high robustness regarding variations from the targeted design and layer sequence. These DBMDs were still functional after NbO_x was subsequently replaced by HfO_x .

Nevertheless, only devices within a central area of a 100 mm wafer were functional (inner red area in Fig. 9a). To understand the reason for this, plasma and TEM characterizations were carried out and revealed distinct differences. In particular, the reactive plasma process to deposit NbO_x layers with metal targets using an Ar/O_2 gas atmosphere leads to variable oxidation states of Nb atoms across a whole wafer [13]. The subsequently patterned devices are showing a distinct memristive functionality in the center of the wafer, while the devices at the edge of the wafer exhibit no memristive $I(V)$ characteristics. These findings indicate a strong correlation between local plasma parameters and device performance and may pave the way to plasma-engineered memristive components in the future. In case of NbO_x , the ELNES of the O-K edge varies between different wafer positions. In the functioning device from the center of the wafer, the ELNES could be assigned to either Nb(IV) or Nb(V) oxide, with slight deviations to both. A mixture of both oxides is probable. However, from a non-functional device from the edge of the wafer (see $I(V)$ -characteristics in Fig. 9d), the O-K edge could not be assigned to a certain oxidation state. The lower signal-to-noise ratio in comparison to the other oxides with the same acquisition parameters, however, suggests a lower Leerschnitt O-K edge signal and with this a lower oxidation state of the Nb atoms. In addition, niobium oxide with a low valence state is known to be a conductor and cannot be used to build a Schottky type barrier as observed from the $I(V)$ characteristics [13]. In summary, it could be proven that the EEL spectroscopy of the niobium oxide O-K edge could explain the different at the specific locations on the wafer.

Independent from the top-most oxide (i.e., NbO_x or HfO_x), the AlO_x layer underneath could be assigned to amorphous Al_2O_3 by the O-K edge. However, exactly in the middle of the AlO_x thin film a pre-peak feature appears as shown in Fig. 10a, which clearly differed from the pre-peak from previous studies as shown in Fig. 8c. Pre-peaks of the O-K edge occupied researchers for a long time, but still, the origin and meaning are under debate. The pre-peak can be part of the main O-K edge in some minerals, instead of being an additional defect feature, as shown by Winkler et al. using DFT calculations, e.g., for Hambergite ($\text{Be}_2\text{BO}_3(\text{OH})$). They have further demonstrated that the “pre”-peak originates not only from the OH-, but also from all BO_3 -groups [28]. An older hypothesis from [29], which claimed that the hydrogen content of minerals can be determined from the pre-peak feature, can therefore be

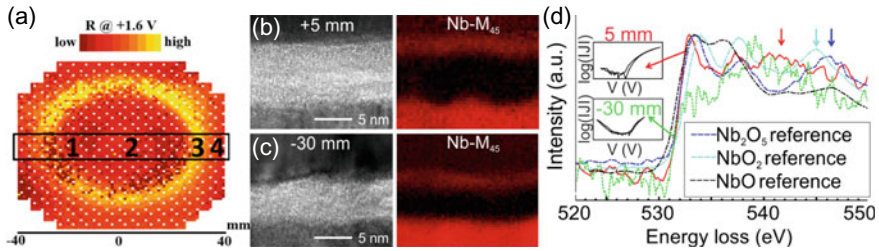


Fig. 9 **a** Resistance map of a whole wafer with memristive devices deposited on top measured at 1.6 V. **b** and **c** HRTEM micrographs from positions comparable with 2 and 3, respectively, indicated in **a**. The STEM-EELS maps of the Nb-M45 edge showing the Nb distribution. **d** illustrates the O-K EEL spectra in the niobiumoxide of the different positions overlaid with reference spectra. The $I(V)$ -characteristics of these devices at the respective positions are shown in the insets. Adapted with permission from American Vacuum Society [13]

ruled out. The same conclusion was stated by Garvie, who assigned the pre-peak of the O-K edge in OH- and H₂O-bearing minerals to beam damage, which could, however, be prevented by low dose-rates [30]. The pre-peak appearing as a result of beam damage originates from the transition of core electrons to unoccupied π^* orbitals of molecular oxygen [23, 31–33], which develops, e.g., due to the decomposition of hydroxides [24] or the reduction of the oxide itself [34, 35]. As indicator for beam damage serves its development over time, as the damage process occurs gradually and the resulting O₂ gas is emitted into the vacuum. A more detailed look on beam-induced damage is provided in the following Chap. 5. A recent study from Ayoola et al. proposed that peroxy O-O bonds developed at the surface of γ -Al₂O₃ in the areas damaged by electron irradiation. However, these O-O bonds could be prevented by employing cryogenic-EELS although beam damage in form of hole drilling still occurred [36]. Peroxy groups were also found in amorphous Al₂O₃ by stochastic quenching simulations which best explains a similar pre-peak of the O-K edge measured by Near Edge X-ray Absorption Fine Structure [37]. Further, single superoxide ions (O_2^-) were detected in sapphire with electron paramagnetic resonance after neutron-irradiation, proving their possible existence in Al₂O₃ [38]. It seems likely though, that peroxy groups in Al₂O₃ are unstable in the electron beam and directly decompose into molecular oxygen. Fritz et al. reported on an electron beam stable pre-peak at the grain boundaries of polycrystalline Al₂O₃, and it was interpreted as chemisorbed molecular oxygen [39].

In our own measurements, the pre-peak in the amorphous Al₂O₃ in the DBMD (Fig. 10a) does not emerge or increase over time, but disappears with electron beam irradiation (see Fig. 10a), which could be an indicator for adsorbed oxygen molecules released to vacuum. This hypothesis is supported by the difference spectrum before and after electron irradiation in Fig. 10a, which resembles the O-K edge of molecular oxygen gas. The pre-peak does not only disappear in the vicinity of the electron beam, but also in distances further than 50 nm, which can be seen in the pre-peak spectral image in Fig. 10e with an exposure time of 100 ms. Here, the oxygen

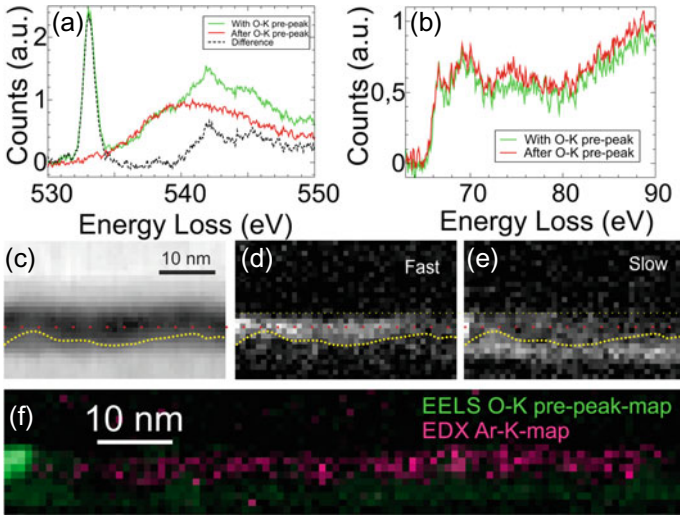


Fig. 10 **a** Time-series of the background-subtracted O-K edge in Al_2O_3 with pre-peak feature (green) in the beginning and after the pre-peak vanished (red). The difference of these spectra (black dashed line), after alignment of the onset of the main O-K edge, resembles the O-K edge of oxygen gas. The Al-L₂₃ edges in **b** recorded simultaneously with **a**, show no significant difference, indicating no valance change of Al_2O_3 and no metallic fractions. The STEM-HAADF image in **c** provides an overview for the sequential spectral images of the pre-peak of the O-K edge in **d** and **e**. **d** was recorded first with an exposure time of 9.8 ms and **e** was recorded second with an exposure time of 100 ms. In fast measurements the pre-peak appears spread throughout the entire Al_2O_3 , while in slow measurement, the pre-peak concentrates at the left corner where the scan started. The yellow dashed line indicates the interface to the Nb bottom electrode and the orange dashed line indicates the center of the Al_2O_3 where the pre-peak is located. The oxidized Nb bottom electrode shows intensity in the oxygen pre-peak spectral image in **e**, because the onset of the main O-K edge of niobium oxide is located at the same energy. In the compound map **f**, the O-K pre-peak (similar to **e**, but different position) and the Ar EDX signal are anticorrelated. No Ar is detected where the O-K pre-peak concentrates but elsewhere Ar does not vanish

pre-peak concentrates at the left corner, from where the scan was initiated, while a fast measurement beforehand (Fig. 10d) proved that the pre-peak spreads over the entire Al_2O_3 . This long-range effect could be explained by two scenarios. In the first, secondary electrons spreading in all directions from the focused electron beam, which causes the pre-peak to vanish before it could be measured in EELS. These secondary electrons have lower energy and a shorter mean free path and may produce stronger beam damage by radiolysis (for more details see Sect. 5 about beam damage). In the second scenario, the focused electron beam positively charges [35] the Al_2O_3 , which could attract negatively charged chemisorbed [40, 41] oxygen-gas molecules.

An argument for the second scenario is that the intensity of the concentrated pre-peak appears higher, when comparing the second, slower scan to the initial fast scan. This could be explained by the attraction of negatively charged chemisorbed

oxygen, not only from the right side of the scanning window in the Al_2O_3 but also from the left. In a time-series (see Fig. 10a), no further increase in the pre-peak was observed but only a decrease, which argues against a development over time and therefore against a pre-peak produced by beam-damage. In the pre-peak-map in Fig. 10e, intensity can be observed below the yellow dashed line, roughly indicating the interface to the Nb bottom electrode. However, this intensity is produced from the main O-K edge of the oxidized Nb, which is located at the same energy and therefore mapped as well. In the fast measurement in Fig. 10d the noise level is too high to detect the Nb O-K edge.

The presence of Ar gas inside the Al_2O_3 further proves the possibility for gas inclusions and thus for incorporated oxygen gas. Ar gas was detected at the same position as the molecular oxygen in the middle of Al_2O_3 by EDX and EELS (compare Fig. 6). The Ar signal disappears, however, where the O-K pre-peak concentrates, even though Ar gas is present where the O-K pre-peak has vanished by the long-range effect (see Fig. 10f). This may indicate a physical displacement of the Ar atoms from the oxygen molecules, as the Ar signal is stable in the electron beam. However, molecular oxygen could have also been formed due to the preparation, where the Ar atoms could destabilize the Al_2O_3 as proposed by Nufer [15]. In this case, the molecular oxygen should be the product of the reduction of Al_2O_3 to metallic Al, but no temporal changes at the Al-L₂₃ edges, indicative for metallic Al [34], could be observed in the EELS time series of the Al-L₂₃ edges in Fig. 10b. Since the TEM sample preparation was done without Ar ions, the origin of the incorporated Ar can be attributed to the sputter gas of the deposition process. We assume, Ar atoms implanted during reactive HfO_2 sputtering, where the Al and the top-most Nb back electrode were oxidized, as no traces of Ar could be found in the metallic Al thin film.

The extensive analysis of a memristive device examined various aspects. These include the oxidation and roughness of the Nb bottom electrode as well as different oxidation states of the NbO_y in dependence on the position of the device on the wafer which lead to malfunctions. However, the switching mechanism could not be resolved with TEM due to its interfacial character with very small dimensions and amorphous structure. Last but not least, gas inclusions inside the 7 nm Al_2O_3 consisting of Ar and most probably molecular O_2 were indicated by EDX and EELS, whether these are important for the resistive switching will be investigated in future.

5 Invasive TEM Measurements

The TEM utilizes the interaction of accelerated electrons with solid matter to facilitate imaging with sub-Å spatial resolution. Due to the highly energetic impact, sensitive materials can be severely damaged by prolonged irradiation. Especially at high dose-rates, the induced beam damage can limit the resolution and available time for TEM analysis. The following chapter introduces a broad overview of different mechanisms

of beam damage, which is followed by practical examples from memristive devices, and reasons, why they require special care.

5.1 *Beam Damage Overview*

Electron beam damage typically describes irreversible changes in the investigated sample during TEM analysis. A variety of phenomena can occur, like surface sputtering, phase transformations, (re)crystallization or amorphization, valence changes and gas bubble formation. In general, beam damage can occur via two different mechanisms, which may develop simultaneously and often cannot be clearly distinguished. The first mechanism is the so-called knock-on damage, where collisions of electrons cause an elastic displacement of atoms. The second mechanism occurs due to the inelastic interaction of beam electrons with core electrons, leading to the formation of an excited state. This effect is called radiolysis. Electron beam damage in the TEM was reviewed by Egerton [42] and Jiang [43]. This chapter discusses important aspects of beam damage in relation to the materials used in memristive devices and how it may be prevented.

Knock-on damage occurs preferentially at the surface due to the lower binding energy of surface atoms, and preferential sputtering is expected in compounds of the elements with the lower binding energy [42]. As a 'rule of thumb', knock-on damage increases with increasing beam energy while the effects of radiolysis decrease. Radiolytic damage occurs predominantly in dielectrics with low conductivity, as the excited state has a longer lifetime to produce point defects compared to conductive materials. In the vicinity of the beam the excited states accumulate and create a charged area - typically positively charged in TEM due to secondary and Auger electron emission. This positive charge can lead to cation migration away from the irradiated area and, e.g., the creation of molecular oxygen [23, 34].

Different parameters influence the beam damage. As mentioned before, the acceleration voltage determines which of the two damage mechanisms dominates. With increasing acceleration voltage an increased maximum energy can be transferred during interaction with an atomic nucleus leading to displacement and thus knock-on damage. On the other hand, inelastic scattering is reduced which reduces charging in dielectrics and the number of atoms in an excited state. Dielectrics typically charge up positively in the TEM [44], as most fast electrons are transmitted through the sample and secondary (SE) and Auger electrons are created. In thicker samples or especially in the SEM this positive charge can be compensated by absorbed electrons from the primary beam leading to negative charging [45]. The cross-section for SE emission has a maximum at several keV and decreases with the acceleration voltage, as the mean free path of the electrons increases [46]. Accordingly, charging and thereby radiolysis can be reduced with higher acceleration voltages and particularly thin specimens provide a further means of mitigation. Beam damage is further affected by the dose (accumulated electrons per unit area) as well as the dose-rate (accumulated electrons per unit area and unit time or also called current density of

the electron beam). The excited state requires sufficient energy for atomic displacements, while the life-time and localization need to be high enough so that an atom can move after the energy is converted into momentum [44]. Accordingly, radiolysis mainly affects insulators, as lifetime (and localization) are too short in conductors, where the excited state is rapidly relaxing. However, in insulators excited atoms may change their valence state or start to move. For reversible processes (in the sense of damage and recovery), the dose-rate determines the damage. Significant damage only occurs if the activation is faster than the recovery. In this case, it is sufficient to reduce the dose-rate to prevent beam damage, although the accumulated dose after prolonged measurements can be high [32]. However, it is not necessarily the case that the recovered and the original state are identical. The simplest way to reduce the dose-rate is to analyze the sample at low magnification. In case, high resolution is relevant, the beam current (spot size) needs to be reduced until no damage occurs. The following chapter describes examples of beam damage occurring in different material classes used for memristive devices.

5.2 *Thin Film Metal Layers in TEM*

Metals mainly suffer from knock-on damage, which can be easily reduced by employing low acceleration voltages. However, most metals are commonly covered with a thin oxide layer, and in the case of silver even with a sulfide layer. These compounds can be decomposed by the electron beam, leaving mobile silver nanoparticles on the surface. TEM experiments were performed at the tip of an intact atomic force microscope (AFM) apex. The apex is coated with a gold thin film as electrode, a silver thin film as active metal and a SiN_x -matrix as insulating layer. This so-called Memtip-approach enables the investigation of the very same conducting filament with different counter-electrodes [47]. The TEM experiments revealed the invasive character of the electron beam making ex situ measurements before and after conductive AFM (c-AFM) not feasible. After 15 min of electron beam irradiation with 200 kV acceleration voltage silver protrusions and morphological as well as chemical changes in the SiN_x -matrix had developed (see Fig. 11a, b). The matrix became denser and thinner, and in EDX, no nitrogen could be detected. Instead, oxygen was found, which suggests that the pristine matrix was porous and did not entirely protect the silver from the atmosphere. It is likely that atmospheric sulfur formed some Ag_2S under the matrix, which is known to build Ag protrusions under electron beam irradiation due to mobile Ag^+ ions inside Ag_2S and injected electrons [48–50]. Accordingly, samples containing silver should be transferred and stored with particular care prior to analysis, e.g., by using a vacuum transfer holder. A different approach is to seal the surface in situ with a non-porous matrix to prevent oxidation and surface migration, as reported in [51]. Otherwise, such morphological changes can lead to the misinterpretation of ex situ or in situ biasing TEM experiments since these morphological changes can already be formed by contact with the atmosphere

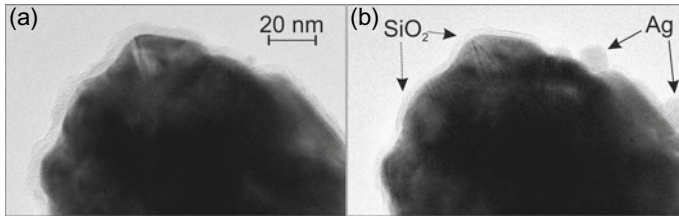


Fig. 11 AFM apex before **a** and after **b** TEM and EDX measurements. In **b**, arrows on the right indicate evolving Ag protrusions and arrows on the left indicate morphological and chemical changes of the SiN_x matrix [47]

or by the electron beam. Accordingly, ex situ or in situ TEM experiments should be reproduced without electrical stimuli.

5.3 Oxides in the TEM and SEM

The key component in a multitude of vertical memristors is a dielectric thin film [52]. Because of this, special care should be taken when analyzing these dielectrics in the TEM, and, of course, also during imaging and preparation in the SEM. The following section describes various phenomena that can occur with oxides in the electron beam, especially when they are produced with high defect concentrations.

5.3.1 Crystallization of Amorphous Oxides

The active oxide in memristors being amorphous can be beneficial, as the amorphous state is isotropic in comparison to the granular microstructure of polycrystalline thin films. Further, the processing costs can be reduced due to the lower deposition temperatures of amorphous films. As a model system to be analyzed, amorphous hafnia was chosen, which represents one of the most studied high-k oxides in many fields, including resistive devices [53, 54] as well as ferroelectric applications [55].

Mittmann et al. observed that radio frequency (RF) sputtering of hafnia leads to sub-stoichiometric, amorphous hafnia thin films, which can crystallize in the ferroelectric orthorhombic ($Pca2_1$) phase after rapid thermal annealing (RTA). However, when sufficient oxygen is added during the deposition, the monoclinic equilibrium phase is obtained [55]. We observed the crystallization of the pristine amorphous hafnia thin films during the TEM measurements, which appeared more pronounced at higher magnification (i.e., with higher dose and dose rate). Figure 12a shows such a crystallized region after 15 min of imaging in high-resolution mode. The crystallization starts from the very beginning of electron beam irradiation, though. The selected area diffraction (SAED) inset indicates a polycrystalline phase-mixture of hafnia. This thin film was RF sputtered with Ar from a hafnia target after a condi-

tioning time of 10 min to reach a steady state on the target surface. However, when the target was pre-treated with a mixed oxygen-argon-plasma to increase the oxygen content, the thin film exhibited few nanocrystallites in the pristine state which could serve as nucleation center for crystallization. Nevertheless, no crystallization caused by the electron beam was observed (see Fig. 12b). We hypothesize that the chemical composition has a significant effect on the beam-induced crystallization of the investigated hafnia. The same is true for the crystalline phase after RTA. Admixture of sufficient oxygen leads to the monoclinic phase while a sub-stoichiometric hafnia can crystallize in the orthorhombic phase [55].

Investigations with different substrates were performed to evaluate their influence on the crystallization, e.g., the influence of charging and oxygen-absorption. These substrates include amorphous carbon, amorphous Si_3N_4 and polycrystalline TiN, which is also serving as back-electrode in metal-insulator-metal (MIM) structures. On the amorphous carbon layer of a standard TEM grid a lot of nanocrystallites were present as-deposited, while the electron beam still induced strong crystallization of the thin film (see Fig. 12c). The same is true for hafnia on polycrystalline TiN thin film (see Fig. 12d). These findings underline that the nanocrystallites are no prerequisite for electron beam crystallization. Of higher importance is the stoichiometry of the thin film.

Further, a simple heating effect as cause for the crystallization in the electron beam can be ruled out. After prolonged heating instead of electron irradiation a pure,

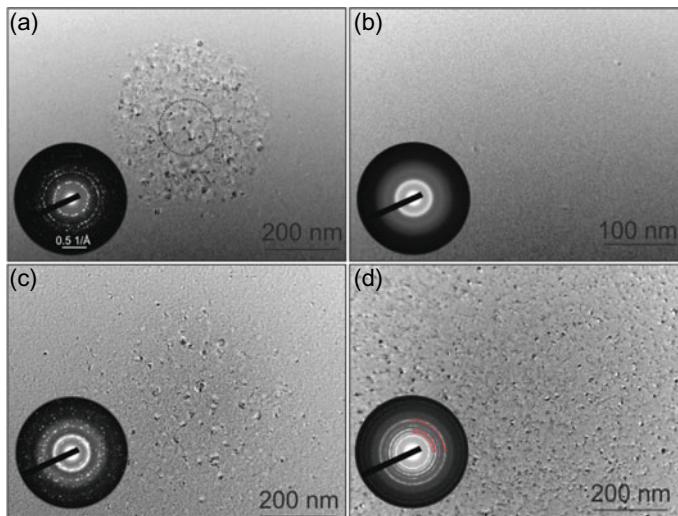


Fig. 12 TEM images of an amorphous, RF-sputtered HfO_x thin film on various substrates after up to 15 min of electron beam irradiation. In **a** and **b**, the thin film was sputtered on a Si_3N_4 membrane, in **c** on amorphous carbon and in **d** on a $\text{Si}_3\text{N}_4/\text{TiN}$ thin film stack. In **b**, the thin film is showing no crystallization in the electron beam, but a few nanocrystals are present in the pristine state. In this case, the ceramic target was pre-sputtered with oxygen admixture. The insets show SAED patterns with diffraction rings of TiN in **d** being indicated with red arcs

thermodynamically stable monoclinic HfO_2 is expected [56]. In our examination, the thin films with higher oxygen content should crystallize as well due to a heating effect, which is not the case. In addition, several studies indicate that the temperature increases by no more than 10°C during normal imaging conditions [42, 44, 57–59].

As a cause for the crystallization, radiolysis seems to be likely, as different studies reported on the occurrence of crystallization when amorphous oxides were reduced in the electron beam [60, 61].

However, this beam crystallization is not affected by whether the substrate is conductive (TiN and amorphous carbon) or insulating (Si_3N_4). Accordingly, a subsequent charging effect does not play a major role for the crystallization.

Taking these investigations into account, we propose that only the initial oxygen content, or, in terms of oxygen vacancies, the defect concentration, determines the crystallization of the amorphous HfO_x in the electron beam and also the resulting phase after annealing. However, an unambiguous statement is hardly possible. Further triggers for crystallization could be, e.g., the amorphous structure, free volume and presence of nucleation centers.

A second example for defect concentration-dependent beam damage was already discussed in Sect. 2. Here, Ar-milling with energies above 3 keV can alter crystalline Al_2O_3 in a way that it becomes susceptible to beam damage [14]. These findings highlight that the defect-concentration has a major impact on the stability of the sample in the TEM. This could be explained by the increasing free volume due to the defects leading to higher diffusion rates due to lower activation energies. On the other hand, beam damage produces defects. One example is a ferroelectric BaTiO_3 thin film that can be converted into a memristive thin film with configurable resistive states by electron beam irradiation [62].

5.3.2 Hole-Drilling and Oxide Reduction

In STEM mode, the focused electron beam has a high dose-rate (current-density), which can be more destructive in comparison to the TEM mode. In Fig. 13a, the electron beam produced a line of holes after an EDX line scan across a HfO_2 thin film with a dwell time of 6 s at each position. Close to the dark area of the holes in the STEM image, high intensity in the vicinity gives an indication for the formation mechanism of these holes. Cazaux proposed a possible mechanism, starting with the accumulation of positive charges, which occur due to secondary and Auger electron emission in the vicinity of the electron beam [35]. These positive charges repel cations, creating an excess of anions which can react to molecular oxygen [23, 34]. Following this mechanism, the bright Z-contrast next to the holes in Fig. 13a may hint towards a thicker crater rim with Hf-excess.

Such a reduction mechanism is dose-rate-dependent so that below a certain, material-dependent threshold, no visible beam damage occurs [32] and recovery processes can compensate the damage [33].

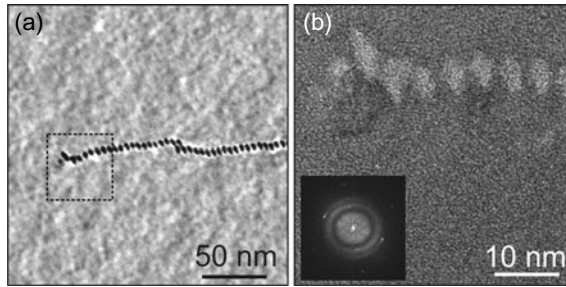


Fig. 13 **a** STEM image showing hole drilling in an amorphous HfO_2 thin film after a line scan. **b** HRTEM micrograph and FFT-inset reveal crystalline regions around the holes

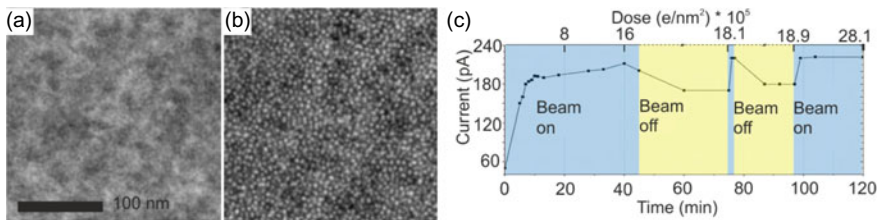


Fig. 14 **a** and **b** EFTEM (at 17 eV energy loss) images of a PECVD-deposited SiO_2 thin film after doses of 1.3 and 22 Me/nm^2 , respectively. In **b**, silicon-rich nanoparticles develop at the bright dots. **c** Electron beam-induced conductivity changes of the same SiO_2 film. Current was measured at 3 V. Adapted with permission from Cambridge University Press [63]

The following section describes a tool which can be used to measure this threshold in situ in the TEM or SEM. This threshold becomes particularly important when defect concentrations or defect variations are to be measured (e.g., using EELS). The (partial) reduction of oxides is equivalent to the generation of oxygen vacancies, which are commonly discussed point defects involved in the memristive switching mechanism. In addition, the thresholds for stoichiometric oxides reported in literature [42] will vary significantly from defective oxides. For this purpose, in situ electrical measurements on several thin film oxides were carried out, revealing reversible changes in conductivity attributed to a reduction in the electron beam and subsequent re-oxidation of the materials (see Fig. 14c). In case of SiO_2 thin films, the electron dose-induced formation of Si-rich particles leads to an increase in conductivity [63]. This effect can be reversed upon exposure to the atmosphere and even in the vacuum of the TEM when the beam is blanked. In the blue regions of Fig. 14c, the conductivity is increased during electron beam irradiation, while in the yellow areas, the conductivity is decreased upon blanking the beam. The chemical changes are indicated by energy-filtered (EF)TEM dose series showing Si nanoparticle formation by filtering the Si plasmon peak at 17 eV. Applying a dose of 22 Me/nm^2 white dots have evolved gradually in Fig. 14b indicating the formation of particles. The conductivity change caused by irradiation can be utilized to determine the critical dose for beam

damage, thus providing data for the inhibition of such effects. This can be used prior to in situ or *in operando* TEM and SEM experiments with ceramics to prevent beam effects on the I(V) characteristics and morphology. Due to this reversible behavior of the electron beam on the I(V) characteristics we labeled this experiment as the TEMristor. In these measurements, the beam damage was dominated by radiolysis, as indicated by the dependence on the acceleration voltage. The critical dose decreases with the acceleration voltage. Especially in the SEM with voltages of 30 kV and lower, e.g., TiO₂ exhibits a lower critical dose compared to 300 kV in the TEM.

This observation reconnects to the TEM preparation in Sect. 2 and highlights that dielectrics may have already been damaged by the electron beam during FIB-preparation. An effect that should be reduced by applying low beam-dose as well as a conductive coating on the sample.

5.4 Conclusions of Electron Beam Effects

In this chapter different mechanisms for beam damage in the TEM were described. Metallic samples are more prone to knock-on damage while dielectrics are more prone to radiolysis. In the context of memristive devices with high defect concentrations, either specially-designed or produced on purpose through a formation step, beam damage is a major concern: The included defects facilitate the beam damage, especially in the regions of interest where the memristive switching occurs. Low-dose rates should therefore be chosen to investigate memristive devices, either to prevent or at least reduce the beam damage. In situ current measurements can be used to determine the threshold below which no significant damage occurs. Further aids and best practices will be described in the next chapter.

6 Solutions

Materials can be sensitive to damage induced by either TEM sample preparation via FIB and Ar ion milling, or even the TEM measurement itself – this holds especially true for thin films designed for memristive devices, i.e., in non-equilibrium states and with high defect concentrations. In this section, different approaches for reliable TEM analysis of memristive devices are proposed.

6.1 Ex situ Measurements on Horizontal Devices

The best way to verify the reliability of TEM measurements of memristive devices, and to exclude severe alterations by the preparation or the electron beam is to electrically characterize the TEM specimen before and after in an ex situ manner. This

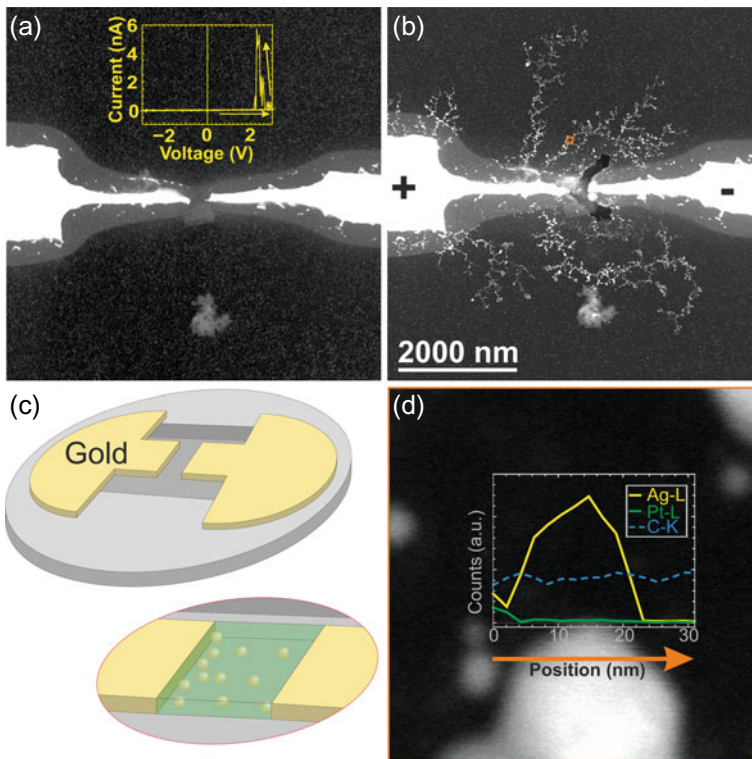


Fig. 15 Overview STEM image before **a** and after **b** ex situ electrical probing. Bimetallic AgPt nanoparticles were deposited by a Haberland cluster source on top of the electrodes and these were covered with a SiON-matrix. The inset in **a** shows the $I(V)$ -characteristics of the dielectric break down. The schematic design of the ex situ grids with structured Au contracts on Si₃N₄ membranes is illustrated in **c**. An EDX linescan in **d** along the orange arrow reveals that the fractal dendrites consist mainly of Ag. The position of **d** is indicated by the orange square in **b**. Adapted by permission from [Springer Nature Customer Service Centre GmbH]: [Springer] [64]

approach can be realized via a horizontal electrode design which can be lithographically structured on electron-transparent Si₃N₄ membranes. Such an electrode design is illustrated in Fig. 15c. The detailed procedure is described in reference [64]. By means of electron beam lithography electrode distances of 40 nm and below were achieved, which enable a good observation of the region of interest at high resolution. The corresponding prototype, deposited via, e.g., sputtering, can subsequently be analyzed in the TEM. Following this approach, it was possible to observe fractal growth in an AgPt nanoparticle film below the percolation limit, encapsulated in a SiON-matrix which suffered a dielectric breakdown during electrical probing. EDX measurements in Fig. 15c revealed that these fractal dendrites consist of silver and leave behind Pt-rich nanoparticles after electrical biasing. This observation provides information for the switching mechanism of a similar, but vertical nanoparticle-based

diffusive memristor [65]. In bimetallic nanoparticles, the less noble metal mainly contributes to the filament formation while the noble counter part remains behind and can serve as an anchor point for the filament to enhance memristive switching [21]. However, this ex situ approach may be insufficient, as a (partially) reduced oxide might re-oxidize during transfer to the ex situ measurement setup and electron beam damage could go unnoticed. To be sure, in situ or *in operando* experiments are necessary, and suitable approaches are described in the following section.

6.2 FIB Preparation for In situ Experiments

A horizontally-arranged prototype device may serve well for TEM measurements and fundamental research, but is not comparable to real devices, as the geometry of the electrodes has a major impact on the properties of the device [66]. For this purpose, the real devices have to be analyzed and to exclude preparation artefacts, the functionality of the TEM sample should be checked and compared to the device. TEM sample preparation for *in operando* experiments on vertical devices requiring FIB turned out to be a major obstacle due to Pt and Ga contamination and material re-deposition, all creating possible short circuits. To overcome these problems, the sample geometry and a specially adapted MEMS chip for in situ electrical biasing were designed by Zintler et al. [67]. The preparation improvements start with a special stub, mounting both the sample and the MEMS chip orthogonal to each other. This stub enables venting-free preparation including lift-out, transfer, contacting of the electrodes and thinning. Top and bottom electrodes are contacted with Pt trenches from the gas injection system in the FIB instrument. To prevent short circuits and memristive switching apart from the electron transparent region, the counter electrodes are cut next to the Pt trenches by means of Ga-ion etching. The whole preparation is described in detail in [67]. After successful preparation, the lamella can be electrically characterized in situ or even *in operando* in the TEM and the comparison of the lamella device with the macro device shows striking similarities (compare Fig. 16b, c with f, e). Further, the influence of the electron beam can be measured electrically.

Despite these striking similarities, the approach also presents some drawbacks. First, the preparation is time-consuming and is likely to end up in a short circuit. Secondly, the lamella device is quite large, and the resistive switching therefore difficult to observe, as it may occur at any point in the lamella. This second point is not only a drawback, as the geometry is not altered significantly and a larger area can be observed. Further, such a lamella has two free surfaces which may dominate the switching when diffusion or migration is involved. And last but not least, the small dimensions make Joule heating more critical which may lead to a faster device degradation and could limit TEM observations [67].

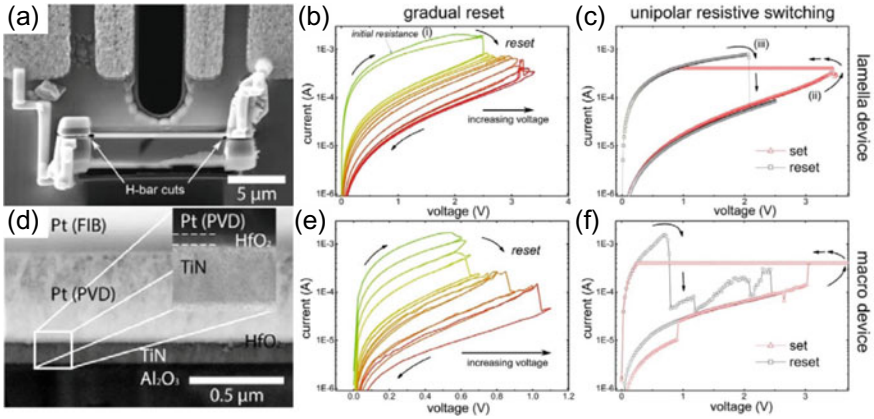


Fig. 16 **a** FIB lamella mounted and contacted by Pt trenches to contacts of the MEMS chip. **d** STEM image of the device for the in situ measurements. The inset shows a Bright-field STEM image with higher magnification. **b** and **c** show $I(V)$ -characteristics of the lamella device while **f** and **e** represent data for the macroscopic device. **b** and **f** show the gradual reset and **c** and **e** the unipolar resistive switching. Adapted with permission from Elsevier [67]

6.3 Conclusions and Best Practice

In conclusion, reliable TEM experiments on memristive devices can be realized by adhering to the following guidelines:

- Surface diffusion processes must be prevented, e.g., by the application of barrier layers on the top and bottom of the devices, cf. [51].
- Contamination of the samples must be prevented, e.g., by using suitable device transfer systems from the sputtering devices to the TEM.
- If possible, check for functionality before and after preparation and after the TEM-measurements. For this purpose, horizontal instead of vertical devices should be produced and analyzed. These systems do not require sophisticated TEM preparation, since they can be fabricated directly for TEM analysis.
- For real device characterization, electrically contacted FIB lamellas should be analyzed *in operando* and the $I(V)$ -characteristics should be critically compared to the real device.
- In *ex situ*, *in situ* or *in operando*, experiments, current measurements should be used to determine the threshold for electron beam damage. If this is not possible, the usage of low dose rates but high acceleration voltages is advisable, if dielectric memristive devices are to be characterized.

7 Experimental

TEM measurements were performed on different microscopes. High-end EELS for the DBMD characterization was done on a FEI Titan³ provided by ESTEEM3 facilities. Acceleration voltages of 300 kV and 80 kV were used in conjunction with a monochromator, XFEG, GIF quantum energy analyzer and SuperX EDX detector. The analysis of the electron beam crystallization of hafnia thin films was conducted on a JEOL JEM-2100 with 200 kV and a LaB₆ cathode. Hole-drilling experiments, also with hafnia thin films were performed using a Tecnai F30 G² microscope with 300 kV and a field emission gun.

Acknowledgements This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823717 – ESTEEM3. The financial support by the German Research Foundation (Deutsche Forschungsgemeinschaft—DFG) through FOR 2093 is gratefully acknowledged. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – ProjectID 434434223 – SFB 1461.

References

1. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7**, 011305 (2020). <https://doi.org/10.1063/1.5129306>
2. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R., Eleftheriou, E.: Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020). <https://doi.org/10.1038/s41565-020-0655-z>
3. Strobel, J., Neelisetty, K.K., Chakravadhanula, V.S.K., Kienle, L.: Transmission electron microscopy on memristive devices: an overview. *Appl. Microsc.* **46**, 206–216 (2016). <https://doi.org/10.9729/AM.2016.46.4.206>
4. Yang, Y., et al.: Probing electrochemistry at the nanoscale: in situ TEM and STM characterizations of conducting filaments in memristive devices. *J. Electroceramics* **39**, 73–93 (2017). <https://doi.org/10.1007/s10832-017-0069-y>
5. Sun, W., et al.: Understanding memristive switching via in situ characterization and device modeling. *Nat. Commun.* **10** (2019). <https://doi.org/10.1038/s41467-019-11411-6>
6. Xu, H., et al.: Dynamic structure-properties characterization and manipulation in advanced nanodevices. *Mater. Today Nano* **7**, 100042 (2019). <https://doi.org/10.1016/j.mtnano.2019.100042>
7. Yang, X., et al.: A review of in situ transmission electron microscopy study on the switching mechanism and packaging reliability in non-volatile memory. *J. Semicond.* **42**, 013102 (2021). <https://doi.org/10.1088/1674-4926/42/1/013102>
8. Ielmini, D., Waser, R. (eds.): *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*. Wiley (2015). <https://doi.org/10.1002/9783527680870>
9. Du, H., et al.: Nanosized conducting filaments formed by atomic-scale defects in redox-based resistive switching memories. *Chem. Mater.* **29**, 3164–3173 (2017). <https://doi.org/10.1021/acs.chemmater.7b00220>
10. Kwon, D.-H., et al.: Atomic structure of conducting nanofilaments in TiO₂ resistive switching memory. *Nat. Nanotechnol.* **5**, 148–153 (2010). <https://doi.org/10.1038/nnano.2009.456>
11. Li, C., et al.: Dynamic observation of oxygen vacancies in hafnia layer by in situ transmission electron microscopy. *Nano Res.* **8**, 3571–3579 (2015). <https://doi.org/10.1007/s12274-015-0857-0>

12. Jang, J.H., et al.: Investigation of oxygen-related defects and the electrical properties of atomic layer deposited HfO₂ films using electron energy-loss spectroscopy. *J. Appl. Phys.* **109**, 023718 (2011). <https://doi.org/10.1063/1.3544039>
13. Zahari, F., et al.: Correlation between sputter deposition parameters and I-V characteristics in double-barrier memristive devices. *J. Vac. Sci. Technol. B* **37**, 061203 (2019). <https://doi.org/10.1116/1.5119984>
14. Nufer, S., et al.: Quantitative atomic-scale analysis of interface structures: transmission electron microscopy and local density functional theory. *Phys. Rev. Lett.* **86**, 5066–5069 (2001). <https://doi.org/10.1103/PhysRevLett.86.5066>
15. Nufer, S.: Spatially Resolved Analytical Electron Microscopy at Grain Boundaries of α -Al₂O₃ (2001)
16. Kang, H.-J., Kim, J., Oh, J., Back, T., Kim, H.: Ultra-thin TEM sample preparation with advanced backside FIB milling method. *Microsc. Microanal.* **16**, 170–171 (2010). <https://doi.org/10.1017/S1431927610054474>
17. Gurvitch, M., Washington, M.A., Huggins, H.A.: High quality refractory Josephson tunnel junctions utilizing thin aluminum layers. *Appl. Phys. Lett.* **42**, 472–474 (1983). <https://doi.org/10.1063/1.93974>
18. Ishitani, T., Yaguchi, T.: Cross-sectional sample preparation by focused ion beam: a review of ion-sample interaction. *Microsc. Res. Tech.* **35**(4), 320–333 (1996). [https://doi.org/10.1002/\(SICI\)1097-0029\(19961101\)35:4<320::AID-JEMT3>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-0029(19961101)35:4<320::AID-JEMT3>3.0.CO;2-Q)
19. Strobel, J., et al.: In depth nano spectroscopic analysis on homogeneously switching double barrier memristive devices. *J. Appl. Phys.* **121**, 245307 (2017). <https://doi.org/10.1063/1.4990145>
20. Kiss, K.: Problem Solving with Microbeam Analysis. OCLC: 1058661531. Elsevier Science, Amsterdam (2014)
21. Vahl, A., et al.: Single target sputter deposition of alloy nanoparticles with adjustable composition via a gas aggregation cluster source. *Nanotechnology* **28**, 175703 (2017). <https://doi.org/10.1088/1361-6528/aa66ef>
22. Hansen, M. et al.: A double barrier memristive device. *Sci. Rep.* **5** (2015). <https://doi.org/10.1038/srep13753>
23. Jiang, N., Spence, J.C.: Interpretation of Oxygen K pre-edge peak in complex oxides. *Ultramicroscopy* **106**, 215–219 (2006). <https://doi.org/10.1016/j.ultramic.2005.07.004>
24. Klimenkov, M., Möslang, A., Lindau, R.: EELS analysis of complex precipitates in PM 2000 steel. *Eur. Phys. J. Appl. Phys.* **42**, 293–303 (2008). <https://doi.org/10.1051/epjap:2008084>
25. Nigo, S., et al.: Conduction band caused by oxygen vacancies in aluminum oxide for resistance random access memory. *J. Appl. Phys.* **112**, 033711 (2012). <https://doi.org/10.1063/1.4745048>
26. Bruley, J., et al.: Investigations of the chemistry and bonding at niobiumsapphire interfaces. *J. Mater. Res.* **9**, 2574–2583 (1994). <https://doi.org/10.1557/JMR.1994.2574>
27. Imamura, T., Hasuo, S.: Cross-sectional transmission electron microscopy observation of Nb/AlO_x-Al/Nb Josephson junctions. *Appl. Phys. Lett.* **58**, 645–647 (1991). <https://doi.org/10.1063/1.104556>
28. Winkler, B., et al.: Oxygen K-edge electron energy loss spectra of hydrous and anhydrous compounds. *J. Phys. Condens. Matter* **25**, 485401 (2013). <https://doi.org/10.1088/0953-8984/25/48/485401>
29. Wirth, R., Wunder, B.: Characterization of OH-containing phases by TEM using electron energy-loss spectroscopy (EELS): Clinohumite-OH, chondrodite-OH, phase A, and the (F, OH)-solid solution series of topaz. *J. Trace Microprobe Tech.* **18**(1), 35–50 (2000)
30. Garvie, L.A.: Can electron energy-loss spectroscopy (EELS) be used to quantify hydrogen in minerals from the O K edge? *Am. Mineral.* **95**, 92–97 (2010). <https://doi.org/10.2138/am.2010.3290>
31. Kosugi, N., Shigemasa, E., Yagishita, A.: High-resolution and symmetry-resolved oxygen K-edge spectra of 0 2. *Chem. Phys. Lett.* **190**, 8 (1992)
32. Jiang, N., Spence, J.C.H.: On the dose-rate threshold of beam damage in TEM. *Ultramicroscopy* **113**, 77–82 (2012). <https://doi.org/10.1016/j.ultramic.2011.11.016>

33. Mkhoyan, K.A., Silcox, J., Ellison, A., Ast, D., Dieckmann, R.: Full recovery of electron damage in glass at ambient temperatures. *Phys. Rev. Lett.* **96** (2006). <https://doi.org/10.1103/PhysRevLett.96.205506>
34. Berger, S.D., Salisbury, I.G., Milne, R.H., Imeson, D., Humphreys, C.J.: Electron energy-loss spectroscopy studies of nanometre-scale structures in alumina produced by intense electron-beam irradiation. *Philos. Mag. B* **55**, 341–358 (1987). <https://doi.org/10.1080/13642818708208619>
35. Cazaux, J.: Correlations between ionization radiation damage and charging effects in transmission electron microscopy. *Ultramicroscopy* **60**, 411–425 (1995). [https://doi.org/10.1016/0304-3991\(95\)00077-1](https://doi.org/10.1016/0304-3991(95)00077-1)
36. Ayoola, H.O., et al.: Origin and suppression of beam damage-induced oxygen-K edge artifact from gamma-Al₂O₃ using Cryo-EELS. *Ultramicroscopy* **219**, 113127 (2020). <https://doi.org/10.1016/j.ultramic.2020.113127>
37. Arhammar, C., et al.: Unveiling the complex electronic structure of amorphous metal oxides. *Proc. Natl. Acad. Sci.* **108**, 6355–6360 (2011). <https://doi.org/10.1073/pnas.1019698108>
38. Seeman, V., et al.: Atomic, electronic and magnetic structure of an oxygen interstitial in neutron-irradiated Al₂O₃ single crystals. *Sci. Rep.* **10**, 15852 (2020). <https://doi.org/10.1038/s41598-020-72958-9>
39. Fritz, S., et al.: Correlating the nanostructure of Al-oxide with deposition conditions and dielectric contributions of two-level systems in perspective of superconducting quantum circuits. *Sci. Rep.* **8**, 7956 (2018). <https://doi.org/10.1038/s41598-018-26066-4>
40. Malek, A., Eikerling, M.H.: Chemisorbed Oxygen at Pt(111): a DFT study of structural and electronic surface properties. *Electrocatalysis* **9**, xs370–379 (2018). <https://doi.org/10.1007/s12678-017-0436-0>
41. Tan, E., Mather, P.G., Perrella, A.C., Read, J.C., Buhrman, R.A.: Oxygen stoichiometry and instability in aluminum oxide tunnel barrier layers. *Phys. Rev. B* **71** (2005). <https://doi.org/10.1103/PhysRevB.71.161401>
42. Egerton, R.: Radiation damage to organic and inorganic specimens in the TEM. *Micron* **119**, 72–87 (2019). <https://doi.org/10.1016/j.micron.2019.01.005>
43. Jiang, N.: Electron beam damage in oxides: a review. *Rep. Prog. Phys.* **79**, 016501 (2016). <https://doi.org/10.1088/0034-4885/79/1/016501>
44. Hobbs, L.W.: Murphy's law and the uncertainty of electron probes. *Micro-anal. Microlithogr.* **4**, 14 (1990)
45. Zhang, H.-B., Feng, R.-J., Ura, K.: Utilizing the charging effect in scanning electron microscopy. *Sci. Prog.* **87**, 249–268 (2004). <https://doi.org/10.3184/003685004783238490>
46. Williams, D.B., Carter, C.B.: *Transmission Electron Microscopy: A Textbook for Materials Science*, 2nd edn., 4 pp. Springer, New York (2008)
47. Carstens, N., et al.: Enhancing reliability of studies on single filament Memristive switching via an unconventional cAFM approach. *Nanomaterials* **11**, 265 (2021). <https://doi.org/10.3390/nano11020265>
48. Liang, C. et al.: Ionic-electronic conductor nanostructures: template-confined growth and non-linear electrical transport. *Small* **1**, 971–975 (2005). <https://doi.org/10.1002/sml.200500155>
49. Liang, C., Terabe, K., Hasegawa, T., Aono, M.: Resistance switching of an individual Ag₂S/Ag nanowire heterostructure. *Nanotechnology* **18**, 485202 (2007). <https://doi.org/10.1088/0957-4484/18/48/485202>
50. Liao, Z.-M., et al.: Resistive switching and metallic-filament formation in Ag₂S nanowire transistors. *Small* **5**, 2377–2381 (2009). <https://doi.org/10.1002/sml.200900642>
51. Alissawi, N., et al.: Tuning of the ion release properties of silver nanoparticles buried under a hydrophobic polymer barrier. *J. Nanoparticle Res.* **14**, 928 (2012). <https://doi.org/10.1007/s11051-012-0928-z>
52. Li, Y., Wang, Z., Midya, R., Xia, Q., Yang, J.J.: Review of memristor devices in neuromorphic computing: materials sciences and device challenges. *J. Phys. D: Appl. Phys.* **51**, 503002 (2018). <https://doi.org/10.1088/1361-6463/aade3f>

53. Gritsenko, V.A., Perevalov, T.V., Islamov, D.R.: Electronic properties of hafnium oxide: a contribution from defects and traps. *Phys. Rep.* **613**, 1–20 (2016). <https://doi.org/10.1016/j.physrep.2015.11.002>
54. Grossi, A., Zambelli, C., Olivo, P., Crespo-Yepes, A., Martin-Martinez, J., Rodriguez, R., Nafria, M., Perez, E., Wenger, C.: Electrical characterization and modeling of 1T-1R RRAM arrays with amorphous and poly-crystalline HfO₂. *Solid-State Electron.* **7**, 187–193 (2016). <https://doi.org/10.1016/j.sse.2016.10.025>
55. Mittmann, T. et al.: Origin of ferroelectric phase in undoped HfO₂ films deposited by sputtering. *Adv. Mater. Interfaces*, 1900042 (2019). <https://doi.org/10.1002/admi.201900042>
56. Narasimhan, V.K., et al.: In situ characterization of ferroelectric HfO₂ during rapid thermal annealing. *Physica Status Solidi (RRL) – Rapid Res. Lett.* **15**, 2000598 (2021). <https://doi.org/10.1002/pssr.202000598>
57. Zhang, Y., et al.: Ion-induced damage accumulation and electron-beam-enhanced recrystallization in SrTiO₃. *Phys. Rev. B* **72**, 094112 (2005). <https://doi.org/10.1103/PhysRevB.72.094112>
58. Jencic, I., Bench, M.W., Robertson, I.M., Kirk, M.A.: Electron-beam-induced crystallization of isolated amorphous regions in Si, Ge, GaP, and GaAs. *J. Appl. Phys.* **78**, 974–982 (1995). <https://doi.org/10.1063/1.360764>
59. Lulli, G., Merli, P.G., Antisari, M.V.: Solid-phase epitaxy of amorphous silicon induced by electron irradiation at room temperature. *Phys. Rev. B* **36**, 8038–8042 (1987). <https://doi.org/10.1103/PhysRevB.36.8038>
60. Kern, P., Müller, Y., Patscheider, J., Michler, J.: Electron-beam-induced topographical, chemical, and structural patterning of amorphous titanium oxide films. *J. Phys. Chem. B* **110**, 23660–23668 (2006). <https://doi.org/10.1021/jp0642589>
61. Kern, P., Widmer, R., Gasser, P., Michler, J.: Local tuning of conductivity in amorphous titanium oxide films by selective electron beam irradiation. *J. Phys. Chem. C* **111**, 13972–13980 (2007). <https://doi.org/10.1021/jp0711803>
62. Molinari, A., et al.: Configurable resistive response in BaTiO₃ ferroelectric memristors via electron beam radiation. *Adv. Mater.* **32**, 1907541 (2020). <https://doi.org/10.1002/adma.201907541>
63. Neelisetty, K.K., et al.: Electron beam effects on oxide thin films-structure and electrical property correlations. *Microsc. Microanal.* **25**, 592–600 (2019). <https://doi.org/10.1017/S1431927619000175>
64. Gronenberg, O., Carstens, N., Vahl, A., Faupel, F., Kienle, L.: In: Tiginyanu, I., Sontea, V., Railean, S. (eds.) 4th International Conference on Nanotechnologies and Biomedical Engineering. IFMBE Proceedings, pp. 71–74. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-31866-6_15
65. Vahl, A., Carstens, N., Strunskus, T., Faupel, F., Hassanien, A.: Diffusive memristive switching on the nanoscale, from individual nanoparticles towards scalable nanocomposite devices. *Sci. Rep.* **9** (2019). <https://doi.org/10.1038/s41598-019-53720-2>
66. Xia, Q., et al.: Impact of geometry on the performance of memristive nanodevices. *Nanotechnology* **22**, 254026 (2011). <https://doi.org/10.1088/0957-4484/22/25/254026>
67. Zintler, A., et al.: FIB based fabrication of an operative Pt/HfO₂/TiN device for resistive switching inside a transmission electron microscope. *Ultramicroscopy* **181**, 144–149 (2017). <https://doi.org/10.1016/j.ultramic.2017.04.008>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Modeling and Simulation of Silver-Based Filamentary Memristive Devices



Sven Dirkmann, Jan Trieschmann, and Thomas Mussenbrock

Abstract Memristive devices based on filamentary switching are investigated by 3D kinetic Monte Carlo simulations. The electrochemical metallization device under consideration consists of a stack of Ag/TiO_x/Pt thin layers. By modeling the Ag ion transport within the solid state electrolyte, driven by the electric field and thermal diffusion, the dynamics of resistive switching and conducting filament growth/dissolution are studied. The model allows to resolve the macroscopic time scale of consecutive growth and dissolution cycles. It provides realistic current-voltage relations as observed experimentally. Simultaneously, it grants a detailed characterization of the influence of the electric field and the thermal heat on the local resistive switching dynamics. It finally provides insight into the microscopic physical mechanisms involved in the set and reset kinetics during switching. It is concluded that the force due to electromigration on Ag in the closed filament may not be negligible during reset process of the device.

Keywords Memristive devices · Electrochemical metallization cells · Filamentary resistive switching · Kinetic Monte Carlo · Modeling and simulation

1 Introduction

Today's predominant memory technology is based on the storage of electrical charge. In these devices relatively high energy barriers are required to prevent the loss of its memory state through electron tunneling currents. Due to approximately 1000 times larger mass of atoms compared to electrons, the tunneling of atoms to neighboring

S. Dirkmann · T. Mussenbrock (✉)
Ruhr University Bochum, Chair of Applied Electrodynamics and Plasma Technology, D-44780
Bochum, Germany
e-mail: Thomas.Mussenbrock@rub.de

J. Trieschmann
Kiel University, Faculty of Engineering, Theoretical Electrical Engineering, Kaiserstraße 2,
D-24143 Kiel, Germany
e-mail: jt@tf.uni-kiel.de

© The Author(s) 2024
M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series
on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_6

states is negligible. In contrast, memristive devices are based on the principle of resistive switching, whereby information is stored in the device resistance. These memory devices are based on the change of the atomic arrangement. Therefore, they do not require high energy barriers and could replace currently used memory concepts, especially with respect to scalability and energy efficiency.

The first memristive devices were introduced as early as the 1960s [1–3]. In 1976, specifically the emergence of Ag filaments as a possibility for resistive switching was recognized [4]. In the 1970s, Leon O. Chua further presented a theory of memristive systems [5, 6]. However, especially due to the rapid progress of silicon-based, integrated circuit technology, interest in memristive devices decreased from the end of the 1970s until new computer architectures were considered in the 1990s. In 2008, Strukov et al. demonstrated resistive switching in the context of the theory of memristive systems presented by Chua [7]. Due to their versatile usability, for example as artificial synapses in neural networks for new types beyond von Neumann computer architectures or as future, non-volatile storage elements, research on memristive devices is steadily increasing to this day [8–10].

Memristive devices are often used in pulsed mode operation. However, direct current operation and its related current-voltage (IV) characteristics provide important insights into the physical behavior of individual devices. For instance, on the basis of the IV characteristics a distinction can generally be made between unipolar and bipolar devices. In the case of bipolar switching devices, the voltage polarity in the set and reset process differs, whereas for unipolar devices it is the same. In addition to this very general classification, memristive devices can also be differentiated based on their physical behavior. For example there are memristive devices based on magnetic effects (such as magnetic tunnel devices [11]), on electrostatic effects (such as the trapping effect of electrons [12]). The memristive devices particularly considered in this work are based on a change in the atomic configuration. Even within this type of memristive systems, the exact physical mechanisms of resistive switching can be very different and range from diffusion and migration of ions [13] through chemical reactions to thermal effects [14] caused by Joule heating [15]. The exact switching behavior is often not understood and thus prevents a reliable integration of these devices in electronic circuits. As a theoretical approach, this chapter is henceforth dedicated to the modeling and simulation of electrochemical metallization (ECM) cells that exhibit filamentary switching behavior.

2 Electrochemical Metallization Cells

Typically, ECM cells consist of a solid electrolyte with poor electrical conductivity, embedded between a chemically active and an inert electrode. The active electrode usually consists of Ag, Cu, or Ni. In addition, resistive switching has already been demonstrated with active electrodes made of alloys containing Ag, Au, Cu, or Ni, like Au/Ag alloys [16]. In addition to the chemical inactivity, another requirement for the inert electrode is the poor miscibility with atoms of the active electrode. These

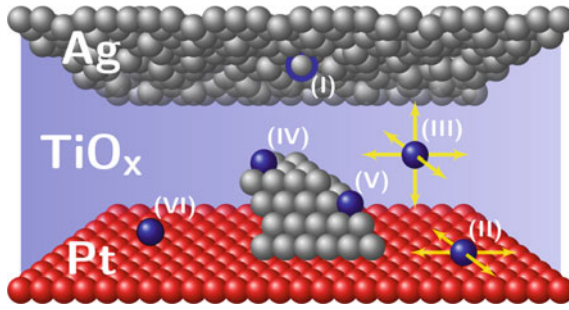


Fig. 1 Scheme of the ECM cell with incipient filament growth and illustration of the processes implemented in the simulation: (I) oxidation, (II) surface diffusion on the Pt electrode, (III) diffusion in the TiO_x , (IV) reduction on Ag, (V) reduction on a step, (VI) nucleation. Reprinted from [14], with the permission of AIP Publishing

requirements are met by materials like Pt, Ir, and W. A wide variety of materials have been demonstrated as possible electrolyte materials, such as TiO_2 , SiO_2 , Ge_xS_y [17].

In this work, the ECM devices based on the $Ag/TiO_x/Pt$ materials system is investigated with the help of simulations [18]. Its general functionality is schematically presented in Fig. 1. Assume that the ECM cell is in the high-resistance state (HRS) initially. In the set process, the chemically active Ag electrode is subject to a positive voltage (with respect to the opposing grounded electrode). This leads to oxidation of the Ag electrode following the reaction,



with monovalent Ag^+ ions [see Fig. 1 (I)] resulting in a partial dissolution of the active Ag electrode. Along with a related transfer of charge through the electrode, the Ag^+ ions are subject to the electric field in the solid electrolyte.

Due to this electric field, the Ag^+ ions move through the TiO_x toward the Pt electrode [see Fig. 1 (III)]. When the Ag^+ ions reach the Pt/ TiO_x interface, they can undergo different processes. Since Ag and Pt are almost immiscible at room temperature (see phase diagram Ag-Pt), surface diffusion of the Ag^+ ions occurs at the Pt/ TiO_x interface [see Fig. 1 (II)]. In addition, Ag^+ ions may form stable nuclei at the Pt electrode after reduction [19]. At these nuclei, additional Ag^+ ions may be reduced based on the following reaction [see Fig. 1 (IV)–(V)]:



This reaction leads to the growth of Ag filaments from the Pt electrode through the TiO_x electrolyte towards the Ag electrode. The two-step process of (i) formation of small Ag clusters and (ii) the reduction of Ag^+ ions on these clusters by electron transfer has been proven experimentally in 1999 [20].

If the distance between the filament and the Ag electrode decreases, the Ohmic resistance also decreases (increases the probability of electron tunneling). This leads

to an increase in the current through the device, which is often intentionally limited to a maximum value. This value of maximum current signifies whether there is a gap between the filament and the electrode, or the filament has formed up to the Ag electrode and begun to develop a galvanic contact. In this state, the device is in the low-resistance state (LRS). To reset the device to the HRS, the voltage polarity must be reversed. Depending on the current through the filament, it partially dissolves due to the electric field, supported by an increased temperature by Joule heating and by electromigration. The distance between the source and the sink of ions is determined by the thickness of the solid electrolyte and may therefore be very small. That may lead to desired properties like fast switching times and a very good scaling potential (<10 nm). In addition, only small voltages are required for device operation due to the small electrolyte thickness and the correspondingly very large electric field strengths. A very small energy consumption ($< \text{pJ}$) results. Despite these advantages and the technical feasibility of these devices, the randomness of the inherent physical processes remains an issue.

To gain a reliable understanding of the resistive switching of ECM cells, simulations on the atomic scale are indispensable. The multitude of different simulation models proposed so far range from concentrated models to molecular dynamics models. This also includes continuous models and kinetic Monte Carlo (kMC) models in 1D and 2D [21–26]. These models aim to describe resistive switching on various time scales. Since filament growth is a three-dimensional phenomenon, however, a 3D kMC simulation model is proposed as part of this work. The following central questions should be addressed: (i) How does filament growth proceed in ECM cells, (ii) what are the central influences that lead to the reset of ECM cells, and (iii) how does the distribution of atoms affect the resistance of the ECM cells?

3 Simulation Scenario and Simulation Methods

To simulate the characteristics of a real ECM cell, only a representative section of the device is considered in the model. The simulation box consists of a square base with an area of $40 \text{ nm} \times 40 \text{ nm}$, a 10 nm thick TiO_x solid electrolyte and a 3 nm thick section of an active Ag electrode. The Ag electrode consists of 38,400 individual atoms, whereas the opposite inert Pt electrode was modeled as a boundary for Ag^+ ion motion to compensate for the poor miscibility of Ag and Pt. At all other interfaces (i.e., the sides), periodic boundary conditions were used for ion motion. Occupiable, stable ion positions were represented by a cubic primitive lattice with a lattice constant of 5 \AA . The periodic lattice is reasoned by the corresponding short range order, which exists also in amorphous materials.

In titanium oxide, silver is preferably present as a monovalent ion at interstitial sites. It can diffuse through the bulk TiO_x under the influence of electric fields [27, 28]. The processes presented in Fig. 1 were taken into account also in the simulation, as these are identified as the primarily important processes. The rate of diffusion of Ag^+ ions through TiO_x is given by

Table 1 Parameters for the simulation of ECM cells

Symbol	Quantity	Value	References
σ_{Ag}	Conductivity of Ag	$6.3 \times 10^7 \text{ Sm}^{-1}$	
σ_{TiO_x}	Conductivity of TiO_x	$1.0 \times 10^2 \text{ Sm}^{-1}$	
$E_{\text{a,ox}}$	Activation energy for oxidation	0.65 eV	
$E_{\text{a,red}}$	Activation energy for reduction (surface/step)	0.62/0.58 eV	
$E_{\text{a,nuc}}$	Activation energy for nucleation	0.81 eV	[29]
$E_{\text{a,diff}}$	Activation energy for diffusion of Ag^+	0.61 eV	
$E_{\text{a,surf}}$	Activation energy for surface diffusion of Ag^+	0.59 eV	
ρ_{Ag}	Mass density of Ag	$10,490 \text{ kgm}^{-3}$	[30]
ρ_{TiO_x}	Mass density of TiO_2	$4,230 \text{ kgm}^{-3}$	[31]
c_{Ag}	Heat capacity of Ag	$235 \text{ Jkg}^{-1}\text{K}^{-1}$	[32]
c_{TiO_x}	Heat capacity of TiO_2	$700 \text{ Jkg}^{-1}\text{K}^{-1}$	[33]
λ_{Ag}	Thermal conductivity of Ag	$429 \text{ Wm}^{-1}\text{K}^{-1}$	[34]
λ_{TiO_x}	Thermal conductivity of TiO_2	$7 \text{ Wm}^{-1}\text{K}^{-1}$	[35]

$$k_{\text{diff}} = \nu_0 \exp\left(\frac{E_{\text{a,diff}} + 0.5z_i e(\phi_j - \phi_i)}{k_b T(\mathbf{r}, t)}\right). \quad (3)$$

Here, ν_0 is the phonon frequency, $E_{\text{a,diff}}$ is the activation energy for diffusion of Ag^+ ions through TiO_x , ϕ_i and ϕ_j are the potential at positions indicated by i and j , z_i is the charge number, e is the elementary charge, T is the local temperature, and k_b is the Boltzmann constant. Since diffusion at the material interfaces usually proceeds at higher velocities than in the volume, a smaller activation energy $E_{\text{a,surf}}$ is used for surface diffusion compared to volume diffusion. All relevant simulation parameters are summarized in Table 1.

The rates for reduction and oxidation can be derived from the Butler-Volmer equation [24]. One finds

$$k_{\text{red}} = \nu_0 \exp\left(-\frac{E_{\text{a,red}} - \alpha_0 z_i e \Delta\phi}{k_b T(\mathbf{r}, t)}\right) \quad (4)$$

$$k_{\text{ox}} = \nu_0 \exp\left(-\frac{E_{\text{a,ox}} - (1 - \alpha_0) z_i e \Delta\phi}{k_b T(\mathbf{r}, t)}\right) \quad (5)$$

with the charge transfer coefficient $\alpha_0 = 0.5$ and the activation energies $E_{\text{a,red}}$ and $E_{\text{a,ox}}$ for reduction and oxidation.

The potential difference $\Delta\phi$ at the electrode/ TiO_x and filament/ TiO_x interface is given as $\Delta\phi = \phi_{\text{TiO}_x} - \phi_{\text{electrode}}$ and $\Delta\phi = \phi_{\text{TiO}_x} - \phi_{\text{filament}}$, respectively. The process of heterogeneous nucleation depends on its environment. Thus, the time constant for nucleation can be given as a function of the Ag concentration. Here, the nucleation of Ag^+ ions is modeled analogously to [26] as a one-particle process where the nucleation process is independent of the material concentration. This ensures a

reduction in simulation time for the very time-consuming 3D simulations. The fact that a critical Ag concentration must first be reached so that nucleation becomes probable and the resulting slowing down of the process is represented here by an increased activation energy. In these simulations the following nucleation rate was used:

$$k_{\text{nuc}} = \nu_0 \exp\left(-\frac{E_{\text{a,nuc}} - \alpha_0 z_i e \Delta\phi}{k_b T(\mathbf{r}, t)}\right) \quad (6)$$

with the activation energy $E_{\text{a,nuc}}$ for nucleation. The correctness of this modeling approach may be confirmed experimentally afterwards from the activation energy for nucleation. The time it takes to set a filament (t_{set}) decreases exponentially for high voltages with the applied voltage. This is reflected in the equations of the process rates in Arrhenius form. t_{set} becomes infinite for small voltages [36]. Therefore, for every oxidation step additionally the voltage-dependent condition

$$\zeta < -\log[\cos(\xi V_{\text{appl}})] \quad (7)$$

must be satisfied, with the uniformly distributed random number $\zeta \in [0, 1]$, the fitting parameter ξ , and the applied voltage V_{appl} .

As an aside, it is well known that in TiO_x , vacancies move in the electrical field and can form conductive filaments. Depending on the kinetics of silver ion migration and vacancy migration, the latter may have an effect on the electric field distribution and hence on the formation of silver filaments. However, this mechanism is neglected in this work.

All mentioned processes depend on the electric field. It can be calculated based on the continuity equation $\nabla \cdot \mathbf{j} = 0$. Here, the displacement current has been neglected. With the help of a simple scale analysis, it can be shown that this assumption is justified. It turns out that the displacement current scales with $\mu\epsilon L^2/c^2 T^2$. Therein, L and T are the typical length and time scale of the system, respectively. With $L \propto 1$ nm and $T \propto 1$ s one finds that $\mu\epsilon L^2/c^2 T^2 \ll 1$.

Electron transport through the device is dominated by Ohmic conduction. In the literature, the tunneling current between the filament and the electrode is often reported as the dominant transport mechanism [19]. To show that electron transport is not dominated by tunneling electrons here, the Ohmic current is compared to the tunneling current at the smallest possible distance resolved in the simulation of 0.5 nm. This distance is chosen because the tunneling current scales exponentially with the thickness of the tunnel barrier, but the Ohmic current only linearly. Thus, for the minimum possible distance, the largest ratio between tunneling current and Ohmic current $I_{\text{tun}}/I_{\text{ohm}}$ is expected. The tunneling current density may be calculated using the Simmons equation [37]. The height of the tunnel barrier was chosen to be $\Phi_{\text{TB}} = \Phi_{\text{TiO}_x/\text{Ag}} = 0.36$ eV [38]. The effective mass of electrons was chosen to be $m_e^* = 9m_e$ [39].

The Ohmic current density was calculated from $j = \sigma E$ with $\sigma = \sigma_{\text{TiO}_x} = 100$ S/m. Due to the high conductivity of the TiO_x matrix, the Ohmic current den-

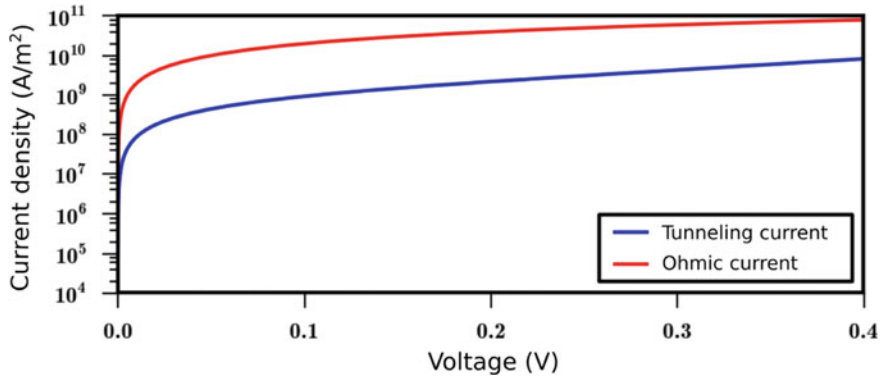


Fig. 2 Comparison between tunneling current and Ohmic current density for a distance between electrode and filament of 0.5 nm [40]

sity is significantly larger than the tunneling current density even in the case of the smallest possible distance between filament and electrode (see Fig. 2).

Consequently, a generalized Ohm's law,

$$\mathbf{j} = \sigma(\mathbf{r})\mathbf{E}(\mathbf{r}), \quad (8)$$

with the local conductivity $\sigma(\mathbf{r})$ of the respective material and the local electric field $\mathbf{E}(\mathbf{r})$ is used to describe the electron current density. Since the dynamics of the system is quasi-stationary, the electric field can be written as $\mathbf{E} = -\nabla\Phi$. The differential equation for the electrostatic potential is thus given by

$$\nabla \cdot (\sigma(\mathbf{r})\nabla\Phi) = 0 \quad (9)$$

To solve this differential equation, Dirichlet boundary conditions are applied to the upper and lower interfaces. Periodic boundary conditions are applied to all other interfaces. The differential equation is solved numerically using the successive over-relaxation method on a structured grid. The current I through the device is calculated using the integral

$$I = \int_A \sigma(\mathbf{r})\mathbf{E}(\mathbf{r}) \cdot \mathbf{n}dA. \quad (10)$$

This can be evaluated at any vertical position, since the continuity equation is implicitly satisfied everywhere.

All rates for the chemical and physical processes in the ECM cell are exponentially dependent on the local temperature. The temperature is therefore a decisive parameter for the memristive behavior of ECM cells. Due to the fact that in the LRS a metallic filament forms through the high resistance electrolyte matrix and connects the top and bottom electrodes, the current flow through the ECM cell is concen-

trated to the filament. This leads to a significant current density through the filament and thus potentially to a large temperature due to Joule heating. Consequently, the assumption of a constant temperature may not be justified and therefore the temperature in the device is calculated. The material dependent parameters of the heat conduction equation are also listed in Table 1. To investigate the effect of temperature on the memristive behavior, the temperature in two different materials, Ag and TiO_x , is calculated. The time constants of temperature evolution in these materials can be estimated based on a scale analysis to $\tau_{\text{Ag}} \propto 5.7 \times 10^{-15}$ s and $\tau_{\text{Ag}} \propto 4.2 \times 10^{-13}$ s with a typical length scale of the system of 1 nm. The typical time scale of ion movement accordingly results from the largest jump rate k_{ij} of the ions of the respective iteration,

$$\tau_{\text{hop}} = \frac{\ln \zeta}{k_{ij}} = \ln \zeta v_0^{-1} \exp\left(\frac{E_a}{k_b T}\right) \approx 10^{-2} \text{s}, \quad (11)$$

with the uniformly distributed random number $\zeta \in]0, 1]$. The timescale of memristive behavior is thus significantly larger than the time constant of the temperature development. Therefore, the steady-state heat equation,

$$-\nabla \cdot (\lambda \nabla T) = \mathbf{j} \cdot \mathbf{E}, \quad (12)$$

is solved to calculate the temperature. $\lambda(\mathbf{r})$ is the material-dependent thermal conductivity, whereas the right hand side represents the Joule heating source term. The upper and lower boundary surfaces are set to room temperature throughout the simulation. Periodic boundary conditions are again used at all other interfaces. This differential equation is solved numerically on the grid, which was also used for the potential. To account for the influence of temperature on the ion motion, the calculated temperature is inserted into the rate equations of the chemical and physical process rates.

4 Results and Discussion

The simulation model was used to calculate the IV characteristics of the Ag/ TiO_x /Pt ECM cell as well as the atomic state and switching kinetics. An ideal voltage source was applied to the Ag electrode, whereas the inert Pt electrode was set to the constant potential of 0 V. The source voltage and the voltage applied to the device, as well as the calculated IV characteristic are shown in Fig. 3. The calculated IV characteristics is in excellent agreement with experimental measurements as reported by Yang [41]. Five instances of time characteristic of the applied voltage are indicated by the numbers (1)–(5) and six different instances of time characteristic of the reset process are indicated by the letters (a)–(f).

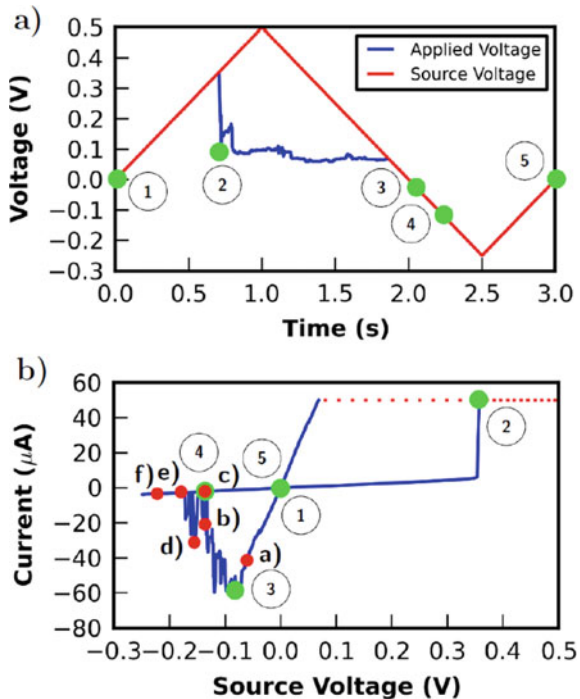
The voltage ramp is applied as follows: The voltage source initially applies a voltage ramp with the slope 0.5 V/s to the device. To prevent an electrical breakdown,

the current is limited to a maximum value of $I_{CC} = 50 \mu\text{A}$. If the current through the device exceeds this value, the voltage source is replaced by an ideal current source which provides exactly I_{CC} . Here, the linear relationship between current and voltage through the device can be exploited. Using a test voltage and a corresponding test current, the resistance of the device, R_{ECM} , can be obtained. From this resistance and the current I_{CC} it is possible to calculate $V_{CC} = R_{ECM}I_{CC}$, i.e., the actual voltage applied to the device. This applied voltage differs from the source voltage.

When the (original) source voltage reaches 0.5 V (this value was chosen to ensure resistive switching), the slope of the voltage ramp is inverted to a value of -0.5 V/s . If as a result the current drops below I_{CC} , the current source is again replaced by a voltage source. When the voltage reaches -0.25 V , the slope of the voltage ramp is reversed to 0.5 V/s until the voltage returns to 0 V.

The calculated temperature distribution (right) and the corresponding atomic state (left) for the five selected instances of time (1)–(5) from Fig. 3 are depicted in Fig. 4. In the original state, where the applied voltage is 0 V, no current flows through the device. All Ag atoms are within the electrode [time (1)]. Since no current flows through the device, the temperature in the device is at room temperature everywhere. When a positive voltage is applied to the Ag electrode, the oxidation process starts at the electrode due to the electric field induced by the externally applied voltage. The oxidized, positive Ag^+ ions move through the solid-state electrolyte, forming a

Fig. 3 Input/output behavior of the ECM cell. Top: Ramped source voltage (red) and voltage applied to the simulation region (blue) as a function of time. Bottom: IV characteristics plotted versus source voltage. Five important instances of time during the voltage ramp (green dots) and six important instances of time during the reset process (red dots) are marked. Reprinted from [14], with the permission of AIP Publishing



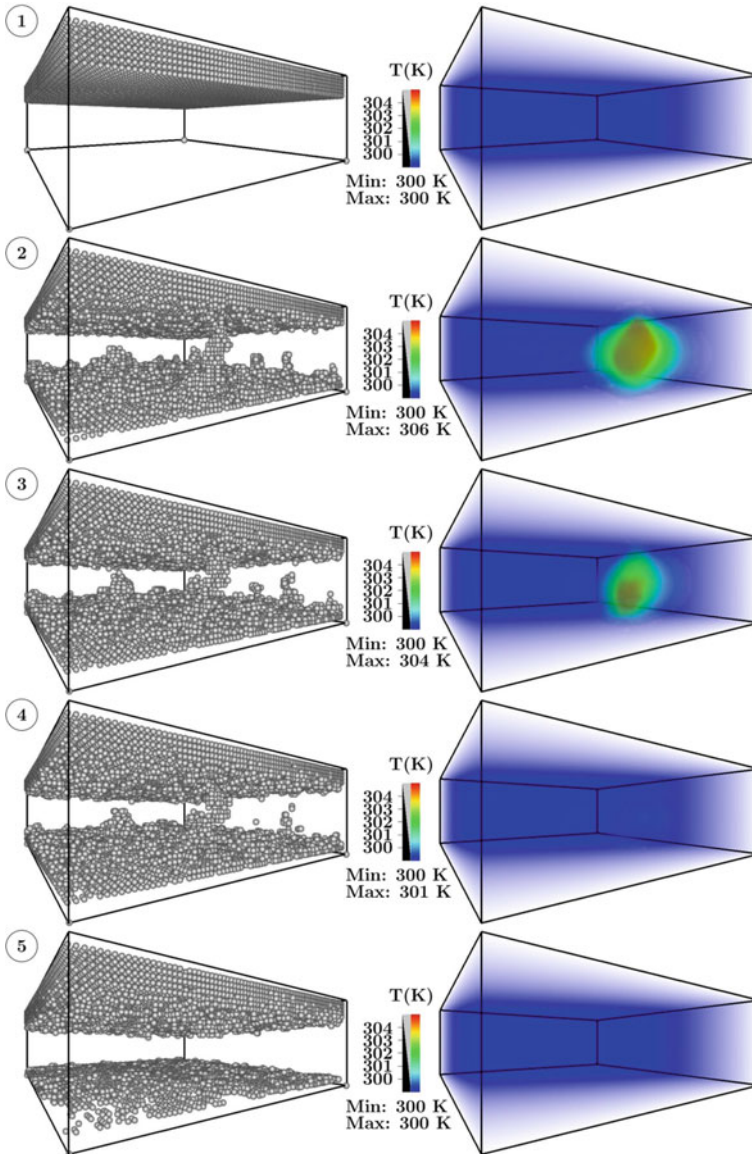


Fig. 4 Status of ECM cell shown for the five marked instances of time from Fig. 3. Left: Status of filament growth. Right: Corresponding temperature distribution. Reprinted from [14], with the permission of AIP Publishing

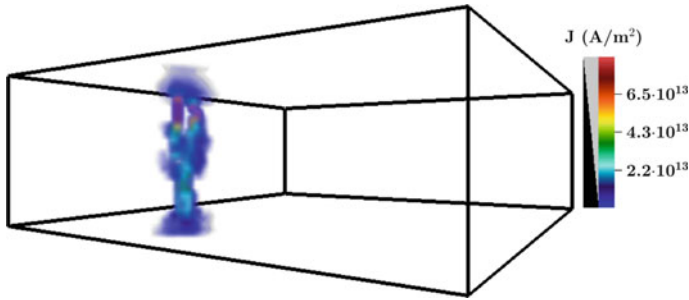


Fig. 5 Calculated current density distribution through the simulation area at time (2) from Fig. 3 immediately after electroforming of the conductive filament. Reprinted from [14], with the permission of AIP Publishing

stable nucleus at the Pt electrode. They reduce at this nucleus and form a conductive filament through the solid-state electrolyte. Once the conductive filament connects the two electrodes, a significant current flow through the filament is observed [time (2)].

At this point it is important to mention that resistive switching does not necessarily require a connection between the two opposing electrodes. In addition to a connection, it is also possible for a gap to remain between the filament and the Ag electrode. Either a tunneling current flows across this gap, if the resistance of the solid electrolyte is correspondingly high, or an Ohmic current, as presented here. In both cases, the current through the device is significantly smaller than in the case of a connection through the filament, which leads to a lower temperature inside the device for the same applied voltage [42]. Therefore, the retention of a gap can only be guaranteed by a correspondingly small current limit.

Figure 5 shows the current density shortly after the filament has connected the upper and the lower electrode [time (2)]. This current density is used as a source term for the temperature calculation. As expected, the current mainly flows through the conductive filament and leads to a local heating of the device. The temperature of the device increases to a maximum value of 320.9 K due to Joule heating. When the current reaches the limiting current I_{CC} , the voltage applied to the device decreases due to the reduced resistance (blue line on the right side of Fig. 3). Due to this reduced voltage during current limiting, further growth of the filament is significantly slowed down.

If the voltage polarity is reversed, the current increases again [time (3)]. Therefore, the temperature in the device also rises again and reaches values around 305 K. As soon as the filament breaks and thus the connection between the two opposing electrodes is dissolved [time (4)], the temperature drops again to room temperature. Due to the electric field within the solid state electrolyte, the filament is degraded [time (5)].

Since no nucleation seed was set in this simulation, nucleation occurs at random positions of the Pt electrode. At these nucleation sites, reduction occurs preferentially

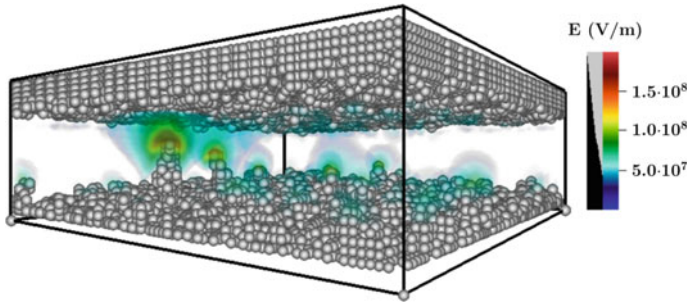


Fig. 6 Calculated magnitude of the electric field in the simulation area at time point $t = 0.66$ s directly before electroforming of the conductive filament. Reprinted from [14], with the permission of AIP Publishing

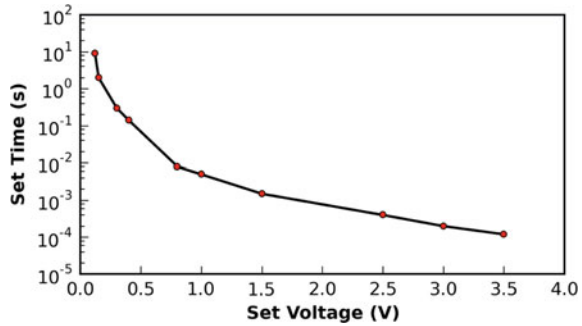
and stable clusters are formed. Figure 6 shows the magnitude of the electric field at time $t = 0.66$ s. At the positions of the clusters, the electric field is increased due to sharp edges and because of a reduced distance between cluster and Ag electrode. Therefore, the growth of the filament is strongly accelerated at these positions. Due to the influence of the inhomogeneous electric field within the solid electrolyte, large clusters grow faster than small ones and eventually form a conductive filament.

Due to the periodic boundary conditions, a copy of the simulation box can be imagined attached to the simulation box in x and y direction for the electric field calculation. The electric field in these copies of the simulation box influences the electric field in the simulation box. This is correct in so far, since the real device extends much further in x and y direction than the simulation box. Since the surface roughness caused by the inhomogeneous distribution of the Ag in the simulation box is randomly distributed, periodic boundary conditions are a good assumption. It limits the interpretability to distinguish single or multi filament behavior, however, as coupling may occur.

Due to the fact that filament growth strongly depends on the electric field, the set time varies with the applied voltage. To investigate the switching kinetics of the ECM cell, different constant voltages were applied to the device and the time until electroforming of a filament was calculated. Figure 7 shows the result of these calculations. With set times in the range of 0.1 ms to 10 s depending on the applied voltage, the calculated switching kinetics of the device are comparable to corresponding experimental measurements of ECM cells [43].

Another aspect of particular interest in the context of the presented simulations are the kinetics of the reset process. At the time of maximum negative current, the maximum temperature is 304.6 K. Even though higher temperatures are expected for higher power densities within the device, this simulation result means that a critical temperature rise is not an issue for the typical use of ECM cells in integrated circuits [44]. Thus, it is clear that the dissolution and filament reset process is predominantly caused by the electric field.

Fig. 7 Switching kinetics of the ECM cell. Calculated time to form of the conductive filament for different applied voltages. Reprinted from [14], with the permission of AIP Publishing



Unlike the set process, which is relatively abrupt, the reset process is often gradual. In addition, the reset process is stochastic and varies from cycle to cycle [45]. Figure 8 shows the atomic state of the conductive filament for the six instances of time of the reset process (a)–(f) selected in Fig. 3. This figure is instructive to explain the reset process in detail. At time (a), the filament is completely established and the reset process has not yet started. Due to the electric field, Ag atoms of the filament oxidize at random positions depending on the potential drop between filament and solid electrolyte and move away from the filament.

Figure 9 shows the distribution of the electrostatic potential over a cross-section of the ECM cell at time $t = 2.2$ s, shortly before the conductive filament breaks. Due to the connection through the conductive filament, the potential is nearly linear from the top to the bottom electrode. A slight variation is observed as the magnitude of the electric field increases at constrictions of the filament. The filament is subject to the electric field approximately uniformly in position and provides oxidation processes. The resulting thinning of the filament leads to several weak connections between the upper and lower electrode [time (b)], shown by red circles in Fig. 8. In consequence, the resistance of the device grows successively, subject to the random dissolution and ion motion. This leads to the stochastic behavior of the reset process. At these weak connection spots, both the temperature and the electric field increase, due to the enhanced current density and resistance. The breakup of the conductive filament occurs preferentially at these narrow spots [time (c)]. After the filament has split. It can nevertheless reconnect [time (d)], which again leads to a drop in resistance and thus to an increase in current. The final breakup of the conductive filament is shown at time (e). As the filament recedes, isolated islands are formed. Additionally, it can be observed that the growth of the filament also starts at the active Ag electrode. This growth mode can also be shown in other simulations [21] as well as in experiments [46].

At this point, two more important points should be discussed. First, it is important to note that the presented simulation model is only valid if the conductivity of the electrolyte matrix is such that Ohmic current dominates over tunneling currents and ionic current conduction. If this is not the case, these conduction mechanisms would have to be accounted for in the continuity equation.

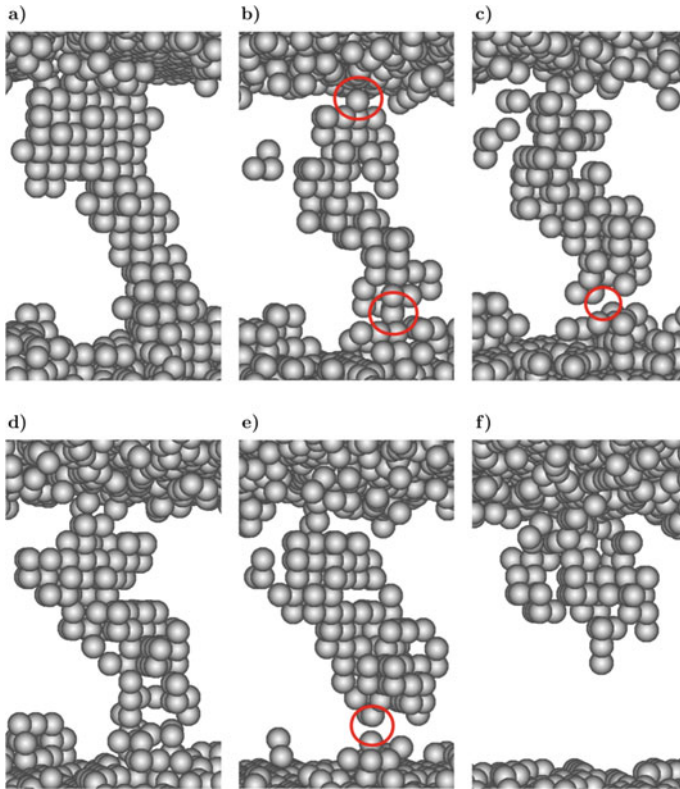


Fig. 8 Atomic configuration of the conductive filament at the six times marked in Fig. 3. Marked instances of time of the reset process. Reprinted from [14], with the permission of AIP Publishing

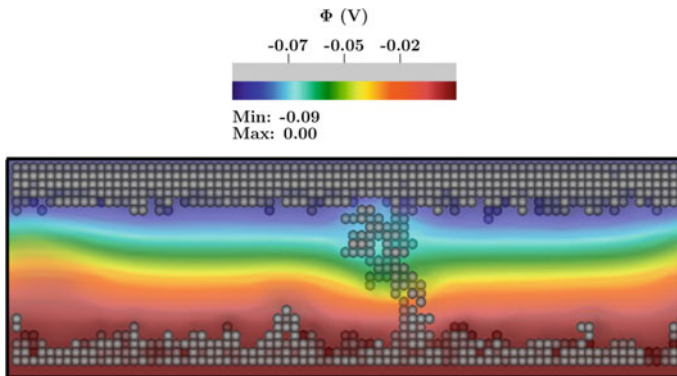


Fig. 9 Potential distribution at time $t = 2.2$ s directly before electroforming of the conductive filament in a section through the simulation area at the position of the filament. Reprinted from [14], with the permission of AIP Publishing

Second, it should be pointed out that this model focuses especially on the influence of temperature and electric field on the reset process. Another mechanism of electromigration was considered in this model only indirectly, through the process of re-oxidation. In the literature, this mechanism is hardly discussed in the context of the reset process of ECM cells [47, 48]. Considering electric fields in a good electrical conductor, the force on an ion can be expressed as a function of an effective charge number. For silver, the effective charge number is -26 [49]. Thus, the force of the electron flow dominates over the force of the electric field on Ag in the filament. It becomes particularly large (but disappears in quantum point contacts) at the narrow spots of the filament. Electromigration can thus facilitate movement of Ag out of the filament. Once they dissolve from the filament, they have a monovalent charge and are no longer affected by electromigration, but by the electric field, since the electron current is confined to the filament according to Fig. 5. Consequently, electromigration may be a crucial mechanism for the initial breakup of the filament. Especially due to the magnitude of the force on Ag in the filament, this mechanism should be considered in future extensions of this model.

5 Conclusions and remarks

In this chapter, the memristive behavior of ECM cells was investigated using a typical Ag/TiO_x/Pt ECM cell as an example. The focus was on the phenomenon of resistive switching, the physical processes of filament growth, and the driving forces of the reset process. A simulation model was developed to address these aspects. The ion motion was described using the kMC method. The continuity equation assuming purely Ohmic behavior was solved to calculate the current density and the electric field. Additionally, the steady-state heat equation was solved to calculate the temperature within the device, subject to Joule heating. First, it was shown that the tunnelling currents are negligible compared to Ohmic currents in the devices at all times. The formation and dissolution of a conductive Ag filament could be identified as the main cause of resistive switching. In addition, the main chemical and physical processes leading to the filament growth could be described, such as oxidation, diffusion, reduction, and nucleation. It was shown that the main driving force for filament growth is the electric field. A large electric field results when the distance between filament and electrode becomes very small. Therefore, growth occurs predominantly at the tip of pre-formed (incomplete) filaments. Furthermore, it was argued that the calculated IV characteristics of the device is in very good agreement with measurements.

To investigate the different physical mechanisms that lead to the reset process, the distribution of the current density in the device and the resulting increase in temperature were calculated. Accordingly, the maximum current density is about 8×10^{13} A/m² and, as expected, flows mainly through the conductive filament. The maximum calculated increase in temperature during the reset process is approximately 5 K. Accordingly, it was concluded that the main driving force of the reset process is the electric field, not the temperature. In addition, it was discussed that

the force due to electromigration on Ag in the closed filament may not be negligible. Up to now, this process has only been indirectly included in the simulation by the process of re-oxidation from the filament. Finally, the reset process of the atomic scale was described in detail.

Acknowledgements The authors gratefully acknowledge financial support provided by the German Research Foundation in the frame of Collaborative Research Centre SFB 1461 (Project-ID 434434223), Research Unit FOR 2093 (Project-ID 239767484), and Research Grants MU 2332/4-1, MU 2332/7-1, and MU 2332/10-1.

References

1. Dearnaley, G., Stoneham, A.M., Morgan, D.V.: Electrical phenomena in amorphous oxide films. *Rep. Prog. Phys.* **33**(3), 1129 (1970)
2. Gibbons, J.F., Beadle, W.E.: Switching properties of thin NiO films. *Solid-State Electron.* **7**(11), 785–790 (1964)
3. Hickmott, T.W.: Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **33**(9), 2669–2682 (1962)
4. Hirose, Y., Hirose, H.: Polarity-dependent memory switching and behavior of Ag dendrite in Ag-photodoped amorphous As₂S₃ films. *J. Appl. Phys.* **47**(6), 2767–2772 (1976)
5. Chua, L.O.: Memristor-the missing circuit element. *IEEE Trans. Circuit Theory* **18**(5), 507–519 (1971)
6. Chua, L.O., Kang, S.M.: Memristive devices and systems. *Proc. IEEE* **64**(2), 209–223 (1976)
7. Strukov, D., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. *Nature* **453**(80), 80–83 (2008)
8. Thomas, A.: Memristor-based neural networks. *J. Phys. D: Appl. Phys.* **46**(9), 093001 (2013)
9. Yang, J.J., Strukov, D.B., Stewart, D.R.: Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13–24 (2013)
10. Yang, J.J., Strukov, D.B., Stewart, D.R.: Pattern recognition with TiO_x-based memristive devices. *AIMS Mater. Sci.* **2**, 203–216 (2015)
11. Krzysteczko, P., Münchenberger, J., Schäfers, M., Reiss, G., Thomas, A.: The memristive magnetic tunnel junction as a nanoscopic synapse-neuron system. *Adv. Mater.* **24**(6), 762–766 (2012)
12. Shao, X.L., Zhou, L.W., Yoon, K.J., Jiang, H., Zhao, J.S., Zhang, K.L., Yoo, S., Hwang, C.S.: Electronic resistance switching in the Al/TiO_x/Al structure for forming-free and area-scalable memory. *Nanoscale* **7**, 11063–11074 (2015)
13. Dirkmann, S., Hansen, M., Ziegler, M., Kohlstedt, H., Mussenbrock, T.: The role of ion transport phenomena in memristive double barrier devices. *Sci. Rep.* **6**, 35686 (2016)
14. Dirkmann, S., Mussenbrock, T.: Resistive switching in memristive electrochemical metallization devices. *AIP Adv.* **7**(6), 065006 (2017)
15. Dirkmann, S., Kaiser, J., Wenger, C., Mussenbrock, T.: Filament growth and resistive switching in hafnium oxide memristive devices. *ACS Appl. Mater. Interfaces* **10**, 14857 (2018)
16. Kuo, C.C., Chen, I.C., Shih, C.C., Chang, K.C., Huang, C.H., Chen, P.H., Chang, T.C., Tsai, T.M., Chang, J.S., Huang, J.C.: Galvanic effect of Au-Ag electrodes for conductive bridging resistive switching memory. *IEEE Electron Device Lett.* **36**(12), 1321–1324 (2015)
17. Pan, F., Gao, S., Chen, C., Song, C., Zeng, F.: Recent progress in resistive random access memories: materials, switching mechanisms, and performance. *Mater. Sci. Eng. R: Rep.* **83**, 1–59 (2014)
18. Tsunoda, K., Fukuzumi, Y., Jameson, J.R., Wang, Z., Griffin, P.B., Nishi, Y.: Bipolar resistive switching in polycrystalline TiO₂ films. *Appl. Phys. Lett.* **90**(11), 113501 (2007)

19. Menzel, S., Tappertzhofen, S., Waser, R., Valov, I.: Switching kinetics of electrochemical metallization memory cells. *Phys. Chem. Chem. Phys.* **15**, 6945–6952 (2013)
20. Henglein, A., Giersig, M.: Formation of colloidal silver nanoparticles: capping action of citrate. *J. Phys. Chem. B* **103**(44), 9533–9539 (1999)
21. Dirkmann, S., Ziegler, M., Hansen, M., Kohlstedt, H., Trieschmann, J., Mussenbrock, T.: Kinetic simulation of filament growth dynamics in memristive electrochemical metallization devices. *J. Appl. Phys.* **118**(21), 214501 (2015)
22. Gergs, T., Dirkmann, S., Mussenbrock, T.: Integration of external electric fields in molecular dynamics simulation models for resistive switching devices. *J. Appl. Phys.* **123**, 245301 (2018)
23. Jameson, J.R., Gilbert, N., Koushan, F., Saenz, J., Wang, J., Hollmer, S., Kozicki, M.: Effects of cooperative ionic motion on programming kinetics of conductive-bridge memory cells. *Appl. Phys. Lett.* **100**(2), 023505 (2012)
24. Menzel, S., Kaupmann, P., Waser, R.: Understanding filamentary growth in electrochemical metallization memory cells using kinetic Monte Carlo simulations. *Nanoscale* **7**, 12673–12681 (2015)
25. Onofrio, N., Guzman, D., Strachan, A.: Atomic origin of ultrafast resistance switching in nanoscale electrometallization cells. *Nat. Mater.* **14**, 440–446 (2015)
26. Qin, S., Liu, Z., Zhang, G., Zhang, J., Sun, Y., Wu, H., Qian, H., Yu, Z.: Atomistic study of dynamics for metallic filament growth in conductive-bridge random access memory. *Phys. Chem. Chem. Phys.* **17**, 8627–8632 (2015)
27. Kulczyk-Malecka, J., Kelly, P., West, G., Clarke, G., Ridealgh, J., Almqvist, K., Greer, A., Barber, Z.: Investigation of silver diffusion in TiO₂/Ag/TiO₂ coatings. *Acta Materialia* **66**, 396–404 (2014)
28. Prada, S., Rosa, M., Giordano, L., Di Valentin, C., Pacchioni, G.: Density functional theory study of TiO₂/Ag interfaces and their role in memristor devices. *Phys. Rev. B* **83**, 245314 (2011)
29. Yoo, J., Park, J., Song, J., Lim, S., Hwang, H.: Field-induced nucleation in threshold switching characteristics of electrochemical metallization devices. *Appl. Phys. Lett.* **111**(6), 063109 (2017)
30. Enghag, P.: *Encyclopedia of the Elements: Technical Data - History - Processing - Applications*. Wiley (2004)
31. Kharisov, B.I., Kharisova, O.V., Ortiz-Mendez, U.: *CRC Concise Encyclopedia of Nanotechnology*. CRC Press (2016)
32. Abu-Eishah, S.I., Haddad, Y., Solieman, A., Bajbouj, A.: A new correlation for the specific heat of metals, metal oxides and metal fluorides as a function of temperature. *Lat. Am. Appl. Res.* **113**(4), 257–264 (2004)
33. Saeedian, M., Mahjour-Shafiei, M., Shojaee, E., Mohammadzadeh, M.R.: Specific Heat Capacity of TiO₂ Nanoparticles. *J. Comput. Theor. Nanosci.* **9**(4), 616–620 (2012). <https://doi.org/10.1166/jctn.2012.2070>
34. Ho, C.Y., Powell, R.W., Liley, P.E.: Thermal conductivity of the elements. *J. Phys. Chem. Ref. Data* **1**(2), 279–421 (1972)
35. Lu, Y.M., Noman, M., Picard, Y.N., Bain, J.A., Salvador, P.A., Skowronski, M.: Impact of joule heating on the microstructure of nanoscale TiO₂ resistive switching devices. *J. Appl. Phys.* **113**(16), 163703 (2013)
36. Russo, U., Kamalanathan, D., Ielmini, D., Lacaita, A.L., Kozicki, M.N.: Study of multilevel programming in programmable metallization cell (PMC) memory. *IEEE Trans. Electron Devices* **56**(5), 1040–1047 (2009)
37. Simmons, J.G.: Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**(6), 1793–1803 (1963)
38. Haus, J.W., Li, L., Katte, N., Deng, C., Scalora, M., de Ceglia, D., Vincenti, M.A., Buranasiri, P.: Nanowire metal-insulator-metal plasmonic devices. *Proc. SPIE* **8883** (2013)
39. Enright, B., Fitzmaurice, D.: Spectroscopic determination of electron and hole effective masses in a nanocrystalline semiconductor film. *J. Phys. Chem.* **100**(3), 1027–1035 (1996)
40. Dirkmann, S.: Ph.D. Thesis, Ruhr University Bochum (2018)

41. Yang, L.: Resistive Switching in TiO₂ Thin Films. Forschungszentrum Jülich (2011)
42. Di Martino, G., Tappertzhofen, S., Hofmann, S., Baumberg, J.: Nanoscale plasmon-enhanced spectroscopy in memristive switches. *Small* **12**(10), 1334–1341 (2016)
43. Lübben, M., Menzel, S., Park, S.G., Yang, M., Waser, R., Valov, I.: SET kinetics of electrochemical metallization cells: influence of counter-electrodes in SiO₂/Ag based systems. *Nanotechnology* **28**(13), 135205 (2017)
44. Menzel, S., Valov, I., Waser, R., Adler, N., van den Hurk, J., Tappertzhofen, S.: Simulation of polarity independent RESET in electrochemical metallization memory cells. In: 5th IEEE International Memory Workshop, pp. 92–95 (2013). <https://doi.org/10.1109/IMW.2013.6582106>
45. Valov, I., Linn, E., Tappertzhofen, S., Schmelzer, S., van den Hurk, J., Lentz, F., Waser, R.: Nanobatteries in redox-based resistive switches require extension of memristor theory. *Nat. Commun.* **4**, 1771 (2013)
46. Celano, U., Goux, L., Belmonte, A., Opsomer, K., Franquet, A., Schulze, A., Detavernier, C., Richard, O., Bender, H., Jurczak, M., Vandervorst, W.: Three-dimensional observation of the conductive filament in nanoscaled resistive memory devices. *Nano Lett.* **14**(5), 2401–2406 (2014)
47. Celano, U.: Metrology and Physical Mechanisms in New Generation Ionic Devices. Ph.D. thesis, Catholic University of Leuven (2016)
48. Ielmini, D., Waser, R.: Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications. Wiley (2016)
49. D’Heurle F., Rosenberg, R.: Physics of Thin Films in: Advances in Research and Development. Academic Press (1974)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Integration of Memristive Devices into a 130 nm CMOS Baseline Technology



Mamathamba Kalishettyhalli Mahadevaiah, Marco Lisker, Mirko Fraschke, Steffen Marschmeyer, Eduardo Perez, Emilio Perez-Bosch Quesada, Christian Wenger, and Andreas Mai

Abstract The two main features of the memristive devices which makes them the promising candidates for neuromorphic applications are low power consumption and CMOS compatibility. The monolithic integration of memristive devices with CMOS circuitry paves the way for in-memory computing. This chapter focuses on the factors governing the CMOS integration process. Firstly, the influence of CMOS baseline technology selection on the memristor module is briefly discussed. Secondly, the selection of metal level interconnects and their effect on the memristive device performance is explained. Further, the widely used deposition technique for the CMOS compatible memristive switching layers is presented. Finally, the implementation of the optimized process for the fabrication of the memristive module and its influence on the device performance is presented in terms of electrical characterization results.

Keywords CMOS compatibility · CMOS integration · Memristive devices · Pristine current · Resistive switching · Thermal budget · Encapsulation

M. K. Mahadevaiah · M. Lisker · M. Fraschke · S. Marschmeyer · E. Perez · E. P.-B. Quesada · C. Wenger · A. Mai

IHP Leibniz-Institut fuer innovative Mikroelektronik, Frankfurt/Oder, Germany

C. Wenger (✉)

BTU Cottbus-Senftenberg, 01968 Cottbus, Germany

e-mail: wenger@ihp-microelectronics.com

M. Lisker · A. Mai

University of Applied Science Wildau, Hochschulring 1, 15745 Wildau, Germany

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_7

1 Introduction

The well know and widely used contemporary computer hardware architectures namely, von Neumann, Harvard, Graphics Processing units (GPU) etc. have their memory and logic fabricated in different process nodes [1, 2]. The scalability gap existing between logic and memory components is the main reason for their fabrication divergence. Such hardware configuration with separate memory and logic leads to a latency delay in moving the data between the two, which is called as von Neumann bottleneck (VNB) [3]. As a result, the throughput of the hardware architectures becomes limited and several approaches had been proposed to overcome this issue. However, a change of the hardware architecture become unavoidable to solve the VNB problem. Memristive devices can be easily fabricated in the complementary metal-oxide semiconductor (CMOS) baseline technologies which are used for the logic circuitry. This possibility of monolithic integration of memory and logic pave the way for the future in-memory computing hardware architectures [2].

Moreover, memristive devices are considered as one of the potential candidates in the field of neuromorphic and edge computing applications due to their fast switching, multi-level conduction, low power consumption, high scalability and CMOS compatibility [4]. In order to obtain a reliable and reproducible performance of memristive devices, integrating them with the CMOS transistors is very essential. Firstly, the transistor which is fabricated in series with the memristor, prevents the sneak path currents in case of memristive arrays by acting as a selector device [5]. Secondly, the transistor limits the current through the memristive devices by setting the current compliance with a specific gate voltage bias [6], thus preventing the hard breakdown of the devices. Finally, the integration of memristive devices with CMOS platform reduces undesired parasitics [7].

2 General Technological Aspects of Integrated Memristive Devices

In order to integrate the memristive module into the CMOS baseline technology, the below aspects need to be considered from the technology point of view:

- CMOS baseline technology node.
- Metal level selection in back-end-of-line (BEOL) interconnects.
- CMOS compatible memristive switching layer.
- Optimization of memristor module fabrication process steps.

2.1 CMOS Baseline Technology Node

The first step towards integrating the memristive devices with CMOS transistors is to choose a compatible CMOS baseline technology. The smaller the technology nodes, the lower the power consumption and the faster the resistive switching [8]. The size of the memristive devices integrated in the BEOL process can be scaled down to the size of $10 \times 10 \text{ nm}^2$ [9]. However, scaling down the lateral dimension of the CMOS transistors is a challenge, due to the voltage levels required to perform the forming and reset operation [10, 11]. It is in the best interest of the memristive device applications to scale down the transistor size while maintaining a memory window of at least 10 [9]. In this work, the small transistors of gate length 130 nm and gate width 150 nm from the 130 nm CMOS baseline technology of IHP is chosen for memristive module integration. Figure 1a shows the transmission electron microscopy (TEM) cross section with energy dispersive X-ray (EDX) analysis of the 1T-1R integrated test structure.

In general, the memristive module consists of a metal-insulator-metal (MIM) stack as shown in Fig. 1a. The memristor module is integrated by using only one additional MIM mask as shown in the Fig. 1b. The dimensions of the via which comes on top of the memristive device in the BEOL interconnects, governs its dimensions. In general, the memristive device size has to be larger than the size of the via as shown in Fig. 1b. The contact to the top electrode of the memristive device is made through this via. Additionally, a predefined layout design rule ‘enclosure’ margin of memristive layer (MIM layer) with respect to the metal layer beneath (Enc.1 in Fig. 1b) and the via layer above (Enc.2 in Fig. 1b) has to be maintained. This enclosure margins are aimed to prevent possible dry etch related damages, obtain reliable contacts between the metal layers and to achieve reproducible process parameters over the entire wafer. The smallest possible size a memristor could take is the size of the via. In order to

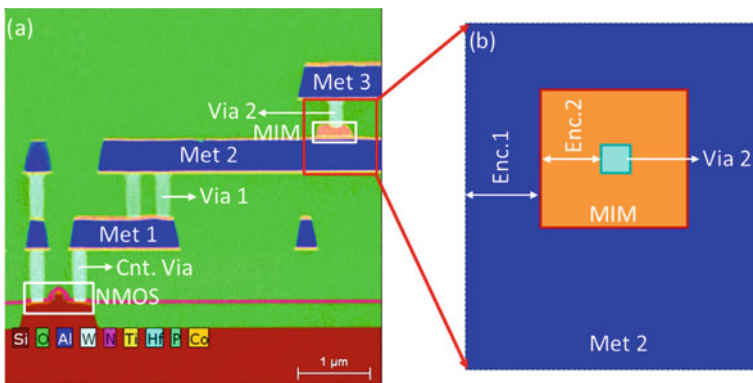


Fig. 1 **a** TEM cross section with EDX based chemical composition maps of memristive module integrated into a 130 nm CMOS technology of IHP and **b** topview layout design rules for the memristive module

reduce the size of the memristive device, intensive process developments are going on to fabricate the device on top of the via [12]. However, this approach of fabrication is out of the scope of this work.

2.2 Metal Level Selection in BEOL Interconnects

In order to have a better control over the performance of the memristive devices, it is important to fabricate them close to the transistor. However, it should not be placed so close that it affects the performance of the transistor. The memristor is a two terminal device and, in principle, the module could be placed between any two metal layers in the BEOL interconnects. We avoid to place the cell on Metal1 in order to maintain the normal functionality of the CMOS transistors and to reduce the undesired parasitic and substrate capacitance [13, 14]. The selection of the BEOL metal level for the integration of the memristive devices depends on the technology used for fabrication, field of application and the material used for the memristive switching layers [15–18]. Additionally, the desired thermal budget is one of the reasons behind choosing the specific metal level for the memristor module integration [19]. Further, the dimensions of the memristor devices and the allowed size of the via on top of the device according to the defined design rules in CMOS baseline technology also plays a role.

The BEOL thermal budget mainly consists of the thermal energy transferred to the wafer from the fabrication process steps including metallization, via-interconnect and inter layer dielectric (ILD) depositions. Additionally, the thermal budget consists of a forming gas passivation step in a diffusion anneal furnace towards the end of line (EOL) fabrication process at BEOL compatible temperature. The passivation step neutralizes the dangling bonds present at the interfaces and improves the contact between silicon and metal and thereby improving the electrical properties of the fabricated devices [20]. The thermal energy encountered by the memristor module is chosen carefully through the selection of the metal layer on which the module is fabricated [12]. The BEOL thermal budget steps activate the properties of the oxygen scavenging layer (OSL) present in the memristor module and thereby reduces the oxygen content in the memristor switching layer [21].

The thermal budget plays a significant role in selecting the phase of the memristive switching layer. The phase of the memristive layer illustrates the distribution of the oxygen vacancies and in turn the performance of the devices [22]. In polycrystalline layers, the oxygen vacancies are concentrated mainly at the grain boundaries whereas, they are homogeneously distributed in amorphous layers [23]. The grain boundaries present in the polycrystalline memristive films results in higher device-to-device (D2D) variability. Whereas, the amorphous memristive films results in reduced D2D variability [24, 25].

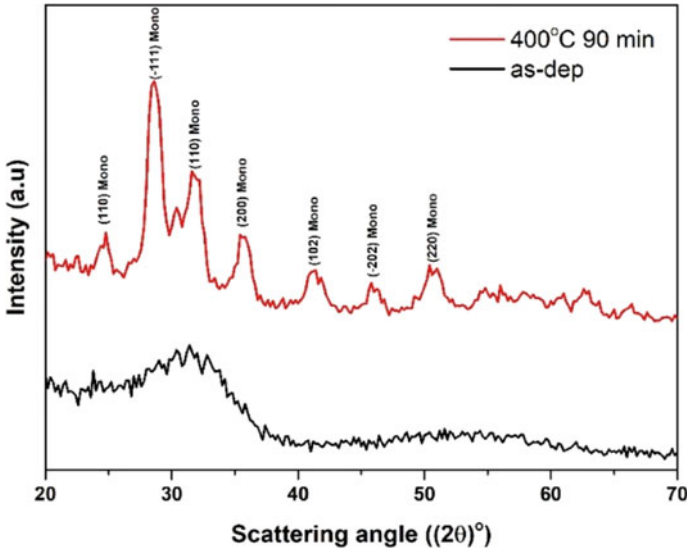
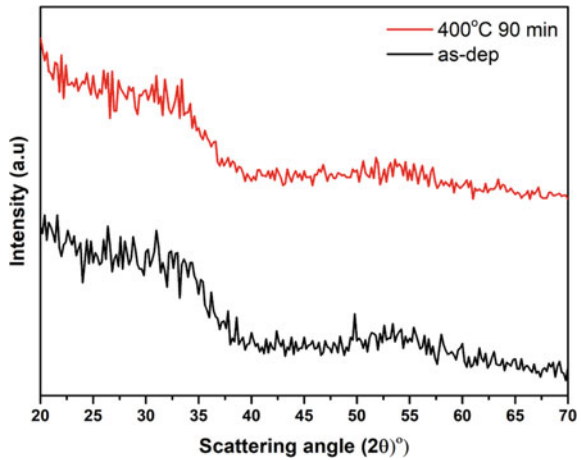


Fig. 2 XRD of as-deposited and sinter annealed HfO₂ layers

Fig. 3 XRD of as-deposited and sinter annealed Al doped HfO₂ layers



It can be clearly seen from the X-ray diffraction (XRD) patterns shown in Figs. 2 and 3 that the as-deposited layers of HfO₂ and Al doped HfO₂ are amorphous in nature. Both the layers were annealed at 400 °C for 90 m which is the BEOL thermal budget in 130 nm CMOS technology of IHP. The HfO₂ layers turns polycrystalline, whereas the Al doped HfO₂ stays amorphous after the annealing process.

2.3 CMOS Compatible Memristive Switching Layer

The successful integration of the memristive module into the CMOS baseline technology requires a CMOS compatible memristive switching layer. Several CMOS compatible memristive layers namely, TiO_2 [26], Al_2O_3 [27], SiO_2 [28], ZrO_2 [29], Ta_2O_5 [30], HfO_2 [31] etc. have been studied in the literature for various applications. Among others, HfO_2 is widely studied and used as a memristive switching layer due to its CMOS compatibility, simple precursor chemistries used for the deposition process, high dielectric constant and well developed deposition processes [32]. Along with the selection of the appropriate top and the bottom electrodes and their areas, the memristor stack configuration decides the type of switching behavior, namely, interface based switching or filamentary based switching [33–35]. The former is out of scope of this work.

Atomic layer deposition (ALD) is one of the prevalent methods used for the deposition of HfO_2 memristive layers. The deposition temperature is compatible with the BEOL thermal budget. The ALD process produces stoichiometric, highly uniform and conformal layers. The soft breakdown in a fully stoichiometric HfO_2 layer for the first time to create a conductive filament (CF) by using an inert metal electrode like TiN is almost impossible [36]. Hence, an active metal electrode like Ti which has a high affinity towards oxygen needs to be deposited on top of the ALD HfO_2 layer. The Ti metal electrode acts as an oxygen scavenging layer (OSL) which scavenges the oxygen from HfO_2 layer during the electro-forming operation and thus enables the creation of a CF [37]. The CF can be disrupted and formed multiple times depending on the direction of the electric field applied. This determines the high resistance state (HRS) and low resistance state (LRS) of the device, respectively.

The HfO_2 memristive layer used in this work is deposited by using thermal atomic layer deposition (TALD) process at 300°C in the Pulsar[®] module from ASM. The layers are grown by using a self-terminating reaction technique. The deposition takes place by alternate pulsing of HfCl_4 and H_2O precursors. Halide based precursors are preferred over the metal organic based precursors in order to reduce the impact of carbon atoms on the performance of the memristive devices [19]. Further, it is possible to incorporate the dopants into HfO_2 layers by replacing the Hf pulse with the dopant pulses at regular intervals. The doping percentage could be altered with the number of dopant pulses. Mostly, HfO_2 memristive layers are doped with Al atoms in order to keep the memristive layers in amorphous state due to the aforementioned benefits.

Finally, HfO_2 and Al_2O_3 layers could also be deposited in a stack without vacuum breakage. Initially, the Al_2O_3 layers of thickness 1 and 2 nm are deposited on the TiN metal layer by using Trimethylaluminum (TMA) and H_2O as precursors at 300°C . The deposition of Al_2O_3 is immediately followed by the HfO_2 deposition without vacuum breakage. The TEM cross section with EDX based chemical composition maps clearly shows the difference between the memristive layers with and without Al_2O_3 layers in Fig. 4. Substantial research is going on in the area of $\text{HfO}_2/\text{Al}_2\text{O}_3$ bi-layer memristive devices. Addition of a thin Al_2O_3 layer below the HfO_2 layer has

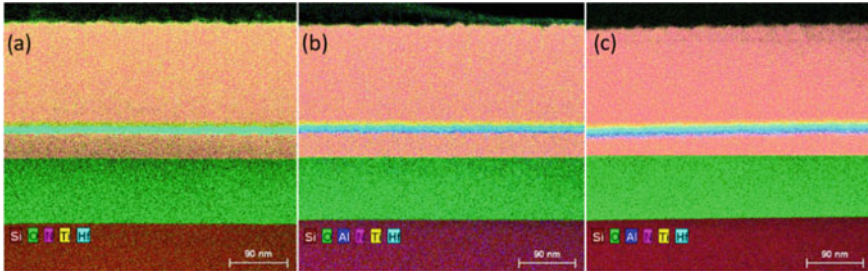


Fig. 4 EDX based chemical compositional maps of the memristor layers **a** HfO_2 **b** $\text{HfO}_2:\text{Al}_2\text{O}_3$ (8 nm: 1 nm) **c** $\text{HfO}_2:\text{Al}_2\text{O}_3$ (8 nm: 2 nm)

demonstrated an improved performance of the memristive devices in terms of HRS reliability [15, 38]. Further, the analog behavior is observed in the bi-layer memristive devices which improved the accuracy of pattern recognition in neuromorphic computing systems [39].

2.4 Optimization of Memristor Module Fabrication Process Steps

In spite of their low power consumption, fast switching, CMOS compatibility and many other potential features, the D2D and cycle-to-cycle (C2C) variability, reliability and performance of the memristive devices still remain a challenge [40]. Many approaches with respect to electrical characterization and fabrication technology, have been studied in the literature to reduce the variability and improve the performance of the devices. One of the ways is to program the memristive devices by using write and verify schemes like the incremental step pulse with verify algorithm (ISPVA) [41]. The small incremental voltage steps with read current operation in between allow the scheme to control the formation/disruption of the CF and hence result in reduced variability and supports the multi-bit operation [5]. Further, in terms of technology, the carbon content in the memristive layers is reduced by using inorganic precursors to obtain the promising results in case of D2D and C2C variabilities [19].

One of the ways to improve the memristive device performance from technology point of view is to optimize the device fabrication technique in the BEOL integration process. In this work, we use the $\text{TiN}/\text{HfO}_2/\text{Ti}/\text{TiN}$ memristive devices fabricated in the SiGe-BiCMOS technology of IHP to illustrate the three different fabrication approaches type 1, 2 and 3 and their impact on the electrical performance of the devices [42]. The memristor device stack consists of 150 nm sputter deposited TiN top electrode (TE) and bottom electrode (BE), 8 nm ALD deposited HfO_2 layer and 7 nm sputter deposited Ti on top of memristive HfO_2 layer. The reactive-ion etching

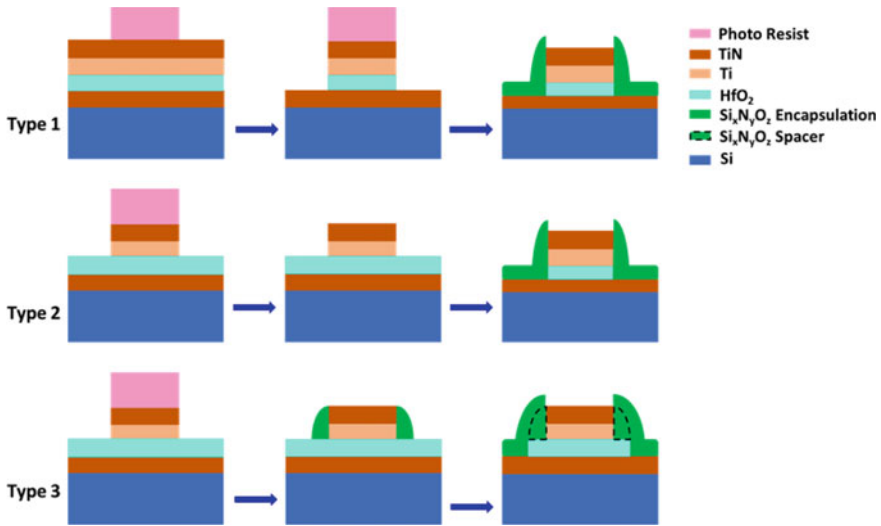


Fig. 5 Process flow of the memristive devices fabricated using three different approaches. Reprinted from [42] (License ID: 1252381-1; <https://marketplace.copyright.com/>)

(RIE) of the stack is one of the crucial steps in the memristive module integration. The RIE etching of the stack is challenging owing to the fact that, the stack consists of the Ti metal layer which could get oxidized during the subsequent process steps. Additionally, the etch residuals from the top electrode if prevail on the sidewalls of the devices degrade the performance of memristive devices [43, 44].

The process flow of the devices fabricated using three different fabrication approaches type 1, 2 and 3 are shown in the Fig. 5. In case of type 1 approach, the Ti/TiN and HfO₂ layers are RIE etched in one single step by using photoresist as the mask. In case of type 2 approach, only the Ti/TiN layers are etched by using photoresist as the mask. The HfO₂ layer is etched by using TiN as the hard mask. Type 3 approach follows similar process flow as type 2 until the Ti/TiN TE etch. However, the Ti/TiN layers in type 3 approach are covered with Si_xN_yO_z spacers on the sidewalls before the HfO₂ memristor layer is etched. Finally, the devices fabricated in all the three approaches are encapsulated by depositing Si_xN_yO_z layers. The RIE etch steps in all the three fabrication approaches are followed by a plasma ashing step and a wet etch step called neutral oxide etch (NOE) which removes the etching residuals and polymer impurities from the wafer surfaces, respectively. Further, in case of type 1 approach, due to the patterning of HfO₂ layer using photoresist, the residuals of the resist prevail even after the ash and NOE clean steps. Hence, an oxygen RIE plasma step is carried out after the memristor layer etch, which removes the residual resist and thereby assists the subsequent ash and NOE clean steps [45].

The performance of the memristor devices is tested initially on the dedicated process control monitor (PCM) structure as shown in Fig. 6a. The PCM structure consists of 2457 memristor devices of size $10 \times 10 \mu\text{m}^2$ connected in parallel.

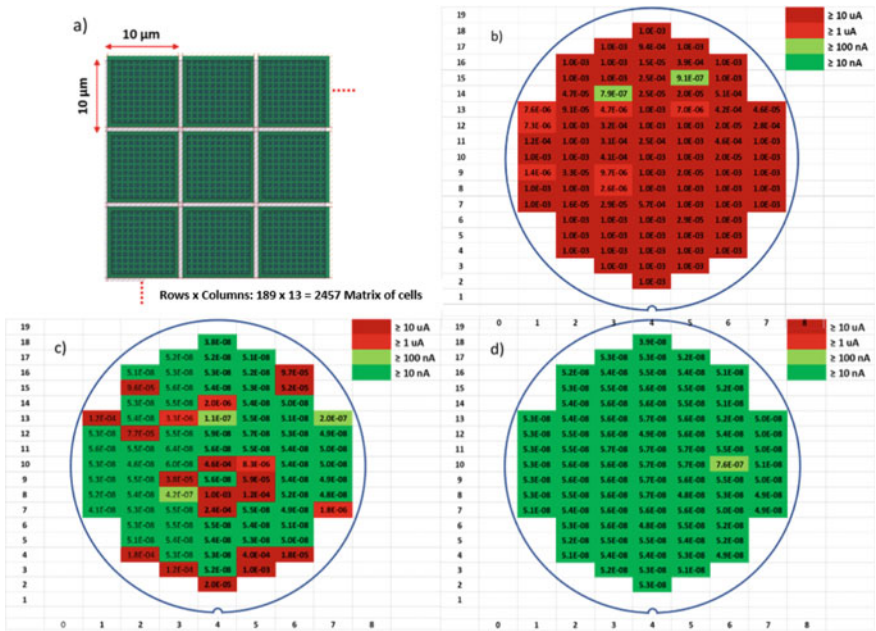


Fig. 6 The pristine state currents measured on **a** a dedicated PCM structure of memristive module fabricated using **b** type 1 **c** type 2 and **d** type 3 approaches. Reprinted from [42] (License ID: 1252381-1; <https://marketplace.copyright.com/>)

The structure demonstrates sensitive device topologies with 9828 corners and 98,280 μm of perimeter. Further, the quality of the memristor devices, edge patterning and their electrical performance could be checked from the in-line tester measurements inside fab. In case of out-of-spec electrical results, the memristor fabrication process steps could be reworked.

The pristine state currents of the memristive devices can be considered as a direct manifestation of their switching performance. The wafer-scale level pristine state currents of the memristive devices fabricated using three different approaches are as shown in the Fig.6. The pristine state currents are measured at a voltage of 0.1 V. It is clearly shown that the devices fabricated in type 3 approach using the spacers and encapsulation techniques exhibit lower pristine state currents compared to the memristive devices fabricated in type 1 and 2 approach [42].

The equivalent circuit of a memristor always consists of a resistor and a capacitor in parallel, to account for the parasitic losses [7]. The parallel resistor in this case is the side walls of the memristor device. In contrast to type 1 and 2 devices, the type 3 devices have their side walls of the TE protected by $\text{Si}_x\text{N}_y\text{O}_z$ spacers before etching the HfO_2 layer. Due to this, the possibility of Ti containing polymers reaching the side walls of the memristive devices is considerably reduced. Hence, the value of parallel resistance is kept high which eventually results in low leakage current values in type 3 devices [42].

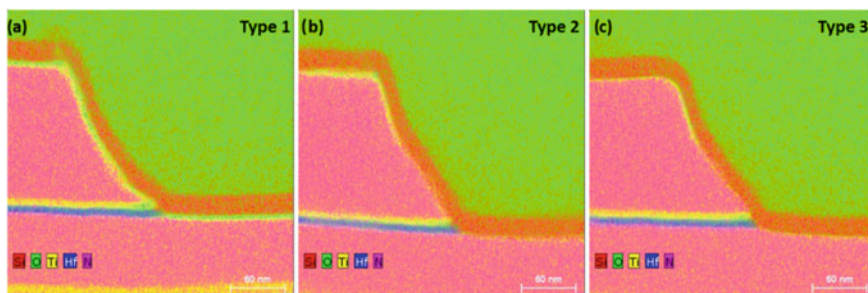
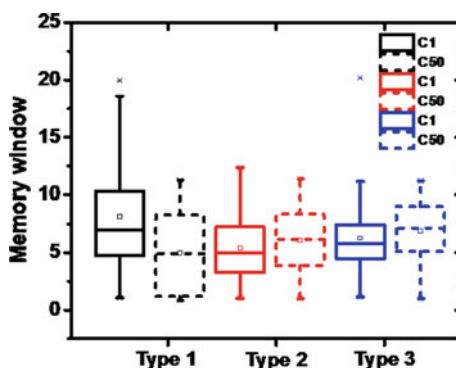


Fig. 7 TEM cross section with EDX analysis of the side walls of the memristive devices fabricated using **a** type 1, **b** type 2 and **c** type 3 approaches. Reprinted from [42] (License ID: 1252381-1; <https://marketplace.copyright.com/>)

Fig. 8 The MW extracted from the 1^{st} t and 50^{th} h set cycles of memristive devices fabricated using three different approaches. Reprinted from [42] (License ID: 1252381-1; <https://marketplace.copyright.com/>)



Further, to analyze the effect of three different fabrication approaches on the elemental composition of the memristive stack, TEM cross section with EDX based chemical composition analysis is performed (Fig. 7). Memristive devices of size $75 \times 75 \mu\text{m}^2$ are chosen for the analysis. In case of type 1 devices, Ti is oxidized at the side walls and there is a formation of $\text{Ti}_x\text{O}_y\text{N}_z$ on the top and bottom electrodes. The oxygen RIE plasma step which assists the removal of residual resist, might possibly have caused this oxidation. The type 2 devices show the formation of $\text{Ti}_x\text{O}_y\text{N}_z$ mainly on the TE, while it is considerably reduced in case of type 3 devices. Differentiating between the spacer and the encapsulation layers is hard due to the usage of the same material and the RIE etching angles. Even though, it is hard to notice the Ti impurities on the side walls of the devices, the electrical measurement results support the claim [42].

The $75 \times 75 \mu\text{m}^2$ memristor devices are further characterized for their resistive switching operation. The type 3 devices exhibited less variability in terms of forming, reset and set voltages. The devices are cycled for 50 cycles of reset and set operations. The memory window (MW) of the devices is extracted from the 50 set cycles. It can be clearly noticed from the Fig. 8 that the MW decreases with cycling in type 1

devices while it increases in type 2 and 3 devices. Further, the type 3 devices show slightly higher MW with less variability compared to type 1 and 2 devices [42].

3 Conclusion

The necessity and importance of memristive device integration into a CMOS baseline technology is discussed. The technological aspects of memristive module integration are presented under four different sections: (1) CMOS baseline technology (2) Metal level selection in BEOL interconnects (3) CMOS compatible memristive switching layer (4) Optimization of memristor module fabrication process steps. Under the first section, the factors governing the size of the memristive devices and, the layout design rules for the memristive module integration are briefly discussed. In the second section, the metal level selection in terms of CMOS performance and thermal budget of the BEOL processes is presented. Additionally, the phase of the memristive switching layers and their influence on the performance of the devices in terms of D2D and C2C variability is presented. In the third section, the ALD and the reason behind its wide usage as a memristive layer deposition process is briefly discussed. In the fourth section, the memristor module is fabricated using three different approaches. The optimized fabrication process steps are studied systematically in terms of TEM cross section with EDX based chemical analysis and correlated with the device performances. The devices fabricated using type 3 approach with spacer and encapsulation technique exhibited lower pristine state currents on a wafer scale analysis. Finally, the extracted MW of type 3 devices exhibited higher values with slightly less variability in comparison with type 1 and type 2 devices.

References

1. Yin, S., et al.: Monolithically integrated RRAM-and CMOS-based in-memory computing optimizations for efficient deep learning. *IEEE Micro*. **39**, 54–63 (2019)
2. Staff, E.: Monolithic embedded RRAM presents challenges, opportunities kernel description. <https://www.embedded.com/monolithic-embedded-rram-presents-challenges-opportunities/> (2021)
3. Zhang, W., et al.: Neuro-inspired computing chips. *Nat. Electron*. **3**, 371–382 (2020)
4. Kim, S.G., Han, J.S., Kim, H., Kim, S.Y., Jang, H.W.: Recent advances in memristive materials for artificial synapses. *Adv. Mater. Technol*. **3**, 1800457 (2018)
5. Mahadevaiah, M.K., Perez, E., Wenger, C.: Influence of specific forming algorithms on the device-to-device variability of memristive Al-doped HfO₂ arrays. *J. Vac. Sci. & Technol. B Nanotechnol. Microelectron. Mater. Process. Meas. Phenom*. **38**, 013201 (2020)
6. Yu, S.: Resistive Random Access Memory (RRAM). In: *Synthesis Lectures on Emerging Engineering Technologies*, vol. 2, pp. 1–79 (2016)
7. Li, H., et al.: A SPICE model of resistive random access memory for large-scale memory array simulation. *IEEE Electron Device Lett*. **35**, 211–213 (2013)

8. Tang, X., Gaillardon, P.-E., De Micheli, G.: A high-performance low-power near-Vt RRAM-based FPGA. In: 2014 International Conference on Field-Programmable Technology (FPT), pp. 207–214 (2014)
9. Govoreanu, B., et al.: $10 \times 10\text{nm}^2$ Hf/HfO_x crossbar resistive RAM with excellent performance, reliability and low-energy operation. In: 2011 International Electron Devices Meeting, pp. 31–6 (2011)
10. Fackenthal, R., et al.: 19.7 A 16Gb ReRAM with 200MB/s write and 1GB/s read in 27nm technology. In: 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), pp. 338–339 (2014)
11. Yu, S., Shim, W., Peng, X., Luo, Y.: RRAM for compute-in-memory: from inference to training. *IEEE Trans. Circuits Syst. I Regul. Pap.* (2021)
12. Lv, H., et al.: BEOL based RRAM with one extra-mask for low cost, highly reliable embedded application in 28 nm node and beyond. In: 2017 IEEE International Electron Devices Meeting (IEDM), pp. 2–4 (2017)
13. O'Connell, B., Thibeault, T., Chaparala, P.: Plasma damage considerations involving metal-insulator-metal (MIM) capacitors. In: 2004 International Conference on Integrated Circuit Design and Technology (IEEE Cat. No. 04EX866), pp. 123–126 (2004)
14. Wang, Z., et al.: Plasma-charging damage of floating MIM capacitors. *IEEE Trans. Electron Devices*. **51**, 1017–1024 (2004)
15. Azzaz, M., et al.: Benefit of Al₂O₃/HfO₂ bilayer for BEOL RRAM integration through 16kb memory cut characterization. In: 2015 45th European Solid State Device Research Conference (ESSDERC), pp. 266–269 (2015)
16. Diokh, T., et al.: Investigation of the impact of the oxide thickness and RESET conditions on disturb in HfO₂-RRAM integrated in a 65nm CMOS technology. In: 2013 IEEE International Reliability Physics Symposium (IRPS), pp. 5E–4 (2013)
17. Portal, J.-M., et al.: Design and simulation of a 128 kb embedded nonvolatile memory based on a hybrid RRAM (HfO₂)/28 nm FDSOI CMOS technology. *IEEE Trans. Nanotechnol.* **16**, 677–686 (2017)
18. Ito, S., et al.: ReRAM technologies for embedded memory and further applications. In: 2018 IEEE International Memory Workshop (IMW), pp. 1–4 (2018)
19. Grossi, A., et al.: Impact of the precursor chemistry and process conditions on the cell-to-cell variability in 1T–1R based HfO₂ RRAM devices. *Sci. Rep.* **8**, 1–11 (2018)
20. Tsai, T.-L., Chang, H.-Y., Jiang, F.-S., Tseng, T.-Y.: Impact of post-oxide deposition annealing on resistive switching in HfO₂-based oxide RRAM and conductive-bridge RAM devices. *IEEE Electron Device Lett.* **36**, 1146–1148 (2015)
21. Walczyk, C., et al.: On the role of Ti adlayers for resistive switching in HfO₂-based metal-insulator-metal structures: top versus bottom electrode integration. *J. Vac. Sci. & Technol. B Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.* **29**, 01AD02 (2011)
22. Milo, V., et al.: Multilevel HfO₂-based RRAM devices for low-power neuro-morphic networks. *APL Mater.* **7**, 081120 (2019)
23. Lanza, M., et al.: Grain boundaries as preferential sites for resistive switching in the HfO₂ resistive random access memory structures. *Appl. Phys. Lett.* **100**, 123508 (2012)
24. Grossi, A., et al.: Electrical characterization and modeling of 1T–1R RRAM arrays with amorphous and poly-crystalline HfO₂. *Solid State Electron.* **128**, 187–193 (2017)
25. Grossi, A., Perez, E., Zambelli, C., Olivo, P., Wenger, C.: Performance and reliability comparison of 1T–1R RRAM arrays with amorphous and polycrystalline HfO₂. In: 2016 Joint International EUROSIOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSIOI-ULIS), pp. 80–83 (2016)
26. Carta, D., et al.: Investigation of the switching mechanism in TiO₂-based RRAM: a two-dimensional EDX approach. *ACS Appl. Mater. & Interfaces.* **8**, 19605–19611 (2016)
27. Sarkar, B., Lee, B., Misra, V.: Understanding the gradual reset in Pt/Al₂O₃/Ni RRAM for synaptic applications. *Semicond. Sci. Technol.* **30**, 105014 (2015)

28. Wang, Z., et al.: Engineering incremental resistive switching in TaO_x based memristors for brain-inspired computing. *Nanoscale*. **8**, 14015–14022 (2016)
29. Wang, S.-Y., Lee, D.-Y., Huang, T.-Y., Wu, J.-W., Tseng, T.-Y.: Con-trollable oxygen vacancies to enhance resistive switching performance in a ZrO₂-based RRAM with embedded Mo layer. *Nanotechnol.* **21**, 495201 (2010)
30. Lee, T.S., et al.: Compliance current-controlled conducting filament formation in tantalum oxide-based RRAM devices with different top electrodes. *ACS Appl. Electron. Mater.* **2**, 1154–1161 (2020)
31. Mahadevaiah, M., et al.: Reliability of cmos integrated memristive HfO₂ arrays with respect to neuromorphic computing. In: 2019 IEEE International Reliability Physics Symposium (IRPS), pp. 1–4 (2019)
32. Long, S., et al.: Cycle-to-cycle intrinsic RESET statistics in HfO₂-based unipolar RRAM devices. *IEEE Electron Device Lett.* **34**, 623–625 (2013)
33. Baeumer, C., et al.: Spectroscopic elucidation of ionic motion processes in tunnel oxide-based memristive devices. *Faraday Discuss.* **213**, 215–230 (2019)
34. Sassine, G., et al.: Interfacial versus filamentary resistive switching in TiO₂ and HfO₂ devices. *J. Vac. Sci. & Technol. B Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.* **34**, 012202 (2016)
35. Peng, H., et al.: Electrode dependence of resistive switching in Mn-doped ZnO: filamentary versus interfacial mechanisms. *Appl. Phys. Lett.* **96**, 192113 (2010)
36. Niu, G., et al.: Geometric conductive filament confinement by nanotips for resistive switching of HfO₂-RRAM devices with high performance. *Sci. Rep.* **6**, 1–9 (2016)
37. Niu, G., et al.: Material insights of HfO₂-based integrated 1-transistor-1-resistor resistive random access memory devices processed by batch atomic layer deposition. *Sci. Rep.* **6**, 1–11 (2016)
38. Kalishettyhalli Mahadevaiah, M., et al.: Modulating the filamentary-based resistive switching properties of HfO₂ memristive devices by adding Al₂O₃ layers. *Electron.* **11**, 1540 (2022)
39. Woo, J., et al.: Improved synaptic behavior under identical pulses using AlO_x/HfO₂ bilayer RRAM array for neuromorphic systems. *IEEE Electron Device Lett.* **37**, 994–997 (2016)
40. Wong, H.-S.P., et al.: Metal-oxide RRAM. *Proc. IEEE*. **100**, 1951–1970 (2012)
41. Pérez, E., et al.: Reduction of the cell-to-cell variability in Hf_{1-x}Al_xO_y based RRAM arrays by using program algorithms. *IEEE Electron Device Lett.* **38**, 175–178 (2016)
42. Mahadevaiah, M.K., et al.: Optimized HfO₂-based MIM module fabrication for emerging memory applications. *ECS Trans.* **92**, 211 (2019)
43. Gu, P.-Y., et al.: Scalability with silicon nitride encapsulation layer for Ti/HfO_x pillar RRAM. In: Proceedings of 2010 International Symposium on VLSI Technology, System and Application, pp. 146–147 (2010)
44. Jousseume, V., et al.: Back-end-of-line integration approaches for resistive memories. In: 2009 IEEE International Interconnect Technology Conference, pp. 41–43 (2009)
45. Shamiryan, D., Baklanov, M., Vanhaelemeersch, S., Maex, K.: Comparative study of SiOCH low-k films with varied porosity interacting with etching and cleaning plasma. *J. Vac. Sci. & Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.* **20**, 1923–1928 (2002)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Wave Digital Approach Towards Bio-inspired Computing Using Memristive Networks



Dennis Michaelis and Karlheinz Ochs

Abstract Bio-inspired computing is a promising approach to tackle problems that are too complex for state-of-the-art computing approaches. It is a highly interdisciplinary field of research as it requires expertise, which ranges from understanding biological organisms (psychology, physiology zoology, biology, bio-chemistry) to building electrical circuits (electrical engineering, physics). This chapter focuses on theoretical circuits designs and concepts to abstractly map biological behavior to idealized circuits. We utilize a special digital emulation technique as a tool to digitally replicate circuits and hence bridge the gap from biological behavioral observation to hardware circuit designs. Application examples include parameter optimization of a neuronal oscillator, gait pattern generation, multi-neuron communication, neuroplasticity and optical illusions. The fundamental circuit elemental to realize several bio-inspired circuit designs is the memristor, which is essentially a resistor with a memory. Since memristors are hard to fabricate with the current state of technology, digital emulators are a useful tool to accelerate development cycles and investigate the circuits of the next generation.

Keywords Memristor · Electrical circuits · Emulation · Neuromorphic engineering · Self-organization

1 Introduction

It is worth noting that a digital emulator is different compared to a hardware emulator and naturally comes with some benefits, although it is generally harder to construct. Hardware emulators use electrical devices to emulate the behavior of a desired electrical system. Due to the intrinsic computational parallelism in a signal processing sense of voltage and current in an electrical circuit, such an emulation

D. Michaelis (✉) · K. Ochs
Ruhr-University Bochum, Bochum, Germany
e-mail: dennis.michaelis@rub.de

K. Ochs
e-mail: karlheinz.ochs@rub.de

© The Author(s) 2024
M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_8

circuit is a powerful instrument. However, such emulators are often hardwired and a change of parameters in the system of interest requires new designing of the electrical hardware [1, 2]. This makes hardware emulators somewhat inflexible, as a physical exchange of electrical devices in the emulation circuit is required, sometimes the circuit even needs to be partially or completely redesigned. Software emulation on the contrary can change parameter values digitally. This enables not only uncomplicated and time-saving parameter changes, but also in-operando parameter manipulation, meaning that parameters can even be changed during run-time. Consequently, procedures such as parameter optimization and sensitivity analyses are possible, as demonstrated in [3]. It should however be mentioned that a digital emulation comes at a cost. First, appropriate interfaces with the help of analogue-to-digital (A/D) and digital-to-analogue converters (D/A) need to be implemented [4]. Second, digital emulators are typically more sophisticated to design compared to hardware emulators because the stability of the overall system must not be altered by the digital emulator which runs under finite precision arithmetic.

2 Digital Emulation Technique

The wave digital concept [5] takes into account all of the above and is here exploited as a digital emulation technique as it is known to preserve energetic properties such as passivity in a digital signal processing sense [6, 7]. It leads to a computationally massive parallel algorithm, which can be implemented on DSPs or FPGAs and ASICs for real-time applications in integrated circuits [8, 9]. A short overview of the wave digital concept is given below.

Voltage u and current i of Fig. 1 are related to incident wave a and reflected wave b via an arbitrary positive constant R that is called the port resistance and the bijective transformation

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & R \\ 1 & -R \end{bmatrix} \begin{bmatrix} u \\ i \end{bmatrix}, \quad R > 0. \quad (1)$$

Consequently, to obtain the wave digital model based on a reference circuit, one has to translate all electrical devices port-wise and their Kirchhoff interconnection

Fig. 1 Definition of a port with electrical quantities u , i (left) and wave quantities a , b with port resistance R (right)

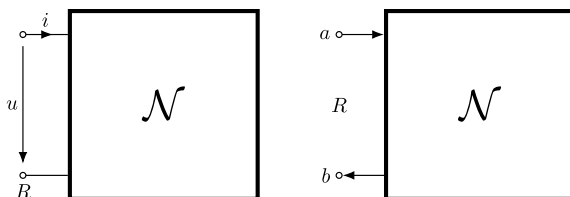


Table 1 Wave digital sources

Element	Electrical circuit	Wave flow graph
Voltage source		
Current source		

network separately in the wave digital domain and then connect them port-wise again subsequently. For example, resistive voltage and current sources described by

$$u = e - R_0i \quad \text{and} \quad i = j - G_0u, \tag{2}$$

with source resistance R_0 or source conductance G_0 and inputs e, j , respectively. Using the bijective transformation (1), their wave digital flow graphs are described by

$$a = [1 - \varrho]e + \varrho b \quad \text{and} \quad a = [1 + \varrho]Rj + \varrho b, \tag{3}$$

respectively, with $\varrho = [R_0 - R]/[R_0 + R]$, cf. Table 1. In the case of ideal sources, the inner resistances vanish and hence

$$R_0 = 0 \Leftrightarrow \varrho = -1 \quad \text{and} \quad G_0 = 0 \Leftrightarrow \varrho = 1, \tag{4}$$

resulting in the simplified expression

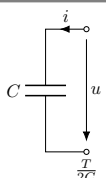
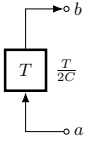
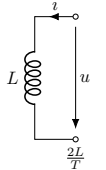
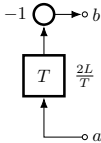
$$a = 2e - b \quad \text{and} \quad a = 2Rj + b. \tag{5}$$

Other selected nonreactive elements include the resistor, which is characterized by

$$u = Zi \quad \Leftrightarrow \quad b = \varrho a, \tag{6}$$

with $\varrho = [Z - R]/[Z + R]$. After discretization and deploying the trapezoidal integration rule due to its beneficial features regarding passivity and stability [6], the

Table 2 Selected wave digital reactive elements

Element	Electrical circuit	Wave flow graph
Capacitor		
Inductor		

capacitor and the inductor become simple delay elements in the wave digital domain. By defining the step size $T = t_k - t_{k-1}$, one obtains

$$b(t_k) = a(t_{k-1}) \quad \text{and} \quad b(t_k) = -a(t_{k-1}), \tag{7}$$

for a capacitor and inductor, respectively, after their port resistances have been chosen to $R = T/[2C]$ and $R = [2L]/T$, respectively. Their corresponding wave digital flow graphs are shown in Table 2.

Now that selected electrical devices and their corresponding wave flow graphs have been presented, their port-wise connection is discussed. Since all wave quantities are always related to their respective port resistance, only elements with the same port resistances can be connected immediately. While certain elements allow for an arbitrary (but positive) choice of the port resistance, such as the resistor, others do not offer this degree of freedom, such as the capacitor or inductor. To still be able to connect elements with different port resistances, parallel and series adaptors representing Kirchhoff parallel and series interconnections are utilized. The underlying equations describing a Kirchhoff series interconnection are

$$\mathbb{1}^T \mathbf{u} = 0 \quad \text{and} \quad \mathbf{i} = \mathbb{1}i_0, \tag{8}$$

with $\mathbb{1} = [1 \dots 1]^T$ being the all one vector of appropriate dimension. When the electrical quantities are replaced by wave quantities, cf. (1), the relation

$$\mathbf{b} = [\mathbf{1} - \boldsymbol{\gamma} \mathbb{1}^T] \quad \text{with} \quad \boldsymbol{\gamma} = \frac{2\mathbf{R}\mathbb{1}}{\mathbb{1}^T \mathbf{R} \mathbb{1}} \tag{9}$$

Table 3 Wave digital adaptors

Element	Electrical circuit	Wave flow graph
Four-port parallel connection		
Four-port series connection		

emerges, where $\mathbf{1}$ is the identity matrix of suitable dimensions and $\mathbf{R} = \text{diag}_{\mu=1}^n R_{\mu}$, with R_{μ} being the port resistance of the μ -th port. Likewise, a Kirchoff parallel interconnection is described by

$$\mathbb{1}^T \mathbf{i} = 0 \quad \text{and} \quad \mathbf{u} = \mathbb{1} u_0 \tag{10}$$

in the electrical domain, which yields the wave relationship

$$\mathbf{b} = [\mathbb{1} \boldsymbol{\gamma}^T - \mathbf{1}] \quad \text{with} \quad \boldsymbol{\gamma} = \frac{2\mathbb{1}^T \mathbf{G}}{\mathbb{1}^T \mathbf{G} \mathbb{1}}, \tag{11}$$

with $\mathbf{G} = \mathbf{R}^{-1}$. The respective wave digital symbols of Kirchoff series and parallel adaptors are shown in Table 3.

3 Memristive Neuronal Oscillator

The well-known Morris-Lecar model is a second-order model describing a neuron's membrane potential behavior [10]. The structure of its electrical circuit is shown in Fig. 2 (top left) and the underlying equations are written as

$$C \frac{du}{dt} = i_{\text{App}} - g_{\text{Ca}}(u)[u - E_{\text{Ca}}] - zG_{\text{K1}}[u - E_{\text{K}}] - G_{\text{L}}[u - E_{\text{L}}], \quad (12)$$

with the nonlinear resistor and memristor described by

$$g_{\text{Ca}}(u) = G_{\text{Ca1}} \sigma \left(\frac{u - U_{\text{Ca1}}}{U_{\text{Ca2}}} \right), \quad (13)$$

$$\frac{dz}{dt} = [z_{\infty}(u) - z]F_{\text{K}} \cosh \left(\frac{u - U_{\text{K1}}}{2U_{\text{K2}}} \right), \quad \text{with } z_{\infty}(u) = \sigma \left(\frac{u - U_{\text{K1}}}{U_{\text{K2}}} \right), \quad (14)$$

respectively, where $\sigma(\cdot)$ is sigmoid function, the capacitor models the changing membrane potential, E_{L} and G_{L} model the leakage current behavior. E_{Ca} is the voltage offset to achieve biologically meaningful calcium (Ca) concentration values and the nonlinear resistor $g_{\text{Ca}}(u)$ described by (13) models the opening and closing behavior of the Ca-channels. Likewise, voltage offset E_{K} aims at accomplishing biologically meaningful potassium (K) concentration values and the memristor $W_{\text{K}}(z, u)$ described by (14) models the opening and closing behavior of K-channels [11]. The nonlinear resistor here models the opening and closing behavior of the Calcium channels, whereas the memristor accounts for opening and closing behavior of the Potassium channels.

A fundamentally important phase of an action potential is the hyperpolarization phase, in which the membrane potential temporarily falls below the resting potential. This is relevant for several learning mechanisms, e.g. spike-timing dependent plasticity [12]. The problem with the parameter set given in [11] is that for several values of the applied current i_{App} the hyperpolarization phase is missing. In order to possibly find a parameter set, where the hyperpolarization phase is present, a parameter optimization on the basis of a wave digital emulation is conducted [3]. The wave digital model of the circuit is shown in Fig. 2 (bottom right) and is obtained from the electrical circuit by a port-wise translation of its electrical components and Kirchhoff interconnection structures. For this reason, the structural similarities between the electrical circuit and the wave digital flow graph are immediately observable. There, implicit relationships between certain wave quantities are highlighted in orange. To deal with these, fixpoint iterations are utilized [13]. Because the wave digital model is real-time capable, *in-operando* parameter manipulations enable a parameter optimization during run-time.

For the emulation results, an ideal current source is placed at the left-hand port to mimic the applied current i_{App} and the resulting membrane potential is observed at the right-hand port. The action potential emergence as a reaction to the constant input $i_{\text{App}} = 69 \mu\text{A}$ with the state-of-the-art parameters of [11] as the reference and in comparison with the in [3] optimized parameters is given in Fig. 3.

The following is gathered. The membrane potential of the reference parameter set never fall below its resting potential $\Phi_{\text{ref}} = -60.86 \text{ mV}$. The highlights that indeed the important hyperpolarization phase is missing. On the contrary, the optimized parameter cause the membrane potential to fall deeper while the resting potential

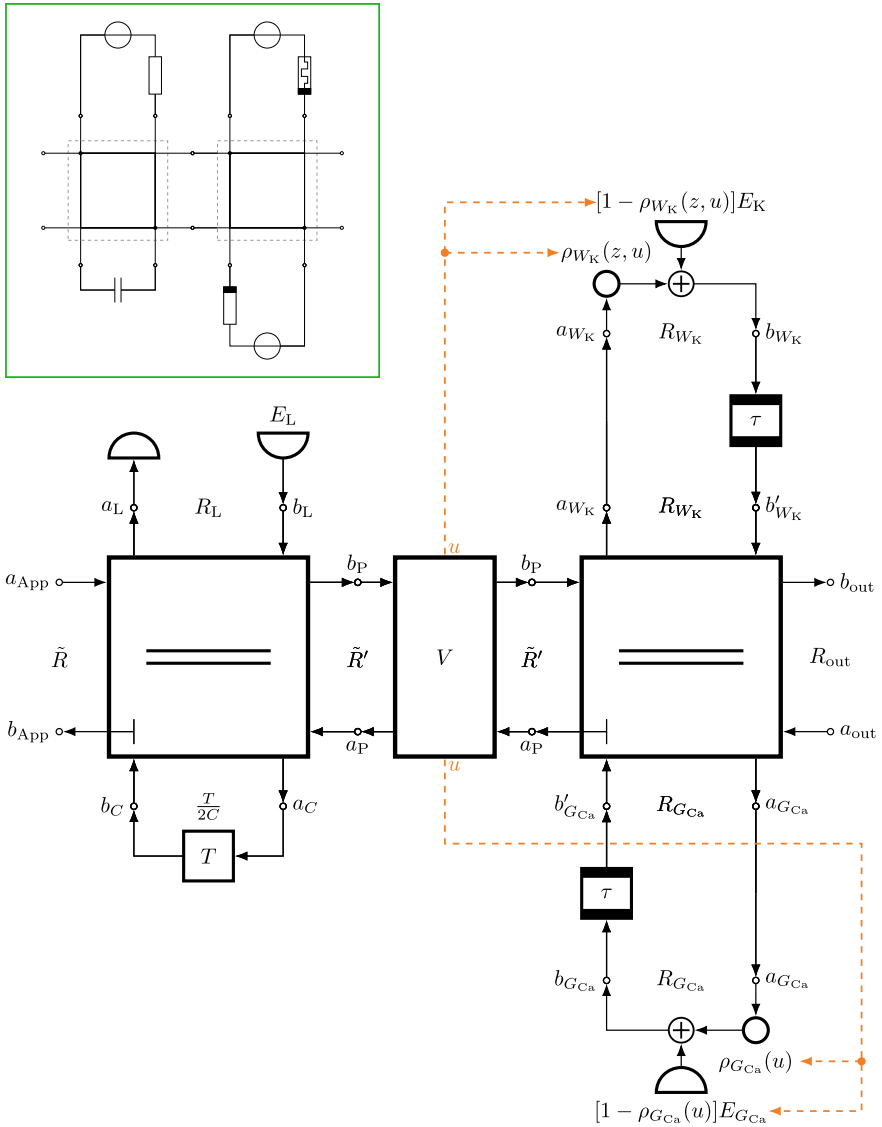


Fig. 2 Wave digital flow graph of the Morris-Lecar neuron model. Its structural, port-wise correspondence to the electrical circuit (top left) becomes apparent

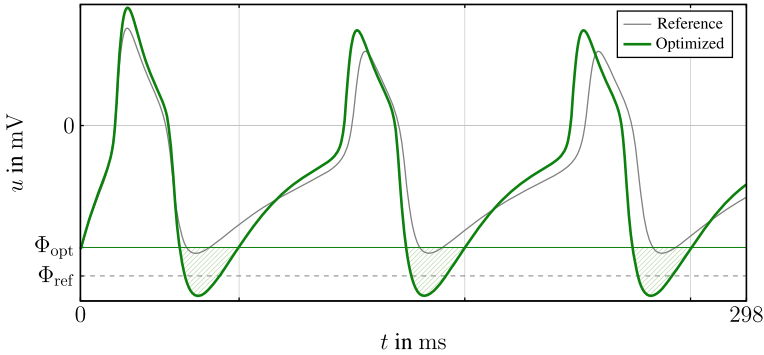


Fig. 3 Constant current of the Morris-Lecar model with optimized parameters from [3] in comparison to parameters as in [11]

$\Phi_{\text{opt}} = -49.35$ mV is increased. This results in a clear hyperpolarization phase, indicated by the green areas in Fig. 3. In summary, a sophisticated *in-operando* parameter optimization has been conducted via the wave digital model, which resulted in a biologically more meaningful behavior.

4 Stimulus-Driven Topology Formation

4.1 Supervised Topology Formation

Memristors are a suitable electrical device for supervised topology formation and first preinvestigations in this regard have been done in [14]. For an extension of this concept, descriptive sketches of neurons and their coupling strengths like in Fig. 4(i) are considered. They are used especially in biology to visualize neuronal setups [15]. In this context, the nodes symbolize the neurons and the arrows represent the (unidirectional) synapses. The numbers indicate the couplings strength between 0 and 1, where the + and – sign stands for excitatory and inhibitory couplings, respectively. The synthesis of an electrical circuit based on these descriptions is challenging for multiple reasons. The probably biggest questions arise when aiming to realize the unidirectional synaptic coupling, as effects like propagation of voltages and currents in an electrical circuit are generally not unidirectional and hence are not free from feedback.

In the following, it is explained how the wave digital model of Fig. 4(ii) achieves these tasks. To do so, the wave digital concept is here exploited as a modeling tool, since this process is more intuitive in the wave domain rather than the electrical domain as it will become apparent subsequently [16]. Lastly, the appropriate electrical circuit is obtained due to its direct correspondence with the final wave digital model. The model is subsequently verified in the context of a dog’s gait patterns.

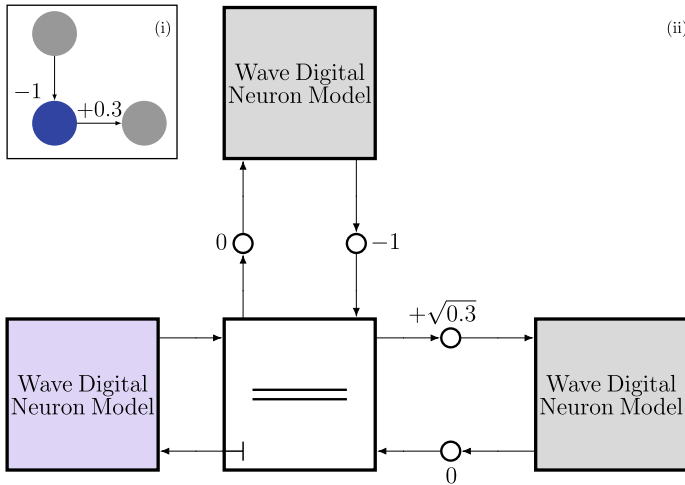


Fig. 4 (i): Conceptual neuronal network with coupling strengths and inhibitory/excitatory coupling represented by negative/positive signs. (ii): Synthesized wave flow diagram

Figure 5 shows the proposed synapse model to synthesize the scenario of Fig. 4(i). It consists of a four port circulator with two attached resistances R_1 and R_2 that account for the coupling strength and coupling type (excitatory for positive reflection coefficients or inhibitory for negative coefficients). To explain its functioning, the reflection coefficients are expressed by

$$\varrho_\mu = \frac{R_\mu - R}{R_\mu + R}, \quad \mu \in \{1, 2\}. \tag{15}$$

The special cases of these reflection coefficients are of interest here in this context. Specifically it holds that

$$\begin{aligned} R_\mu \rightarrow \infty &\Leftrightarrow \varrho_\mu = 1, \\ R_\mu \rightarrow 0 &\Leftrightarrow \varrho_\mu = -1, \\ R_\mu = R &\Leftrightarrow \varrho_\mu = 0. \end{aligned}$$

It can be seen that the whole spectrum $\varrho_\mu \in [-1, 1]$ is experienced and that the sign of the reflection coefficients is determined by whether the resistance value R_μ is smaller or bigger than the circulation resistance R . Of special interest is the case, where $R_\mu = R$, as this results in $\varrho_\mu = 0$. This case is special, because it allows for unidirectional coupling. This can best be seen when the power flow is calculated to

$$\begin{bmatrix} p_{\text{pre,in}} \\ p_{\text{post,in}} \end{bmatrix} = \begin{bmatrix} 0 & \varrho_2^2 \\ \varrho_1^2 & 0 \end{bmatrix} \begin{bmatrix} p_{\text{pre,out}} \\ p_{\text{post,out}} \end{bmatrix}, \tag{16}$$

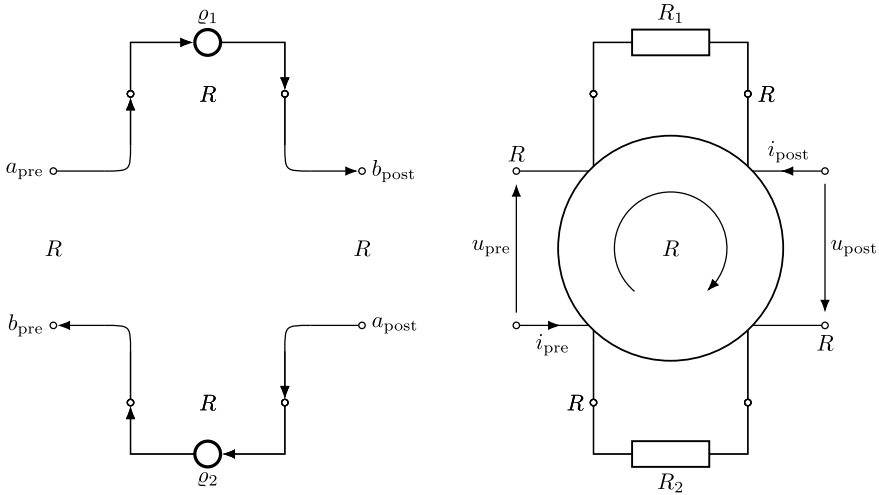


Fig. 5 Left: Proposed synapse model designed as a wave digital model. Right: Its corresponding electrical circuit

where

$$p_{pre,in} = \frac{b_{pre}^2}{4R}, \quad p_{pre,out} = \frac{a_{pre}^2}{4R}, \quad p_{post,in} = \frac{b_{post}^2}{4R}, \quad p_{post,out} = \frac{a_{post}^2}{4R}$$

and the total power consumption of the pre- and postsynaptic ports

$$p_{pre} = p_{pre,in} - p_{pre,out}, \quad p_{post} = p_{post,in} - p_{post,out}.$$

It becomes evident that $\varrho_1 = 0$ prohibits any power to flow from pre- towards post-synaptic neuron and vice versa for $\varrho_2 = 0$. Consequently, this enables unidirectional coupling if either one of the reflection coefficients is zero or complete decoupling if both are zero. With these insights it is possible to synthesize an electrical circuit corresponding to the scenario of Fig. 4(i) as shown in Fig. 4(ii). It becomes evident that the nature of the scattering parameter depiction is much more similar to the conceptual graph compared to a depiction in electrical quantities. Therefore, depending on the context, the wave digital concept can not only be a powerful emulation technique but is also a versatile modeling tool by enabling perspectives that are quite unintuitive in the electrical domain, but are easily accessible in the wave domain.

To verify the proper functioning of the concept presented above, a simple application example in the context of a dog’s gait patterns is examined. More specifically, the gait pattern running, pacing and trotting are investigated. To analyze these patterns, two subgroups \mathcal{S}_1 and \mathcal{S}_2 with the following features are created

1. Every component within a subgroup is synchronized to every other component within the same subgroup (intra-group synchronization),

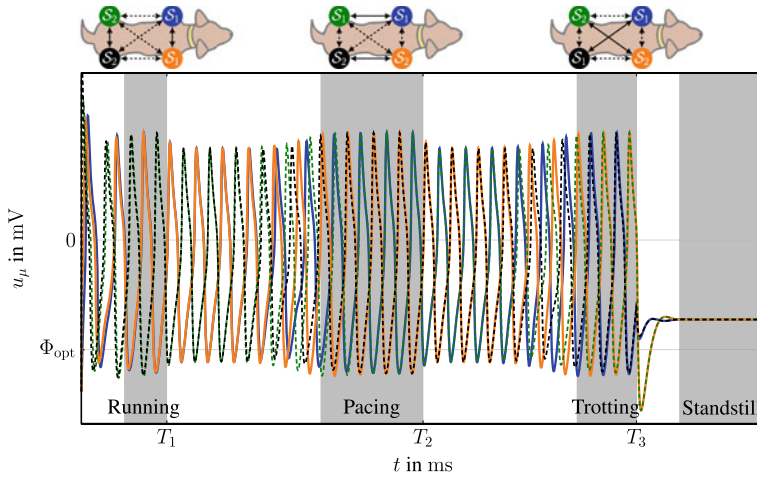


Fig. 6 Output signals of four interconnected, optimized Morris-Lecar neuronal oscillators mimicking the gait of a dog changing from running ($t \leq T_1$) to pacing ($T_1 < t \leq T_2$) to trotting ($T_2 < t \leq T_3$) to standstill ($t > T_3$). Each color represents the motion of a dog’s leg as indicated at the top

2. Every component of subgroup \mathcal{S}_1 is asynchronous to every component of subgroup \mathcal{S}_2 (inter-group asynchronization).

By these categorizations, the three above mentioned gait patterns can be established by assigning the four legs of a dog to the subgroups $\mathcal{S}_1, \mathcal{S}_2$.

For the emulation scenario, four optimized wave digital neuronal oscillators, cf. Sects. 3 and [3], are associated with a leg of a dog each as indicated by the top sketches of Fig. 6. They represent a neuron population that accounts for the movement of a leg. First results in this context have been published in [16]. The neuronal oscillator model of choice is the optimized Morris-Lecar neuron model of Fig. 2 with its parameters as mentioned in [3]. The sketches on top of Fig. 6 indicate intra-group synchronization (solid arrows) and intra-group asynchronization (dashed arrows). The intra-group synchronization is here realized by a bidirectional coupling with edge weights $+1$ while intra-group asynchronization is achieved by bidirectional couplings with edge weights -1 . Although arbitrary but different initial conditions are chosen for the four neuron’s membrane potentials, the desired pattern is achieved after just two oscillations and indeed represents the gait pattern of a running dog, cf. Fig. 6. At $t = T_1$, the edge weights are abruptly changed to the pacing gait pattern. Note that although the topology changes are abrupt, the transition behavior is actually smooth and steady. In fact, almost no changes with respect to the phases are visible during the first four oscillations after the switch, only the amplitude is temporarily reduced. Note that in the transition phase the underlying synaptic interconnection of the neurons does not match their relative behavior. For this reason, they interact destructively until this transition phase is completed. This is the case after approximately 10 oscillations

after $t = T_1 = 250$ ms. A similar behavior is observable when the synaptic coupling weights are changed once again to mimic a trotting gait pattern at $t = T_2 = 1000$ ms. During the transition phase, where the constellation of neuronal activity and underlying synaptic interconnection is inconsistent, the amplitudes are damped. Then, after around 6 oscillations the relative positions of the membrane potentials start to change until completion after around 14 oscillations. Note that here the change in synaptic coupling weights are again abrupt, but the transition is still smooth. This is highly desirable because this translates to the dog being able to change gaits without tumbling. At $t = T_3 = 1625$ ms, the current exciting the Morris-Lecar neurons is switched off and consequently a standstill is achieved after just one oscillation but in a smooth fashion. To summarize, a slow down process from a running to a still standing dog was successfully emulated, where its multiple gait patterns were mimicked in the process. An important detail worth mentioning is the fact that in this application example the hyperpolarization phase is present at all times, which is observable by the action potential minima always surpassing the resting potential $\Phi_{\text{opt}} = -49.35$ mV. This once again highlights the improved Morris-Lecar neuron model [3].

It is also apparent that the here presented gait pattern generator is closely related to the synchronization of neuronal activity to form a biologically meaningful behavior. More general investigations on synchronization of linear(ized) circuits have been conducted in [17, 18] and also in the context of an Ising machine [19]. Furthermore, the first results in this context have sparked more sophisticated results incorporating axon growth in the case of gait pattern generation [20] and gait pattern classification [21].

4.2 *Self-organized Topology Formation*

4.2.1 **Optimal Two-Neuron Communications**

It is desirable to investigate how information from an origin neuron can be transferred to a destination neuron optimally, i.e. with the least amount of delay. This problem can be formulated by graph theory, where nodes represent neurons and edges mimic synapses. Doing so, the task is comparable to solving a maze, where the entry of the maze accounts for the origin neuron and the exit accounts for the destination neuron. The here exploited analog parallelism is different from computational parallelism and enables the solution to the problem to outperform all known graph-theoretical algorithms [22]. The structure of the maze represents the structure of the neural network, where it is assumed that only nearest neighbor synaptic couplings exist and that an obstacle, e.g. a wall, means that no synaptic coupling exists between these specific neighboring neurons. First preinvestigations have been done in [23]. The concept can also be used to determine the longest rather than the shortest path, which is also a computationally complex problem [24].

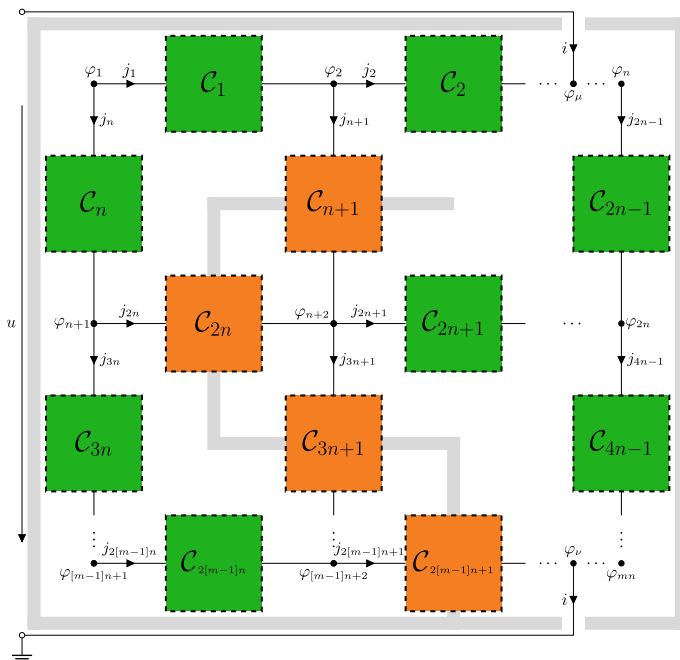


Fig. 7 Setup of the maze to mimic two-neuron communication, which consists of nodes φ_μ and their interconnections C_μ

Stacking all neuron potentials $\varphi_{\mu\nu}$ and all synaptic currents j_κ in vectors $\boldsymbol{\varphi}$ and \boldsymbol{j} , respectively, allows for a condensed notation of the maze topology, cf. Fig. 7. The synapse voltages \boldsymbol{v} are obtained by the respective neuron potential differences and can be expressed as

$$\boldsymbol{v} = \boldsymbol{N}^T \boldsymbol{\varphi}(u), \tag{17}$$

where \boldsymbol{N} is the incidence matrix and voltage u is the potential difference between origin neuron potential $\varphi_\mu = u$ and destination neuron potential $\varphi_\nu = 0$. Consequently, by Tellegen's theorem, the relationship between synaptic currents \boldsymbol{j} and neuronal currents \boldsymbol{i} is

$$\boldsymbol{i} = \boldsymbol{N}\boldsymbol{j}, \tag{18}$$

where $\boldsymbol{i} = -\boldsymbol{e}_\mu^T \boldsymbol{i} = \boldsymbol{e}_\nu^T \boldsymbol{i}$.

The synapse model mimics bidirectional (both switches closed) or unidirectional (only one switch closed) connection or no connection at all (both switches open) and is depicted in Fig. 8. This is because the switches decide whether a current can flow through the memductors W'_κ , W''_κ or not. Here, a wave digital model with physically

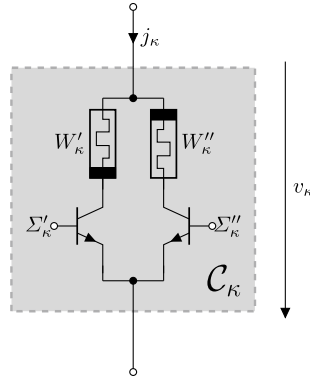


Fig. 8 Synapse model for two-neuron communication

Table 4 Maze memristor emulation parameters

$i_0 = 38.74 \mu\text{A}$	$R_0 = 10 \text{ k}\Omega$
$u_0 = 0.5 \text{ V}$	$S_p = -50 \text{ Hz/V}$
$N = 1.5$	$S_n = -0.4 \text{ Hz/V}$
$\beta = 0.1$	$U_{tp} = 1 \text{ V}$
$\gamma = 3.8$	$U_{tn} = -0.5 \text{ V}$

meaningful parameters of a RRAM-cell, developed in [25], has been utilized. This device distinguishes itself by its rapid switching behavior, which here is exploited to achieve the final state of the maze as quickly as possible. Consequently, synaptic currents \mathbf{j} and voltages \mathbf{v} are related by $\mathbf{j} = \mathbf{W}(\mathbf{z}, \mathbf{v}) \mathbf{v}$, which in combination with the synaptic voltages and the neuronal currents, cf. (17) and (18), yields a nonlinear relationship between i and u

$$i = e_v^T N \mathbf{W}(\mathbf{z}, N^T \boldsymbol{\varphi}(u)) N^T \boldsymbol{\varphi}(u), \tag{19}$$

which is numerically solvable in the wave digital domain by using fixpoint iterations for the implicit relationships between u , i and \mathbf{z} [23]. By Chua’s closure theorem, the memristive network can be summarized by single memristor, see [26]. With the emulation parameters as shown in Table 4, there are two emulation scenarios investigated in the following.

Scenario 1 involves a maze of size 20×20 as shown in Fig. 9 (top left). There are only three possible communication paths for the origin neuron (entry) to transmit information to the destination neuron (exit). The course of the current as a response to the input voltage ramp e is displayed to the right. Every notch represents the switching of all RRAM-cells along a certain path, which can be exploited as a detection mechanism. Since here the current is highest among the path with the least total memristance, the possible communication paths are found in order of ascending path

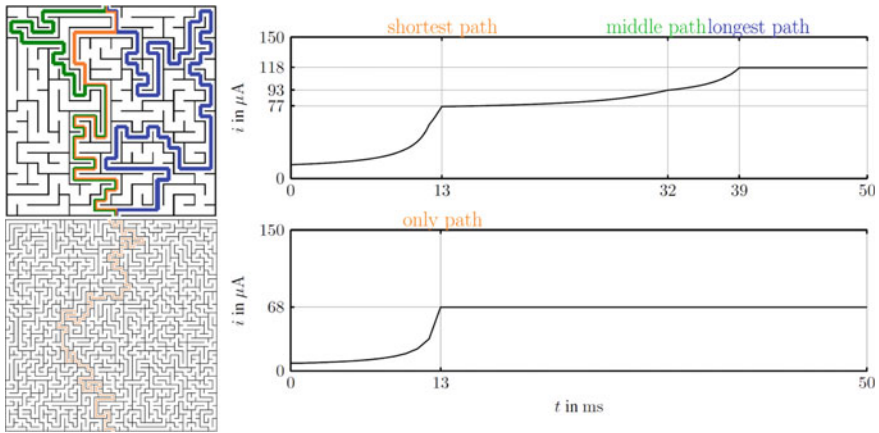


Fig. 9 Comparison of a 20×20 maze with three possible paths from entry to exit, highlighted in orange, green and blue (top left) and a 50×50 maze with only a single path (bottom left). The course of the currents for the 20×20 and 50×50 maze are shown at the top right and bottom right, respectively

length at $t = 13$ ms, $t = 32$ ms and $t = 39$ ms, respectively. In scenario 2 a bigger maze of size 50×50 , cf. Fig (bottom left), is investigated with the course of the current again illustrated to its right. A single notch indicates that there is only a single communication path possible. Note that the point in time, in which the communication path is discovered by the setup, is also $t = 13$ ms. This is the same amount of time required in the first scenario with the significantly smaller maze to find the shortest communication path. Remember that the first scenario requires 361 connection elements and consequently 722 RRAM-cells, while the second scenario requires more than six times that many (2401 connection elements, 4802 RRAM-cells).

This leads to the counterintuitive conclusion that the size of the maze has a negligible influence on the convergence time. This is due to the inherently massive analog parallelism and is the reason why computationally complex problems as the one investigated in this work can be solved efficiently by memristive circuits. While in the first scenario an input voltage $e = 180$ V has been utilized, in comparison to the second scenario, where an input voltage $e = 500$ V has been deployed, roughly three-times more voltage was required to compensate the time requirement, it is still remarkable that independent of the maze size, the presented setup requires an identical convergence time to find the shortest communication path. Furthermore, due to the immense amount of memristive devices present in the setup, a run-time comparison between the here presented wave digital emulation and the circuit simulation software LTspice was investigated in [23]. Although the scenarios were different, a run-time advantage by the factor of 300 in favor of the wave digital emulation was discovered on an identical computational device.

4.2.2 Mimicking Neuroplasticity

Here, one neuron of interest aims to find the optimal, e.g. shortest, communication pathways to *all* other neurons being part of the neural network. The basic concept was presented in [27], extended in [24] to solve the longest path problem and in [28] to mimic neuroplasticity. A fundamental scenario of optimal communication pathways with regards to multiple neurons is shown in Fig. 10.

There, three interconnected neurons are depicted with different time delays that can be associated for example with their axon lengths. For instance, an action potential can travel from neuron \mathcal{N}_A to neuron \mathcal{N}_B with a time delay of τ . The same holds true for a communication between neurons \mathcal{N}_B and \mathcal{N}_C . However, a communication between neurons \mathcal{N}_A and \mathcal{N}_C requires triple the time, namely 3τ . From the perspective of \mathcal{N}_A , there are two possibilities to transmit information towards \mathcal{N}_C : First, the direct path, which would require a time delay of 3τ . Second, the path over \mathcal{N}_B , which would require 2τ in total. The second choice is obviously better than the first with respect to time delay and should hence be used. This would make the synaptic connection between \mathcal{N}_A and \mathcal{N}_C obsolete. Because keeping this path (axon, axon terminals, synapse) alive requires energy, it would make sense from an energy perspective to let this path shut down, as indicated in Fig. 11 (right).

The ambition of the following steps is to synthesize an electrical circuit that inherently incorporates the above mechanisms. First, the problem is abstractly represented as a graph, where a neuron is mimicked by a node and a synaptic connection by a directed edge. The edge weights then represent the according time delay of the path. This is visualized by way of example as a nine neuron setup $\mathcal{N}_A, \dots, \mathcal{N}_I$ in Fig. 11 (left). By doing so, the problem of finding the minimal communication paths from neuron \mathcal{N}_A to all other neurons $\mathcal{N}_B, \dots, \mathcal{N}_I$ becomes the problem of finding the minimum spanning tree, where \mathcal{N}_A is the root. Previously, non-circuit approaches to solve this problem exist, such as the Dijkstra, A*, Bellman-Ford or Floyd-Warshall algorithms [29]. Among other features, these algorithms differ in their access to information on the topology and costs in the graph. Apart from the inherent favorable convergence time [30], a self-organizing circuit solution to the minimum spanning tree problem is also desirable in the context of unsupervised

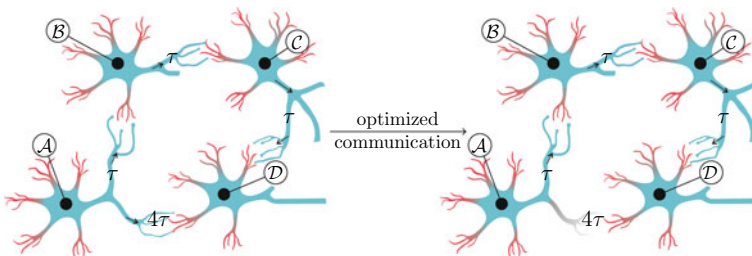


Fig. 10 A fundamental scenario what adaption and optimization means: The setup determines the paths of least delay self-organizing and lets all other paths die off to increase energy efficiency

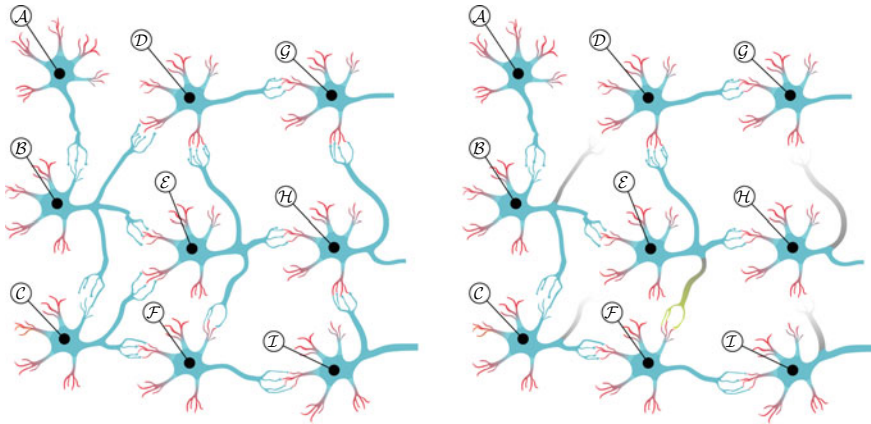


Fig. 11 Left: Initial neuronal network with dense synaptic connection. Right: Only synaptic connections enabling the least delay communication between neuron \mathcal{A} and all other neurons in the network remain, all others have died off. The synapse between neurons \mathcal{E} and \mathcal{F} is formed after the synapse between neurons \mathcal{C} and \mathcal{F} was damaged

learning, as the circuit itself discovers optimal communications structures for information transport. Unsupervised learning is also of interest in other domains, such as pattern recognition, where circuit-based solutions are applied [31]. The circuit to inherently find this minimum spanning tree is designed by synthesizing the neurons (nodes) and synapses (edges) as proposed by Fig. 12.

There, two different types of memductors are utilized. The neurons are represented by the μ -indexed memductor, which is described by equations

$$i_\mu = W_\mu(z_\mu)u_\mu, \tag{20a}$$

$$W_\mu(z_\mu) = \frac{W_0^N W_1^N}{W_1^N + z_\mu [W_0^N - W_1^N]} \tag{20b}$$

where $W_\mu(z_\mu)$ is the memductance and z_μ is the inner state. Here, $z_\mu = 0$ indicates the neuron memductor is in the high conductance state W_0^N and $z_\mu = 1$ indicates the low conductance state W_1^N . The state equation is described by

$$\begin{aligned} \dot{z}_\mu &= g(u_\mu) [\sigma(u_\mu)\sigma(z_\mu) + \sigma(-u_\mu)\sigma(1 - z_\mu)], \\ g(u_\mu) &= S_N [u_\mu - u_p] \sigma(u_\mu - u_p) + S_N [-u_\mu + u_n] \sigma(-u_\mu + u_n), \end{aligned} \tag{21}$$

where $\sigma(\cdot)$ is the Heaviside function, S_N is the steepness of the change from W_0^N to W_1^N and vice versa, u_p is the set voltage and u_n is the reset voltage. The synapse between neuron μ and neuron ν is described by the double-indexed memductor model

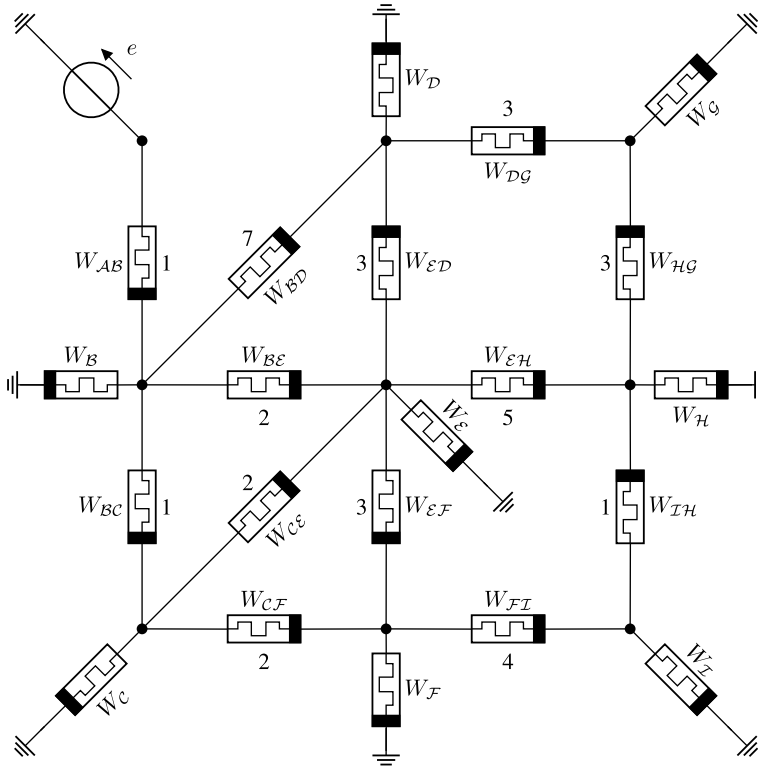


Fig. 12 Circuit concept to mimic neuroplasticity in the neuronal network of Fig. 11

$$j_{\mu\nu} = W_{\mu\nu}(z_{\mu\nu})v_{\mu\nu}, \tag{22a}$$

$$W_{\mu\nu}(z_{\mu\nu}) = \frac{W_0^S W_1^S}{W_1^S + z_{\mu\nu} [W_0^S - W_1^S]} \tag{22b}$$

and its state equation

$$\begin{aligned} \dot{z}_{\mu\nu} &= g(v_{\mu\nu}) [\sigma(v_{\mu\nu})\sigma(z_{\mu\nu}) + \sigma(-v_{\mu\nu})\sigma(1 - z_{\mu\nu})] + S_R\sigma(z_{\mu\nu} - 0.1), \\ g(v_{\mu\nu}) &= S_S [v_{\mu\nu} - v_p] \sigma(v_{\mu\nu} - v_p) + S_S [-v_{\mu\nu} + v_n] \sigma(-v_{\mu\nu} + v_n), \end{aligned} \tag{23}$$

where S_R is the amplitude of the retention characteristic, S_S is the steepness of the change from W_0^S to W_1^S and vice versa, v_p is the set voltage and v_n is the reset voltage. Additionally, in order for the circuit to function properly, it was explained in [27] in detail that

$$W_0^S \gg W_0^N \gg W_1^S \gg W_1^N. \tag{24}$$

Table 5 Simulation parameters

$W_0^S = 1 \text{ S}$	$v_p = 0 \text{ V}$	$u_p = 6 \text{ V}$
$W_0^N = 1 \text{ mS}$	$v_n = -1 \text{ V}$	$u_n = -0.74 \text{ V}$
$W_1^S = 1 \text{ } \mu\text{S}$	$S_S = -1.2 \text{ Hz/V}$	$S_N = -100 \text{ kHz/V}$
$W_1^N = 1 \text{ nS}$	$S_R = 0.1 \text{ Hz/V}$	

must hold. By keeping the above mentioned specifics in mind, the setup of Fig. 11 (left) is synthesized by the electrical circuit presented in Fig. 12.

The mechanisms work as follows. Assuming all neuron memductors are initially in the high conductance state and all synapse memductors in the low conductance state, the root neuron \mathcal{N}_A is attached to an ideal voltage source e mimicking its membrane potential activity. When the synapse memductors directly attached to the root memductor experienced enough magnetic flux, they transition to the high conductance state. At this point, both the root memductor and the newly reached neuron memductor are approximately at the the same voltage potential. Therefore, this neuron memductor experiences almost all of the voltage from e . Since the neuron memductors are oriented in a way such that they transition in the low conductance state when experiencing a negative voltage, i.e. during the hyperpolarization phase, it is exploited as a mechanism to signal that the path of minimal delay towards this neuron has been found. To further illustrate this procedure, the initial neuron population of Fig. 11 (left) is considered in a state where it has not yet adapted to its environment. The corresponding memristive circuit of Fig. 12 is simulated for this purpose with its parameters given in Table 5. The initial states of synapse and neuron memductors are chosen to $z_0^S = 0.7$ and $z_0^N = 0$, respectively, and the input signal is given by a sequence of action potentials with $U_1 = -0.75 \text{ V}$, $U_2 = 4 \text{ V}$ and period $T_e = 1 \text{ s}$.

The behavior of all synapse memductance states are shown in Fig. 13, where all synapses' states being part of the optimal synaptic structure formation are visualized on top and all others at the bottom.

In particular, path \mathcal{N}_A to \mathcal{N}_B is found first at $t_1 = 0.45 \text{ s}$, followed by path \mathcal{N}_B to \mathcal{N}_C at $t_2 = 1.35 \text{ s}$, path \mathcal{N}_B to \mathcal{N}_E at $t_3 = 1.55 \text{ s}$, path \mathcal{N}_C to \mathcal{N}_F at $t_4 = 2.45 \text{ s}$, path \mathcal{N}_E to \mathcal{N}_D at $t_5 = 4.25 \text{ s}$, path \mathcal{N}_E to \mathcal{N}_H as well as path \mathcal{N}_F to \mathcal{N}_I at $t_6 = 5.45 \text{ s}$, and finally path \mathcal{N}_D to \mathcal{N}_G at $t_7 = 6.35 \text{ s}$. At this point, all neurons $\mathcal{N}_B, \dots, \mathcal{N}_I$ in the setup can be reached from the root neuron \mathcal{N}_A . There are three things worth highlighting. First, notice that path E to H and path F to I are found simultaneously, which shows that no overwriting of previous states is required. Second, note that the communication paths are found in order of ascending total delay time from the perspective of \mathcal{N}_A . Lastly, all synapses associated with the memductance states shown in Fig. 13 (bottom) are not part of the optimal communication structure and hence must die off from the perspective of energy efficiency. The circuit identifies those synapses by all memductances whose states are not in the high conductance state. The proper interpretation is visualized by the blue communication paths in Fig. 11 (right).

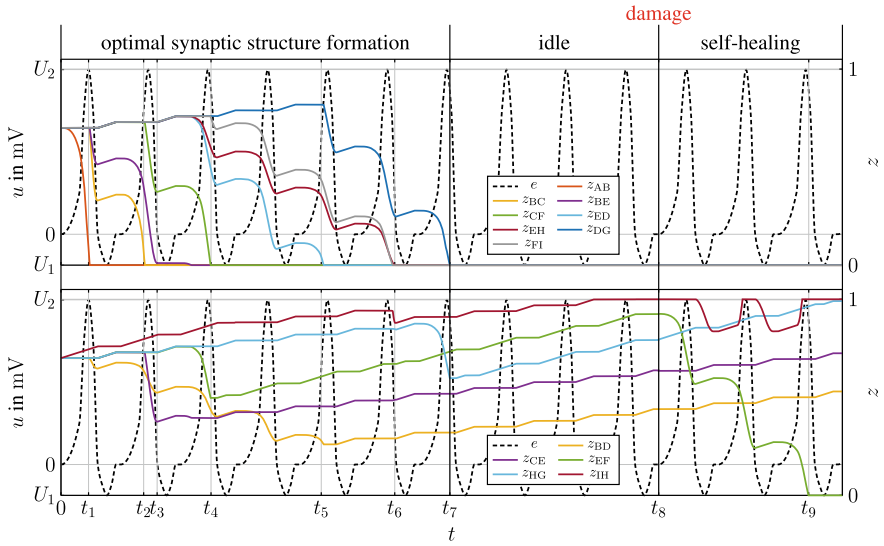


Fig. 13 Simulation results of all synaptic memristors which are part of the optimal communication structure (top) and are not part of the optimal communication structure (bottom)

However at $t_8 = 9.76$ s, the memductor W_{CF} is removed from the electrical circuit to account for some damage to the neural network, maybe induced by an accident or a disease. Because the path from \mathcal{N}_C to \mathcal{N}_F was part of the previously optimal communication structure, \mathcal{N}_A can not communicate with all neurons of the neural network any more. Specifically, neurons \mathcal{N}_F and \mathcal{N}_I are affected. However, the circuit is able to self-heal this damage. After a short period of time at $t_9 = 12.21$ s, the memductor W_{EF} transitions to the high conductance state, indicating that the previously died synapse between \mathcal{N}_E and \mathcal{N}_F is now created anew. Now the root neuron \mathcal{N}_A can communicate again with neurons \mathcal{N}_F and \mathcal{N}_I , highlighted by the green arrow in Fig. 11 (right). Note that the circuit will always inherently find a way to self-heal the communication structure with the minimal amount of changes compared to the already established, but incomplete structure. By this circuit synthesis and the subsequent simulation results it was shown that a structure can be simultaneously energy-efficient and fault-tolerant. Both are features frequently associated with the human brain and typically are in a trade-off with one another, where more energy needs to be invested in order to achieve a redundancy, which in return guarantees the fault tolerance.

5 An Elementary Decision Problem via Optical Illusions

Optical illusions are closely linked to associative learning. They also showcase that we permanently make unconscious decisions and it is for example when we encounter an optical illusion that we become aware of it, because something does not seem to be right. This is where the limits of our decision making is revealed. *Associative learning* or *Hebbian learning* was developed by Hebb, and is often summarized as “Cells that fire together, wire together” [32]. It describes a rule for change in synaptic coupling strength between mutually active pre- and postsynaptic neurons. In general, it attempts to explain associative learning, but is a very simplistic explanation that lacks details. For example, time delays caused by the axon are not considered. One famous example of associative learning is Pavlov’s dog [33]. There, it was observed that a dog produces saliva whenever he was given food. After a time, Pavlov discovered that the saliva was already produced when the lab’s assistant entered the room with the food. He developed the theory that the dog had learned to associate the assistant’s foot steps with the food and hence the foot steps caused the saliva production. He then tested the theory by ringing a bell shortly before the dog was given food. And sure enough, after repeating the process several times, a ring of the bell produced saliva in the dog’s mouth, even if there was no food given after the bell’s sound. He discovered that the two stimuli (sound of bell and presenting of food) both have to be presented within short period of time in order for the dog to associate the two events with one another. Since then, associative learning has been tried to replicate in electrical circuits as one branch of bio-inspired behavior. In this context, a simple form of associative learning is synthesized by a simple memristive network with neuronal oscillators [34]. Such oscillators are generally nonlinear models, cf. Sect. 3. A well-known, well-investigated nonlinear oscillator model is the Kuramoto model, which describes a set of multiple interconnected systems [35]. It is here utilized to model such a neuronal oscillator. The overall set Σ_K is described by

$$\Sigma_K : \begin{cases} \dot{\varphi}_\mu &= \omega + \frac{1}{m} \sum_{v=1}^m \kappa_{\mu v} \sin(\varphi_v - \varphi_\mu) \\ z_\mu &= \varphi_\mu \\ y_\mu &= \sin(\varphi_\mu) \end{cases}, \quad (25)$$

where $\mu = 1, \dots, m$, φ_μ is the angle of the μ -th oscillator, ω is the intrinsic radian frequency, $\kappa_{\mu v}$ is a nonlinear coupling factor, z_μ is an output providing the neighboring subsystems with information about φ_μ and y_μ is the output from which the oscillations can be observed. Note that the phases $\dot{\varphi}_\mu$ itself are only integrated and the oscillation of the system is only due to the nonlinear output function y_μ . To incorporate the Hebbian learning rule, the coupling strength $\kappa_{\mu v}$ between oscillator μ and v has a memory

$$\dot{\kappa}_{\mu v} = \varepsilon[\alpha \cos(\varphi_v - \varphi_\mu) - \kappa_{\mu v}], \quad (26)$$

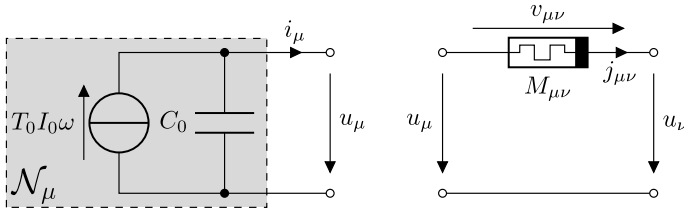


Fig. 14 Integrator circuit representing the Kuramoto model (left) and a memristive coupling element to describe the input signal of integrator μ from integrator ν (right)

where ϵ is the learning rate, α determines the weighting of new information. Here, the cosine function seeks to model the associative learning rule, as it is maximizes when the phases φ_ν and φ_μ are similar and minimized when the phase shift is π [36]. For the circuit synthesis process, details are elaborated in [37]. The two circuits of Fig. 14 are synthesize the Kuramoto model as in (25). The underlying equation of the left circuit is

$$\dot{u}_\mu = \frac{T_0 I_0}{C_0} \omega - i_\mu, \quad (27)$$

where T_0 , I_0 , C_0 are normalization constants with the units of a time, a current and a capacitance, respectively. A detailed analysis on a Kuramoto model with adaptive coupling is provided in [38]. It can already be seen that (27) synthesizes the first equation of (25) under the condition that

$$i_\mu = \frac{1}{m} \sum_{\nu=1}^m \kappa_{\mu\nu} \sin\left(\frac{u_\mu - u_\nu}{U_0}\right), \quad (28)$$

where U_0 is a normalization constant with the unit of a voltage and u_μ is associated with φ_μ . Note that the negative sign has been incorporated in the argument of the sine function.

The right circuit of Fig. 14 possesses the underlying equation

$$j_{\mu\nu} = W_{\mu\nu}(z_{\mu\nu}, v_{\mu\nu})v_{\mu\nu}, \quad v_{\mu\nu} = u_\mu - u_\nu, \quad (29)$$

with $W_{\mu\nu} = 1/M_{\mu\nu}$. The goal of the memductive interconnection is to model the ν -th element of the sum in (25). This is achieved when the coupling strength is associated with the state of the memductor and the phase difference with the voltage over the memductor, such that $\kappa_{\mu\nu} = z_{\mu\nu}$ and $\varphi_\mu - \varphi_\nu = v_{\mu\nu}$. For the memductance it follows that

$$W_{\mu\nu} = \frac{j_{\mu\nu}}{v_{\mu\nu}} = \frac{1}{m} z_{\mu\nu} \text{si}\left(\frac{v_{\mu\nu}}{U_0}\right) G_0, \quad (30)$$

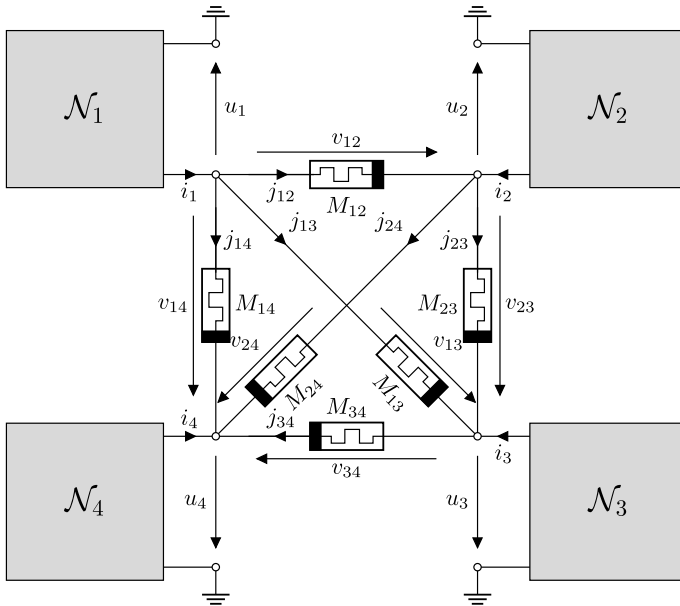


Fig. 15 Setup of four synthesized and fully coupled Kuramoto oscillators with the memductances incorporating the Hebbian learning rule

where G_0 is a normalization constant with the unit of a conductance. It must be mentioned however that (30) allows for $W_{\mu\nu} \leq 0$ and hence is not passive in general. Such memristors are called *active* memristors [26]. The Hebbian learning rule (26) is then synthesized by the proper modeling of the memductor’s state equation

$$\dot{z}_{\mu\nu} = \varepsilon \left[\alpha \cos \left(\frac{v_{\mu\nu}}{U_0} \right) - z_{\mu\nu} \right]. \quad (31)$$

Figure 15 shows a setup synthesizing four Kuramoto oscillators with a fully coupled interconnection network [37].

In the following, a condensed description of the memductive interconnection network will be given before the wave digital model is established. To obtain a general description of an arbitrarily large number of connected Kuramoto oscillators m , the voltages and currents are summarized by vectors $\mathbf{u} = [u_1, \dots, u_m]^T$ and $\mathbf{i} = [i_1, \dots, i_m]^T$, respectively. Notice in Fig. 15 that the orientation of the currents $j_{\mu\nu}$ is chosen such that they originate at the lower level subsystem and are directed towards to the higher level subsystem. This allows for a representation of the interconnection structure by the incidence matrix N . Kirchoff’s node and current laws are then given by

$$\mathbf{i} = N\mathbf{j}, \quad \mathbf{u} = N^T\mathbf{v}, \quad (32)$$

where j is the vector of all currents $j_{\mu\nu}$ and v is the vector of all voltages $v_{\mu\nu}$ of the interconnections with $\mu, \nu = 1, \dots, m, \mu < \nu$. Utilizing this notation, (29) is rewritten as

$$j = W(z, v)v, \quad W(z, v) = \text{diag} \left(W_{\mu\nu}(z_{\mu\nu}, v_{\mu\nu}) \right)_{\mu, \nu=1, \mu < \nu}^m. \quad (33)$$

With the help of (32), the relationship between the oscillator voltages and currents u, i and the memductors of the interconnection network is described by

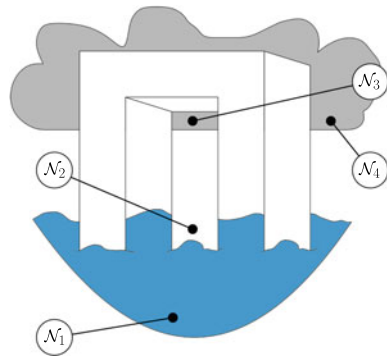
$$i = \hat{W}u, \quad \text{with} \quad \hat{W} = NW(z, v)N^T, \quad (34)$$

where $\hat{W} = \hat{W}(z, v)$ contains both the information about the memductances and the topology. With these insights it is possible to not only form a vector-based representation of the synthesized Kuramoto model, but to also obtain the corresponding wave digital model, cf. [37]. There, $R_C = T/2C_0\mathbf{1}$, $R_W = R_W\mathbf{1}$ and consequently $R_S = [R_C^{-1} + R_W^{-1}]^{-1}$. The value of R_W can essentially be chosen arbitrarily, as long as $R_W > 0$, but it should have a value close to the middle of $W_{\mu\nu}$'s range. Doing so typically results in a good numerical condition of the scattering matrix S_W , where

$$i = \hat{W}u \Leftrightarrow a_W = S_W b_W, \quad \text{with} \quad S_W = [R_W^{-1} + \hat{W}]^{-1}[R_W^{-1} - \hat{W}]. \quad (35)$$

In the following, an application example of the synthesized Kuramoto model in the context of an optical illusion is presented. For this, the optical illusion depicted in Fig. 16 is considered. Here, the setup of Fig. 15 models the four relevant aspects of the optical illusion, where $\mathcal{N}_1, \mathcal{N}_3$ account for the water and cloud (background), respectively. \mathcal{N}_2 accounts for the bottom part of middle pillar of the object and \mathcal{N}_4 accounts for both the top part of the middle pillar and the cloud background. It becomes apparent that the state of \mathcal{N}_4 and whether it is associated with the cloud or the middle pillar is the key factor of why the whole setup is an optical illusion. As the bottom part of the middle pillar and the ocean are spatially close, it appears

Fig. 16 An optical illusion being modeled by the setup of Fig. 15



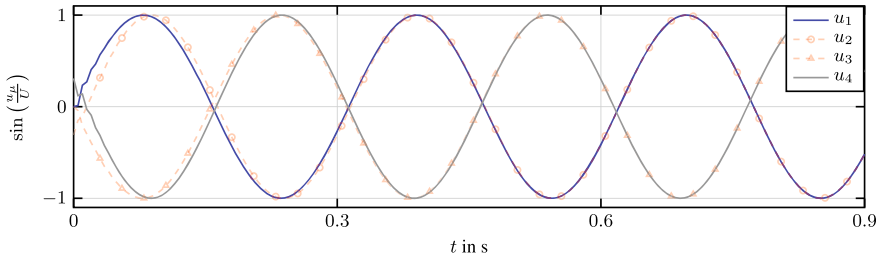


Fig. 17 Four oscillators form two synchronized groups with the latter being in an anti-phase configuration, explaining the contradiction of the optical illusion

natural to initialize the phases of the corresponding Kuramoto oscillators close as well. The same holds true for the archway and the cloud. Since the ocean and cloud are spatially far away from each other, their phase difference is comparatively big initially. Hence, $\varphi_1 = 0, \varphi_2 = -0.1\pi, \varphi_3 = \pi$ and $\varphi_4 = 0.9\pi$ was chosen. Moreover, the learning rate was chosen $\epsilon = 0.1$, the weighting of new information $\alpha = 0.1$ and the intrinsic radian frequency $\omega = 1s^{-1}$. The according emulation results are shown in Fig. 17. It can be observed that due to the initial conditions, the oscillators \mathcal{N}_1 and \mathcal{N}_2 representing the ocean background and the lower part of the middle pillar fire together. And as expected, due to “cells that fire together, wire together” [32], these two oscillators synchronize. This is interpreted as being an association of the bottom part of the middle pillar with the ocean. The same association takes place with with the top part of the pillar and the cloud, represented by oscillators \mathcal{N}_3 and \mathcal{N}_4 . This example might indicate that the optical illusion works due to the spatial correlation of the contradictory association between object and background. This contradiction is represented by a large phase shift between the two synchronized subgroups in Fig. 17. Of course, a human is able to detect the contradiction after a short period of time and to determine the origin of the optical illusion.

6 Conclusion

Due to their inherent massive parallelism in a signal processing sense, electrical circuits in general are candidates to replicate the computing power as well as energy and spatial efficiency of biological systems. In order to do so, the underlying biological systems must be understood and then abstractly transferred to hardware circuits, making bio-inspired computing a highly interdisciplinary field. Before circuits can be manufactured, theoretical circuits must be designed to abstractly mimic biological behavior and serve as a proof of concept. This purpose has been fulfilled in a variety of application presented in this chapter. The wave digital concept has been used as an emulation technique to digitally replicate these theoretical circuits. Since memristors are the centerpieces of all these circuit concepts and currently hard to fabricate with

the current state of technology. Hence, software-based preinvestigations are useful to shorten development cycles, especially since the wave digital concept is real-time capable and offers *in-operando* parameter manipulation. For example, this has been exploited to optimized parameters of the well-known Morris-Lecar model in order to show a more biologically meaningful behavior. Further applications included a gait pattern generator, for which a unidirectional synapse has been developed. Also, two- and multi-neuron communications has been investigated on an abstract level, where the circuit concept for the latter also exhibited self-healing to repair synaptic damage and hence showed a rich behavior in the context of neuroplasticity. Lastly, an optical illusion was investigated to mimic an elementary decision problem.

In summary, the field of bio-inspired computing has the potential to tackle problems, which are too complex for current computing concepts. The memristor is a novel electrical circuit element that enables computations with and in memory and could be the key component to enable bio-inspired circuits. The bottlenecks, which need to be overcome to manufacture high performance circuits that mimic biological behavior is twofold: First, the understanding of the way biological systems function must be increased to identify its general and abstract mechanism, which then need to be modeled in an electrical circuit. Second, the manufacturing process of memristors must become more reliable and precise.

References

1. Biolek, D., Biolková, V., Kolka, Z., Biolek, Z.: Passive fully floating emulator of memristive device for laboratory experiments. *Adv. Electr. Comput. Eng.*, 112–116 (2015)
2. Sánchez-López, C., Aguila-Cuapio, L.: A 860kHz grounded memristor emulator circuit. *AEÜ Int. J. Electron. C.* **73**, 23–33 (2017)
3. Ochs, K., Michaelis, D., Jenderny S.: An Optimized Morris-Lecar Neuron Model Using Wave Digital Principles. In: 2018 IEEE 61st MWSCAS, Windsor, ON, Canada, pp. 61–64 (2018)
4. Solan, E., Ochs, K.: Wave digital emulation of general memristors. In: *Int. J. Circuit Theory Appl.* **46–11**, 2011–2027 (2018)
5. Fettweis, A.: Wave digital filters: theory and practice. *Proc. IEEE.* **74–2**, 270–327 (1986)
6. Ochs, K.: *Passive Integrationsmethoden*. Universität Paderborn, Paderborn, Germany, At (2001)
7. Meerkötter, K.: On the passivity of wave digital networks. *IEEE Circuits Syst. Mag.* **18–4**, 40–57 (2018)
8. Fettweis, A., Nitsche, G.: Numerical integration of partial differential equations using principles of multidimensional wave digital filters. In: Nossek, J.A. (ed.) *Parallel Processing on VLSI Arrays*. Springer, Boston, MA (1991)
9. Hetmanczyk, G.: Exploiting the parallelism of multidimensional wave digital algorithms on multicore computers. *Multidim Syst Sign Process.* **21–45** (2010)
10. Morris, C., Lecar, H.: Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35–1**, 193–213 (1981)
11. Sah, M.P., Kim, H., Eroglu, A., Chua, L.: Memristive model of the barnacle giant muscle fibers. *Int. J. Bifurc. Chaos Appl. Sci. Eng.* **26–01** (2016)
12. Sheng, M., Sabatini, B., Sudhof, T.: *The Synapse*, vol. 1. Cold Spring Harbor Laboratory Press (2012)

13. Schwerdtfeger, T., Kummert, A.: Nonlinear circuit simulation by means of Alfred Fettweis' wave digital principles. *IEEE Circuits Syst. Mag.* **19–1**, 55–65 (2019)
14. Ochs, K., Michaelis, D.: Neural Network Topology Formation Using Memristive Jaumann Structures. In: 2018 IEEE 61st MWSCAS, Windsor, ON, Canada, pp. 660–663 (2018)
15. Levitan, I.B., Kaczmarek, L.K.: *The Neuron: Cell and Molecular Biology*, vol. 4. Oxford University Press, New York (2015)
16. Ochs, K., Michaelis, D., Jenderny, S.: Mimicking Gait Pattern Generators. In: 2018 IEEE 61st MWSCAS, Windsor, ON, Canada, pp. 73–76 (2018)
17. Ochs, K., Michaelis, D., Solan, E., Feketa, P., Schaum A., Meurer, T.: Synthesis, design, and synchronization analysis of coupled linear electrical networks. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67–12**, 4521–4532 (2019)
18. Feketa, P., Schaum, A., Meurer, T., Michaelis, D., Ochs, K.: Synchronization of nonlinearly coupled networks of Chua oscillators. In: 2019 IFAC 11th NOLCOS, Vienna, Austria, pp. 1–4 (2019)
19. Ochs, K., Al Beattie, B., Jenderny, S.: An Ising Machine Solving Max-Cut Problems based on the Circuit Synthesis of the Phase Dynamics of a Modified Kuramoto Model. In: 2021 IEEE 64th MWSCAS, Lansing, MI, USA (2021)
20. Michaelis, D., Jenderny, S., Ochs, K.: Self-Organizing Gait Pattern Generator Exploiting an Electrical Circuit for Axon Growth. In: 2021 IEEE 64th MWSCAS, Lansing, MI, USA (2021)
21. Michaelis, D., Ochs, K., Jenderny, S.: A Memristive Circuit for Gait Pattern Classification Based on Self-Organized Axon Growth. In: 2021 IEEE 64th MWSCAS, Lansing, MI, USA (2021)
22. Pershin, Y.V., Di Ventra, M.: Self-organization and solution of shortest-path optimization problems with memristive networks. *Phys. Rev. E.* **88–1**, 013305 (2013)
23. Ochs, K., Solan, E., Michaelis, D., Hilgers, L.: Emulation of a Navigation Processor with Physical Memristors Models. In: 2019 30th ISSC, Maynooth, Ireland, pp. 1–4 (2019)
24. Ochs, K., Michaelis, D., Solan E.: Solving the Longest Path Problem using a HfO₂-based Wave Digital Memristor Model. In: 2019 IEEE 62nd MWSCAS, Dallas, TX, USA, pp. 355–358 (2019)
25. Solan, E., Pérez, E., Michaelis, D., Wenger, C., Ochs, K.: Wave digital model of a TiN/Ti/HfO₂/TiN memristor. *Int. J. Numer. Model. El.* e2588 **32–5**, 1–12 (2019)
26. Chua, L.: Everything you wish to know about memristors but are afraid to ask. *Radioengineering* **24–2**, 319–368 (2015)
27. Ochs, K., Michaelis, D., Solan, E.: Wave Digital Emulation of a Memristive Circuit to Find the Minimum Spanning Tree. In: 2019 IEEE 62nd MWSCAS, Dallas, TX, USA, pp. 351–354 (2019)
28. Ochs, K., Michaelis, D., Jenderny, S., Kohlstedt, H.: Mimicking Neuroplasticity by Memristive Circuits. In: 2020 IEEE 63rd MWSCAS, Springfield, MA, USA, pp. 448–451 (2020)
29. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: *Introduction to Algorithms*, vol. 3. The MIT Press, Cambridge, Mass (2009)
30. Pershin, Y.V., Di Ventra, M.: Self-organization and solution of shortest-path optimization problems with memristive networks. *Phys. Review. E.* **88–1**, 013305 (2013)
31. Hansen, M., Zahari, F., Ziegler, M., Kohlstedt, H.: Double-barrier memristive devices for unsupervised learning and pattern recognition. *Front. Neurosci.* **11** (2017)
32. Shatz, C.J.: The developing brain. *Sci. Am.* **267–3**, 60–67 (1992)
33. Pavlov, I.P.: *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press, Oxford, England (1927)
34. Ochs, K., Michaelis, D., Jenderny, S., Szymendera, M.: Kuramoto Model with Hebbian Learning Mimics Spatial Correlations Causing an Optical Illusion. In: 2021 IEEE 64th MWSCAS, Lansing, MI, USA (2021)
35. Kuramoto, Y.: *International Symposium on Mathematical Problems in Theoretical Physics*. In: *Lecture Notes in Physics*. Springer (1975)
36. Timms, L., English, L.Q.: Synchronization in phase-coupled Kuramoto oscillator networks with axonal delay and synaptic plasticity. *Phys. Rev. E.* **89–3**, 032906 (2014)

37. Ochs, K., Michaelis, D., Roggendorf, J.: Circuit Synthesis and Electrical Interpretation of Synchronization in the Kuramoto Model. In: 2019 30th ISSC, Maynooth, Ireland, pp. 1–5 (2019)
38. Feketa, P., Schaum, A., Meurer, T.: Synchronization and multicluster capabilities of oscillatory networks with adaptive coupling. *IEEE Trans. Automat. Contr.* **66–7**, 3084–3096 (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Memristive Switching: From Individual Nanoparticles Towards Complex Nanoparticle Networks



Niko Carstens, Maik-Ivo Terasa, Pia Holtz, Sören Kaps, Thomas Strunskus, Abdou Hassanien, Rainer Adelung, Franz Faupel, and Alexander Vahl

Abstract Novel hardware concepts in the framework of neuromorphic engineering are intended to overcome fundamental limits of current computer technologies and to be capable of efficient mass data processing. To reach this, research into material systems which enable the implementation of memristive switching in electronic devices, as well as into analytical approaches helping to understand fundamental mechanisms and dynamics of memristive switching is inevitable. In this chapter, memristive switching based on Ag metal filament formation is discussed throughout different scales, providing insights on the stability of metal filaments and the onset of collective behaviour. An unconventional cAFM approach, which intends to integrate the memristive system directly on the apex of the cantilever instead of usual contacting is presented. This facilitates the nanoscale probing of filamentary memristive switching dynamics on long time scales for the purpose of basic research, which is demonstrated by an archetypical electrochemical metallization (ECM) based system consisting of Ag/Si₃N₄/Au. Further, the application of AgAu and AgPt noble metal alloy nanoparticles (NPs) for memristive devices is discussed with special focus on the device scalability. For the smallest scale it is shown, that a single AgPt-NP encapsulated in SiO₂ operates via stable diffusive switching. Finally, two concepts for the self-assembled fabrication of NP-based memristive switch networks are evaluated regarding to collective switching dynamics: A sub-percolated CNT network decorated with AgAu-NPs and a Ag-NP network poised at the percolation threshold. The hybrid CNT/AgAu-NPs networks exhibit a mixed form of diffusive and bipolar switching, which is very interesting for tailoring the retention time, while

N. Carstens · T. Strunskus · F. Faupel · A. Vahl (✉)

Institute for Materials Science, Kiel University, Chair for Multicomponent Materials, Kaiserstr. 2, 24143 Kiel, Germany
e-mail: alva@tf.uni-kiel.de

M.-I. Terasa · P. Holtz · S. Kaps · R. Adelung

Institute for Materials Science, Kiel University, Chair for Functional Nanomaterials, Kaiserstr. 2, 24143 Kiel, Germany

A. Hassanien

Department of Condensed Matter Physics, J. Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_9

219

the networks dynamics of percolated Ag-NP networks are governed by ongoing transitions between a multitude of metastable states, which makes them interesting for reservoir computing and other neuromorphic computation schemes.

Keywords Memristive devices · Diffusive switching · Self-assembled networks · Nanoparticles · Neuromorphic engineering · Dynamic networks

1 Introduction

For several decades, the increasing demand for computational data processing was satisfied by miniaturization of the transistor, which plays the role of the fundamental building unit in modern computer technology, through extreme manufacturing techniques and engineering of materials. Alongside with current developments like Internet of Things or autonomous driving for which massive data processing is required, the demand for computational power presumably cannot be met by conventional computer technologies in the near future [1]. One reason for this relates to physical limits as the size of a single transistor approaches atomic scale like a fundamental integration density limit or information lost via inadvertent tunneling currents [2]. Another reason is the strict separation of information processing and storage units predefined by the von-Neumann architecture of modern computer systems, since the duration for massive data transfer between both units evolved into a bottleneck. In foreseeable future, a gap in the computer technology must be filled in order to remain capable of dealing with an increasing massive amount of data. The field of neuromorphic engineering comes along with unconventional computation schemes, which aim to overcome the aforementioned limits by implementation of biological computation principles in novel kind of hardware [3, 4]. Memristive devices play a fundamental role in the realization of bio-inspired hardware. They are two-terminal passive circuit elements, whose internal state (represented by the resistance of the device) is sensitive to external voltage or current stimuli [5]. This property allows the implementation of key features for neuromorphic computation like parallelism, in-memory computing and plasticity [2]. Moreover, the functionality of memristive devices can be expanded towards memsensing, which means a combination of memristive and sensitive properties [6]. Tremendous work was dedicated to the development of material systems showing memristive behavior and understanding of nanoscale processes which are responsible for the memristive switching property, which resulted in a multitude of memristive material systems exhibiting diverse memristive switching dynamics [7]. One example is analog non-volatile switching, which is a potential electrical characteristic for long-term memory in neuromorphic circuits and enables the hardware implementation of artificial neural networks. Diffusive switching possesses a volatile threshold-like dynamic beneficial for short-term memory realization and incorporation of non-linearity in neuromorphic circuits, which are both inherent for brain functionality [8, 9].

Besides valence change [10] and phase change mechanisms [11], the electrochemical metallization (ECM) is of special interest to engineer materials with memristive properties. ECM relies on redox-reactions and ion transport of active metallic species, like Ag or Cu, in a nanoscale switching gap upon application of an external electrical field [12]. The switching dynamics in ECM devices is dominated by the reconfigurable formation and rupture of a thin filament composed of the active metallic species. The operation of ECM devices can be constrained into a diffusive regime by utilization of a strict limitation of current flowing through the device during its operation or by limiting the amount of the active metallic species [13]. The understanding and control of the complex processes governing the filament evolution during device operation is still a current challenge. Especially, the inherent stochasticity in the switching dynamics (in terms of variation of switching voltages from cycle to cycle) is merely insufficiently studied. The stochasticity may impede the development of certain applications like selector devices, but it may also be exploited like in the design of true random number generators [14, 15].

A critical issue in the development of neuromorphic hardware is the challenge of upscaling from individual memristive units to networked assemblies with an enormous degree of complexity. Promising approaches under circumvention of vast wiring procedures rely on self-assembly of memristive switching units. Reports have shown, that nanoparticles (NPs) of Au [16, 17] or Sn [18] or Ag-nanowires [19, 20] can be applied as fundamental memristive building units for self-assembled networks. Such networks are capable of responding via complex dynamical switching patterns featuring short-term memory and temporal correlations, which are both key requirements for concepts like reservoir computing [21]. A feasible strategy to maximize the degree of complexity in self-assembled networks of switching units, and therefore to approach neural complexity, is to make use of criticality, where the network dynamics are governed by complex correlated patterns, similar to those observed in neural tissues [22, 23]. Networks of memristive switching units can be operated at critical dynamics, when the connectivity inside the system is poised at the percolation transition [24]. Developments in the design of material systems and deeper characterization of emergent properties in networks of memristive units could be a significant advancement for bio-inspired computation strategies.

In this chapter, we discuss basic studies on Ag-based filamentary switching and approaches for device implementation on fundamentally different scales, ranging from single switching junctions to complex network behavior. Firstly, an unconventional conductive atomic force microscopy (cAFM) approach giving deep insights into the long-term switching dynamics and inherent stochasticity of a single Ag-filament is presented. Moreover, the application of Ag-based bimetallic noble metal NPs as individual memristive switching units is discussed. Major emphasis is put on implications which arise, when the memristive system is scaled up from a single switching unit towards macroscopic assemblies of switching units, i.e. up to which scale does the characteristics of individual NPs retain and when does collective behavior emerge. Therefore, discussions on macroscopic assemblies realized by two different approaches in which Ag-based noble metal NPs are incorporated into complex networks are given in the last section: The first approach comprises a

randomly assembled CNT network with Ag-based NPs as localized switching units while the second approach involves self-assembled networks of Ag-NPs poised at the percolation threshold.

2 Deep Insight into Single Filament Switching Via an Unconventional cAFM Approach

Progress in the development of novel neuromorphic hardware is directly reliant on the understanding of the fundamental units causing the memristive action. In ECM-based devices, as well as in self-assembled memristive switch networks, this fundamental unit is a single metallic filament, which defines an either conductive (low resistive state, LRS) or insulating (high resistive state, HRS) state of a nanogap. In ECM-based devices, there is in most cases solely one single filament, which dominates the behavior of the whole device [25]. Considering the typical diameter of an active filament in the order of 10 nm, this means that all memristive action during device operation takes place highly localized on the nanometer scale and is independent of the devices' geometrical dimensions defined through the fabrication [26]. The analysis of inherent stochastic features in the switching dynamics of single filaments are of special interest, because they set fundamental constraints on possible applications. On the contrary, in self-assembled memristive switch networks, filamentary switching is spatially distributed over a huge number of nanogaps throughout the network and the network response results from a complex interplay between the states of different filaments. In such networks, studies on single filamentary switching are likewise inevitable to understand the role of single elements in the collective network dynamics.

CAFM raised to an essential tool probing electrical properties on the nanoscale and was commonly used to make localized studies on switching filaments [27]. However, a prominent drawback of the conventional cAFM method is associated to the thermal drift of the cantilever, because it imposes uncertainties regarding the position of the cantilever relative to the filament under investigation. The extent of the thermal drift is usually more than 10 nm h^{-1} , which exceeds the typical dimensions the filament. This is especially problematic when the long-term electrical response of a single filament shall be measured. To circumvent this issue, an unconventional cAFM approach to probe the long-term response of the memristive switching of a single filament can be applied [28]. Instead of conventional contacting, the memristive device is integrated directly on the apex of a conductive cantilever. In this way, the functionalized cantilever can be brought in reliable contact using the cAFM instrumentation to any surface which acts as an inert counter electrode. This approach is sketched in Fig. 1.

Archetypical memristive systems consisting of a few nm thin Ag layer covered by a dielectric layer on Si_3N_4 were deposited on the apex of cantilevers by conventional physical vapor deposition (PVD) techniques. The basic coating of the cantilever

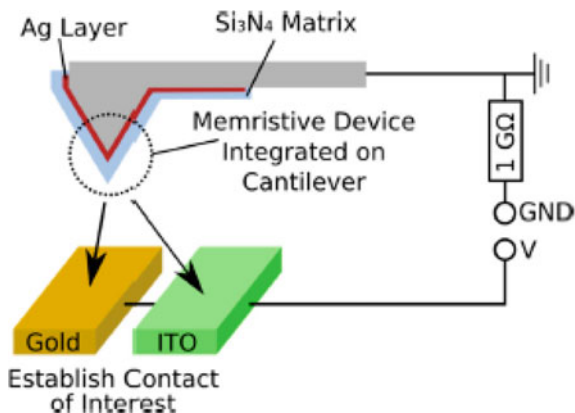


Fig. 1 Unconventional cAFM approach in the present study. Instead of conventional contacting, the memristive system is integrated on the apex of the cantilever via PVD methods. The functionalized cantilever is brought into contact with inert counter electrode surfaces (Au and ITO in this work) resulting in an ECM-based memristor configuration. As the filamentary switching is directly constrained to the apex, this approach allows more reliable long-term measurements on the nanoscale via conventional cAFM instrumentation. Reproduced under CC-BY 4.0 license from [28]

was chosen as Au to guarantee a good conductivity and electrochemical passivity of the cantilever. The configuration of functionalized cantilever contacted to the inert counter electrode surface defines a complete ECM-device to be probed. Cantilevers functionalized with $\text{Ag/Si}_3\text{N}_4$ were investigated under two different configurations: Having either Au or indium tin oxide (ITO) as inert counter electrode. This approach offers two decisive advantages for the purpose of basic research. Firstly, the memristive action through the filament is constrained directly to the apex of the cantilever. Consequently, thermal drift causing lateral motion of the cantilever over the counter electrode surface does not lead to any separation of probe and switching locality. As thermal drift effects are mitigated, reliable nanoscale long-term measurements of the filamentary switching activity are enabled using conventional cAFM instrumentation. Another advantage of this approach concerns the flexibility in choosing the experimental configurations, because an identical functionalized cantilever can be brought into contact to diverse counter electrodes. This enables strategies to investigate the switching activity of the very same filament on different interfaces.

In the following, the long-term switching of a single filament is discussed based on data gathered from a cantilever functionalized with $\text{Ag/Si}_3\text{N}_4$ contacted to a Au surface. The contact was held at a well-controlled force of 1.2 nN to provide non-invasive measurement conditions. The counter electrode was biased via voltage sweeps between -2.6 and 2.6 V and current responses were recorded continuously over 18 h resulting in over 12,000 switching cycles. During the measurement a $1\text{ G}\Omega$ serial resistance was applied to limit the current and protect the cantilever

from Ag over-diffusion. The filamentary switching mechanism taking place sharply localized at the apex and an exemplary switching cycle are shown in Fig. 2a and 2b, respectively. Initially, under zero-bias conditions, the sample is in its HRS since no stable filament can exist between probe and counter electrode. Upon biasing the counter electrode (towards either polarity), electrochemical oxidation of Ag is triggered at anodic sites in the sample, leading to a release of Ag^+ -ions into the dielectric layer. The Ag^+ -ions are mobile and can migrate in the Si_3N_4 layer until they become reduced at cathodic sites in the sample where they aggregate as metallic silver. Through this mechanism, metallic Ag is reconfigured in the dielectric layer until a continuous filament bridges the Si_3N_4 layer, which causes a switching event to the LRS. This moment is denoted as a SET event in the current response and the current level (highlighted in blue) is solely limited by the external $1\text{ G}\Omega$ serial resistance. When the voltage bias is reduced coming from the LRS, a RESET event is observed before reaching zero-bias. The physical mechanism behind this is the spontaneous disintegration of the Ag filament into individual clusters caused by interface energy minimization [13]. The fact that the filament is only stable above a certain threshold voltage is the most prominent feature of diffusive memristive switching. Among all cycles in the long-term measurement, 95% of them showed clear diffusive switching, which indicates the stable and non-invasive measurement conditions of this approach. It should be noted, that the incorporated $1\text{ G}\Omega$ resistance is the reason why the switching is constrained into the diffusive regime. On the one hand side, the limited current only allows for the formation of thin filaments which are more prone to spontaneous disintegration. On the other hand, in the moment of SET switching the major part of the externally applied electrical potential shifts to the serial resistor, which significantly reduces the field across the filament.

In another experiment, the consequences of having Au or ITO as counter electrode interface are investigated. For this purpose, the Ag/ Si_3N_4 system of an identical cantilever was brought into contact firstly with Au and subsequently with ITO. The decisive advantage of this approach is that an identical active filament was probed in consecutive measurements under different ECM-based device configurations. Therefore, all influence coming with the active electrode or dielectric layer (such as defect structure or filament morphology), which might disturb the comparability of different configuration, are ultimately kept constant. In contrast to that, in conventional cross-point devices, an investigation of the variation of different counter electrodes is only possible at the expense of losing the direct comparability in terms of the exact same filament [29, 30]. Figure 2c depicts the current response during operation on Au (black) and ITO (red) over 50 consecutive cycles for each electrode. It can be seen that interfacing to Au drives the switching window to lower voltages. A representative switching cycle (scan speed: 4 s per cycle) drawn in bold line shows a switching window between a RESET voltage of 0.3 V and a SET voltage of 1.9 V in the positive regime. Further, it was detected that upon interfacing to Au the LRS follows a linear characteristic, which reveals that contact between filament and counter electrode surface is purely ohmic. Switching the contact to ITO in the subsequent measurement resulted in fundamentally different electrical characteristics. In this configuration, the switching window is shifted to higher voltages (above

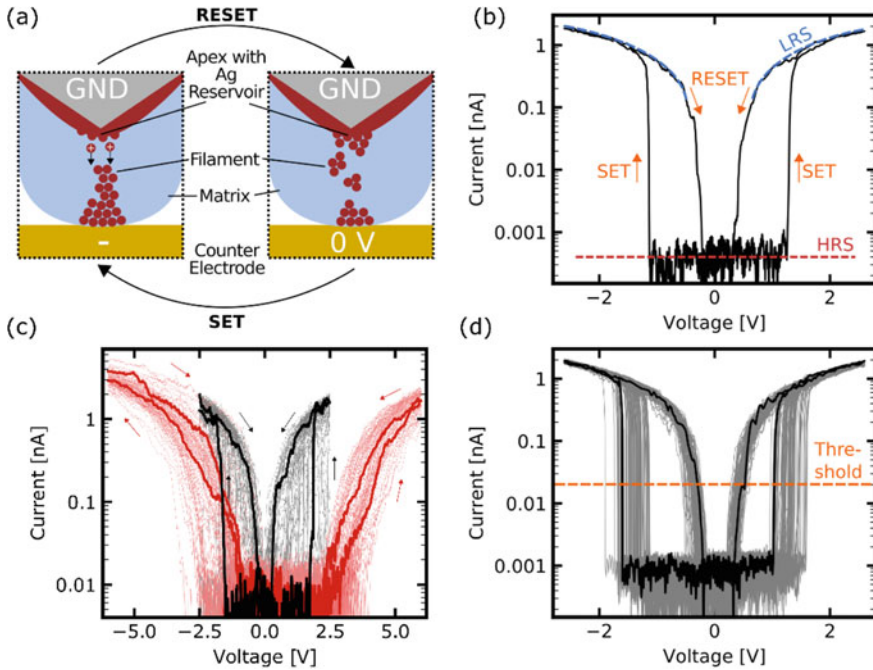


Fig. 2 **a** Physical mechanisms taking place during filamentary switching in the sample. **b** Representative cycle from the long-term measurement. The operation of the Ag/Si₃N₄ system integrated on a cantilever and contacted to a Au surface resulted in diffusive switching characteristics due to the 1 GΩ serial resistance. The distinctive feature of diffusive switching is, that the LRS is only stable above a certain threshold voltage and volatile at zero-bias conditions. **c** Operation of an identical filament from one functionalized cantilever on different counter electrode surfaces: Au (black) and ITO (red). A non-linearity in the current response was detected for ITO, which indicates the existence of an energetic barrier which was not observed during operation of the very same filament on Au. **d** Representation of 100 consecutive cycles among the long-term measurement. It can be seen, that SET and RESET threshold voltages have a statistical nature. Reproduced under CC-BY 4.0 license from [28]

2.5 V in the positive regime) and the current–voltage dependency exhibited a strong non-linearity. This indicates the occurrence of an energetic barrier at the interface between filament and ITO surface.

The long-term measurements of an identical functionalized cantilever open up the possibility to study the inherent statistical nature of diffusive switching. For this purpose, a long-term characterization on Au was conducted. Figure 2d depicts 100 consecutive cycles among the long-term measurement. It becomes apparent, that SET and RESET events are not triggered by fixed threshold voltages, but underlie a certain degree of stochasticity. Although great effort was invested to engineer filamentary systems with minimized switching variability, it is difficult to suppress it completely. Therefore, a deeper understanding of the statistical nature is inevitable. In Fig. 3a, all four threshold voltages per cycle (SET and RESET at either polarities) are

plotted against cycle number. A notable observation is, that the evolution of threshold voltages shows a significant trend ranging over hundreds of cycles in some regimes (highlighted as high correlation in the plot). However, in other regimes (highlighted as low correlation), the threshold voltages are subject to strong randomness. This observation suggests that there is partially a correlation in the switching characteristics in consecutive cycles and that the degree of correlation varies over the long-term measurement. To elucidate this, the linear correlation in the switching statistics was quantified by means of the Pearson coefficient. For both indicated regimes, the individual threshold voltages were related to the respective former value from the cycle before. This quantification is visualized in Figs. 3b and 3d for SET and RESET voltages with respect to positive polarity in the high correlation regime, respectively, and in Figs. 3c and 3e for the same characteristic voltages but in the low correlation regime. The values for the Pearson coefficient amount to 0.879 and 0.889 for SET and RESET in the high correlation regime, respectively, whereas they are 0.435 and 0.581 for SET and RESET in the low correlation regime, respectively. This supports the observation, that the threshold voltages are not statistically independent on each other, but are subject to correlations in the switching dynamics, which puts fundamental constraints in the design of real devices.

3 Noble Metal Alloy Nanoparticles for Diffusive Switching

An interesting approach to confine memristive action to the nanoscale lies in the transition from conventional bulk planar active electrodes towards nanostructured active electrodes or NPs. Using nanostructured electrodes (e.g. nanocones) showed a beneficial effect on the switching uniformity in Cu-based filamentary memristive devices [31]. NPs have recently attracted attention as building blocks for memristive switching devices and are readily employed, either to enhance the reproducibility by tailoring the electrical field inside the dielectric layer [32] or to act as a source of mobile metallic cations [8, 33]. Using AgAu and AgPt alloy NPs, the benefits of predefined electrical field enhancement and restricted mobile cation reservoirs were combined, and diffusive memristive action was studied at different length scales, ranging from individual NPs to multistack nanocomposites. Restriction of the mobile species reservoir is crucial to achieve stable diffusive switching as the formation of non-volatile filaments is inseparably linked to larger mass transport of the mobile species. The use of bimetallic NPs is particularly promising, because the amount of mobile species at the switching location can be effectively limited through NP composition and size [33]. Figure 4 gives an overview of different device setups which were studied in this context. AgPt and AgAu-NPs fabricated from a Haberland-type gas aggregation source serve as fundamental memristive building units. To obtain alloy NPs such as AgAu or AgPt, the concept of gas phase synthesis in a magnetron-based Haberland-type gas aggregation source [34–36] was expanded by the introduction of the multicomponent target approach. Here, instead of a planar bulk target, a custom-made target is applied, which consists of a Ag target with a

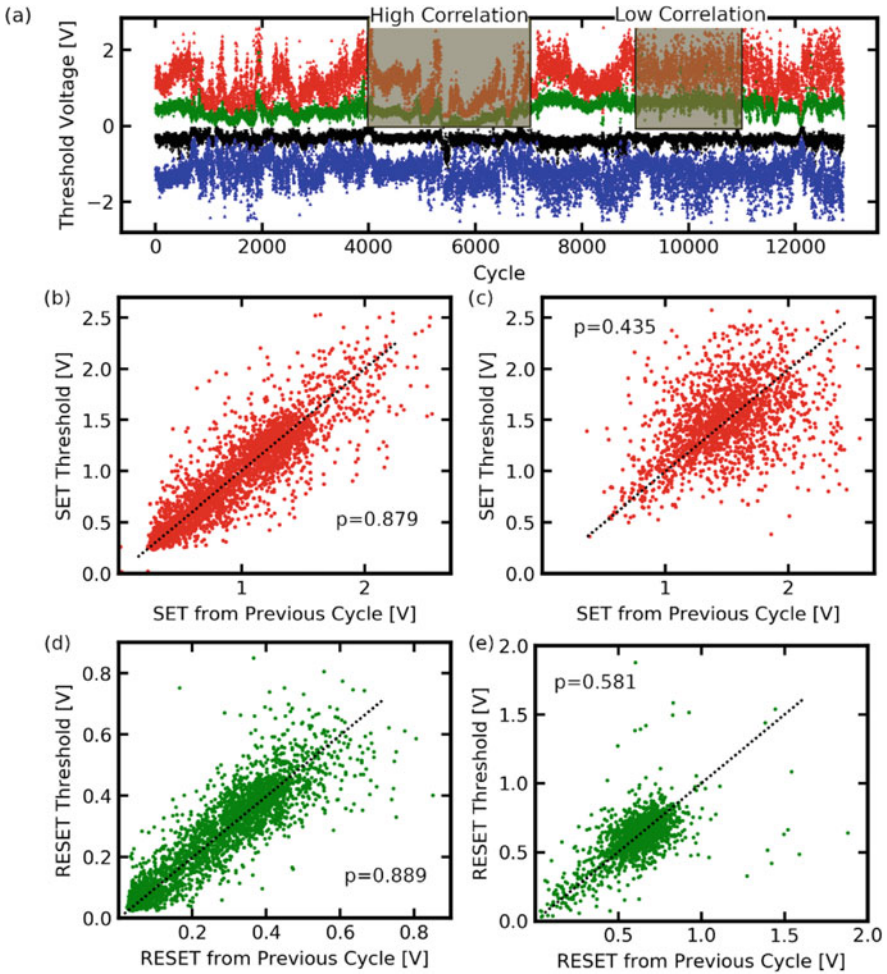


Fig. 3 Switching statistics extracted from the long-term measurement of a $\text{Ag/Si}_3\text{N}_4$ system on Au. **a** Characteristic threshold voltages (SET (red) and RESET (green) at positive polarity, SET (blue) and RESET (black) at negative polarity) plotted versus cycle number. Two regimes are highlighted: One showing systematic trends in the evolution of threshold voltages (labeled as high correlation) and one showing strong randomness (low correlation). The linear correlation was quantified by the Pearson coefficient in both regimes by relating the individual threshold voltages to the respective former value from the cycle before for SET and RESET in the high correlation regime (**b**, **d**) and SET and RESET in the low correlation regime (**c**, **e**). Reproduced under CC-BY 4.0 license from [28]

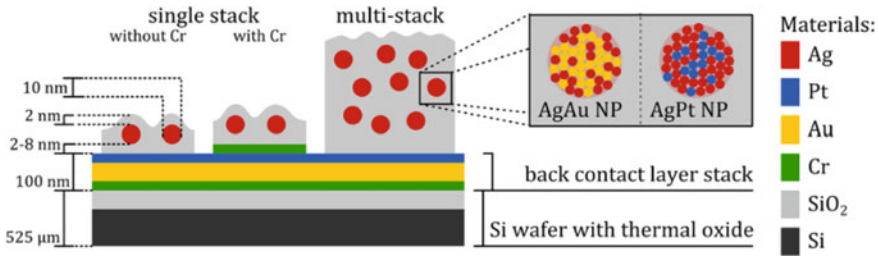


Fig. 4 Device setups under investigation. AgAu and AgPt-NPs act as fundamental memristive building units and are encapsulated into SiO_2 for a complete ECM-based memristive device. Left: Experimental setup to test the switching performance of individual NPs via cAFM. Right: The upscaling capabilities of the NP-based approach was tested on multiple stacks of NPs and SiO_2 layers via a macroscopic probe. Reproduced under CC-BY 4.0 license from [33]

trench and embedded Au or Pt wires. Using this methodology, the material usage of the more precious metal was significantly enhanced and the NP alloy composition can be tailored and controlled *in-operando*. More details regarding the NP deposition methodology can be found elsewhere [35, 36]. Encapsulation of the alloy NPs in a dielectric layer consisting of SiO_2 (deposited by reactive pulsed DC sputtering from a Si target in the presence of an Ar/O_2 atmosphere) completes the memristive device setup.

In contrast to the vast majority of ECM-based devices, the reservoir of active species (Ag in this case) has not the form of a continuous electrode, but is embedded into a dielectric layer between two inert electrodes. A major focus of this study was to explore the scalability of this concept which is one key aspect concerning device implementation. Therefore, this concept was tested on two different scales: Firstly, the switching performance of an individual alloy NP was tested via cAFM, which embodies the lowest limit of the scalability range. Secondly, macroscopic devices consisting of multiple stacks of NPs and SiO_2 layers were fabricated to characterize the upscaling capabilities.

The NP-based approach used to implement memristive devices on the nanoscale offers several advantages. The location of NPs inside the dielectric layer efficiently predefines the path where the filament is formed, which is essential for the device integration on the lower nanoscale. The strong localization of memristive action is additionally enhanced by the inherent electrical field concentration at the surface of a NP. Since the mechanisms leading to filament formation are driven by electrical fields, a substantial facilitation of memristive switching by using NPs is expected. Further, application of alloy NPs instead of the elemental counterparts provides additional degrees of freedom to tune the memristive action. For the alloy system presented here, AgAu and AgPt, Ag is the active species, whereas the more noble component, Au or Pt, is completely inert and will remain at its location inside the dielectric layer, which is expected to facilitate the long-term anchoring of the filament path.

To understand the memristive action of alloy NPs as building block for memristive devices, first the switching at the level of an individual NP has to be understood.

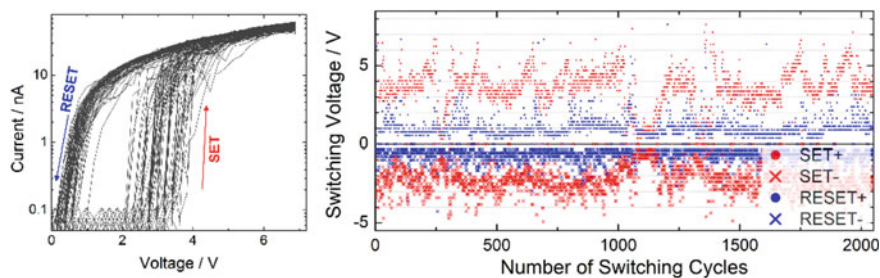


Fig. 5 Essential results from the memristive performance characterization of a single AgPt-NP encapsulated in SiO₂ via voltage sweeps between -7 and $+7$ V probed by cAFM. **a** shows a plot of 70 consecutive cycles which are representative for the diffusive switching behavior of a single NP. **b** shows the threshold voltage statistics for over 2000 consecutive cycles. Adapted under CC-BY 4.0 license from [33]

For this purpose, a stacked system consisting AgPt-NPs encapsulated in SiO₂ was prepared through consecutive deposition steps of 8 nm SiO₂, AgPt-NPs at sparse filling factor and 2 nm SiO₂. A cAFM setup was applied to identify the position of a NP and to measure the current response at the location of an individual AgPt-NP via driving voltage sweeps between -7 and $+7$ V through the conductive tip, while the common Au thin film served as back electrode. A serial resistance of 101 M Ω was applied to this measurement to limit the current flowing through the system. In this measurement, stable diffusive memristive switching was observed for over 2000 cycles. Essential results of the measurement are depicted in Fig. 5.

The memristive behavior of the single AgPt-NP matches diffusive switching and consequently follows the identical qualitative characteristics as described for single filament switching in the previous section. This is represented in Fig. 5a, where 70 consecutive cycles among the measurement are plotted. The threshold voltage statistics for over 2000 consecutive cycles are given in Fig. 5b. Also in the case of NPs, the switching characteristics underlie a considerable degree of stochasticity. Interestingly, no initial electroforming step had to be performed to achieve a stable switching regime. The origin of the diffusive switching behavior in the single AgPt-NP system is not merely determined by the external current limitation (like in the functionalized cantilever devices in the previous section), but rather by the absence of bulk Ag electrodes. The alloy NP approach allows for a severe limitation of the Ag amount (i.e. by small NP diameter or low ratio between Ag and noble alloy component) and therefore provides degrees of freedom to constrain the system into diffusive switching regime by supplying not enough Ag to build a non-volatile filament.

Further, the scalability of the NP-based approach was tested through studies on macroscopic devices. For this purpose, samples consisting of multiple stacks of SiO₂ layers and either AgAu or AgPt-NPs were fabricated and characterized in analogy to the cAFM measurements. To account for effects arising from the thickness of the SiO₂ separation layer, an effective SiO₂ thickness of 2 nm between the NP depositions was chosen for the AgAu sample whereas it was chosen as 4 nm for the

AgPt sample. A soft Pt probe having a diameter of 125 μm was applied to contact the samples. The samples were stressed again via voltage sweeps under current limitation with a 1 M Ω serial resistor. Figure 6a and 6b exhibit the current responses and threshold switching statistics measured for the macroscopic devices based on AgAu or AgPt-NPs, respectively. Generally, it can be observed, that the diffusive switching characteristics as known from a single NP is preserved, although multiple switching junctions inside the sample are expected to contribute to the filament formation. Comparison of the different alloy systems show, that the HRS of the AgAu sample allows a measurable conduction whereas the HRS of the AgPt sample is hidden in the noise level of the experimental instrumentation. These observations could be connected to the different thicknesses of the SiO₂ separation layer. In the AgAu sample, the lower SiO₂ thickness possibly allows significant leakage currents whereas the 4 nm SiO₂ layer thickness in the AgPt sample results in effective insulation.

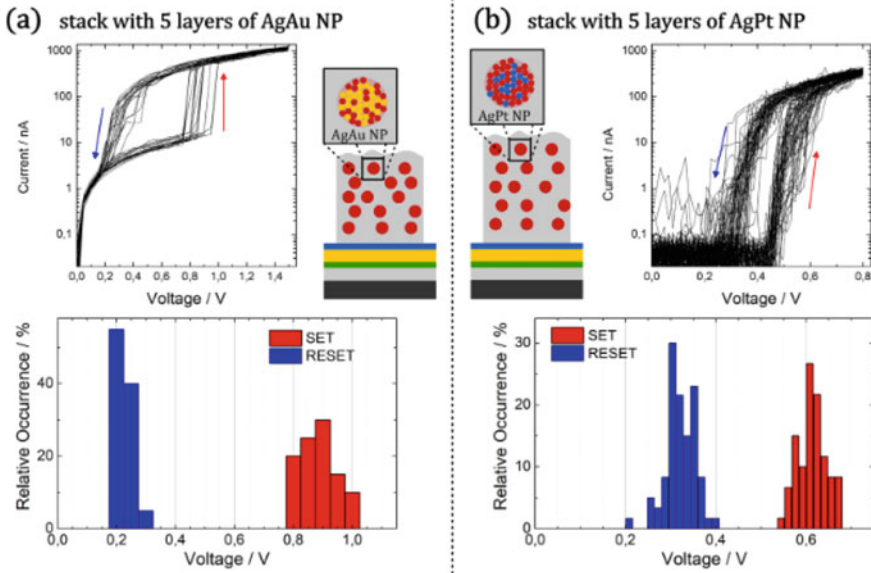


Fig. 6 Upscaling of the NP-based approach. Macroscopic devices were fabricated by multiple stacking of SiO₂ layers and AgPt or AgAu-NPs. The samples were contacted by a macroscopic Pt probe (diameter of 125 μm) and characterized via voltage sweeps. **a** and **b** show switching cycles at positive bias and extracted switching statistics for systems with AgAu and AgPt-NPs, respectively. Reproduced under CC-BY 4.0 license from [33]

4 Distributed Ag-Based NP Switching in Memristive Networks

Alongside with the profound characterization and device engineering of single memristors, further sophisticated challenges arise from the question how to build functional networks with a magnitude of memristive switching units. These challenges can be tackled by two different paradigms: Either by a regular and precisely engineered arrangement of hardwired memristive devices, like it is done in cross-bar arrays, or by reliance on self-assembly mechanisms to achieve a random dispersion of memristive switching units throughout a network. In the last section, two strategies for the fabrication of Ag-NP based memristive switch networks under self-assembly principles will be discussed.

4.1 Sparse CNT Networks with Implanted AgAu Nanoparticles

From investigating single NPs or vertical stacks of only a few NPs, in this section, the transition of NP-based memristive switching to the horizontal orientation is discussed. For achieving the nanoscaled distances required for memristive switching phenomena vertically, a variety of thin film deposition techniques are already well established. However, in a lateral orientation only sophisticated techniques like e-beam lithography [37, 38] offer the necessary resolution to produce nanoscaled gaps between electrodes, that allow for memristive switching. In this section, a new approach for the fabrication of lateral nanogaps as well as the memristive switching, when combined with bimetallic NPs described in the previous section, is discussed [39].

This approach is based on depositing carbon nanotubes (CNTs) between electrodes fabricated by conventional UV-lithography where the spacing between the electrodes is of several micrometers. The deposition is controlled such, that the CNTs form a sparse network and a subsequent Joule heating step removes any existing conductive pathways between the electrodes. The network then exhibits nanoscaled gaps between CNTs allowing for memristive switching phenomena. To fulfill its expected functionality, the CNT network has to meet certain requirements:

- The network must fill the space between the electrodes, so that all electrodes are in electrical contact with the CNT network.
- The CNTs must be finely dispersed, so that there are no dense agglomerations of CNTs, as those would impede the subsequent Joule heating step.
- The network must be below the percolation point but dense enough, so that the distance between individual tubes is in the nanometer range.

- The CNTs must not be heavily coated by surfactants or other additives, which would hinder removing short-circuiting paths during the Joule heating step.

Based on these requirements, a custom-made CNT dispersion has been developed by mixing ethanol with dry, pristine CNTs and a small amount of PEDOT:PSS. The mixture has been ultra-sonicated with a high-power sonicator to break up the CNT bundles and disperse them, whereas the PEDOT:PSS prevented a re-agglomeration of the CNTs and kept them finely dispersed. Despite the PEDOT:PSS acting as a surfactant, the low amount used is not detrimental to the Joule heating step. Using ethanol as a solvent allowed for a quick spin coating deposition in which single drops of dispersion are dropped subsequently onto the substrate under constant rotation, making the density of CNTs in the network well controllable. A sparse network of CNTs is shown in Fig. 7a. Conductive pathways in the CNT network, which have been short-circuiting the electrodes, were removed via Joule heating by applying voltages of up to 30 V.

AgAu NPs have been deposited onto the CNT networks with a Haberland-type gas aggregation source, which has been discussed in the previous section. The percolation point of the NPs deposition process has been determined by *in-operando* current measurements. It has been used to adjust the NP density to be just below the percolation point, so that the distances between particles are in the lower nanometer range, which is indicated in Fig. 7b. CNT networks with implanted AgAu NPs showed ECM-type memristive switching with a switching behavior exhibiting a hybrid of diffusive and bipolar characteristics. Current–voltage measurements are shown in Fig. 7c and d. Figure 7c depicts the diffusive switching mode, where the LRS is reset to the HRS when falling below a voltage threshold (cf. Figure 5). Additionally, Fig. 7d shows the capabilities for resetting the resistance state when cycled quickly into the reverse voltage regime, which is a feature of bipolar memristors.

A time resolved current measurement is shown in Fig. 7e, which indicates the retention time i.e. the time until the HRS is reached after reducing the applied voltage. Below the voltage threshold, diffusion and surface tension lead to a collapse of the silver filament. However, the thickness of the filament, and thus the amount of silver in it, defines its lifetime [13]. Therefore, it is proposed that the deposited NPs provide a limited silver reservoir for filament formation. The amount of silver atoms is limited such that no stable filament can be formed. But there is still sufficient silver to form filaments that are thick enough to show a substantial lifetime.

The retention time can also be described as the “memory span” of the device, as it remembers its resistance state. While in non-volatile memory applications, the retention time is supposed to be as high as possible to prevent data loss, a second-scale retention can be used for a short-term memory effect. Short-term memories are an efficient way of storing information that is only required for a limited amount of time, as it automatically forgets that information by itself again and thus does not have to be explicitly reset.

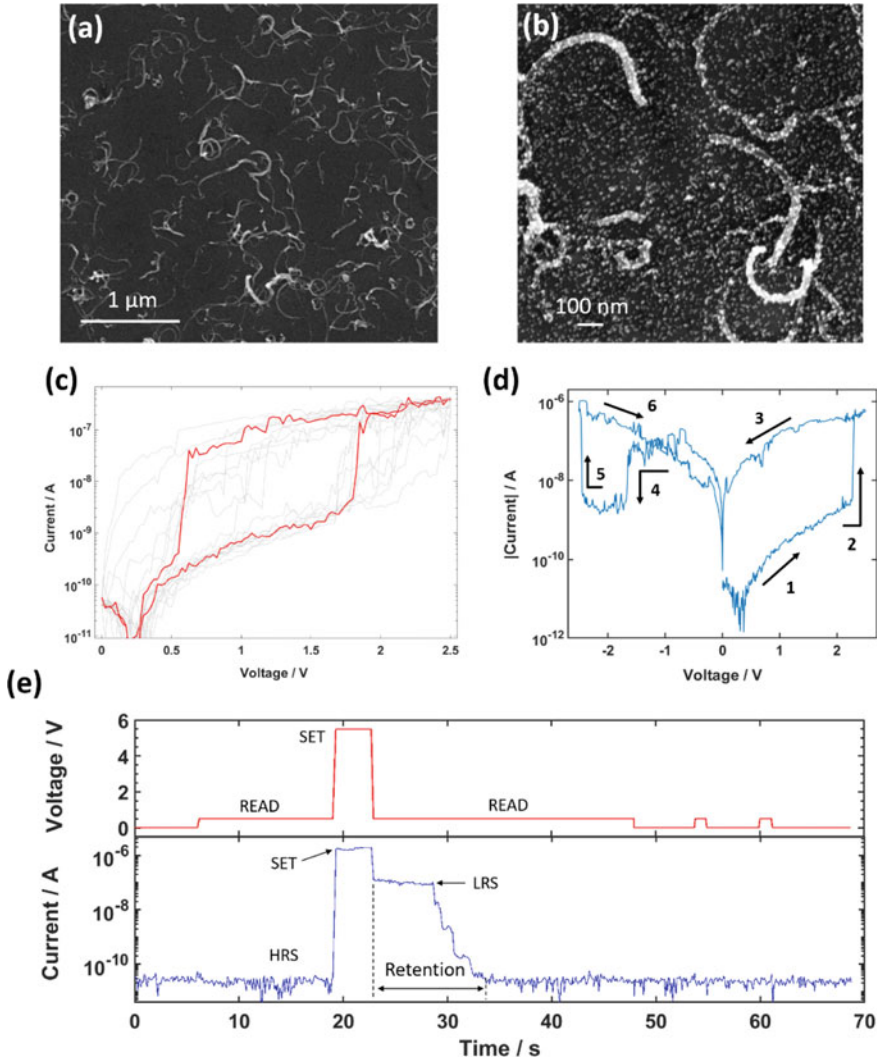


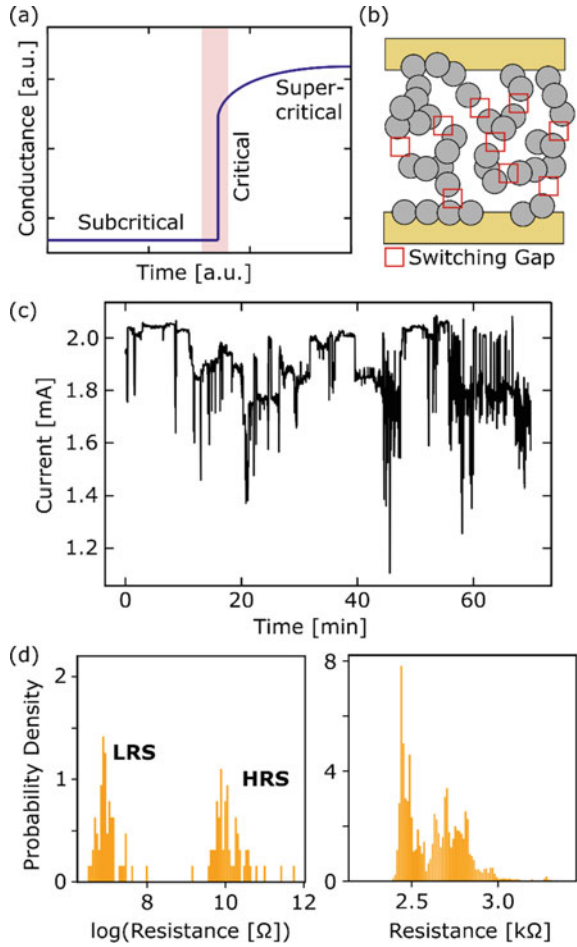
Fig. 7 Morphology and switching behavior of CNT networks with implanted AgAu NPs. **a** + **b** SEM micrographs of sparse CNT networks without **a** and with **b** deposited AgAu NPs. **c** Current–voltage cycles of a CNT network with NPs showing diffusive memristive switching behavior. **d** A current–voltage cycle showing the bipolar reset mechanism by applying reverse voltages (see arrow 4). **e** Time-resolved current measurement showing the retention of the LRS. Adapted under CC-BY 4.0 license from [39]

4.2 Nanoparticle Networks

In addition to the networks of the previous section, another approach for large-scale memristive switch network implementation, which solely makes use of metallic NPs for building the network, is discussed in this section. Elemental Ag-NPs synthesized from Haberland-type gas aggregation source, which are similar to the memristive bimetallic Ag-based NPs as discussed in Sect. 3, are applied as fundamental memristive building units in this type of network. Network formation is done by random assembly (realized by the statistical nature of the deposition process) of the Ag-NPs between chemically inert electrodes. The resulting network consists of a complex arrangement of clusters of electrically connected Ag-NPs and potential switching gaps between different clusters. The system scale is defined by the electrode separation and can be easily set to large-scale systems in horizontal electrode geometry. A NP network system with maximized complexity (in terms of maximized amount of metastable network configurations) can be realized through careful tuning of the NP filling factor at the percolation transition [24]. A schematic setup of a percolated NP network is shown in Fig. 8a. In the subcritical regime, no network response is possible, because of missing current paths having the scale of the overall system. In addition to that, in the supercritical regime, the network complexity is expected to be strongly reduced due to the existence of stable current paths that dominate the network response. However, it should be noted that NP networks in the supercritical regime potentially are still able to show a dynamic response which results from the re-configuration of interface defects (such as grain boundaries) between connected NPs, as it is reported for Au-NP networks above the percolation threshold [17]. Exactly at the percolation threshold, which marks the critical regime of a phase transition, the network configuration is located at some point between none and full connectivity and allows for ongoing transitions between a multitude of different metastable network configurations [40]. Figure 8b shows schematically a Ag-NP network under indication of potential switching gaps. The state of the network is accessed by the conductance between both electrodes, which is determined by the configuration of all switching gaps across the network. Upon an external disturbance, e.g. application of voltage stimuli across the electrodes, localized memristive switching at individual switching gaps may be triggered, which lead to a redistribution of potential differences across other switching gaps. As a consequence, memristive switching at a certain location could induce switching at other gaps in the network, resulting in a complex interplay of switching gaps distributed over the whole network. Figure 8c depicts the network response of a percolated Ag-NP network to a DC bias of 5 V. It can be seen, that the network dynamics are governed by ongoing transitions between a multitude of metastable conductance states over more than 1 h of operation. There is no indication that the network converges to a definable conductance state, which is in contrast to the lower-scale vertical NP-based devices in Sect. 3. Figure 8d shows the probability density of the observed conductance states for the vertical AgPt-NP devices from Sect. 3 (cf. Figure 6, right) and the percolated Ag-NP network. This comparison stresses out, that for the small-scale NP-based memristive devices well

differentiable conductance states (i.e. LRS and HRS) dominate the operation, which is not the case for the large-scale NP network. Instead, for the percolated NP network, the well differentiable conductance states are fading, which can be seen as an indicator for collective behavior of a multitude of switching junctions. Such a behavior is potentially interesting for neuromorphic computational concepts like reservoir computing, which require complex networks of interacting non-linear elements able to map input information into temporal dynamics in the network [41].

Fig. 8 Behavior of networks comprised of Ag-NPs and poised at the percolation threshold. **a** A schematic percolation curve showing classification of the network states into sub- and supercritical and critical regime depending on the filling factor. **b** is showing schematically a critical network state, where no global connectivity persists, but the network state is defined by the configuration of a multitude of switching gaps. **c** exhibits the network response upon application of a constant external bias of 5 V. **d** shows the distribution of distinct resistive states for a AgPt-multistack device as discussed in Sect. 3 (left) and extracted from the Ag-NP network response shown in (c) (right)



5 Conclusion

In this chapter, Ag-mediated reconfiguration of resistance states have been observed throughout different device implementations and system scales, ranging from individual filaments to complex assemblies of switching gaps. Fundamental research into the long-term switching dynamics of an individual Ag-filament via cAFM can be considerably facilitated through direct implementation of the memristive material system at the cantilever apex instead of conventional contacting, because it mitigates loss in data significance by thermal drift effects. This approach was exploited to study the cycle-to-cycle variability of filamentary diffusive switching dynamics. It was observed that the variability in the switching dynamics of an individual filament cannot be seen as purely stochastic, but potential correlation effects must be taken into account. Further, AgAu and AgPt alloy NPs have been found to be promising and versatile building blocks for diffusive memristive devices with a broad range of switching properties. Investigations via cAFM on individual Ag-based alloy NPs have shown, that already a single NP can act as a fundamental memristive building unit. Application of NPs for memristive devices opens up several design opportunities, such as inherent field enhancement at the surface of the NP, high localization of the memristive action at the location of the NP in contrast to bulk electrodes and precise regulation of the number of active species through NP size and composition, which is important to stabilize the diffusive switching regime of the device. Finally, two approaches to build large-scale memristive switch networks by self-assembly principles with a different degree of sparseness are discussed. The first approach is comprised of a sparse CNT network, which establishes a static network topology, and AgAu-NPs, which enable memristive switching between adjacent CNT sites in the network. Such hybrid CNT/AgAu-NPs networks were shown to exhibit a mixed form of diffusive and bipolar switching. The second approach treats networks only comprised of Ag-NP with a filling factor around the percolation threshold. The Ag-NP networks differ from their lower-scale counterparts regarding the non-existence of stable and definable conductance states. Instead of a clear differentiation into LRS or HRS, the networks dynamics of percolated Ag-NP networks are governed by ongoing transitions between a multitude of metastable states, which makes them interesting for neuromorphic computation schemes, where spatio-temporal mapping of information into complex patterns is required.

Acknowledgements The authors thank Dr. Sven Dirkmann and Prof. Thomas Mussenbrock for fruitful and stimulating scientific discussions and kinetic Monte-Carlo simulations, which deepened the understanding on electrical field distributions and resistive switching phenomena in nanoparticle-based memristive systems.

References

1. Jeong, D.S., Kim, K.M., Kim, S., Choi, B.J., Hwang, C.S.: Memristors for energy-efficient new computing paradigms. *Adv. Electron. Mater.* **2**, 1–27 (2016). <https://doi.org/10.1002/aelm.201600090>
2. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7** (2020). <https://doi.org/10.1063/1.5129306>
3. Zhang, T., Yang, K., Xu, X., Cai, Y., Yang, Y., Huang, R.: Memristive devices and networks for brain-inspired computing. *Phys. Status Solidi - Rapid Res. Lett.* **13**, 1–21 (2019). <https://doi.org/10.1002/pssr.201900029>
4. Yang, J.Q., Wang, R., Ren, Y., Mao, J.Y., Wang, Z.P., Zhou, Y., Han, S.T.: Neuromorphic engineering: from biological to spike-based hardware nervous systems. *Adv. Mater.* **32**, 1–32 (2020). <https://doi.org/10.1002/adma.202003610>
5. Zidan, M.A., Strachan, J.P., Lu, W.D.: The future of electronics based on memristive systems. *Nat. Electron.* **1**, 22–29 (2018). <https://doi.org/10.1038/s41928-017-0006-8>
6. Vahl, A., Carstensen, J., Kaps, S., Lupan, O., Strunskus, T., Adelung, R., Faupel, F.: Concept and modelling of memsensors as two terminal devices with enhanced capabilities in neuromorphic engineering. *Sci. Rep.* **9**, 1–9 (2019). <https://doi.org/10.1038/s41598-019-39008-5>
7. Wang, J., Zhuge, F.: Memristive synapses for brain-inspired computing. *Adv. Mater. Technol.* **4**, 1–20 (2019). <https://doi.org/10.1002/admt.201800544>
8. Wang, Z., Joshi, S., Savel'ev, S.E., Jiang, H., Midya, R., Lin, P., Hu, M., Ge, N., Strachan, J.P., Li, Z., et al.: Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nat. Mater.* **16**, 101–108 (2017). <https://doi.org/10.1038/nmat4756>
9. Midya, R., Wang, Z., Asapu, S., Joshi, S., Li, Y., Zhuo, Y., Song, W., Jiang, H., Upadhyay, N., Rao, M., et al.: Artificial Neural Network (ANN) to Spiking Neural Network (SNN) converters based on diffusive memristors. *Adv. Electron. Mater.* **5**, 1–7 (2019). <https://doi.org/10.1002/aelm.201900060>
10. Lim, E.W., Ismail, R.: Conduction mechanism of valence change resistive switching memory: a survey. *Electron.* **4**, 586–613 (2015). <https://doi.org/10.3390/electronics4030586>
11. Edwards, A.H., Barnaby, H.J., Campbell, K.A., Kozicki, M.N., Liu, W., Marinella, M.J.: Reconfigurable memristive device technologies. *Proc. IEEE* **103**, 1004–1033 (2015). <https://doi.org/10.1109/JPROC.2015.2441752>
12. Cha, J.H., Yang, S.Y., Oh, J., Choi, S., Park, S., Jang, B.C., Ahn, W., Choi, S.Y.: Conductive-bridging random-access memories for emerging neuromorphic computing. *Nanoscale* **12**, 14339–14368 (2020). <https://doi.org/10.1039/d0nr01671c>
13. Wang, W., Wang, M., Ambrosi, E., Bricalli, A., Laudato, M., Sun, Z., Chen, X., Ielmini, D.: Surface diffusion-limited lifetime of silver and copper nanofilaments in resistive switching devices. *Nat. Commun.* **10**, 1–9 (2019). <https://doi.org/10.1038/s41467-018-07979-0>
14. Wang, Z., Rao, M., Midya, R., Joshi, S., Jiang, H., Lin, P., Song, W., Asapu, S., Zhuo, Y., Li, C., et al.: Threshold switching of Ag or Cu in dielectrics: materials, mechanism, and applications. *Adv. Funct. Mater.* **28**, 1–19 (2018). <https://doi.org/10.1002/adfm.201704862>
15. Jiang, H., Belkin, D., Savel'Ev, S.E., Lin, S., Wang, Z., Li, Y., Joshi, S., Midya, R., Li, C., Rao, M., et al.: A novel true random number generator based on a stochastic diffusive memristor. *Nat. Commun.* **8** (2017). <https://doi.org/10.1038/s41467-017-00869-x>
16. Minnai, C., Bellacicca, A., Brown, S.A., Milani, P.: Facile fabrication of complex networks of memristive devices. *Sci. Rep.* **7**, 1–8 (2017). <https://doi.org/10.1038/s41598-017-08244-y>
17. Mirigliano, M., Decastri, D., Pullia, A., Dellasega, D., Casu, A., Falqui, A., Milani, P.: Complex electrical spiking activity in resistive switching nanostructured Au two-terminal devices. *Nanotechnology* **31** (2020). <https://doi.org/10.1088/1361-6528/ab76ec>
18. Bose, S.K., Mallinson, J.B., Gazoni, R.M., Brown, S.A.: Stable self-assembled atomic-switch networks for neuromorphic applications **64** (2017), 5194–5201
19. Stieg, A.Z., Avizienis, A.V., Sillins, H.O., Martin-Olmos, C., Aono, M., Gimzewski, J.K.: Emergent criticality in complex turing B-type atomic switch networks. *Adv. Mater.* **24**, 286–293 (2012). <https://doi.org/10.1002/adma.201103053>

20. Sillin, H.O., Aguilera, R., Shieh, H.H., Avizienis, A.V., Aono, M., Stieg, A.Z., Gimzewski, J.K.: A theoretical and experimental study of neuromorphic atomic switch networks for reservoir computing. *Nanotechnology* **24** (2013). <https://doi.org/10.1088/0957-4484/24/38/384004>
21. Du, C., Cai, F., Zidan, M.A., Ma, W., Lee, S.H., Lu, W.D.: Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 1–10 (2017). <https://doi.org/10.1038/s41467-017-02337-y>
22. Pike, M.D., Bose, S.K., Mallinson, J.B., Acharya, S.K., Shirai, S., Galli, E., Weddell, S.J., Bones, P.J., Arnold, M.D., Brown, S.A.: Atomic scale dynamics drive brain-like avalanches in percolating nanostructured networks. *Nano Lett.* **20**, 3935–3942 (2020). <https://doi.org/10.1021/acs.nanolett.0c01096>
23. Shirai, S., Acharya, S.K., Bose, S.K., Mallinson, J.B., Galli, E., Pike, M.D., Arnold, M.D., Brown, S.A.: Long-range temporal correlations in scale-free neuromorphic networks. *Netw. Neurosci.* **4**, 432–447 (2019). https://doi.org/10.1162/netn_a_00128
24. Mallinson, J.B., Shirai, S., Acharya, S.K., Bose, S.K., Galli, E., Brown, S.A.: Avalanches and criticality in self-organized nanoscale networks. *Sci. Adv.* **5**, eaaw8438 (2019). <https://doi.org/10.1126/sciadv.aaw8438>
25. Chekol, S.A., Menzel, S., Ahmad, R.W., Waser, R., Hoffmann-Eifert, S.: Effect of the threshold kinetics on the filament relaxation behavior of ag-based diffusive memristors. *Adv. Funct. Mater.* **32** (2022). <https://doi.org/10.1002/adfm.202111242>
26. Valov, I.; Waser, R.; Jameson, J.R.; Kozicki, M.N. Erratum: Electrochemical metallization memories—fundamentals, applications, prospects. *Nanotechnology* **22** (2011). <https://doi.org/10.1088/0957-4484/22/28/289502>
27. Yang, Y., Huang, R.: Probing memristive switching in nanoionic devices. *Nat. Electron.* **1**, 274–287 (2018). <https://doi.org/10.1038/s41928-018-0069-1>
28. Carstens, N., Vahl, A., Gronenberg, O., Strunskus, T., Kienle, L., Faupel, F., Hassanien, A.: Enhancing reliability of studies on single filament memristive switching via an unconventional cafm approach. *Nanomaterials* **11**, 1–16 (2021). <https://doi.org/10.3390/nano11020265>
29. Ke, J.J., Wei, T.C., Tsai, D.S., Lin, C.H., He, J.H.: Surface effects of electrode-dependent switching behavior of resistive random-access memory. *Appl. Phys. Lett.* **109** (2016). <https://doi.org/10.1063/1.4963671>
30. Lübben, M., Menzel, S., Park, S.G., Waser, R., Valov, I.: SET kinetics of electrochemical metallization cells: influence of counter-electrodes in SiO₂/Ag based systems. *Nanotechnology* **28** (2017)
31. Kim, H.J., Park, T.H., Yoon, K.J., Seong, W.M., Jeon, J.W., Kwon, Y.J., Kim, Y., Kwon, D.E., Kim, G.S., Ha, T.J., et al.: Fabrication of a Cu-Cone-Shaped cation source inserted conductive bridge random access memory and its improved switching reliability. *Adv. Funct. Mater.* **29**, 1–13 (2019). <https://doi.org/10.1002/adfm.201806278>
32. Choi, B.J., Torrezan, A.C., Norris, K.J., Miao, F., Strachan, J.P., Zhang, M.X., Ohlberg, D.A.A., Kobayashi, N.P., Yang, J.J., Williams, R.S.: Electrical performance and scalability of Pt dispersed SiO₂ nanometallic resistance switch. *Nano Lett.* **13**, 3213–3217 (2013). <https://doi.org/10.1021/nl401283q>
33. Vahl, A., Carstens, N., Strunskus, T., Faupel, F., Hassanien, A.: Diffusive memristive switching on the nanoscale, from individual nanoparticles towards scalable nanocomposite devices. *Sci. Rep.* **9**, 1–10 (2019). <https://doi.org/10.1038/s41598-019-53720-2>
34. Haberland, H., Karrais, M., Mall, M., Thurner, Y.: Thin films from energetic cluster impact: a feasibility study. *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.* **10**, 3266–3271 (1992). <https://doi.org/10.1116/1.577853>
35. Drewes, J., Vahl, A., Carstens, N., Faupel, F., Strunskus, T.: Enhancing composition control of alloy nanoparticles from gas aggregation source by in operando optical emission spectroscopy 1–11 (2021). <https://doi.org/10.1002/ppap.202000208>
36. Vahl, A., Strobel, J., Reichstein, W., Polonskyi, O., Strunskus, T., Kienle, L., Faupel, F.: Single target sputter deposition of alloy nanoparticles with adjustable composition via a gas aggregation cluster source. *Nanotechnology* **28**, (2017). <https://doi.org/10.1088/1361-6528/a66ef>

37. Puzyrev, Y.S., Shen, X., Zhang, C.X., Hachtel, J., Ni, K., Choi, B.K., Zhang, E.X., Ovchinnikov, O., Schrimpf, R.D., Fleetwood, D.M., et al.: Memristive devices from ZnO nanowire bundles and meshes. *Appl. Phys. Lett.* **111**, (2017). <https://doi.org/10.1063/1.5008265>
38. Le, P.Y., Murdoch, B.J., Barlow, A.J., Holland, A.S., McCulloch, D.G., McConville, C.F., Partridge, J.G.: Electroformed, self-connected tin oxide nanoparticle networks for electronic reservoir computing. *Adv. Electron. Mater.* **6**, 1–7 (2020). <https://doi.org/10.1002/aelm.202000081>
39. Terasa, M.I., Holtz, P., Carstens, N., Kaps, S., Faupel, F., Vahl, A., Adelung, R.: Sparse CNT networks with implanted AgAu nanoparticles: a novel memristor with short-term memory bordering between diffusive and bipolar switching. *PLoS One* **17**, 1–12 (2022). <https://doi.org/10.1371/journal.pone.0264846>
40. Shew, W.L., Plenz, D.: The functional benefits of criticality in the cortex (2013). <https://doi.org/10.1177/1073858412445487>
41. Tanaka, G., Yamane, T., Benoit, J., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., Hirose, A.: Recent advances in physical reservoir computing: a review. *Neural Netw.* **115**, 100–123 (2019). <https://doi.org/10.1016/j.neunet.2019.03.005>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Photocatalytic Deposition for Metal Line Formation



Salih Veziroglu, Moritz Paulsen, Jan Schardt, Blessing Adejube, Cenk Aktas, Alexander Vahl, and Martina Gerken

Abstract In neural systems, plasticity can be found throughout a variety of scales, ranging from local synaptic plasticity between two neurons towards long-range connections and global plasticity within larger neuron assemblies. While memristive devices have attracted a lot of attention as a potential neuromorphic analog to represent local synapses and are regarded as promising building blocks for neuromorphic engineering, long-range connections and globally mediated aspects like homeoplasticity are not yet widely considered for neuromorphic systems. In this chapter, photocatalytic deposition is discussed as an approach to form metallic structures from a global liquid reservoir. In this context, the photocatalytic properties of TiO_2 thin films are employed to reduce metallic species from the surrounding solution. This chapter will elucidate the fundamental process of photocatalytic deposition with photocatalytic TiO_2 thin films and will showcase the applicability towards the formation of metallic structures at the example of arrangements of locally grown metallic Au structures.

Keywords Photocatalysis · Self-assembly · Thin films · Strip conductor · Optical characterization

1 Introduction

Highly parallel systems such as neuronal networks combine processing and memory at the local synaptic level and are capable of completing complex tasks like pattern separation and completion with outstanding efficiency [1]. Neuromorphic engineering takes inspiration from neural networks, pursuing the aim of developing

S. Veziroglu · B. Adejube · C. Aktas · A. Vahl
Institute of Materials Science, Kiel University, Chair for Multicomponent Materials, Kaiserstr. 2, 24143 Kiel, Germany

M. Paulsen · J. Schardt · M. Gerken (✉)
Institute of Electrical Engineering and Information Engineering, Chair for Integrated Systems and Photonics, Kiel University, Kaiserstr. 2, 24143 Kiel, Germany
e-mail: mge@tf.uni-kiel.de

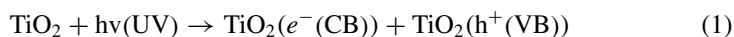
novel, neuron-inspired, efficient computing approaches. Memristive devices, due to their capability of unifying processing and data storage, have been studied broadly in the context of the development of novel computing paradigms and in the field of neuromorphic engineering [2]. Commonly, memristive devices are applied as an electrical representation of synapses and arranged in crossbar arrays, which enable interesting applications such as *in-memory* computing [3]. In the field of neuromorphic engineering however, new paradigms are necessary to capture the entirety of signal processing in neural networks [1, 4]. One important aspect in neural networks is their capability of dynamic on-demand reconfiguration of synaptic connections between the individual neurons. This is impressively shown by the onset of synaptic pruning, which describes the decrease of the number of connections between neurons in a fully developed brain. Accordingly, the development of biological neural networks happens on two different time scales: The fast synaptic plasticity (at the level of local synaptic connections between two neurons) and the slow blooming and pruning (at the global level throughout the neural network). While the first aspect, the synaptic plasticity, has attracted considerable research interest and memristive devices have been readily applied to mimic a variety of synaptic properties, the latter aspect of global plasticity is still under research and novel approaches are necessary to incorporate it into future bio-inspired hardware [4–6]. First approaches to replicate the global interactions of neuron assemblies are considering global connectivity through electrolyte gating, thus in a liquid medium [7]. In this chapter, metal line formation on a templated substrate from a liquid phase is discussed as an interesting approach with the potential to pave the way to achieve long-range, global plasticity.

The formation of a metallic lines on a template substrate requires a method that allows the formation of well-defined, localized metal structures. An effective approach to creating lines with high accuracy is the fabrication by localized reaction on a pre-structured substrate. A suitable compound is needed that can react in an aqueous medium to cause deposition of the metal atom in such a manner to have a localized growth of the metal on the surface of a template. TiO_2 is a metal oxide which is known for its photoreduction ability and high photocatalytic property. In the presence of UV light, photoreduction of an electro positive metal ion will occur when TiO_2 is present in the medium. For example, Dawson et al. showed that HAuCl_4 can be reduced on surface of TiO_2 nanoparticles to form gold capped nanocomposites with TiO_2 core [8]. It was discussed that the concentration of the TiO_2 core influenced the nanocomposite size and stability. Having a thin film of pre-structured TiO_2 on a surface could serve as a base for gold nanoparticle growth when placed in HAuCl_4 aqueous medium. This process is based on photocatalytic reduction of the gold precursor ions by TiO_2 which would then result in the immobilization of the Au nanoparticles on areas patterned with TiO_2 forming a structured metallic growth on the substrate. Thus, understanding the photocatalytic process involved is thus vital.

Semiconductor-based photocatalysis covers a wide range of cutting-edge applications in some important areas such as energy, environment, hygiene and disinfection [9, 10]. Principally, photocatalysis is described as the integration of photoexcitation and surface catalysis. The photoexcitation contains light absorption and charge output, while surface catalysis concerns the utilization of photoinduced charge

carriers (electron (e^-) and hole (h^+)) for reduction or oxidation reactions [11]. This process is highly dependent on various parameters, including the photocatalytic activity of semiconductors, light wavelength, and intensity as well as reaction temperature [12]. Recently, semiconductor-based photocatalysts (TiO_2 , ZnO, WO_3 , etc.) have gained considerable attention due to their high potential for solving environmental and energy problems [13]. However, these photocatalytic materials can be also considered to easily synthesize metallic or metal oxide structures (forming hybrid structures) for different kinds of applications (hydrophobic and hydrophilic coatings, sensor devices, reconfigurable connections, etc.) rather than only using them for environmental and energy applications [14].

In the recent decade, titanium (IV) oxide (TiO_2) is one of the most widely studied semiconductor photocatalysts due to its significant properties such as its strong photo-oxidizing power, non-toxicity, very good stability over a wide range of pH, chemical inertness, low-cost preparation as well as photocatalytic properties resulting from its bandgap energy and positions of the conduction band (CB) and valence band (VB) [15–17]. An overall mechanism for the photocatalytic reaction on TiO_2 surface is categorized as two main processes such as (i) the reduction of oxygen molecules (O_2) and (ii) oxidation of water molecules (H_2O) by photoinduced charge carriers (e^- and h^+) [18]. Initially, the photocatalytic reaction is started when a photoelectron is excited from the filled VB of TiO_2 to the empty CB as a result of irradiation. The absorbed photon must have the energy either equal or greater than the bandgap of TiO_2 (3.0–3.2 eV) to excite the electron in the VB. Then, the excitation process leaves behind a hole in the VB [19]. Here, as a net result, electrons (e^-) and holes (h^+) pair are generated by the Eq. (1) below.



Afterward, the photogenerated holes at the VB react with a water molecule to generate hydroxyl (OH^*) radicals. These hydroxyl radicals are extremely powerful oxidizing agents, which attract adsorbed organic molecules to mineralize them depending upon their structure and stability level [19]. While the photogenerated holes react with water to generate the hydroxyl radical, at the same time, the electrons in the CB are taken up by oxygen (O_2) molecules to produce anionic superoxide radical (O_2^{*-}). These superoxide radicals are formed into hydroperoxyl (HO_2^*) radicals and the subsequent hydrogen peroxide (H_2O_2), which further dissociates into highly reactive hydroxyl (OH^*) radicals. All these oxidation and reduction processes commonly take place on the surface of the TiO_2 photocatalyst [19]. Therefore, the surface area of the TiO_2 is highly important to achieve the high photocatalytic activity for desired application.

Recent studies have been revealed that the particle size of TiO_2 plays a crucial role in photocatalytic activity [20]. Decreasing the particle size to the nanoscale (increasing active surface area) leads to a higher photocatalytic activity. However, the conventional nano-powdered photocatalysts need post-treatment separation in a slurry system after the photocatalytic reaction. This problem can be overcome by

immobilizing TiO₂ particles as a robust and stable thin film on solid surfaces [21, 22]. However, thin films have a limited surface area in comparison to nanoparticle systems and they show limited photocatalytic activity, which has to be improved by further strategies such as (i) tailoring morphology and crystal structure, (ii) decorating/depositing noble metals (Au, Ag, Pt, etc.) and (iii) coupling metal oxide semiconductors (ZnO, CeO₂, etc.) [23].

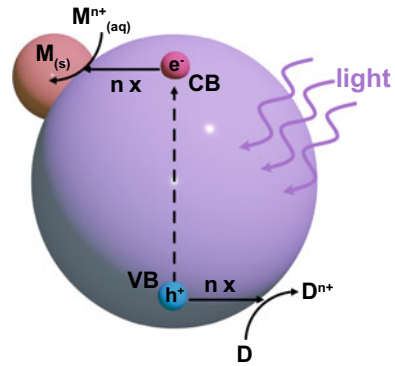
Noble metal nanoparticles (Au, Ag, Pt, etc.) deposited on semiconductor surfaces have been extensively studied due to their unusual (photo)catalytic and optoelectronic properties [24] as well as their potential to tailor the sensor properties of semiconducting metal oxide gas sensors [11]. In particular, TiO₂ decorated by noble metal nanoparticles like Au and Ag has received more and more attention because of its significant improvement in the photocatalytic activity of TiO₂ for specific applications [22]. The size and distribution of these noble metal nanoparticles are the most effective parameters on the activity of photocatalyst [24]. Additionally, the photocatalytic performance of TiO₂ modified by noble metals highly depends on some external factors such as strong contact with metal and support [25]. Noble metals can be deposited on TiO₂ surface by various methods including sputtering, physical mixing, chemical reduction, electrodeposition, photocatalytic reduction, and so on. However, among the mentioned methods, the photocatalytic deposition (sometimes it is called photodeposition or photoreduction) is a facile, simple, and low-temperature process with efficient metal-support interaction especially for TiO₂ thin film as support [26].

The photocatalytic deposition method is based on the photocatalytic properties (bandgap, activity, etc.) of semiconductor materials. Generally, several conditions are needed to allow the deposition of metals on the semiconductor surface. First, the photon energy of the light (solar or artificial light) should be larger than the energy bandgap of the semiconductor. Mostly, photocatalytic deposition of metal to TiO₂ employs under artificial ultra-violet (UV) light illumination due to TiO₂'s wide band gap, which is around 3.0 eV. When the light has proper energy, the light is absorbed by a semiconductor, and electrons in its VB are excited to its CB. These electrons are used for the reduction process. Second, the CB energy level of the semiconductor should be more negative than the reduction potential of the metal ion (M⁺). Third, the efficient charge (e⁻ and h⁺) separation and migration, by electron donor, is necessary to continue the reaction [27]. Finally, the semiconductor acts as a template for metallic structures as shown in Fig. 1.

The depositing of noble metals on semiconductor surface by photocatalytic deposition is possible, if all conditions, which are mentioned above, are available. However, various parameters are also needed to be considered to allow the well-controlled (precise shape, size, and distribution) photodeposition process to occur on semiconductor surfaces such as sacrificial reagent, pH, temperature, metal precursor, light exposure time and intensity, absence, or presence of oxygen in the media [14, 28, 29].

In this chapter, photocatalytic deposition is discussed as an approach to form metallic structures from a global liquid reservoir. In this context, the photocatalytic

Fig. 1 Schematic illustration for photodeposition of metal structures on semiconductor surface (M: Metal and D: Electron donor)



properties of TiO_2 thin films are employed to reduce metallic species from the surrounding solution. This chapter will elucidate the fundamental process of photo-reduction with photocatalytic TiO_2 thin films and will showcase the applicability towards the formation of metallic structures at the example of arrangements of locally grown metallic Au structures.

2 Reactive Sputtering of Photocatalytic TiO_2 Thin Films

TiO_2 is typically characterized as a n-type semiconductor and has three crystalline phases such as anatase (tetragonal), rutile (tetragonal), and brookite (orthorhombic) [30]. In general, anatase shows higher photocatalytic activity than rutile and brookite due to mainly its indirect bandgap structure (low electron–hole recombination rate). Therefore, most of the researchers focused on preparing of mainly anatase TiO_2 structures for achieving high photocatalytic activity [31]. However, recent studies have shown that the mixture of anatase and rutile has a synergetic effect on decreasing the recombination rate of photogenerated electron–hole pairs, which boosts the overall photocatalytic performance [32].

As mentioned before, TiO_2 is extensively employed in various applications due to its unique optical, electronic, and photocatalytic properties. In these applications, it is mainly used as powder form (nanoparticles around 25 nm size), which shows high photocatalytic activity results from the high surface area. However, the separation of these tiny particles from the reaction media after the end of the reaction is highly difficult with conventional filtration systems [33]. Recently, TiO_2 thin films have gradually replaced conventional TiO_2 powders due to separation and agglomeration problems, especially for their long-term use.

Various deposition methods have been utilized to prepare TiO_2 thin films such as sol–gel dip coating, sol–gel spin coating, chemical vapor deposition (CVD), physical vapor deposition (PVD), electron beam, spray pyrolysis method [34]. Among these deposition methods, PVD methods have been extensively used because they

can easily control the composition and morphology of the prepared thin film. Various PVD methods are being utilized for the deposition of TiO_2 thin film such as thermal evaporation, pulsed laser deposition, DC, and RF magnetron sputtering [35]. DC magnetron sputtering is one of the most preferred methods because of its good adhesion, high deposition rate, high quality, and uniform film [36]. In this deposition method, there are many vital parameters (sputtering power, distance between substrate and target, argon/oxygen ratio, working pressure, substrate temperature, and annealing treatment) to optimize the quality of the thin film. However, it is still a challenge to achieve highly photocatalytic active TiO_2 thin films, which compete with TiO_2 nanoparticles, due to their limited surface area [33]. To achieve a high surface area, nowadays an effective process is established, which is called as nano-crack network formation, within sputter-deposited TiO_2 thin films. For example, Henkel et al. reported on thermally controlled nano-crack formation as a possible method to improve the photocatalytic performance of well adhering, reactively DC sputtered TiO_2 thin film [37]. Networks of nanoscopic cracks were produced into tailored columnar TiO_2 thin films by thermal annealing as shown in Fig. 2.

It can be easily seen from Fig. 2a, irregular and narrow columns (underlined by blue and red marks, respectively) dominate the formation of the film at the TiO_2 -substrate interface. According to the author's observations, broader columnar structures start to grow at a distance of 200 nm from the substrate and become more dominant features at TiO_2 thin film surface. However, after thermal annealing, the deep trenches separate into small bundles of TiO_2 columns, which enhance the overall

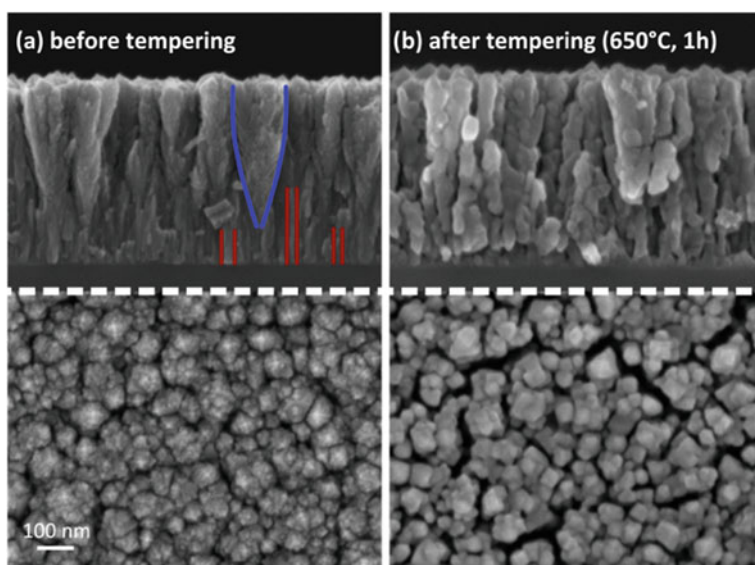


Fig. 2 SEM images of sputtered TiO_2 thin films, **a** before and **b** after thermal annealing (at 650 °C for 1 h), important morphological features are marked by red and blue lines. (Printed with permission from [37]. Copyright (2018) IOP)

photocatalytically active surface area (Fig. 2b). Additionally, they also reported that the thermal annealing step (at 650 °C for 1 h) is highly crucial not only for the formation of the nano-crack network but also the transformation of amorphous phase (as-deposited TiO_2) to the mixture of anatase and rutile phase (anatase is more dominant), which is extremely beneficial for enhancing photocatalytic activity. Similarly, Zhan et al. observed that the correlation between annealing temperature and phase transformation of TiO_2 from amorphous phase to anatase (>300 °C), partially rutile (around 600–800 °C), and fully rutile phase (at 1000 °C) [34, 38].

Nano-crack network formation is a promising approach to achieve a high surface on the thin film surface. It can be one of the biggest features on competition between thin film and nanoparticles photocatalyst. Ghori et al. demonstrated that sputter-deposited TiO_2 film with nano-crack networks shows extremely high photocatalytic activity in comparison to a reference TiO_2 thin film prepared by immobilized TiO_2 (Degussa P25) nanoparticles [21]. Here, the authors reported that sputtering of TiO_2 at high oxygen partial pressure and low deposition rate leads to the formation of high aspect ratio structures with weak inter-columnar bonding, which can be transformed to a crack network by thermal annealing process (Fig. 3a).

As a comparison, they prepared the reference thin film by spin-coating. The prepared thin film has a porous structure composed of commercial spherical TiO_2 nanoparticles, which are well distributed and maintained their original size (25–30 nm) and shape (Fig. 3b). The thickness of the sputter-deposited and the reference TiO_2 thin films are around 400–450 nm for a reliable comparison. The photocatalytic activity of both TiO_2 thin films was investigated by degradation of methylene blue (MB) aqueous solution under UV illumination. The results showed that the sputter-deposited TiO_2 film degrades 90% of MB in 4.95 min while the reference TiO_2 thin film needs about 25 min (Fig. 3c).

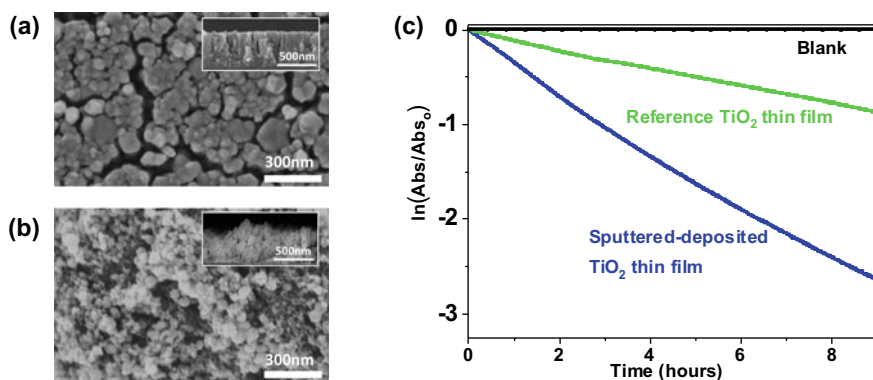


Fig. 3 SEM images of **a** sputter deposited TiO_2 film after thermal annealing and **b** spin-coated reference TiO_2 film. (Insets show cross-sectional SEM micrographs) **c** Normalized degradation (plot) of MB solution in the presence of sputter-deposited and reference TiO_2 thin films under UV illumination. (Quartz substrate (given as blank) is used as the control substrate). (Printed with permission from [21]. Copyright (2018) Elsevier)

The thermal annealing process is an easy and reliable process for phase transformation and formation of a nano-crack network for better photocatalytic performance as aforementioned. However, some additional process parameters can be also controlled for high photocatalytic performance. Generally, crystalline TiO_2 thin film is deposited on the substrate at low operating temperatures. Here, the crystallinity of TiO_2 thin film can be mainly tuned by changing total pressure and oxygen partial pressure [39]. The total gas pressure affects the kinetic energy of sputtered atoms, which increases the probability of collisions and acceleration of particles and subsequently the particle energy [40]. Therefore, the anatase phase forms at high total pressure when the rutile phase forms at low total pressure [40, 41].

During magnetron sputtering, argon (Ar) and oxygen (O_2) are used as the plasma gas and the reactive gas, respectively. Oxygen plays a significant role in the deposition of TiO_2 thin film via magnetron sputtering especially for reactive sputtering. Many researchers have pointed out that the oxygen concentration during the magnetron sputtering might affect the formation of oxygen vacancies in the deposited thin film [41]. Additionally, the oxygen partial pressure (the oxygen/argon ratio) has a huge effect on the discharge parameters, such as plasma potential, discharge voltage, and ion composition of discharge [42]. For example, Zhang et al. reported that the discharge voltage increases with increasing oxygen flow rate up to a certain threshold. Before the threshold is reached, a mixture of metallic Ti and TiO_2 is deposited on the substrate. The uniform TiO_2 thin film forms on the substrate only after the oxygen flow rate reaches this certain threshold [34, 43].

Although numerous attempts have been performed to investigate the relationship between the oxygen/argon ratio and photocatalytic activity of TiO_2 thin film, no clear agreement has been reached so far. However, a lot of properties of TiO_2 thin film are affected by oxygen/argon ratio such as deposition rate, grain size and surface roughness, surface chemical composition, and optical parameters, which might directly or indirectly influence the photocatalytic activity [44, 45]. For example, Vahl et al. reported the availability of oxygen during reactive sputtering is a vital parameter for tailoring of thin film morphology for high photocatalytic performance (Fig. 4) [46].

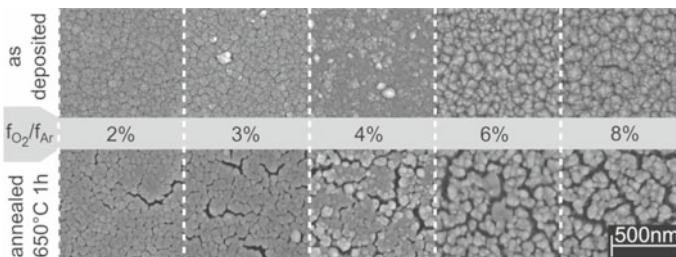


Fig. 4 Top-view SEM images of TiO_2 thin films for different oxygen/argon flow ratios, before (top row) and after (bottom row) thermal treatment (Printed with permission from [46]. Copyright (2019) IOP)

The authors reported that the formation of nano-crack networks is more noticeable for TiO₂ thin films deposited at higher oxygen/argon ratios. The reason for this might be the presence of crystalline anatase seeds for high oxygen partial flows. These seeds grow by thermal annealing process more effectively as a well-defined nano-crack network. Here, the annealed TiO₂ thin films deposited different oxygen/argon ratios showed different activity on degradation of MB under UV illumination. After 7 h UV illumination, the TiO₂ thin film deposited at 3% oxygen/argon ratio degraded almost 81.2% of MB when this value is 96.1% for TiO₂ thin film deposited at 6% oxygen/argon ratio.

3 Deposition of Metal Structures from Photocatalytic Reduction

As discussed in the previous chapters, various strategies have been proposed to improve the volume to surface area ratio of TiO₂ thin films for a high photocatalytic activity. Additionally, TiO₂ thin film incorporated with noble metal nanoparticles can be considered another strategy to enhance the (photo)catalytic properties by means of tuning the Fermi level of TiO₂ and acting as an electron sink [47]. The prepared noble metal/TiO₂ hybrid structure has a huge potential on not only environmental applications but also different kinds of applications including hydrophobic and hydrophilic coatings, sensor devices, reconfigurable connections, etc. [13].

In the literature, there are numerous methods, which can be used to synthesize the noble metal/semiconductor nanocomposites such as hydrothermal, droplet, precipitation, electrochemical methods, and so on [27]. The photodeposition is one of the simplest methods to prepare the noble/semiconductor hybrid structure by using only the bandgap of the semiconductor [48]. This method needs only the irradiation of a light source (artificial or sunlight) rather than requiring additional reducing agents, high temperature, multi-step processing, etc. [49]. Furthermore, the photodeposition method is easily applicable to 2D substrates (such as thin film) by adjusting the concentration of the (noble)metal precursors, irradiation intensity, and duration [50]. For example, Mendoza-Diaz et al. demonstrated that Au nanoparticles were deposited on columnar TiO₂ structure by photodeposition process under UV irradiation (365 nm, 100 W, 30 min) by using an Au precursor solution (Fig. 5) [51].

It can be easily seen from the SEM images in Fig. 5a, the TiO₂ thin film has a columnar morphology with grain sizes from ~20 to 50 nm. This columnar morphology is generally observed in TiO₂ films, which are deposited by chemical and physical vapor deposition techniques [52, 53]. For the Au-on-TiO₂ sample, Au nanoparticles were deposited on top of the TiO₂ thin film surface (Fig. 5b). The photodeposited Au nanoparticles have a spherical shape with a size range of 30–70 nm (inset in Fig. 5b). The authors reported that the size dispersity of the photodeposited nanoparticles is closely related to the geometrical restraints enforced by the roughness of the columnar TiO₂ structure [51].

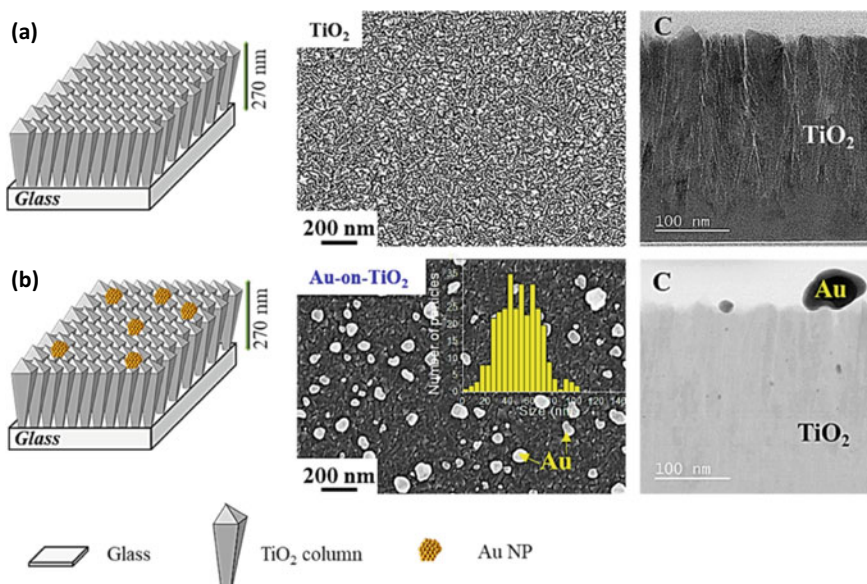


Fig. 5 (Left) Schematic illustrations of TiO₂ and TiO₂/Au hybrid structures. (Middle) SEM images of the sample surface area. (Right) HAADF-STEM cross-sectional images: **a** TiO₂ thin film; **b** Au-on-TiO₂ (inset: Au NP size dispersion in yellow). Letter C corresponds to the top layer of protective sputtered carbon. (Printed with permission from [51]. Copyright (2020) American Chemical Society)

Similarly, Veziroglu et al. showed that not only the TiO₂ morphology but also the composition of the precursor solution has a significant effect on the size and distribution of the photodeposited Au structure as demonstrated in Fig. 6 [28]. Figure 6a shows that only a few Au nanoclusters (surface coverage of 9.2%) were photodeposited on TiO₂ surface water was used as a solvent. However, densely distributed Au nanoclusters (surface coverage of 51.6%) on TiO₂ were obtained when a mixture of 1-hexanol–water (v/v: 20/ 80) was used (Fig. 6d).

Here, one can easily see differences in morphologies given in helium ion microscopy (HIM) images (Fig. 6a–d), the solvent type (or composition of the precursor solution) significantly affects the final morphology of the photodeposited Au nanoclusters. Because, during the photodeposition process, the TiO₂ thin film contact with the Au precursor solution, which forms the solid–liquid interface. This interface plays a crucial role in the electron transferring from a photocatalyst surface to electron-acceptor species (e.g., Au³⁺ ions) in the solution. When the composition of the precursor solution changes, the distribution of the electrical potential also changes [54]. This phenomenon creates various pathways to reduce Au³⁺ ions on TiO₂ as metallic Au nanoclusters, which form different shapes, sizes, and distribution profiles.

Similarly, the pH of the precursor solution also affects the charge distribution in the solid–liquid interface. Unfortunately, according to the author’s best knowledge,

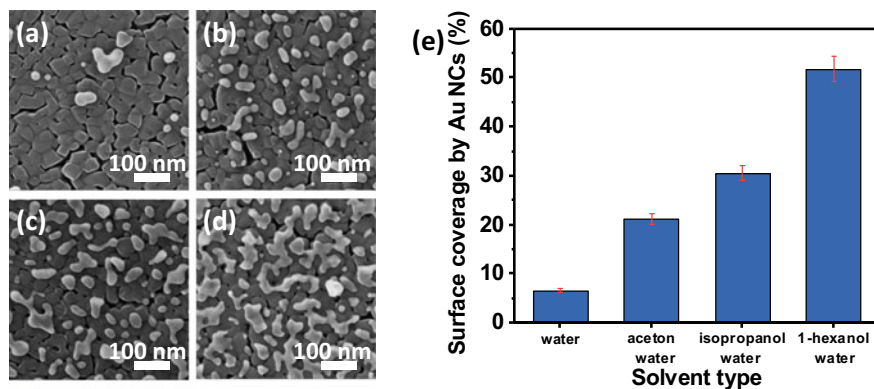


Fig. 6 Helium ion microscopy (HIM) images of photodeposited Au nanoclusters onto TiO_2 thin film surface in **a** water, **b** acetone–water (v/v: 20/80), **c** isopropanol–water (v/v: 20/80), and **d** 1-hexanol–water (v/v: 20/80) mixtures. **e** Surface coverage (%) of photodeposited Au nanoclusters with different solvents. (Printed with permission from [28]. Copyright (2020) American Chemical Society...)

there is not any systematic research related direct effect of the pH on photodeposited Au structure especially on TiO_2 thin film surface, yet. However, similar studies can give us an overall understanding of the pH effect on the photodeposited particles. For example, Guo et al. reported that pH plays a significant role in the nucleation position of the photodeposited Au nanoparticles on to plate-like BiOBr [55]. It can be seen in Fig. 7, SEM images after the photodeposition process show facet-selective deposition of Au particles with pH-dependency. At pH 3, they observed an overall deposition of Au nanoparticles, favorably on the top ([001] facet) as shown in Fig. 7a. However, when increasing the pH of the precursor solution to 3, 5, and 9, the pattern of the photodeposited nanoparticles gradually changes (Fig. 7a–c). For example, Au nanoparticles are founded on every facet ([001] and [102], top and side, respectively) at pH 5 while only on the side ([102] facet) at pH 9.

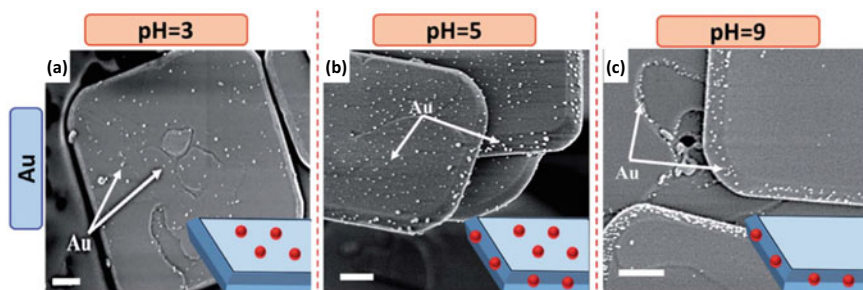


Fig. 7 SEM images of BiOBr showing pH-dependent geometrical deposition of Au nanoparticles. Au/ BiOBr at **a** pH 3; **b** pH 5; **c** pH 9. The scale bar always corresponds to 200 nm. (Printed with permission from [55]. Copyright (2018) Royal Society of Chemistry)

The authors reported that only at a low pH value, the charge density of the semiconductor and precursor interface is weak. Therefore, [001] facets are weakly positively charged, which might significantly reduce the recombination of the photoinduced charge carriers (electron and holes). This phenomenon may enhance the electron/hole mediated reaction on the semiconductor surface during the photodeposition process. However, at a high pH value, all facets are rather strongly negatively charged. Therefore, electrons are no longer attracted towards the [001] facet surface. They are no longer available for the reduction of Au ions on the surface. Hence, the facet-selectivity of the photodeposition reaction of Au ions to metallic Au nanoparticles reverses from the [001] facet to [102] facet at a high pH value [55].

As already mentioned, the composition and pH of the precursor solution have a huge effect on the size, distribution, and position of the photodeposited nanoparticles on the semiconductor surface due to various charge distributions in the solid–liquid interface during the photodeposition process. Additionally, the photocatalytic activity of the semiconductor affects the final morphology of the deposited particles on the surface because it limits how many electrons can be generated by the photocatalyst under light illumination. For instance, Veziroglu et al. demonstrated that hierarchical Au needle clusters (HAuNCs) were deposited on a highly active TiO₂ thin film surface via UV illumination (Fig. 8a–b). Here, the size and the geometry of deposited HAuNCs were controlled by simply altering the photocatalytic activity of the TiO₂ (depends on the crystal structure), UV light intensity, and irradiation time [14]. The basic growth mechanism of a HAuNC is semantically shown in Fig. 8c. First, the photoinduced charge carriers (electrons and holes) in TiO₂ thin film are generated by UV illumination. These electrons reduce the Au³⁺ ions into a stable Au cluster. This first step is almost same with all conventional photodeposition process for different kind of metallic structures [56]. However, a high density of electrons can be generated by the prepared highly photocatalytic active TiO₂ thin film. Therefore, the relative decrease of Au³⁺ ions on the surface, where the first embryonic Au cluster nucleate, occurs in a very short time. This promotes the diffusion of more Au³⁺ ions (from higher to lower concentration regions), which is followed by preferential piling of Au nanoclusters into needle-like structures (Fig. 8a–b). It seems that the high density of electrons triggers the directional growth of Au nanostructures (Growth). These dense electrons, which are generated by TiO₂ thin film, can be trapped by former Au nanoclusters (nuclei) and they act as the further nucleation sites for incoming Au³⁺ ions from the precursor solution (Needle formation).

Additionally, the authors reported that UV light intensity plays a major role in the final size and sharpness of the deposited Au structures during the deposition process. Because the UV light intensity also decide the number of photogenerated electrons by TiO₂, which leads to quick nucleation and a high growth rate. This is a piece of key information about controlling the nucleation point during the photodeposition process. For this purpose, they used the polymer mask between the light source and the sample. Here, patterned HAuNC structures on TiO₂ thin film surface were obtained by selective light illumination as shown in Fig. 8d–e. These patterned HAuNC structures find some applications especially for catalysis, plasmonic, and biomedical technologies.

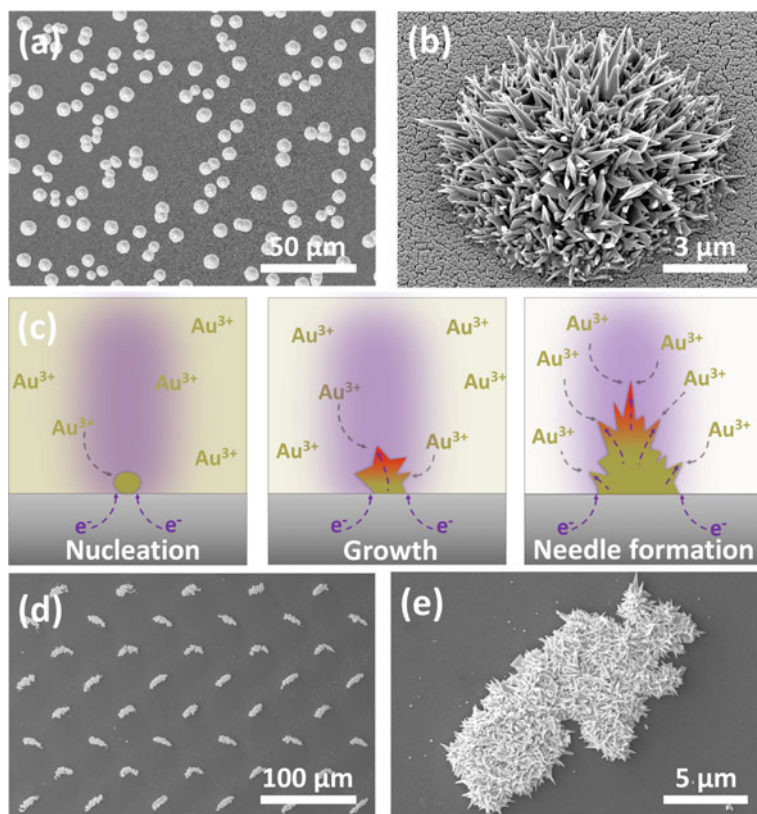


Fig. 8 **a** SEM image of HAuNCs deposited on TiO₂. **b** HIM image of a single HAuNC. **c** Schematic representation of HAuNC growth mechanism. SEM images of **d** periodic HAuNC arrays and **e** a single HAuNC. (Printed with permission from [14]. Copyright (2020) Wiley)

4 Metal Line Formation by Photocatalytic Reduction

As discussed in the previous section, localized irradiation of a continuous TiO₂ thin film with the use of a shadow mask results in the formation of metallic nano- and microstructures from the precursor solution. In the following, a second approach to obtain localized formation of metallic structures will be discussed. In essence, this approach is based on the image reversal lithography process. The general process scheme is depicted in Fig. 9. The process is similar to the positive resist lithography except that in this case, a chemical modification step is included to ensure the cross linking of the resist. Also, the process involves a double exposure of the resist which causes the positive resist to act like a high-resolution negative resist.

The procedure involves: cleaning and priming of the glass wafer, spin coating of the resist, soft bake (110 °C for 50 s), exposure, hard bake (120 °C for 2 min), flood exposure, development, inspection, TiO₂ deposition, lift off, final inspection.

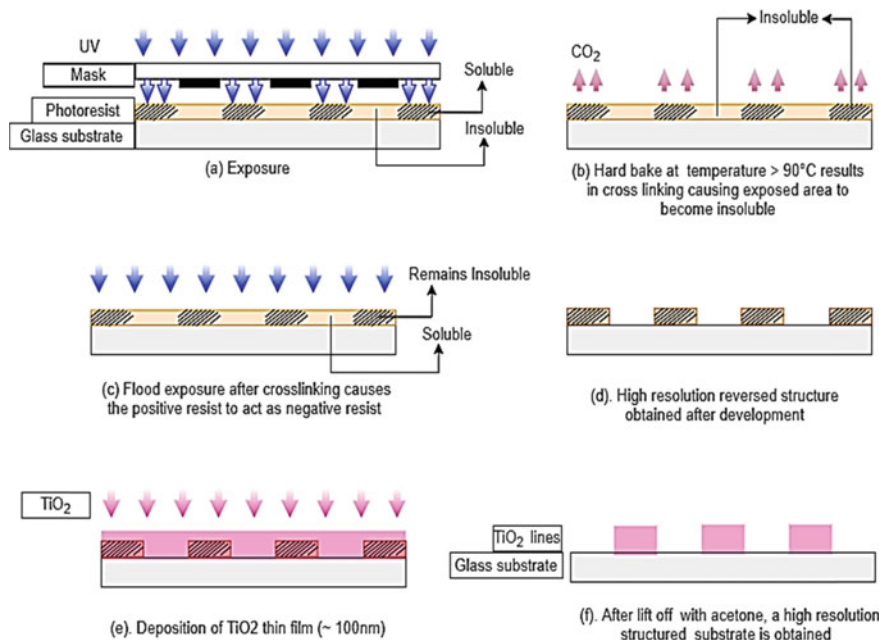


Fig. 9 An illustration of the major steps in the image reversal process necessary to obtain structures with high resolution

The image reversal technique is preferred because the idea is to have the reversed structure of the mask transferred on to the glass substrate. The TiO_2 which is then later deposited directly on to the glass surface takes the pattern of the mask while everywhere else is covered with resist. After the lift off (using acetone) remains on the substrate. With this approach, very thin lines of TiO_2 are transferred to the glass substrate. It allows for fabrication of quite small structures on a substrate. For example, TiO_2 lines of a few microns in width is structured on a glass substrate with sufficient accuracy.

The substrates with lithographically structured TiO_2 thin films are consecutively used as substrates for photocatalytic deposition of metallic gold from precursor solution. For in-situ monitoring, the absorption and reflection of the photocatalytically grown gold on the titanium dioxide surface is taken as a measure. A standard transmission microscope (Leica DMi8) setup is used with an external UV-LED to stimulate the photocatalytic process as shown in Fig. 10.

A spectrograph extension of the setup is used to record the bright-field and UV-LED spectra. The UV-LED has its peak at approximately 365 nm, while the bright field source has no energy in the UV regime and is thus suitable for the in-situ monitoring. The intensity of the UV-light is controlled by the distance to the sample and was determined with a power meter (Newport 2936-C) for different distances. Samples are positioned in a beaker filled with HAuCl_4 solution. Images are taken in regular intervals and are computed to obtain growth dynamic graphs of the process.

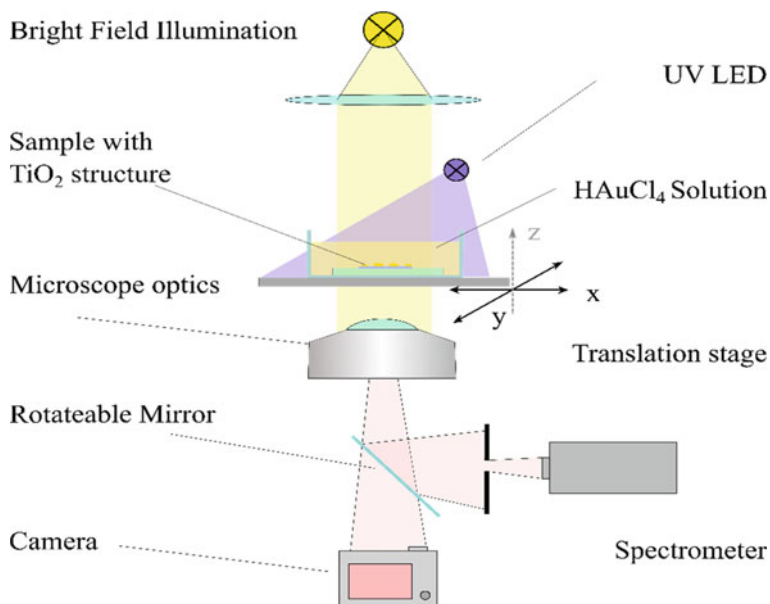


Fig. 10 Transmission microscope setup for in-situ monitoring of photocatalytic gold growth. Photocatalytic gold growth is excited by an external UV-LED. A standard microscope setup (Leica DmI8) with bright-field illumination is used to track the growth by recording images during the growth process

In a first experiment Au is grown using a $0.736 \cdot 10^{-3}$ mol/l HAuCl₄ solution on a sample covered partially with a titanium dioxide layer for different illumination intervals. Figure 11 shows three example transmission images of the sample. On the left side there is no TiO₂, while the right side is covered with titanium dioxide. Due to photocatalytic Au growth on the TiO₂ the transmission intensity is reduced. The total normalized transmission was 0.72 after completion of the total duration of the experiment.

Within the MATLAB environment different areas of the images are selected and evaluated over time. First the RGB channels as well as all pixels are summed up for these areas and are subsequently normalized to the first pictures, which are regarded as full normalized transmission for every measurement. The normalization accounts for different loss mechanisms such as thin-film interferences, absorption and scattering in the solution and the beaker. In the following, the three red reference areas and the three blue areas on the titanium dioxide are considered in more detail. The time evolution of the transmission for these six areas is depicted in Fig. 12. Time intervals of two minutes without illumination are followed by five minutes of UV illumination. Two illumination levels are employed—a lower excitation intensity of 1.8 mW/cm^2 for the first two intervals and a higher intensity of 2.7 mW/cm^2 for the third illumination phase.

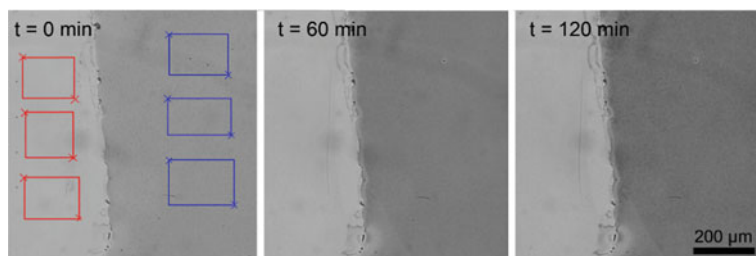


Fig. 11 Bright-field microscope transmission images at three different times during photocatalytic gold growth for a sample half-covered with TiO_2 . The computation areas are depicted in the left picture. In the red reference areas, there is no TiO_2 and negligible growth occurs. The blue areas are on the TiO_2 and become darker as the gold layer on the surface increases during growth. The normalized transmission on the titanium dioxide goes down to 0.72 after two hours

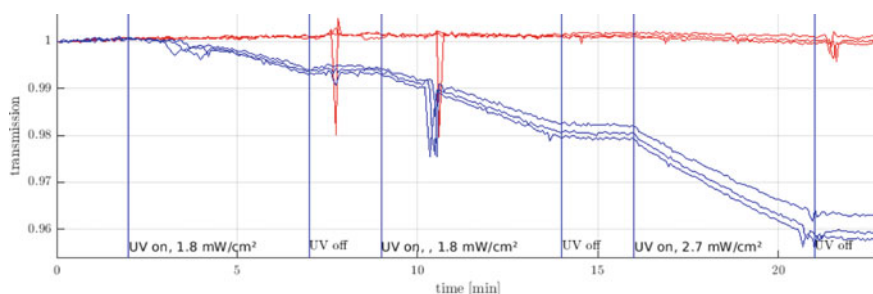


Fig. 12 Normalized transmission measurement of gold growth dynamics on titanium dioxide surface with different illumination powers. The image areas corresponding to the six timelines are indicated in Fig. 11. The blue lines show the transmission for areas on the titanium dioxide surface, the red lines are reference areas on the substrate. Different illumination intervals are depicted

The results presented in Fig. 12 clearly show the suitability of the chosen method for monitoring the gold growth process. In the reference areas (red lines) almost no intensity change is observed indicating that no gold is deposited on the substrate. In the three areas on the titanium dioxide (blue lines) the change in transmission intensity is clearly visible. Furthermore, it is observed that the growth speed is controlled by the illumination intensity.

Figure 13 shows the growth results on a TiO_2 line with $200 \mu\text{m}$ width. Three growth processes were conducted as depicted from stages (a–d). The first growth experiment ($0.736 \cdot 10^{-3} \text{ mol/l HAuCl}_4$, 2.7 mW/cm^2 UV intensity, 21 h) yielded only gold clusters and no full coverage. Repeating the experiment led to a full coverage of the surface, which was then destroyed by drying the sample with nitrogen as seen in Fig. 13c. To have full coverage the sample underwent the growth experiment a fourth time, showing fast and homogeneous growth over the titanium structure. As shown in the graph and Fig. 13d. We found that the growth process is faster and has a better coverage, when the growth is done repeatedly. This effect was also observable

in other experiments within the same setup and is currently under investigation. A picture of the grown gold lines can be seen in Fig. 14. The linewidth varies from 10 to 500 μm .

Figure 15 shows the results for grown metal lines with different line widths, ranging from 50 to 250 μm . The growth showed to be homogenous. Damages within the lines resulted from drying the sample.

The samples were further characterized after the growth process with scanning electron microscopy (SEM) to get a better understanding of the layer composition and the coverage of the TiO_2 with Au particles. After the first growth in HAuCl_4 solution, sparsely distributed Au nanoclusters are observed with large spaces between

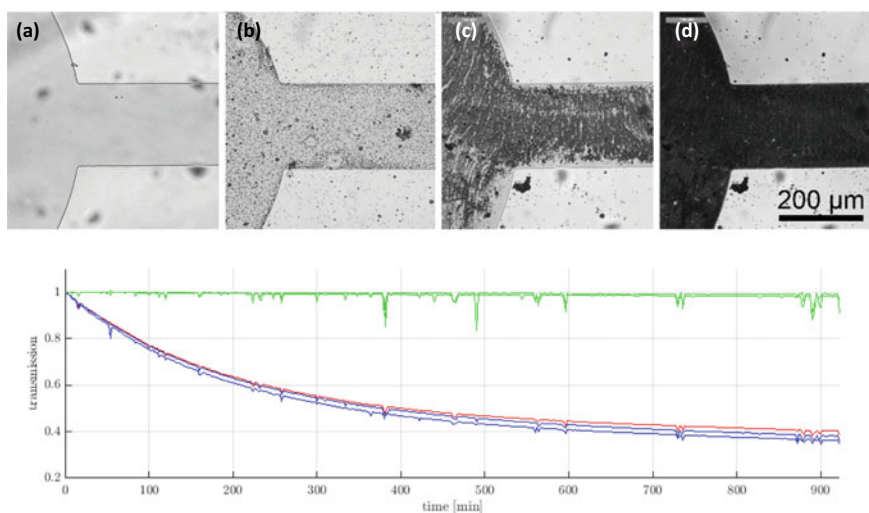
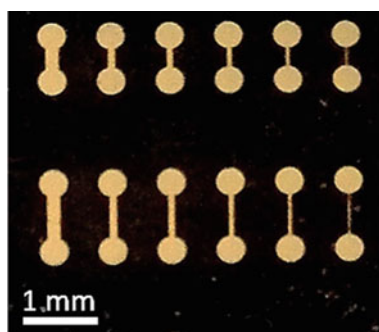


Fig. 13 Different stages of gold growth on a $200\ \mu\text{m}$ TiO_2 line. **a** shows the structure before growth, **b**, **c** and **d** after the first, second and third growth experiments. The graph depicts the normalized transmission measurement for the growth process between (c) and (d). The green lines are off structure reference spots, blue is on the connector dots (here to be seen just the left one) and red is on the line itself

Fig. 14 Reflection microscopy image of the grown Au lines with varying widths and lengths



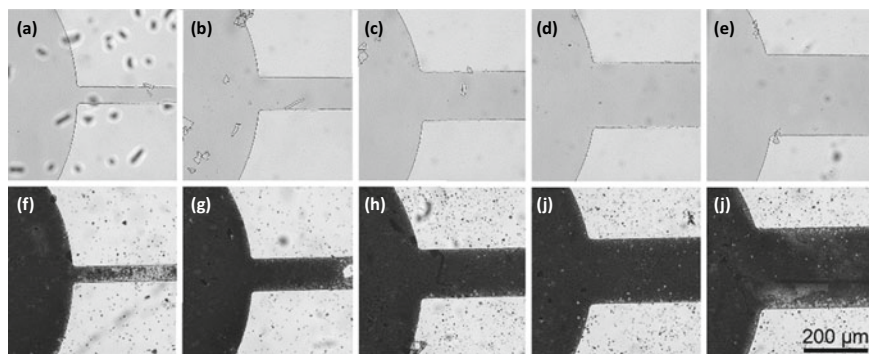


Fig. 15 Gold growth on titanium dioxide lines with different line widths: 50 μm (a, f), 100 μm (b, g), 150 μm (c, h), 200 μm (d, i), 250 μm (e, j)

the growing clusters (compare Fig. 16a). An explanation for this observation could be the rate of the electron transfer at the TiO_2 -solvent interface [28]. The rate of electron transfer is dependent on the potential gradient at the interface which could be inversely proportional to the dielectric constant of the liquid. For water, which has a high dielectric constant, a lower electron transfer rate is expected [28].

After a second and third growth process on the same sample, the surface coverage of the Au nanoclusters increased drastically. It is possible that already deposited Au nanoparticles act as nucleation sites for aggregation, and the nanocluster will tend to grow in the more energetically favorable direction. The SEM image in Fig. 16b shows that large clusters have formed compared to the first growth, supporting the hypothesis.

It shows additionally that photocatalytic deposition of Au^{3+} still occurs on the surface without a nucleation source, leading to new formations of clusters. Their growth depends on how much of the TiO_2 -layer can be still excited by the UV-light despite the presence of Au particles. They reflect the light, which prohibits the

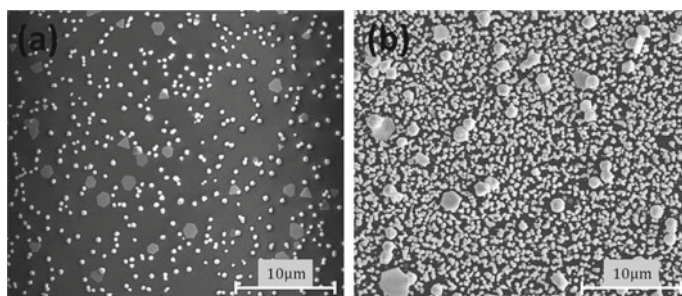


Fig. 16 SEM-images with 8000 \times times magnification of TiO_2 with photocatalytic growth of Au clusters one time (a) and three times successively (b)

generation of electrons in that area. The experiment shows that the surface coverage can be improved a lot by successive photocatalytic gold growth.

5 Conclusion

In this chapter the photodeposition of metallic nano- and microstructures is showcased at the example of locally grown metallic Au structures on photocatalytically active TiO₂ thin films. Among the various deposition methods for TiO₂ thin films, ranging from sol gel synthesis over evaporation to sputtering, in particular reactive DC magnetron sputter deposition has been found to offer the potential to tailor the morphology of the photocatalytic thin films. At the example of introducing nano-crack networks upon heat treatment it is shown how photocatalytic performance of thin films can be tailored.

Using photocatalytic TiO₂ thin films as a substrate, metal structures such as Au micro- and nanoparticles can be deposited on the thin film surface via a photoreduction reaction. Upon the illumination of TiO₂ thin films with UV light, the metal precursor is reduced via reaction with photogenerated electrons/holes, which results in the formation of solid metal structures. The morphology, size and coverage of the photodeposited metal structures can be greatly influenced by the choice of TiO₂ substrate, illumination intensity and the composition of the precursor solution. With high photocatalytic performances of the TiO₂ thin film it is possible to obtain high aspect ratio structures such as nanoneedles. The choice of thin film as well as the illumination intensity determines the morphology of the obtained structures. Lateral selectivity in the deposition of the metallic structures was showcased at two examples: On the one hand, selective illumination (e.g., by using shadow mask) results in localized gold growth. On the other hand, by structuring the TiO₂ thin film (e.g., via lithography and lift-off) the formation of metal structures can be locally restricted. To obtain better insight and control, the formation of metal structures via UV-stimulated photodeposition can be monitored in-situ via changes in the transmission. The density of the grown nano- and microstructures can be varied upon performing consecutive growth steps from individual nanostructures towards a dense coverage. In addition, the surface coverage can be tailored further upon adding additives to the precursor solution. As such, light-stimulated photodeposition of metal structures is an interesting technique for the local on-demand formation of metallic aggregates and has the potential to contribute towards an inclusion of long-range connections and globally mediated aspects like homeoplasticity into the field of neuromorphic systems.

References

1. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7** (2020). <https://doi.org/10.1063/1.5129306>

2. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. *Nature* **453**, 80–83 (2008). <https://doi.org/10.1038/nature06932>
3. Wright, C.D.: Precise computing with imprecise devices. *Nat. Electron.* **1**, 212–213 (2018)
4. Sangwan, V.K., Hersam, M.C.: Neuromorphic nanoelectronic materials. *Nat. Nanotechnol.* **15**, 517–528 (2020). <https://doi.org/10.1038/s41565-020-0647-z>
5. Zhu, X., Lee, S.H., Lu, W.D.: Nanoionic resistive-switching devices. *Adv. Electron. Mater.* **5**, 1–21 (2019). <https://doi.org/10.1002/aelm.201900184>
6. Tang, J., Yuan, F., Shen, X., Wang, Z., Rao, M., He, Y., Sun, Y., Li, X., Zhang, W., Li, Y. et al.: Bridging biological and artificial neural networks with emerging neuromorphic devices: fundamentals, progress, and challenges. *Adv. Mater.* **31** (2019). <https://doi.org/10.1002/adma.201902761>
7. Gkoupidenis, P., Koutsouras, D.A., Malliaras, G.G.: Neuromorphic device architectures with global connectivity through electrolyte gating. *Nat. Commun.* **8**, 1–8 (2017). <https://doi.org/10.1038/ncomms15448>
8. Dawson, A., Kamat, P.V.: Semiconductor–metal nanocomposites. Photoinduced fusion and photocatalysis of gold-capped TiO₂ (TiO₂/Gold) nanoparticles. *J Phys Chem B* **105**, 960–966 (2001). <https://doi.org/10.1021/jp0033263>
9. Chen, X., Shen, S., Guo, L., Mao, S.S.: Semiconductor-based photocatalytic hydrogen generation. *Chem. Rev.* **110**, 6503–6570 (2010). <https://doi.org/10.1021/cr1001645>
10. Neal, L.M., Everett, M.L., Hoflund, G.B., Hagelin-Weaver, H.E.: Characterization of palladium oxide catalysts supported on nanoparticle metal oxides for the oxidative coupling of 4-methylpyridine. *J. Mol. Catal. A Chem.* **335**, 210–221 (2011). <https://doi.org/10.1016/j.molcata.2010.11.036>
11. Jiang, D., Wang, W., Sun, S., Zhang, L., Zheng, Y.: Equilibrating the plasmonic and catalytic roles of metallic nanostructures in photocatalytic oxidation over Au-modified CeO₂. *ACS Catal.* **5**, 613–621 (2015). <https://doi.org/10.1021/cs501633q>
12. Haselmann, G.M., Eder, D.: Early-stage deactivation of platinum-loaded TiO₂ using in situ photodeposition during photocatalytic hydrogen evolution. *ACS Catal.* **7**, 4668–4675 (2017). <https://doi.org/10.1021/acscatal.7b00845>
13. Zulfiqar, A., Temerov, F., Saarinen, J.J.: Multilayer TiO₂ inverse opal with gold nanoparticles for enhanced photocatalytic activity. *ACS Omega* **5**, 11595–11604 (2020). <https://doi.org/10.1021/acsomega.0c00833>
14. Veziroglu, S., Ghori, M.Z., Kamp, M., Kienle, L., Rubahn, H.G., Strunskus, T., Fiutowski, J., Adam, J., Faupel, F., Aktas, O.C.: Photocatalytic growth of hierarchical Au needle clusters on highly active TiO₂ thin film. *Adv. Mater. Interfaces* **5**, 1–7 (2018). <https://doi.org/10.1002/admi.201800465>
15. Yu, Y., He, T., Guo, L.L.L., Yang, Y., Guo, L.L.L., Tang, Y., Cao, Y.: Efficient visible-light photocatalytic degradation system assisted by conventional Pd catalysis. *Sci. Rep.* **5**, 1–7 (2015). <https://doi.org/10.1038/srep09561>
16. Safajou, H., Khojasteh, H., Salavati-Niasari, M., Mortazavi-Derazkola, S.: Enhanced photocatalytic degradation of dyes over graphene/Pd/TiO₂ nanocomposites: TiO₂ nanowires versus TiO₂ nanoparticles. *J. Colloid Interface Sci.* **498**, 423–432 (2017). <https://doi.org/10.1016/j.jcis.2017.03.078>
17. Kusmierek, E.: A CeO₂ semiconductor as a photocatalytic and photoelectrocatalytic material for the remediation of pollutants in industrial wastewater: A review. *Catalysts* **10**, 1–54 (2020). <https://doi.org/10.3390/catal10121435>
18. Bahruji, H., Bowker, M., Davies, P.R., Pedrono, F.: New insights into the mechanism of photocatalytic reforming on Pd/TiO₂. *Appl. Catal. B Environ.* **107**, 205–209 (2011). <https://doi.org/10.1016/j.apcatb.2011.07.015>
19. Ajmal, A., Majeed, I., Malik, R.N., Idriss, H., Nadeem, M.A.: Principles and mechanisms of photocatalytic dye degradation on TiO₂ based photocatalysts: a comparative overview. *RSC Adv.* **4**, 37003–37026 (2014). <https://doi.org/10.1039/C4RA06658H>
20. Kočí, K., Obalová, L., Matějová, L., Plachá, D., Lacný, Z., Jirkovský, J., Šolcová, O.: Effect of TiO₂ particle size on the photocatalytic reduction of CO₂. *Appl. Catal. B Environ.* **89**, 494–502 (2009). <https://doi.org/10.1016/j.apcatb.2009.01.010>

21. Ghori, M.Z., Veziroglu, S., Henkel, B., Vahl, A., Polonskyi, O., Strunskus, T., Faupel, F., Aktas, O.C.: A comparative study of photocatalysis on highly active columnar TiO₂ nanostructures in-air and in-solution. *Sol. Energy Mater. Sol. Cells* **178**, 170–178 (2018)
22. Zhou, J., Cheng, Y., Yu, J.: Preparation and characterization of visible-light-driven plasmonic photocatalyst Ag/AgCl/TiO₂ nanocomposite thin films. *J. Photochem. Photobiol. A Chem.* **223**, 82–87 (2011). <https://doi.org/10.1016/j.jphotochem.2011.07.016>
23. Schneider, J., Matsuoka, M., Takeuchi, M., Zhang, J., Horiuchi, Y., Anpo, M., Bahnemann, D.W.: Understanding TiO₂ photocatalysis: mechanisms and materials. *Chem. Rev.* **114**, 9919–9986 (2014)
24. Ma, J., Guo, X., Zhang, Y., Ge, H.: Catalytic performance of TiO₂ at Ag composites prepared by modified photodeposition method. *Chem. Eng. J.* **258**, 247–253 (2014). <https://doi.org/10.1016/j.cej.2014.06.120>
25. Ayati, A., Ahmadpour, A., Bamoharram, F.F., Tanhaei, B., Mänttari, M., Sillanpää, M.: A review on catalytic applications of Au/TiO₂ nanoparticles in the removal of water pollutant. *Chemosphere* **107**, 163–174 (2014)
26. Tossi, C., Hällström, L., Selin, J., Vaelma, M., See, E., Lahtinen, J., Tittonen, I.: Size- and density-controlled photodeposition of metallic platinum nanoparticles on titanium dioxide for photocatalytic applications. *J. Mater. Chem. A* **7**, 14519–14525 (2019). <https://doi.org/10.1039/c8ta09037h>
27. Lee, Y., Kim, E., Park, Y., Kim, J., Ryu, W.H., Rho, J., Kim, K.: Photodeposited metal-semiconductor nanocomposites and their applications. *J. Mater.* **4**, 83–94 (2018). <https://doi.org/10.1016/j.jmat.2018.01.004>
28. Veziroglu, S., Obermann, A.-L., Ullrich, M., Hussain, M., Kamp, M., Kienle, L., Leibner, T., Rubahn, H.-G., Polonskyi, O., Strunskus, T. et al.: Photodeposition of Au nanoclusters for enhanced photocatalytic dye degradation over TiO₂ thin film. *ACS Appl. Mater. Interfaces* **12**, 14983–14992. <https://doi.org/10.1021/acsami.9b18817>
29. Wenderich, K., Mul, G.: Methods, mechanism, and applications of photodeposition in photocatalysis: a review. *Am. Chem. Soc.* **116**, 14587–14619 (2016)
30. Hanaor, D.A.H., Sorrell, C.C.: Review of the anatase to rutile phase transformation. *J. Mater. Sci.* **46**, 855–874 (2010). <https://doi.org/10.1007/S10853-010-5113-0>
31. Taherniya, A., Raoufi, D.: The annealing temperature dependence of anatase TiO₂ thin films prepared by the electron-beam evaporation method. *Semicond. Sci. Technol.* **31**, (2016). <https://doi.org/10.1088/0268-1242/31/12/125012>
32. Siah, W.R., Lintang, H.O., Shamsuddin, M., Yuliati, L.: High photocatalytic activity of mixed anatase-rutile phases on commercial TiO₂ nanoparticles. *IOP Conf. Ser. Mater. Sci. Eng.* **107**, 012005 (2016). <https://doi.org/10.1088/1757-899X/107/1/012005>
33. Zheng, S.K., Xiang, G., Wang, T.M., Pan, F., Wang, C., Hao, W.C.: Photocatalytic activity studies of TiO₂ thin films prepared by r.f. magnetron reactive sputtering. *Vacuum* **72**, 79–84 (2003). [https://doi.org/10.1016/S0042-207X\(03\)00104-0](https://doi.org/10.1016/S0042-207X(03)00104-0)
34. Wang, Y.-H., Rahman, K.H., Wu, C.-C., Chen, K.-C.: A review on the pathways of the improved structural characteristics and photocatalytic performance of titanium dioxide (TiO₂) thin films fabricated by the magnetron-sputtering technique. *Catalysts* **10**, 598 (2020). <https://doi.org/10.3390/catal10060598>
35. Lee, M.K., Park, Y.C.: Super-hydrophilic anatase TiO₂ thin film in-situ deposited by DC magnetron sputtering. *Thin Solid Films* **638**, 9–16 (2017). <https://doi.org/10.1016/j.tsf.2017.07.046>
36. Wang, Q.M., Zhang, T.F., Kwon, S.H., Kim, K.H.: Fabrication of TiO₂ films on glass substrates by a pulsed DC reactive magnetron sputtering. *Appl. Mech. Mater.* **71–78**, 5050–5053 (2011). <https://doi.org/10.4028/www.scientific.net/AMM.71-78.5050>
37. Henkel, B., Vahl, A., Aktas, O.C., Strunskus, T., Faupel, F.: Self-organized nanocrack networks: a pathway to enlarge catalytic surface area in sputtered ceramic thin films, showcased for photocatalytic TiO₂. *Nanotechnology* **29**, 35703 (2018). <https://doi.org/10.1088/1361-6528/aa9d35>

38. Zhang, W., Zhu, S., Li, Y., Wang, F.: Photocatalytic property of TiO₂ films deposited by pulsed DC magnetron sputtering. *J. Mater. Sci. Technol.* **20**, 31–34 (2004)
39. Kavaliunas, V., Sestakauskaite, A., Sriubas, M., Laukaitis, G.: Influence of deposition parameters on the structure of TiO₂ thin films prepared by reactive magnetron sputtering technique. *Lect. Notes Netw. Syst.* **53**, 49–57 (2019). https://doi.org/10.1007/978-3-319-99834-3_7
40. Bersani, D., Lottici, P.P., Ding, X.-Z.: Phonon confinement effects in the Raman scattering by TiO₂ nanocrystals. *Appl. Phys. Lett.* **72**, 73 (1998). <https://doi.org/10.1063/1.120648>
41. Lee, S.H., Yamasue, E., Okumura, H., Ishihara, K.N.: Effect of oxygen and nitrogen concentration of nitrogen doped TiO_x film as photocatalyst prepared by reactive sputtering. *Appl. Catal. A Gen.* **371**, 179–190 (2009). <https://doi.org/10.1016/J.APCATA.2009.10.011>
42. Zeman, P., Takabayashi, S.: Effect of total and oxygen partial pressures on structure of photocatalytic TiO₂ films sputtered on unheated substrate. *Surf. Coatings Technol.* **153**, 93–99 (2002). [https://doi.org/10.1016/S0257-8972\(01\)01553-5](https://doi.org/10.1016/S0257-8972(01)01553-5)
43. Zhang, W., Li, Y., Zhu, S., Wang, F.: Influence of argon flow rate on TiO₂ photocatalyst film deposited by dc reactive magnetron sputtering. *Surf. Coat. Technol.* **182**, 192–198 (2004). <https://doi.org/10.1016/J.SURFCOAT.2003.08.050>
44. Laha, P., Panda, A.B., Mahapatra, S.K., Barhai, P.K., Das, A.K., Banerjee, I.: Development of rf plasma sputtered Al₂O₃–TiO₂ multilayer broad band antireflecting coatings and its correlation with plasma parameters. *Appl. Surf. Sci.* **258**, 2275–2282 (2012). <https://doi.org/10.1016/J.APSUSC.2011.09.118>
45. Liu, B., Zhao, X., Zhao, Q., Li, C., He, X.: The effect of O₂ partial pressure on the structure and photocatalytic property of TiO₂ films prepared by sputtering. *Mater. Chem. Phys.* **90**, 207–212 (2005). <https://doi.org/10.1016/J.MATCHEMPHYS.2004.10.038>
46. Vahl, A., Dittmann, J., Jetter, J., Veziroglu, S., Shree, S., Ababii, N., Lupan, O., Aktas, O.C., Strunskus, T., Quandt, E., et al.: The impact of O₂/Ar ratio on morphology and functional properties in reactive sputtering of metal oxide thin films. *Nanotechnology* **30**, 235603 (2019). <https://doi.org/10.1088/1361-6528/ab0837>
47. Singhal, N., Kumar, U.: Noble metal modified TiO₂: selective photoreduction of CO₂ to hydrocarbons. *Mol. Catal.* **439**, 91–99 (2017). <https://doi.org/10.1016/J.MCAT.2017.06.031>
48. Veziroglu, S., Ghori, M.Z., Obermann, A.L., Röder, K., Polonskyi, O., Strunskus, T., Faupel, F., Aktas, O.C.: Ag nanoparticles decorated TiO₂ thin films with enhanced photocatalytic activity. *Phys. Status Solidi Appl. Mater. Sci.* **216**, 1–6 (2019). <https://doi.org/10.1002/pssa.201800898>
49. Li, H., Wu, J., Yin, Z., Zhang, H.: Preparation and applications of mechanically exfoliated single-layer and multilayer MoS₂ and WSe₂ nanosheets. *Acc. Chem. Res.* **47**, 1067–1075 (2014). <https://doi.org/10.1021/AR4002312>
50. Veziroglu, S., Ullrich, M., Hussain, M., Drewes, J., Shondo, J., Strunskus, T., Adam, J., Faupel, F., Aktas, O.C.: Plasmonic and non-plasmonic contributions on photocatalytic activity of Au-TiO₂ thin film under mixed UV–visible light. *Surf. Coat. Technol.* **389**, 125613 (2020). <https://doi.org/10.1016/j.surfcoat.2020.125613>
51. Mendoza-Diaz, M.I., Cure, J., Rouhani, M.D., Tan, K., Patnaik, S.G., Pech, D., Quevedo-Lopez, M., Hungria, T., Rossi, C., Estève, A.: On the UV-visible light synergetic mechanisms in Au/TiO₂ hybrid model nanostructures achieving photoreduction of water. *J. Phys. Chem. C* **124**, 25421–25430 (2020). <https://doi.org/10.1021/acs.jpcc.0c08381>
52. Miquelot, A., Debieu, O., Rouessac, V., Villeneuve, C., Prud'homme, N., Cure, J., Constantoudis, V., Papavieros, G., Roualdes, S., Vahlas, C.: TiO₂ nanotree films for the production of green H₂ by solar water splitting: from microstructural and optical characteristics to the photocatalytic properties. *Appl. Surf. Sci.* **494**, 1127–1137 (2019). <https://doi.org/10.1016/J.APSUSC.2019.07.191>
53. Vahl, A., Veziroglu, S., Henkel, B., Strunskus, T., Polonskyi, O., Aktas, O.C., Faupel, F.: Pathways to tailor photocatalytic performance of TiO₂ thin films deposited by reactive magnetron sputtering. *Materials (Basel)*. **12**, 2840 (2019). <https://doi.org/10.3390/ma12172840>
54. Grätzel, M.: *Heterogenous Photochemical Electron Transfer*. CRC Press (2018). ISBN 9781351073202

55. Guo, Y., Siretanu, I., Zhang, Y., Mei, B., Li, X., Mugele, F., Huang, H., Mul, G.: PH-dependence in facet-selective photo-deposition of metals and metal oxides on semiconductor particles. *J. Mater. Chem. A* **6**, 7500–7508 (2018). <https://doi.org/10.1039/c8ta00781k>
56. Wenderich, K., Mul, G.: Methods, mechanism, and applications of photodeposition in photocatalysis: a review. *Chem. Rev.* **116**, 14587–14619 (2016). <https://doi.org/10.1021/acs.chemrev.6b00327>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Smart Sensor Arrays



Maik-Ivo Terasa, Leonard Siebert, Pia Holtz, Sören Kaps, Oleg Lupan, Jürgen Carstensen, Franz Faupel, Alexander Vahl, and Rainer Adelung

Abstract Sensors play a crucial role in our everyday life and will become more and more demanded with the transition towards smart cities and the “Internet of Things” with the result of an ever-increasing energy demand. Thus, the research of improving the energy efficiency of sensor systems has started to move towards edge computing and neuromorphic engineering while at the same time additive manufacturing has gained increasing attention as a means for a rapid, scalable fabrication of functional devices yet with huge design freedom and quick iteration cycles during the development phase. In this chapter the junction of pre-designed components with self-organizing material systems for the facile fabrication of sensor devices via direct ink writing is demonstrated. As examples 3D-printed CuO/Cu₂O/Cu- and CuO/Fe₂O₃

M.-I. Terasa (✉) · L. Siebert · P. Holtz · S. Kaps · O. Lupan · J. Carstensen · R. Adelung
Functional Nanomaterials, Department of Materials Science and Engineering, Kiel University,
Kaiserstraße 2, 24143 Kiel, Germany
e-mail: mate@tf.uni-kiel.de

L. Siebert
e-mail: lesi@tf.uni-kiel.de

P. Holtz
e-mail: piah@tf.uni-kiel.de

S. Kaps
e-mail: ska@tf.uni-kiel.de

O. Lupan
e-mail: ollu@tf.uni-kiel.de

J. Carstensen
e-mail: jc@tf.uni-kiel.de

R. Adelung
e-mail: ra@tf.uni-kiel.de

F. Faupel · A. Vahl
Multicomponent Materials, Department of Materials Science and Engineering, Kiel University,
Kaiserstraße 2, 24143 Kiel, Germany
e-mail: ff@tf.uni-kiel.de

A. Vahl
e-mail: alva@tf.uni-kiel.de

acetone gas sensors are shown, as well as how the decoration of metal-oxide semi-conducting sensor materials with noble metallic nanoparticles can enhance the sensor properties with respect to sensitivity, selectivity and stability. Finally, the assembly of conductive bridges between electrodes stimulated by voltage pulses is introduced as an approach towards facilitating the self-organization of neuromorphic circuits.

Keywords Memensors · Metal-oxide semiconductor sensors · 3D-printing · Direct ink writing · Triggered assembly

1 Introduction

Sensors are a key component of our everyday life. They are used to detect changes in the environment that are hazardous or of interest otherwise. Living organisms use environmental stimuli triggering their sensory systems as inputs to adapt their behaviour to the environmental surrounding, e.g., for finding food or evaluating danger. In this regard biological systems have been evolutionary optimized to be most energy-efficient since this is a crucial advantage for survival. An example is the adaptation to a permanent background stimulus [1]. The sensory system changes such, that the response to an excitation decreases, if the stimulus does not change over a long period of time. For example, a human eye in a dark room contracts its pupil when suddenly the light is turned on, accommodating to the new lighting situation, thus allowing the eye to see normally after adaptation. However, technical advances presented humanity with new challenges surpassing the capacities of humans' natural sensors. Technical sensors have been developed to compensate or extend the capabilities of the human body, especially for security and health applications e.g. monitoring the concentration of harmful radiation or toxic gases, but also for a wide variety of purposes like magnetic field sensors, CCD-sensors or mechanical sensors. With the ever growing number of electronic devices and the development towards the "Internet of Things" and "Smart Cities" the number of deployed sensors will increase significantly in the future and a crucial aspect is the growing energy demand, that accompanies this development [2–4]. However, biological systems are remarkably efficient, when it comes to data acquisition and processing. This has inspired research to mimic the way they achieve their great energy efficiency. A prominent approach for higher energy efficiency to account for the growing energy demand is overcoming the von-Neumann bottleneck by using features of edge computing and neuromorphic engineering, i.e. decentralizing computing logic and memory to reduce the power consumption of communication between them and diminish the need for data transmission to a central server [5–8]. In the following sections, various concepts will be discussed, that advance the research in sensor technology to account for the demand of efficient, scalable, specialized and smart sensor devices. At first the concept of the memsensor will be introduced, which is a novel junction of a sensor with memristive components yielding additional unique abilities, such as adaptation to a permanent background stimulus. It will be discussed with respect to an equivalent circuit model

and a ZnO needle as an example system. When it comes to the fabrication of sensor devices, most processes rely on micro-patterning by lithography and physical or chemical vapour deposition techniques, which are inherently demanding with respect to workplace and material quality [9, 10]. Here, additive manufacturing (AM), which is based on inexpensive and highly available equipment, is discussed as a means of fabrication for sensor devices. The AM approach provides huge design freedom with respect to sensor materials and shape as well as quick iteration cycles making it a valuable tool for research and development of sensor devices. Metal-oxide semiconductor (MOS) microparticles are showcased as sensor materials printable by direct ink writing (DIW) and the enhancement of their sensor performance by decoration with metallic nanoparticles is discussed. In the last section, the dynamic establishment of conductive interconnections between electrodes at the microscale in a liquid matrix is introduced by the example of copper electrodes in an electrolyte of dimethyl sulfoxide (DMSO) and methylene blue (MB), which serves as a first step towards expanding brain-inspired plasticity mechanisms to a larger scale.

2 Smart Sensors

This chapter evolves around sensor devices that belong to the category of semi-conducting metal oxide micro- and nanostructures, such as CuO, TiO₂, ZnO. Such materials are commonly used for sensing applications, for example as photon sensors or gas sensors. While photon sensing involves a considerable fraction of the material, depending on the photon penetration depth in the material, gas sensing is strongly related to the adsorption and desorption of molecular species onto the sensor surface and therefore can be considered as a surface phenomenon. Interestingly, there is a considerable overlap between materials used for MOS sensors and memristive devices that rely on the valence change mechanism upon migration of oxygen vacancies. This implies, that if manufactured in a proper way, a single device can intrinsically unify the inherent functionalities of sensors and memristive devices, namely the stimulus dependent resistance and the controllable switching between resistance states. This section will elucidate the properties of memsensors that arise upon the junction between memristive switching and stimulus sensing. In order to obtain memsensitive properties, an important requirement is that the active material must be accessible to the chosen external stimuli. This requirement strongly depends on the type of stimulus—while gas sensors need a direct contact of the gas species to the sensor surface, other stimuli like photons or magnetic fields may also potentially penetrate through matter. This requires the top electrodes of vertically sandwiched devices to be chosen carefully or to be additionally structured. Another option is the use of lateral structures, like micro- or nanowires that are in direct contact with the surrounding medium. An example for such horizontal structure are ZnO microneedles. A ZnO needle with a diameter in the μm range and a length in the cm range is shown in Fig. 1a. Details on the manufacturing are reported elsewhere [11].

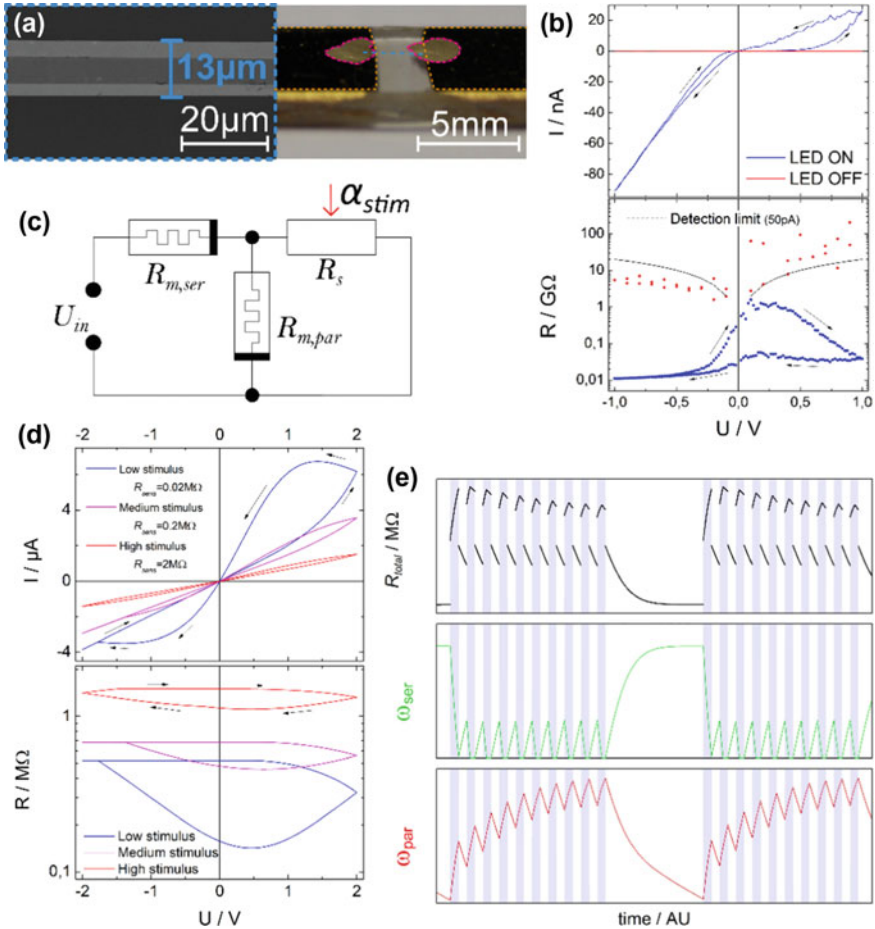


Fig. 1 A ZnO microrod device **a** as a memsensor device prototype is capable of showing stimulus dependent hysteresis. **b** An equivalent circuit **c** with three components was used to model stimulus dependent hysteresis **d** and adaptation of the resistance of a memsensor device to an external stimulus **e**. In case of amplitude adaptation **e**, the resistance response (black) of the modelled memsensor decreases with each consecutive stimulus pulse decreases, which is based on the slower change in the internal state of the parallel resistor (red) than the serial resistor (green) [11]. Reprinted with permission from [11]

The electrical resistance of a ZnO microneedle exhibits a strong dependence on the illumination by UV photons. Interestingly, IV hysteresis measurements of the ZnO needle device show an UV dependent hysteresis, which is depicted in Fig. 1b. While the IV characteristics in the dark state do not show a pinched hysteresis and in general exhibit an overall very high resistance, under UV illumination there is a pronounced pinched hysteresis. Accordingly, the device requires the UV stimulus to show memristive switching. A junction between memristive switching and sensing

has been reported for a variety of devices [12–14]. To understand the phenomenon of stimulus dependent hysteresis and study the potential properties of memsensors, a simple three-component model was proposed (Fig. 1c), that allows to investigate the interplay between memristive switching and sensing [11]. Within this model, in first approximation, the sensor response is assumed to be immediate, while the memristive devices can occupy a smooth transition of resistance states between a fixed LRS and HRS. As such, stimulus dependent hysteresis at a given voltage window was reproduced, as depicted in Fig. 1d. In the context of the model, the occurrence of stimulus-sensitive hysteresis originates from the difference in potential drop over the serial memristive element, that leads to a significantly more pronounced hysteresis in the illuminated state. As an additional memsensor feature that goes beyond the inherent features of sensors and memristive devices, amplitude adaptation to a constant stimulus was modelled (Fig. 1e), mainly governed by the different rate of change in inner state for both memristive devices. This predicts that the junction of sensing and memristive switching has the potential to create smart sensor devices, that show bio-inspired features like adaptation.

3 Additive Manufacturing of Sensors

Self-organized systems are shaped by mechanisms of positive and negative feedback, meaning that occurring “positive” events make them more likely to occur again, while “negative” events become more unlikely. This feedback-driven dynamic makes the system generally robust against fluctuations and will lead it towards its most desirable state, defined by the feedback mechanisms. While self-organization is an important feature of efficient bio-inspired systems, there are many examples for parts of biological systems that are formed by a fixed pattern because they often serve a specific purpose e.g. the eye as a means of perceiving visual information. Such pre-designed patterns generally make up every technical device, though when used in dynamic self-organized systems in an emerging concept like neuromorphic computing it is advisable to have fast iteration speed during the development stage, to be able to adapt and optimize the design quickly. At the same time as much design-freedom as possible should be maintained with respect to the geometry, material or composition of parts. In this context the field of additive manufacturing (AM) has shown to offer beneficial features for rapid design and production of prototypes of both structural and functional components. Traditional manufacturing of parts is mostly subtractive e.g. milling, drilling, etc., where material is removed from the part until it is shaped as intended. This creates a substantial deal of waste while it becomes more difficult and thus expensive the more complex the shape of the part is. In contrast to this top-down approach, additive manufacturing (AM) works “bottom-up” by assembling the final part out of small fractions. The most common way is by stacking thin layers of material on top of each other to create 3D objects (see Fig. 2a). For the deposition of material for each layer there are several different approaches available, depending on the matrix material and specifications of the

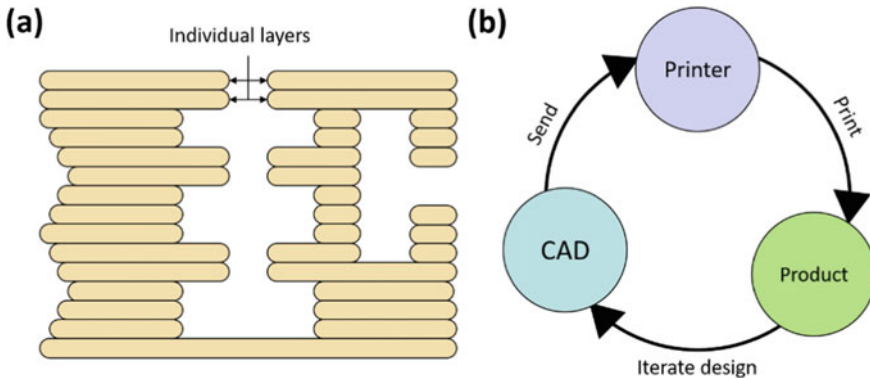


Fig. 2 Benefits of additive manufacturing: **a** The bottom-up approach (i.e. stacking layers of material) facilitates more complex geometries, like undercuts in the center of a piece, that would be difficult to manufacture by traditional machining. **b** Quick iteration cycles are made possible by directly linking the design and fabrication steps

part like minimum feature size. Processes like fused deposition modelling (FDM) melt a thermoplastic filament and extrude it through a nozzle to deposit the material, while stereolithography (SLA) is based on the photopolymerization of thin layers of liquid resin. Metal parts can be produced by selective laser sintering (SLS) or the conceptually similar selective laser melting (SLM), which heat up a metal powder to sinter or melt the grains together [15, 16]. Most commonly object data is provided by digital 3D object files, e.g. in the STL-format, created by computer aided design (CAD) programs. Object design and fabrication are direct subsequent steps, as seen in Fig. 2b. A designer has both large freedom with respect to geometry and materials, while also being able to work through a part's iteration cycles quickly. This led to AM also being considered “rapid prototyping” making it a perfect tool to use in research and development.

3.1 Direct Ink Writing of Microparticles

Direct Ink Writing (DIW) is a subtype of AM, where an ink composed of a carrier fluid and colloidal particles with high filler content is extruded onto a substrate through a moveable nozzle. The carrier fluid consists of a volatile solvent and possibly a polymeric binder. After extrusion the solvent evaporates and the ink solidifies. While this technology already has been around for many years in the form of inkjet printers for digital images it also got more recent attention as means for AM by depositing ceramic particles via the ink [17–19]. DIW allows printing of any materials in powder form, particularly materials that are not suitable for other additive manufacturing methods or that have complex particle morphologies. This also includes functional particles like tetrapodal ZnO (t-ZnO, [20]), which expands the capabilities of AM

from purely structural parts towards functional or composite components. However, the rheological properties of the ink are crucial for the reliable fabrication with DIW. The ink must remain homogeneous and flow evenly to avoid defects or aberrations in the print. Colloidal suspension inks are based on a carrier fluid and one or more species of suspended micro- and/or nanoparticles. In some cases, a polymeric binder has to be added to the solvent increasing the inks viscosity. E.g. if the particles are too large to stay suspended in the pure solvent, a polymeric binder is added to increase the viscosity preventing the particles from sedimenting. Furthermore, complex shaped particles like t-ZnO form highly porous, tangled agglomerates. A high viscosity ink is necessary to exert the shear forces required to move the particles, otherwise the solvent would flow through the pores separating from them. The carrier fluid commonly solidifies either by evaporation of the solvent (e.g. for the systems H₂O/PEG or EtOH/PVB) or by a chemical reaction (e.g. silicones). Once solid it can either persist as a structural matrix for the particles or be removed by thermal or chemical decomposition. Especially for functional microparticles it is commonly desired to remove the binder to create interfaces to other external species, like gases or light, or create percolating or dense structures approximating the properties of the bulk material e.g. the electrical conductivity.

3.2 EtOH/PVB as a Carrier Fluid

Using ethanol (EtOH) as a solvent and polyvinyl butyral (PVB) as a polymeric binder has shown to yield interesting possibilities, which are further enhancing the flexibility of DIW for micro- or nanoparticle based components and can possibly be transferred to other similar systems as well. An EtOH/PVB based ink can be printed on a glass substrate (see Fig. 3a). Since PVB is a hydrophobic polymer, submerging the glass substrate in H₂O makes the printed material detach from the glass (see Fig. 3b). The result is a flexible free-standing structure, which in itself already opens new possibilities towards prints that can be folded to create 2.5-dimensional designs. In addition, if slightly wetted with EtOH again, the print can be re-attached to any suitable surface, including round or complex ones, on which printing and most other deposition techniques (e.g. sputter deposition) are unfeasible or even impossible, dramatically increasing the applicability of printed components (see Fig. 3c). Moreover, with a high-power laser instead of traditional unlocalized heating, the heat treatment process can be controlled precisely in terms of localization and power transfer. Using a 100 W CO₂ laser two main goals have been achieved: (1) With low power output, a gentle milling of a thin portion polymeric binder has been performed on a single layer print. Without damaging the embedded microparticles, the top portion of the layer has been stripped of its polymeric binder, exposing the functional particles, while keeping the bottom portion of the binder as mechanical support Fig. 3d. Even after the laser milling process the print could still be detached from the substrate. By combining inks of t-ZnO and CNTs an all-printed freestanding sensor device has been fabricated with this method, as shown in Fig. 3e. The geometric and mechanical

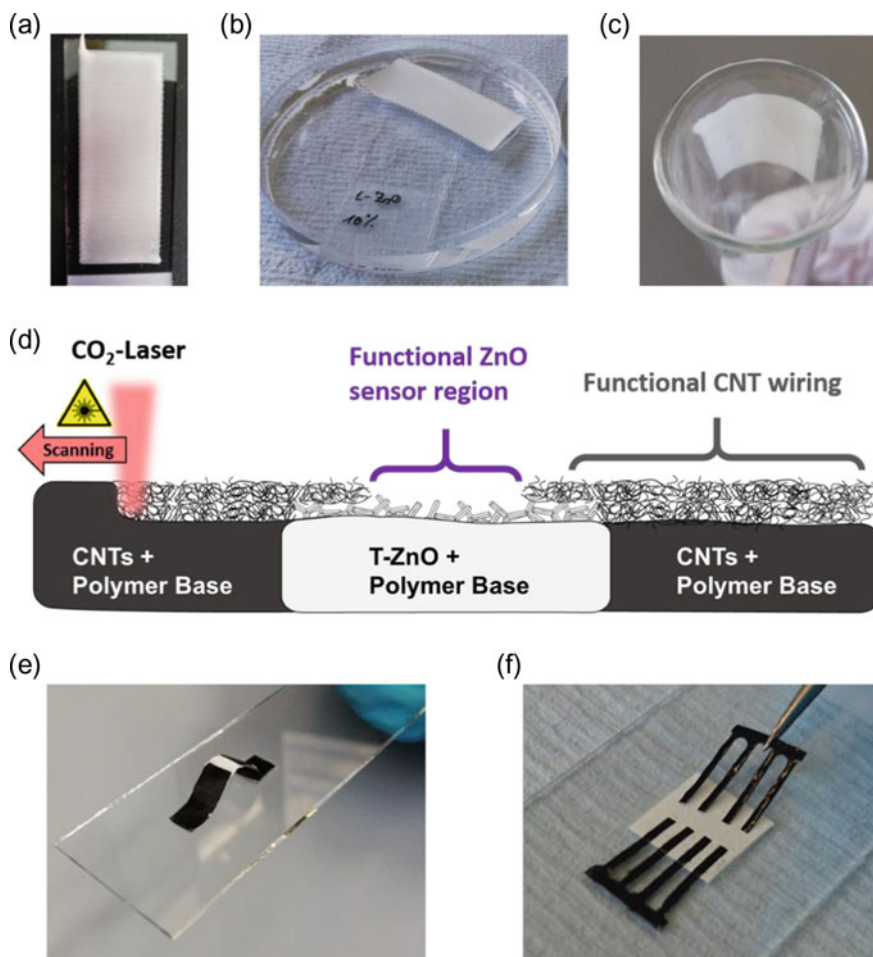


Fig. 3 **a** An as-printed layer of an ink of EtOH/PVB with t-ZnO as functional particles. **b** Submerging the glass substrate in water will make the printed layer delaminate. **c** The delaminated print can be re-attached conformably to round or complex surfaces. **d** Schematic of the gentle laser milling process. It allows for precise stripping of thin layers of polymer to expose the functional particles, while keeping a base of polymer for structural stability. **e** A free-standing all-printed UV-sensor composed of CNT-ink as wiring and t-ZnO as the sensor component. The print has been laser milled to expose the ZnO enabling its sensor functionality. **f** Separately printed and delaminated CNT- and ZnO-prints being assembled freely by slightly wetting the junctions with ethanol

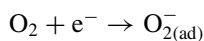
flexibility makes such sensor devices a great candidate for components in electronic skin applications [21, 22]. Additionally separately printed parts can be assembled freely by wetting them at the respective junctions with ethanol (see Fig. 3f) or be stored as well as processed mechanically e.g. by cutting, allowing for production of stock or modular designs. (2) With higher power output a localized oxidation of metal particles has been accomplished, leading to precisely located metal/metal-oxide/metal interfaces, which is a promising starting point for simple and accessible, yet precise and scalable fabrication processes of MOS sensor devices.

3.3 High-Viscosity DIW Setup

The Chair for functional Nanomaterials at Kiel University has established a custom-built DIW setup for printing with inks of high viscosities, which is depicted in Fig. 4a. Its extrusion unit is able to exert high pressures up to 40 bar on the ink via a steel piston driven by a geared stepper motor. The ink reservoir is a commercially available polypropylene (PP) syringe encased in an aluminium shell to prevent expansion of the PP during extrusion. The extrusion nozzle is a conical syringe tip with an orifice diameter of 100 μm to 840 μm . The nozzle size determines the spatial resolution of the print, since a smaller diameter yields thinner lines and thus finer features. But at the same time it limits the size, morphology and fill factor of the printable microparticles. E.g. tetrapodal ZnO particles (t-ZnO) with an arm length of 25 μm or more have shown to clog nozzles with 100 μm diameter at higher fill factors of more than 10 wt% due to the tendency of the tetrapods to become tangled. In addition, the print bed is heatable and a 3 W computer controlled blue laser is pointed at the printing spot for precise localized heating. Figure 4b shows a selection of microparticles, that have been printed with this setup. The setup has been used to print all devices presented here and in Sect. 3.4.

3.4 Gas Sensors by DIW

Semiconducting metal oxides (MOS) have shown to be a superior material class for gas sensing applications, which are essential in the fields of workplace security, healthcare or environmental monitoring [23]. The mechanism of gas sensing is commonly attributed to the ad- or desorption of molecules on the solid-gas interface of the sensing material changing its electronic landscape, where oxygen is discussed as the most popular surface adsorbant species. Oxygen adsorbs to the surface of MOS by binding electrons from the conduction band:



In a n-type semiconductor this creates a high resistance electron depletion region narrowing the conduction channel, whereas in a p-type semiconductor a hole accumu-

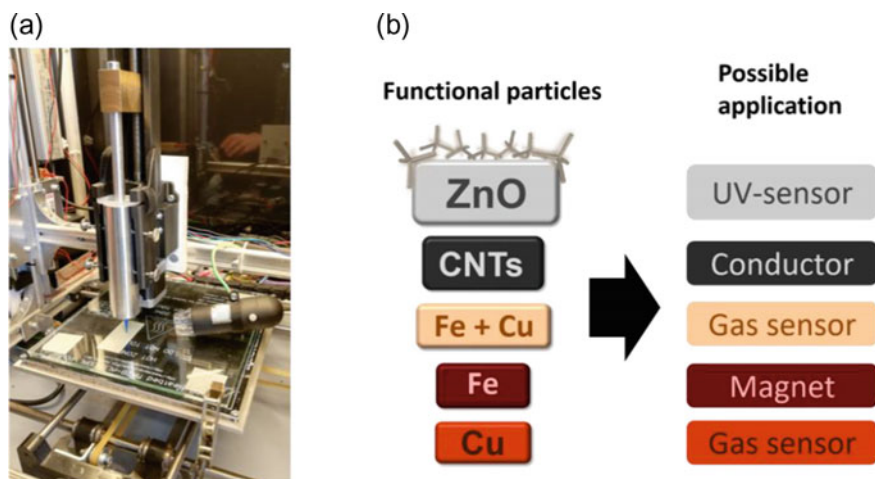


Fig. 4 **a** Photograph of the custom-built high-viscosity ink printing setup. **b** A selection of microparticle species that have been printed with the setup, as well as possible applications for each.

Table 1 Composition of the functional inks for MOS sensor devices

Sensor	CuO/Cu ₂ O/Cu	CuO:Fe ₂ O ₃
Carrier fluid	Deionized H ₂ O	Ethanol
Polymeric binder	Polyethylene glycole	Polyvinyl butyrale
Microparticles	Copper (15–25 μm)	Copper (15–25 μm) + Iron (40–60 μm)

lation layer (HAL) with lower resistance is created. The sensor response is most commonly discussed in the context of chemical reactions or desorption of surface oxygen decreasing either the electron depletion layer or the hole accumulation layer. The acetone content of human breath has been shown to be significantly increased for individuals with diabetic ketoacidosis [24]. Thus acetone breath sensing is a promising tool for non-invasive diagnosis and monitoring of diabetic patients, with the aim for inexpensive, portable and broadly available sensor devices, giving patients cheaper and safer healthcare options, compared to conventional blood tests [25]. Acetone gas sensors based on self-organized MOS nanostructures, namely CuO/Cu₂O/Cu- and Fe₂O₃/Fe-nanospikes, have been manufactured via DIW and subsequent thermal annealing [26, 27]. The composition of the respective inks can be found in Table 1.

Ink lines have been printed on glass substrates with deposited gold contacts (see Fig. 5a–d). The prints have been annealed at 425 °C for 2 h (CuO/Cu₂O/Cu) and 4 h (CuO:Fe₂O₃) respectively, which served as means of binder removal as well as oxidation of the microparticles.

During oxidation self-organized spikes of metal oxide have grown outwards from the particles, significantly increasing their surface area and generating new heterojunctions between particles, as indicated in the micrographs in Fig. 6a–d. Both the

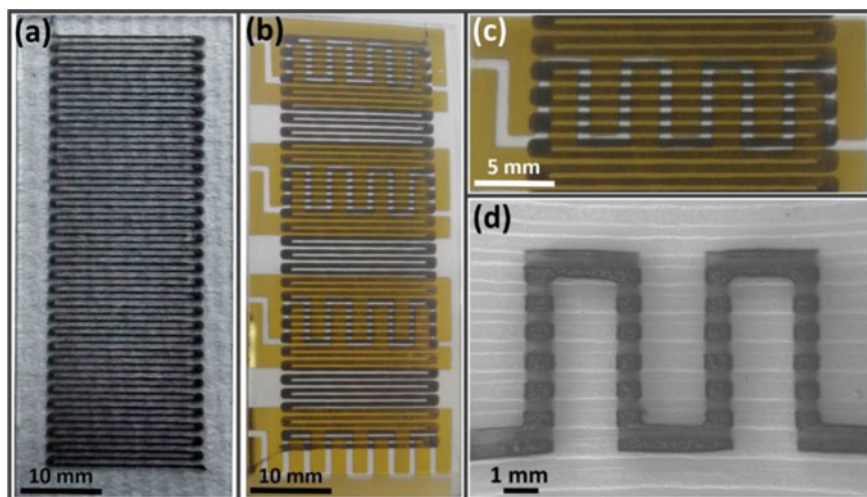
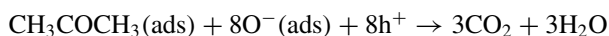


Fig. 5 Images of meandering printed ink lines: **a** Directly after printing. **b** After deposition of gold contacts. **c** Close-up of a single sensor element. **d** SEM micrograph of a sensor element. Reprinted with permission from [27]

CuO and the CuO:Fe₂O₃ sensors showed great selectivity and sensitivity to acetone vapours at an operating temperature of 300 °C or higher (see Fig. 6e+f).

The proposed sensing mechanism is the reaction of acetone with adsorbed surface oxygen:



Since CuO is a p-type semiconductor the sensor response is based on the reduction of the HAL. In the case of CuO:Fe₂O₃ sensors besides the charge accumulation/depletion effect, the interfaces between CuO and Fe₂O₃ nanopikes form p-n-junctions that are proposed to increase in resistance by adsorbed acetone providing an additional potential barrier. These examples showcased that the technology of DIW is perfectly suitable for the design and production of sensing devices. Any species and even combinations of microparticles with sensing properties can be deployed with great geometric freedom, while allowing one to print on pre-designed circuit boards or substrates for immediate integration.

3.5 *Enhancing Sensor Properties by Decoration with Noble Metal Nanoparticles*

The gas sensing performance of MOS sensors with respect to sensitivity, stability or selectivity can be enhanced by doping as well as decoration with metallic nanoparticles. To give an example, ZnO as a H₂ sensor is considered. ZnO nanowires (a

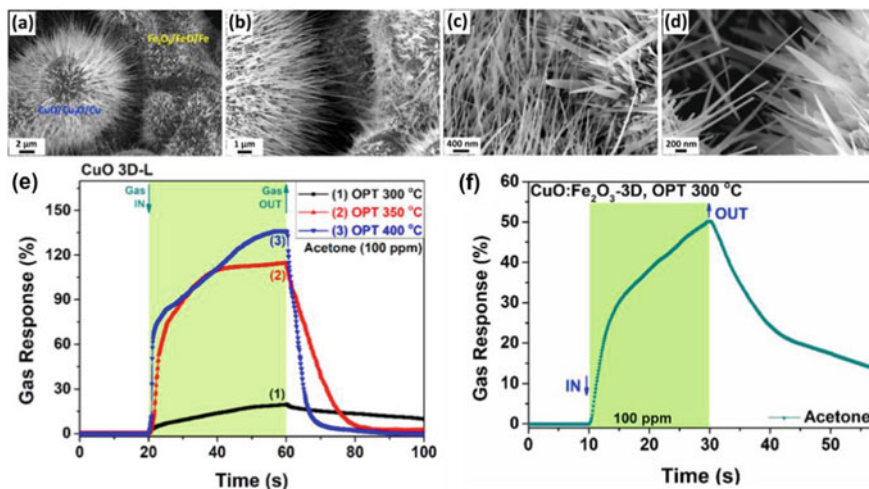
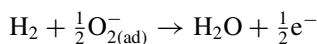


Fig. 6 SEM micrographs of the Cu and Fe microparticles after heat treatment. **a–d** Interface between the CuO/Cu₂O/Cu- and Fe₂O₃/FeO/Fe-nanospikes at different magnifications. **e** Gas response of 3D-printed CuO sensor to 100 ppm acetone vapor. **f** Gas response of CuO:Fe₂O₃ sensor to 100 ppm acetone vapor. Reprinted with permissions from [26, 27]

n-type semiconductor) are showing H₂ detection capabilities. A model proposed for the H₂ sensing mechanism while exposed to ambient air is based on the reaction of incoming H₂ molecules with the adsorbed oxygen to H₂O releasing electrons into the bulk during this process and thus decreasing the electron depletion region: [28]



Though, such sensors are influenced as well by ambient parameters like relative humidity (RH), so that it is beneficial for the sensor performance to alleviate the response to changing conditions [29]. ZnO nanowires that have been decorated with Au nanoparticles (AuNP) have shown a significantly reduced susceptibility to changes in RH between 30% and 85% [30]. The tentatively proposed sensing mechanism for AuNP decorated ZnO nanowires extends the one for pure ZnO wires by taking into account the AuNP/ZnO interface (see Fig. 7). Due to the higher work function of the AuNP ($\Phi_{\text{Au}} = 5.1$ eV, $\Phi_{\text{ZnO}} = 4.5$ eV) electrons flow from the bulk to the interface resulting in Schottky barriers at the Au/ZnO interface thus further enlarging the electron depletion region.

The water desorbs by the heat of the exothermic reaction and by the release of electrons the electron depletion layer is reduced. The higher sensor response for Au NP decorated ZnO is attributed to the larger depletion layer around the Au NP and the induced Schottky barriers, which result in a stronger modulation of the conduction channel [31]. The effect of doping and nanoparticle decoration of MOS sensors has been studied for a wide variety of material systems. An overview of examples is given in Table 2.

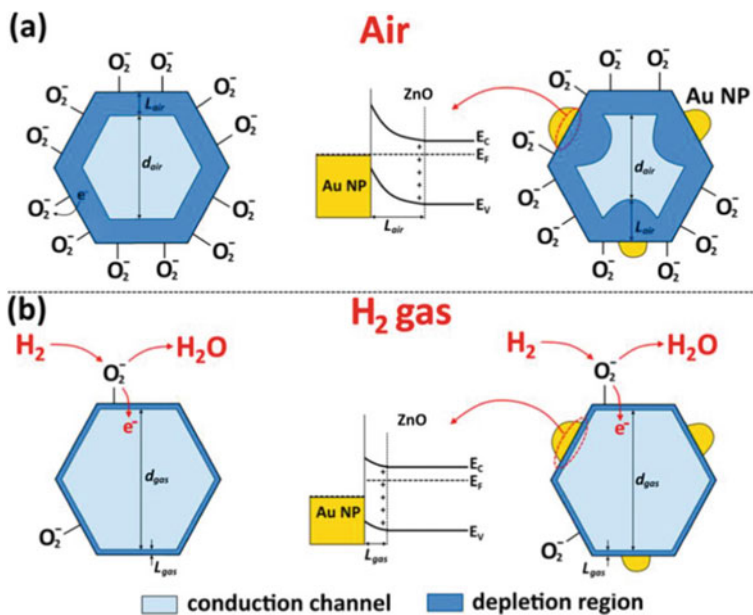


Fig. 7 Proposed mechanism for the gas-sensing of noble nanoparticle decorated ZnO nanowires. **a** Under ambient air. **b** Under exposure to H_2 gas. Reprinted with permission from [31]

Table 2 Overview over the effect of doping and nanoparticle decoration on gas sensor properties for various material systems

Base system	Nanoparticles	Effect	References
ZnO:Fe	AgO/Ag	Enhanced sensitivity to ethanol vapor	[32]
ZnO:Ag	Ag	Enhanced humidity stability	[33]
ZnO:Ag	AgAu	Enhanced sensitivity towards VOCs	[34]
ZnO:Ag	AgPt	Enhanced selectivity for hydrogen	[34]
TiO ₂	Ag, Au, AgAu, AgPt	Enhanced selectivity for various gases	[35, 36]
ZnO:Eu	Pd	Decrease of working temperature for hydrogen sensing to RT	[37]
ZnO:Pd	PdO	Enhanced selectivity and sensitivity to hydrogen gas	[38, 39]
SnO ₂	Zn ₂ SnO ₄	Enhanced sensitivity to ethanol at RT	[40]

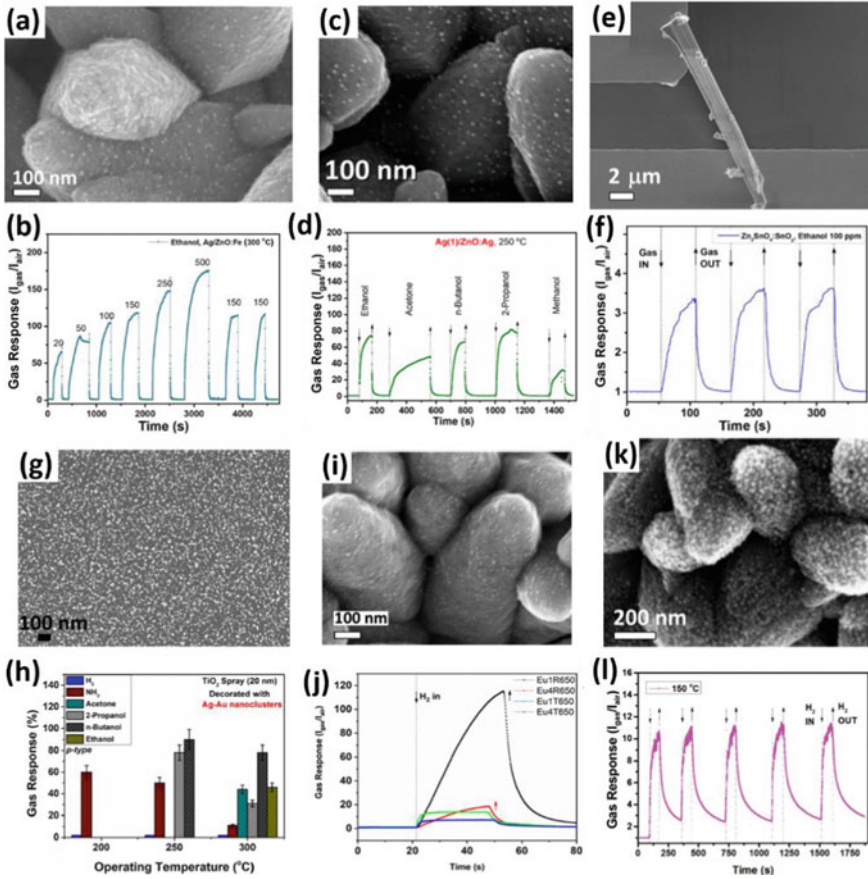


Fig. 8 Collection of doped and/or surface functionalized MOS based gas sensors. For each system an SEM micrograph as well as a characteristic gas sensor measurement is shown: **a+b** Fe doped ZnO decorated with Ag NP [32]. **c+d** Ag doped ZnO decorated with Ag NP [33]. **e+f** SnO₂ decorated with Zn₂SnO₄ NP [40]. **g+h** TiO₂ thin film decorated with AgAu NP [36]. **i+j** Europium doped ZnO decorated with Pd NP [37]. **k+l** Pd doped ZnO decorated with PdO NP [38]. Reprinted with permissions from [32, 33, 36–38, 40]

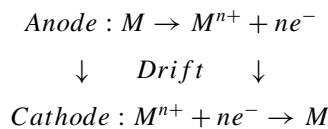
Except for the SnO₂/Zn₂SnO₄ system, all nanoparticles have been deposited with a Haberland-type gas aggregation cluster source, which is described in more detail in the chapter “Memristive Switching: From Individual Nanoparticles Towards Complex Nanoparticle Networks” and allows for precise tailoring of alloy composition during deposition. In Fig. 8 a selection of gas sensors is depicted, with SEM micrographs showing the morphology and nanoparticle decoration, as well as characteristic gas sensor measurements.

The decoration with nanoparticles has shown to enhance the performance of MOS-based sensor devices. A thorough comparison of recent works can be found in [34]. A

deeper understanding of the participating interactions and mechanisms will allow for a tailored design of specialized or optimized material systems for sensor applications. As an example, the concept of the “electronic nose” can be realized by an array of highly selective and stable gas sensors, which is discussed as a valuable tool for the food industry or environmental applications [41, 42].

4 Triggered Assembly

The growing demand for computational power, arising from the development towards AI and the “Internet of Things” [2–4], as well as the resulting energy consumption will require highly efficient computer systems, with one of the main drawbacks of today’s conventional computers being the losses of the von-Neumann architecture. A bio-inspired approach for a computing system based on the mammalian brain uses a low frequency but in a highly parallelized architecture in which memory, computation logic, energy source and cooling are decentralized. One challenge of such a complex system is the communication and thus the interconnections between the parallel components analogous to the connections between cortical regions of the brain [5]. In a brain interconnections are formed by closely timed spiking of neurons, called spike timing-dependent plasticity (STDP), which is, as a Hebbian learning rule, commonly summarized as “what fires together wires together”. Neurons that spike in a window of a few milliseconds before another neuron nearby, potentiate the connection between these improving the efficiency of signal transmission [43]. On a technological level this concept can be transferred to a circuit of initially unconnected electronic components. Conductive interconnections between two nearby terminals are formed, if both are “active” during a narrow timeframe (but not simultaneously). This activity driven formation of conductive interconnections can be established by utilizing electrochemical metallization (ECM), where “activity” in this context is used to describe the potential of a terminal relative to a reference. The working principle of ECM cells is based on applying a voltage to electrodes where at least one of them consists of an electrochemically active material (e.g. Cu or Ag), which leads to a dissolution of metal ions into an electrolyte between the electrodes. The ions drift in the electrical field towards the cathode, where they are reduced and crystallize on the surface:



The crystallized metal forms a filament bridging the electrolyte thus reducing the resistance between the electrodes as long as the filaments holds. ECM is commonly used on the nanoscale as a memristive switching mechanism with an inert counter electrode and a solid electrolyte (e.g. SiO₂) [44, 45]. However, it is also possible to transfer this process to a larger length scale up to a range of several hundreds of

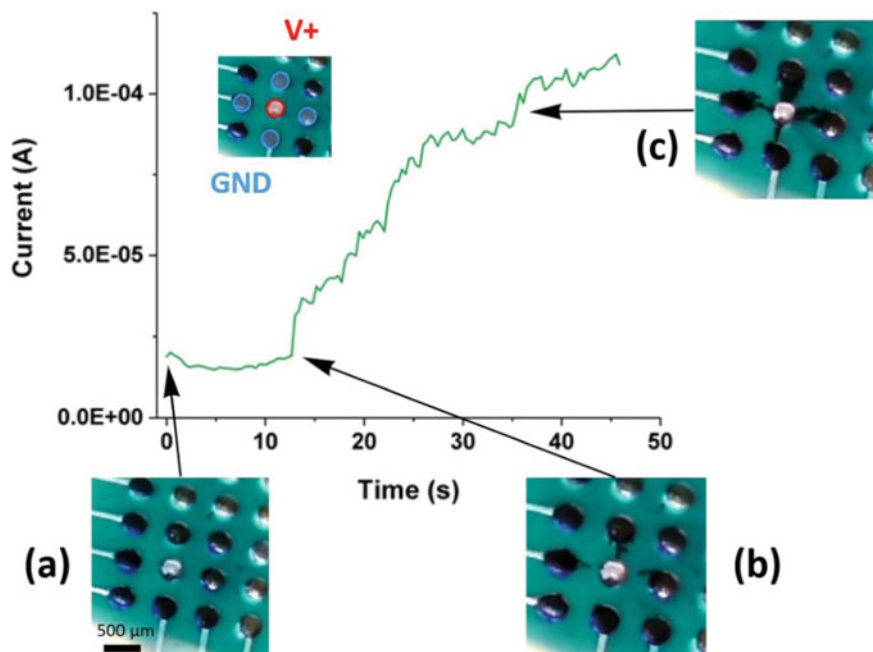


Fig. 9 Photographs and current measurement of an electrochemical metallization process over a distance of $300\ \mu\text{m}$ in an electrolyte of DMSO and MB with copper electrodes under an applied voltage of 15 V. The voltage has been applied to the center electrode. The perpendicular adjacent electrodes have been connected to ground. All remaining ones were left floating. **a** In the beginning the current is determined by the conductivity of the electrolyte. **b** Filaments grow by metallization from cathode towards anode. The first filament establishing contact increases the conductance significantly. **c** Further growth leads to thicker filaments, gradually increasing the current

micrometers. Thus, filament growth by electrochemical metallization occurs as well for Cu electrodes with a spacing of $300\ \mu\text{m}$, when submerged in a liquid electrolyte of dimethyl sulfoxide (DMSO) and methylene blue (MB), which is shown in Fig. 9. By applying a voltage of 15 V to the central electrode and connecting ground to the adjacent ones, a metal filament has formed from cathode to anode directed by the electrical field, significantly increasing the conductance between them. EDX measurements revealed, that the filament formed consists of the electrode material, as seen in Fig. 10a+b. The liquid state facilitates ion mobility and allows for adjustment of electrolyte concentration, which can be used as parameter for the time constants of filament growth (see Fig. 10c).

If a periodic voltage signal is considered for a number of terminals oscillating around a resting potential, filament growth acts as positive feedback for closely but not exactly synchronized oscillators, since a large voltage drop will occur, where repeatedly one terminal is at its maximum while a nearby one is at its minimum. Thus, if the voltage magnitudes are properly set and two nearby terminals spike with similar (but not equal) frequencies over a sufficient period of time, they are

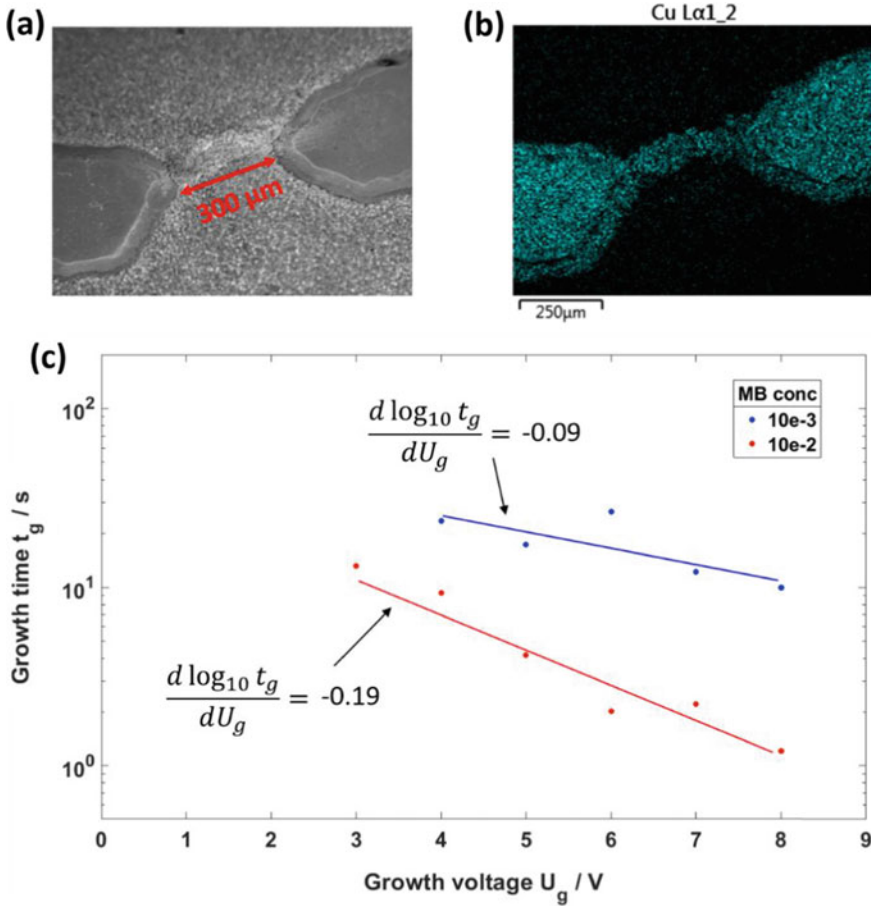


Fig. 10 a+b SEM micrograph and respective EDX map (Cu $L\alpha 1$ signal) of two copper electrodes with grown filament. c Semi-logarithmic plot of time of filament growth over applied voltage for two electrolyte concentrations. The slope has been calculated by linear regression

interconnected by a metal filament drastically increasing the conductance between them. The growth time being in the second- to minute-scale acts as a low-pass filter, so that only those terminals are interconnected, that show a suitable potential drop consistently over a sufficient number of spikes. The larger length scale makes this approach a suitable base for integration into small-world setups and furthermore a combination with sensor devices allows the terminal activity to be driven by external stimuli.

5 Conclusion

The advance towards a data driven society with the “Smart City” and “Internet of Things” paradigms requires an ever growing demand of monitoring and sensing devices as well as imposing new challenges with respect to performance, energy efficiency and data processing but also design, scalability and fabrication of sensor devices. In this chapter various concepts have been discussed, tackling these challenges. A general model for a memsensor, whose equivalent circuit model consists of a stimulus dependent resistor (i.e. sensor) and two memristive components, has shown unique features like adaptation to a permanent background stimulus. It serves as a foundation for the interpretation of experimental findings and prediction of the behaviour of memsensitive devices, facilitating the development of smart, decentralized sensor systems. The unique features of memsensitive devices can be applied e.g. for monitoring the concentration of hazardous gases in a storage cabin. The sensor should not react to the slowly changing permanent background which is based on the number and kind of chemicals in the cabin, but only to quick changes such as when a container is damaged. Furthermore, it has been pointed out that additive manufacturing is a perfect candidate for the inexpensive, scalable fabrication of sensor devices with great design freedom in terms of geometry and material space as well as rapid iteration cycles. Microparticle based MOS-sensors have been shown as examples for sensor devices fabricated by DIW. An overview over various material system has been given, showcasing the enhancement of the performance, selectivity or stability of MOS-sensors by doping and decoration with nanoparticles, which allows for tailoring or designing sensor materials for specific applications. Finally, an approach is introduced for dynamically establishing conductive interconnections between metal electrodes in a liquid matrix. The concept is based on electrochemical metallization stimulated by voltage pulses. It acts as a self-organization mechanism to increase the conductance between active parts of a circuit resembling STDP. Using sensor devices, the “activity” can be driven by external stimuli, which offers a new way of introducing neuromorphic plasticity into sensor circuits.

References

1. Webster, M.A.: Evolving concepts of sensory adaptation. *F1000 Biol. Rep.* **4**, 21 (2012). <https://doi.org/10.3410/B4-21>
2. Zanella, A., et al.: Internet of things for smart cities. *IEEE Internet Things J.* **1**(1), 22–32 (2014). ISSN: 2327-4662. <https://doi.org/10.1109/JIOT.2014.2306328>
3. Bellavista, P., et al.: Convergence of MANET and WSN in IoT urban scenarios. *IEEE Sens. J.* **13**(10), 3558–3567 (2013). <https://doi.org/10.1109/JSEN.2013.2272099>
4. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. *Comput. Netw.* **54**(15), 2787–2805 (2010). ISSN: 13891286. <https://doi.org/10.1016/j.comnet.2010.05.010>
5. Merolla, P.A., et al.: Artificial brains. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science (New York, N.Y.)* **345**(6197), 668–673 (2014). <https://doi.org/10.1126/science.1254642>

6. Khan, W.Z., et al.: Edge computing: a survey. *Future Gener. Comput. Syst.* **97**(99), 219–235 (2019). <https://doi.org/10.1016/j.future.2019.02.050>
7. Shaafie, M., Logeswaran, R., Seddon, A.: Over-coming the limitations of von Neumann architecture in big data systems. In: Bansal, A., Singhal, A., (eds.), *Proceedings of the 7th International Conference Confluence 2017 on Cloud Computing, Data Science and Engineering*, pp. 199–203. IEEE, Piscataway (2017). isbn: 978-1-5090-3519-9. <https://doi.org/10.1109/CONFLUENCE.2017.7943149>
8. Kendall, J.D., Kumar, S.: The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7**(1), 011305 (2020). ISSN: 1931–9401. <https://doi.org/10.1063/1.5129306>
9. Zou, L., et al.: Novel tactile sensor technology and smart tactile sensing systems: a review. *Sensors (Basel, Switzerland)* **17**(11) (2017). <https://doi.org/10.3390/s17112653>
10. Nazemi, H., et al.: Advanced micro- and nano-gas sensor technology: a review. *Sensors (Basel, Switzerland)* **19**(6) (2019). <https://doi.org/10.3390/s19061285>
11. Vahl, A., et al.: Concept and modelling of memsensors as two terminal devices with enhanced capabilities in neuromorphic engineering. *Sci. Rep.* **9**(1), 4361 (2019). <https://doi.org/10.1038/s41598-019-39008-5>
12. Chiolerio, A., et al.: Ultraviolet mem-sensors: Flexible anisotropic composites featuring giant photocurrent enhancement. In: *Nano Res.* **8**(6), 1956–1963 (2015). ISSN: 1998-0124. <https://doi.org/10.1007/s12274-014-0705-2>
13. Wang, X., et al.: Spintronic memristor temperature sensor. *IEEE Electron Device Lett.* **31**(1), 20–22 (2010). ISSN: 0741-3106. <https://doi.org/10.1109/LED.2009.2035643>
14. Li, H., et al.: Light and magnetic field double modulation on the resistive switching behavior in BaTiO₃/FeMn/BaTiO₃ trilayer films. *Phys. Lett. A* **381**(25–26), 2127–2130 (2017). ISSN: 03759601. <https://doi.org/10.1016/j.physleta.2017.04.039>
15. Ngo, T.D., et al.: Additive manufacturing (3D printing): a review of materials, methods, applications and challenges. *Compos. Part B: Eng.* **143**(2), 172–196 (2018). ISSN: 13598368. <https://doi.org/10.1016/j.compositesb.2018.02.012>
16. Kaufui V. Wong and Aldo Hernandez. “A Review of Additive Manufacturing. In: *ISRN Mechanical Engineering 2012.4* (2012), pp. 1–10. ISSN: 2090-5130. <https://doi.org/10.5402/2012/208760>
17. Lewis, J.A.: Direct ink writing of 3D functional materials. *Adv. Funct. Mater.* **16**(17), 2193–2204 (2006). <https://doi.org/10.1002/adfm.200600434>
18. Lewis, J.A.: Direct-write assembly of ceramics from colloidal inks. *Current Opinion Solid State Mater. Sci.* **6**(3), 245–250 (2002). ISSN: 13590286. [https://doi.org/10.1016/S1359-0286\(02\)00031-1](https://doi.org/10.1016/S1359-0286(02)00031-1)
19. Rueschhoff, L., et al.: Additive manufacturing of dense ceramic parts via direct ink writing of aqueous alumina suspensions. *Int. J. Appl. Ceram. Technol.* **13**(5), 821–830 (2016). <https://doi.org/10.1111/ijac.12557>
20. Paulowicz, I., et al.: Zinc oxide nanotetrapods with four different arm morphologies for versatile nanosensors. *Sens. Actuat. B: Chem.* **262**, 425–435 (2018). ISSN: 09254005. <https://doi.org/10.1016/j.snb.2018.01.206>
21. Wagner, S., et al.: Electronic skin: architecture and components. *Phys. E: Low-dimensional Syst. Nanostruct.* **25**(2–3), 326–334 (2004). ISSN: 13869477. <https://doi.org/10.1016/j.physe.2004.06.032>
22. Wang, X., et al.: Recent progress in electronic skin. *Adv. Sci. (Weinheim, Baden-Wuerttemberg, Germany)* **2**(10), 1500169 (2015). ISSN: 2198-3844. <https://doi.org/10.1002/advs.201500169>
23. Ji, H., Zeng, W., Li, Y.: Gas sensing mechanisms of metal oxide semiconductors: a focus review. *Nanoscale* **11**(47), 22664–22684 (2019). <https://doi.org/10.1039/c9nr07699a>
24. Anderson, J.C.: Measuring breath acetone for monitoring fat loss: review. *Obesity (Silver Spring, Md.)* **23**(12), 2327–2334 (2015). <https://doi.org/10.1002/oby.21242>
25. Saasa, V., et al.: Sensing technologies for detection of acetone in human breath for diabetes diagnosis and monitoring. *Diagnostics (Basel, Switzerland)* **8**(1) (2018). ISSN: 2075-4418. <https://doi.org/10.3390/diagnostics8010012>

26. Lupan, O., Siebert, L.: 3D-printed chemiresistive sensor array on nanowire CuO/Cu₂O/Cu heterojunction nets. *ACS Appl. Mater. Interfaces* (2019)
27. Siebert, L., et al.: Facile fabrication of semiconducting oxide nanostructures by direct ink writing of readily available metal microparticles and their application as low power acetone gas sensors. *Nano Energy* **70**(2), 104420 (2020). ISSN: 22112855. <https://doi.org/10.1016/j.nanoen.2019.104420>
28. Lupan, O., et al.: Selective hydrogen gas nanosensor using individual ZnO nanowire with fast response at room temperature. In: *Sens. Actuat. B: Chem.* **144**(1), 56–66 (2010). ISSN: 09254005. <https://doi.org/10.1016/j.snb.2009.10.038>
29. Lupan, O., et al.: Highly sensitive and selective hydrogen single-nanowire nanosensor. *Sens. Actuat. B: Chem.* **173**(3) 772–780 (2012). ISSN: 09254005. <https://doi.org/10.1016/j.snb.2012.07.111>
30. Lupan, O. et al.: Room temperature gas nanosensors based on individual and multiple networked Au-modified ZnO nanowires. In: *Sens. Actuat. B: Chem.* **299**, 126977 (2019). ISSN: 09254005. <https://doi.org/10.1016/j.snb.2019.126977>
31. Lupan, O., et al.: Low-temperature solution synthesis of Au-modified ZnO nanowires for highly efficient hydrogen nanosensors. In: *ACS Applied Materials and Interfaces* (2019). <https://doi.org/10.1021/acsami.9b08598>
32. Postica, V., et al.: Tuning doping and surface functionalization of columnar oxide films for volatile organic compound sensing: experiments and theory. *J. Mater. Chem. A* **6**(46) 23669–23682 (2018). ISSN: 2050-7488. <https://doi.org/10.1039/c8ta08985j>
33. Postica, V., et al.: Tuning ZnO sensors reactivity toward volatile organic compounds via Ag doping and nanoparticle functionalization. *ACS Appl. Mater. Interfaces* **11**(34), 31452–31466 (2019)
34. Vahl, A., et al.: Surface functionalization of ZnO: Ag columnar thin films with AgAu and AgPt bimetallic alloy nanoparticles as an efficient path-way for highly sensitive gas discrimination and early hazard detection in batteries. *J. Mater. Chem. A* **8**(32), 16246–16264 (2020). ISSN: 2050-7488. <https://doi.org/10.1039/d0ta03224g>
35. Lupan, O., et al.: Ultra-thin TiO₂ films by atomic layer deposition and surface functionalization with Au nanodots for sensing applications. *Mater. Sci. Semiconduct. Process.* **87**, 44–53 (2018). ISSN: 13698001. <https://doi.org/10.1016/j.mssp.2018.06.031>
36. Ababii, N., et al.: Effect of noble metal functionalization and film thickness on sensing properties of sprayed TiO₂ ultra-thin films. *Sens. Actuat. A: Phys.* **293**, 242–258 (2019). ISSN: 09244247. <https://doi.org/10.1016/j.sna.2019.04.017>
37. Lupan, C., et al.: Pd-functionalized ZnO: Eu columnar films for room-temperature hydrogen gas sensing: a combined experimental and computational approach. *ACS Appl. Mater. Interfaces* **12**(22), 24951–24964 (2020). <https://doi.org/10.1021/acsami.0c02103>
38. Lupan, O., et al.: PdO/PdO₂ functionalized ZnO: Pd films for lower operating temperature H₂ gas sensing. *Nanoscale* **10**(29), 14107–14127 (2018). <https://doi.org/10.1039/c8nr03260b>
39. Lupan, O. et al.: Functionalized Pd/ZnO nanowires for nanosensors. *Physica Status solidi (RRL) - Rapid Res. Lett.* **12**(1), 1700321 (2018). ISSN: 18626254. <https://doi.org/10.1002/pssr.201700321>
40. Lupan, O. et al.: Properties of a single SnO₂: Zn₂SnO₄ – functional-ized nanowire based nanosensor. *Ceram. Int.* **44**(5), 4859–4867 (2018). ISSN: 02728842. <https://doi.org/10.1016/j.ceramint.2017.12.075>
41. Wilson, A., Baietto, M.: Applications and advances in electronic-nose technologies. *Sensors* **9**(7), 5099–5148 (2009). ISSN: 1424-8220. <https://doi.org/10.3390/s90705099>
42. Boeker, P.: On ‘Electronic Nose’ methodology. *Sens. Actuat. B: Chem.* **204**, 2–17 (2014). ISSN: 09254005. <https://doi.org/10.1016/j.snb.2014.07.087>
43. Caporale, N., Dan, Y.: Spike timing-dependent plasticity: a Hebbian learning rule. In: *Annual Review of Neuroscience*, vol. 31, pp. 25–46 (2008). ISSN: 0147-006X

44. Valov, I., et al.: Electrochemical metallization memories—fundamentals, applications, prospects. *Nanotechnology* **22**(28), 289502 (2011). ISSN: 0957-4484. <https://doi.org/10.1088/0957-4484/22/28/289502>
45. Peng, S., et al.: Mechanism for resistive switching in an oxide-based electrochemical metallization memory. *Appl. Phys. Lett.* **100**(7), 072101 (2012). ISSN: 0003-6951. <https://doi.org/10.1063/1.3683523>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Bio-inspired, Neuromorphic Acoustic Sensing



Claudia Lenk, Kalpan Ved, Steve Durstewitz, Tzvetan Ivanov, Martin Ziegler, and Philipp Hövel

Abstract We present an overview of recent developments in the area of acoustic sensing that is inspired by biology and realized by micro-electromechanical systems (MEMS). To support understanding, an overview of the principles of human hearing is presented first. After the review of bio-inspired sensing systems, we continue with an outline of an adaptable acoustic MEMS-based sensor that offers adaptable sensing properties due to a simple, real-time feedback. The transducer itself is based on an active cantilever, which offers the advantage of an integrated deflection sensing based on piezoresistive elements and an integrated actuation using thermomechanical effects. We use a feedback loop, which is realized via a field-programmable gate array or analog circuits, to tune the dynamics of the sensor system. Thereby, the transfer characteristics can be switched between active, linear mode, for which the sensitivity and minimal detectable sound pressure level can be set by the feedback strength (similar to control of the quality factor), and an active nonlinear mode with compressive characteristics. The presented sensing system, which is discussed both from an experimental and theoretical point of view, offers real-time control for adaptation to different environments and application-specific sound detection with either linear or nonlinear characteristics.

Keywords Bio-inspired acoustic sensing · Micro-electromechanical system (MEMS) · Mathematical modelling · Real-time feedback · Adaptive sensing

C. Lenk (✉) · K. Ved · S. Durstewitz · T. Ivanov · M. Ziegler
Department of Microelectronic and Nanoelectronic Systems, Technische Universität Ilmenau,
98684 Ilmenau, Germany
e-mail: claudia.lenk@tu-ilmenau.de

P. Hövel
Theoretical Physics and Center for Biophysics, Universität des Saarlandes, 66123 Saarbrücken,
Campus E2 6, Saarlandes, Germany
e-mail: philipp.hoevel@uni-saarland.de

© The Author(s) 2024
M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series
on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_12

1 Introduction

Current technological sound and speech processing systems typically consist mainly of three parts (schematically shown in Fig. 1): a sensing unit, a pre-processing unit and the sound analysis/processing unit. The sensing unit consists of one or more (typically capacitive) microphone based on micro-electro-mechanical system (MEMS)-technology with linear transfer characteristics and 120 dBA as maximal detectable sound pressure level (SPL), a self-noise floor (corresponding to the minimal detectable SPL) of 20 – 30 dB (absolute minimum at the moment (non-MEMS): ≈ 7 dB), and a frequency range of 20 Hz-20 kHz [1]. The pre-processing unit is applied to tune the sensed signals for easier processing/analysis. It typically includes a pre-amplifier, amplifying the microphones signals, and filter stages, decomposing the signal into different frequency bands, and additional stages to include e.g. weighting functions, correcting the signal and/or digitalizing it. The processing unit is based on signal processing stages and/or on learning systems such as convolutional, recurrent and spiking neural networks. Thereby, the pre-processing of the signals shall improve the processing performance by extracting important sound features, which are then fed into the neural networks (see e.g. [2]). Typical low-complexity features are, for example, the envelope of the sound signal, the RMS energy, the frequency (given by the zero crossing rate or the spectral centroid) or spectral spread or flux. Alternatively, the sound data is fed directly into the NN, known as end-to-end learning, or first transformed into a time-frequency representation, known as spectrogram, before it is fed into the network. The goal is to separate the relevant information like speech from irrelevant one like noise.

The performance of speech processing systems has strongly increased in recent years [3–7]. This can be mainly attributed to the improvements in the pre-processing stage and the sound processing networks, since microphone technology including readout is well developed with easy and cheap fabrication, small size and large dynamic ranges and performance. The pre-processing stage typically includes pre-amplifiers, amplifying the microphone signals, and filter stages, decomposing the signal into different frequency bands, and additional stages to include, e.g., weighting functions, correct and/or digitalize the signal. The development of pre-processing stage was driven mainly by implementing bio-inspired pre-processing (e.g. nonlinear amplification) as well as developing extensive noise cancellation algorithms. Thereby, it was recently shown that nonlinear, bio-inspired pre-processing based on frequency filtering and nonlinear amplification of signals is particularly important for sound processing, since it yields strongly increased success rates for speech classification tasks [8]. The increasing demands in computation power and energy can be addressed by neuromorphic, bio-inspired pre-processing implementations in hardware like silicon cochlea, Hopkins electronic ear, AER-EAR, or the FPGA cochlea [9]. These model sound processing by the cochlea in the inner ear, as schematically shown in Fig. 1 and, compared to software implementations, offer the advantage of low latency, real-time performance, reduction of redundant information and data streaming (due to asynchronous event-based spiking output) and lower power con-

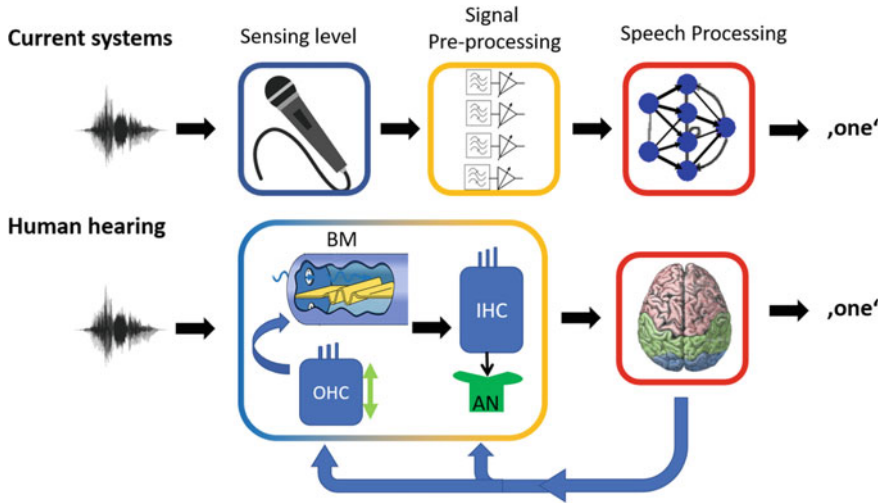


Fig. 1 Comparison of steps in technological speech processing systems and the human hearing system. Detailed description of human hearing is given in chapter 2. BM-basilar membrane, OHC-outer hair cell, IHC-inner hair cell, AN-auditor nerve

sumption. Noise cancellation algorithms can be applied to improve sound processing in more complex hearing situations with multiple sources, but these increase the necessary computing power and energy consumption. In both cases, the idea is to improve the sound feature representation in the audio signals to allow a better sound processing in the neural networks.

Despite these developments, room reverberation (echos), interfering noise, or any other perturbation to the signal can critically affect the underlying feature representation and thus limit sound processing performance, in particular, at low signal-to-noise ratios (SNRs) [10–12]. Further challenges arise from the inability to separate individual sound sources from a mixed acoustic signal, to generalize to unknown acoustic conditions, and to run on low-power embedded devices. Adaptation of the system parameters can help to overcome the issues arising from changing acoustic environments such as room reverberations, noise level etc. Current research focused on adaptation of the pre-processing stage: changing the gain of the pre-amplifier [13], the settings of the filter banks [14], how signals are combined for improved directionality [15], microphone switching to increase dynamic range [16] or automatic gain control in neuromorphic systems [17, 18]. Although methods to adapt the sensing properties of microphones have been published, e.g., adjusting sensitivity by changing the bias voltage or effective microphone area [19], these are not (widely) applied. While the above described developments could improve sound processing performance for standard hearing conditions (combined with higher computing power demands), until now satisfying solutions for reasonable sound processing of signals in noisy environments, i.e., with low signal-to-noise ratios, low volume of important sounds or loud masking sounds, could not be derived.

Thus, improving the sensor stage is the next logical step to bridge this gap. In this chapter we will review (some of) the research in bio-inspired acoustic sensors. For a better understanding, we will shortly summarize what is known about the acoustic sensing in humans until the transduction stage, including the processing and encoding mechanisms occurring at this level. Following this, several bio-inspired sensors will be described in relation to their processing capabilities, in particular frequency decomposition and nonlinear filtering of the sound signals, and their adaptation properties, which are the most important functionalities in acoustic bio-inspired sensing. Finally, we will present a bio-inspired acoustic sensor [20–22] that includes all the above mentioned three functionalities, i.e. frequency decomposition, nonlinear amplification and adaptation.

2 Human Hearing/Auditory Pathway

In this section, we will give a short overview of the hearing process until the transduction of acoustic signals with a particular focus on the pre-processing of the signal taking place before the actual sound analysis in the cortex (see Fig. 1). The description is mainly based on the information given in [23, 24] and is separated according to the three parts of the ear (see Fig. 2): (i) the outer ear, which includes the pinna and the external auditory canal, (ii) the middle ear with the ossicular chain consisting of incus, malleus and stapes, and (iii) the inner ear with the cochlea.

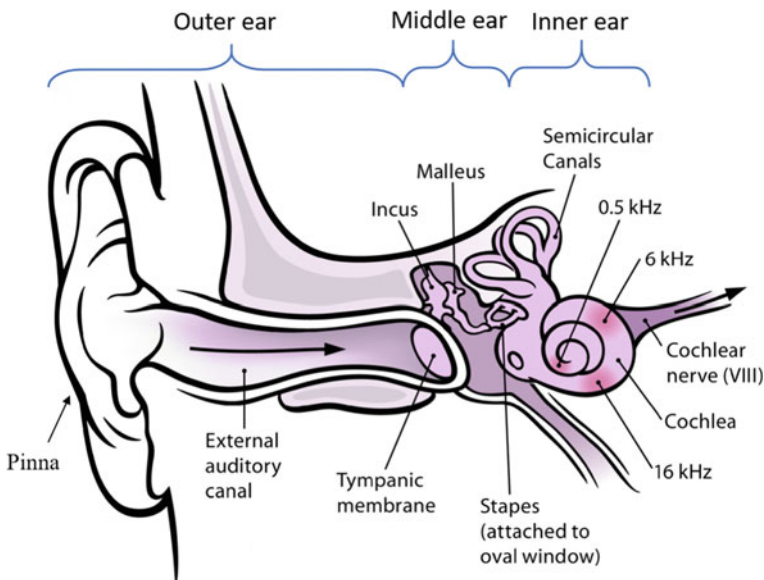


Fig. 2 Schematic illustration of the ear. Modified from [25]

Human hearing starts with the pinna, which funnels the sound signals into the ear canal, further transferring the signal to the tympanic membrane. Even in this very first step of the hearing process, several steps before the actual transduction of sound signals into electric signals (action potentials/spikes), a pre-processing/conditioning of the acoustic signals is present, which shall support later sound processing and sound perception. In detail, the transfer characteristics of the outer ear pinna and ear canal is on the one hand direction-dependent, providing localization cues, and on the other hand includes frequency-selective amplification of sound signals. The latter amplifies the sound signals in the important range of 200Hz to 10kHz, with particularly strong amplification in the range of 2-4kHz, the range of highest sensitivity for hearing [26]. This pre-conditioning is mainly constant and not adaptive.

In the next step of the hearing process, the vibrations of the tympanic membrane due to sound excitation are transferred to the inner ear by the ossicular chain in the middle ear. Its task is to provide an impedance matching between the air-based sound in the outer ear and the fluid-based sound propagation in the inner ear. Without this matching, most of the sound energy would be reflected at the air-fluid interface with a reflectance of approximately 99.9%. The transfer characteristics is almost flat in the range between 200Hz and 8kHz (small peaks around 1kHz-2kHz and 4-5kHz) and decreases outside this range. Important to note here is that the efficiency of sound transfer can be adapted for small frequencies below 2kHz by contraction of the stapedius muscle. Its contraction yields an increase of the stiffness of the ossicular chain, thus reducing the sound energy transferred to the inner ear. The main purpose of this adaptation is a protection of the inner ear from damage due to loud noises and an attenuation of the perception of ones own voice.

The final part in the acoustic sensing process is the inner ear, the cochlea to be accurate. The cochlea itself, a bony, snail shell-resembling structure consists of three fluid-filled chambers: the scala vestibuli, the scala tympani and between both the scala media, hosting the organ of corti (see Fig. 3). Thereby, the scala vestibuli is separated from the scala media by a thin membrane, the Reissner's membrane. On the lower side, scala media is separated from the scala tympani by the basilar membrane (BM).

The pressure wave, initiated by the stapes at the oval window of the cochlea, propagates through the scala tympani up to the apex of the cochlea, where it is transferred to the scala vestibuli through a small connection of both chambers and then propagates in this chamber down to the base of the cochlea until finally reaching the round window (see Fig. 4). In response to the propagating sound wave, the basilar membrane and Reissners membrane oscillate. Thereby, due to a varying stiffness and thickness of the basilar membrane along the cochlea length, resonance occurs at different locations in the cochlea depending on the frequency of the sound wave (traveling wave theory), yielding an effective *frequency decomposition* of the input signal. This encoding of sound frequency into a location/place along the cochlea is termed tonotopy and the frequency-place coding is kept throughout the complete auditory pathway up to the auditory cortex in the brain. Thereby, humans can distinguish frequencies differing only by 0.1% [28]. The sharpness of the frequency

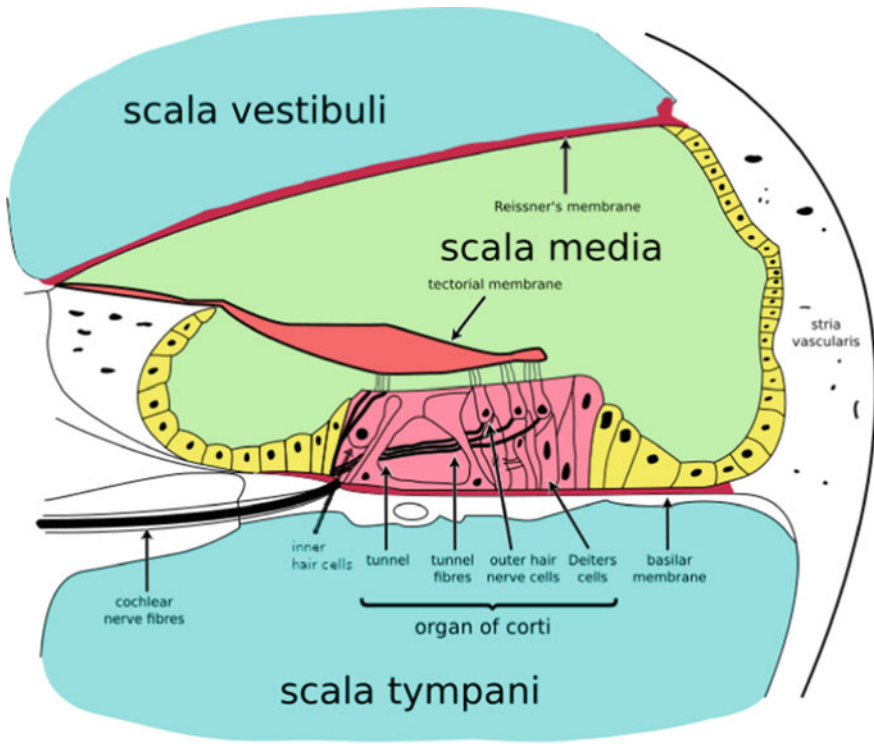


Fig. 3 Schematic showing the cross-section of the cochlea with focus on the organ-of-corti, which is the place of sound transduction. Thereby, the organ of corti is located in the the scala media, one of the three fluid-filler chambers of the cochlea with scala tympani below and scala vestibuli above. The three chambers are separated by Reissner's membrane on top and basilar membrane below, which hosts the organ of corti. From [27]

resolution and cochlear sensitivity is discussed to be further influenced by the longitudinal variation in stiffness of the tectorial membrane (TM) [29].

Situated on the basilar membrane and connected on top with the tectorial membrane in the scala media is the organ of corti (OoC), which hosts the actual acoustic transducers/receptors of human hearing, namely the hair cells. These are cells with cilia on top, giving them their name. In the inner ear, there are two types of hair cells, the inner hair cells (IHC) and the outer hair cells (OHC), with different functionalities, morphology and purpose [23, 31]. Both of them have three rows of differently sized cilia on top, whereby these are V or W-shaped on top of outer hair cells and in a straight line for inner hair cells.

The inner hair cells perform the acoustic sensing/transduction in the following way. The sound pressure wave deflects the cilia on top of the inner hair cell. If the cilia are bent towards the largest ones, potassium channels on top of the cell open. The potassium influx thus polarizes the hair cell and triggers the opening of

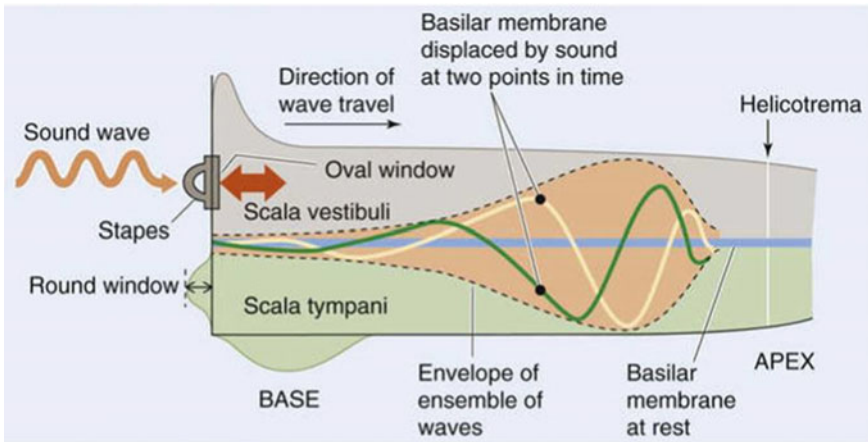


Fig. 4 Schematic of a traveling wave propagation in the cochlea. From [30]

calcium channels at the bottom. This in turn initiates the release of neurotransmitter glutamate into the synapse between hair cell and acoustic nerve. Glutamate then activates the connected acoustic nerve, yielding to an increase of its spike rate. If cilia are unbend/not deflected, a part of the potassium channels is open, which yields a constant small neurotransmitter release and a constant spiking of the acoustic nerve cell. If the cilia are bent into the other direction, i.e., towards the smallest cilia, (during the second half of the sound sine wave) potassium channels close, calcium channels close and no neurotransmitter is released. This yields a reduction of spike rate at the acoustic nerve. Thus, the sound pressure changes are transformed into the changes of the spike rate at the acoustic nerve (electrical signal).

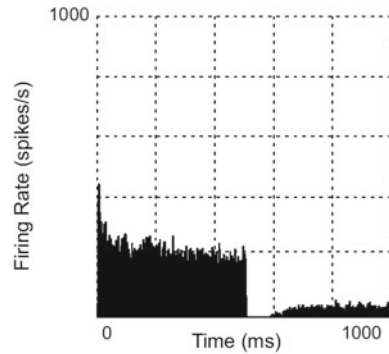
Here, additional encoding of sound information takes place, in particular of sound amplitude, sound frequency and time course of sound signals, as summarized in Table 1.

The above described phase locking of polarization/hyperpolarisation of the hair cell with the sound input is possible only up to frequencies in the range of 2kHz. The phase locking is another encoding of sound frequency, additional to the frequency-place code discussed above. For frequencies larger than 2kHz, the alternating component of the hair cell potential decreases while the DC component, also termed direct component, increases due to incomplete repolarisation of the cell. In response to this, the spike rate on the auditory nerve will increase strongly at first, but then decrease fast and settle at a certain plateau value in response to a constant pure tone input (see Fig. 5). This is termed rapid and *short-term adaptation* and can be modelled using e.g. exponential and power-law adaptation [32]. At the offset of the sound input, the spiking rate is strongly reduced below the spontaneous spiking rate (even up to cessation) and slowly recovers to the spontaneous spiking rate. Besides the duration of the stimuli, the adaptation behaviour depends on previous stimuli, sound amplitude, previous stimulation history, and the spontaneous rate of the fibre [33, 34], making

Table 1 Pre-processing in the sensing process. Abbreviations: BM-basilar membrane, IHC-inner hair cell, OHC-outer hair cell, AN-auditory nerve

Location	Pre-processing functionality	Functional relevance	Realisation of functionality
Outer ear pinna and ear canal	Direction-dependent transfer characteristics	Localisation	Shape of pinna
	Frequency-selective amplification	Amplification of important frequency range	Shape of ear canal
Middle ear ossicular chain	Damping of sounds	Protection of inner ear from loud sounds (incl. own voice) and reducing masking by own voice	Contraction of muscles reduces flexibility of ossicular chain/reduced efficiency in transfer property
Inner ear BM, IHC	Frequency decomposition	Sound analysis	Resonance behaviour of BM (varying stiffness and mass), phase locking of IHC, filtering properties of AN
OHC	Compressive amplification (nonlinear filtering)	Detection of sounds near or below thermal noise levels, dynamic range increase, amplitude-dependent resolution	Active feedback from OHC (by somatic motility and hair bundle motility)
IHC-AN synapse, OHC	Adaptation	Improve perception in noisy environments, highlight sound features, increase efficiency (less spikes)	Efferent feedback, which can damp OHC motility (using calcium-controlled electromotility), and depletion of neurotransmitter at synapse
IHC-AN synapse	Amplitude-encoding	Map large dynamic range to possible range of spike rates	Spike rate encoding by single neuron (sigmoid dependence on amplitude), combination of response from multiple fibres at one IHC with different sensitivities and thresholds

Fig. 5 Short-term adaptation of spike rate observed at synapse between inner hair cell and auditory nerve for a constant amplitude sound input lasting 500ms (from [32]). Shown is the histogram obtained from using the model by Zilany et al. [32]



it rather complex for modeling. The mechanism underlying the adaptation process is not completely understood. Discussed reasons are the depletion of neurotransmitter or the desensitization of post-synaptic receptors (see e.g. [35, 36]). Whether this adaptation is important for the sound processing is unclear. A discussed functional relevance of this adaptation is the shaping of the onset response, highlighting the sound onset and offset, which might be important, e.g., for time-of-arrival detection used in localisation.

Furthermore, the amplitude of the sound is encoded in the spike rate in three different ways. First, for a single fibre of the auditory nerve, the spike rate depends nonlinear (sigmoidal) on the amplitude of the receptor (hair cell) potential, i.e., for potentials below and above a certain threshold the spiking rate is constant and in between these thresholds a correlation between spike rate and amplitude is observed. The range between thresholds is small compared to the dynamic range of hearing, but enables a high resolution of amplitude encoding. To cover a larger dynamic range, multiple fibres with different sensitivities (from low over medium to high) are connected to one IHC. The combination of their spike rates in response to a sound enables the encoding of a larger amplitude range. For low volume sounds, only the highly sensitive fibres are encoding the sound amplitude. For larger sound levels, the high sensitivity fibres enter the saturation regime (higher threshold), while the medium fibres cross their lower threshold, thus encoding the amplitude of the stimulus. For even higher sound levels, medium sensitivity fibres saturate and low sensitivity fibres can encode the amplitude. For very large sound levels, the region of activated sound fibres becomes much larger than expected from the characteristic frequency. In this case, also the number of activated fibres (size of activated area) encodes the sound amplitude. This combination of different sensitivities helps to encode a large dynamic range.

Nevertheless, the hair cell has to respond to a large range of amplitudes (120 dB) and sound pressure levels (SPL). For the lower SPL, only sub-nm deflections of cilia are generated and the energy of the sound wave exhibits energy levels in the range of thermal noise [37, 38]. Reliable sensing of low SPLs and the large dynamic range would not be possible for purely linear sensors. Here another effect enables

the large amplitude range, namely a *compressive (nonlinear) amplification* before the transduction. Thereby, two mechanisms are discussed as main source for the nonlinear amplification, which are hair bundle motility and somatic motility of outer hair cells [39]. The hair bundle is the group of cilia on top of a hair cell. These are linked by filaments, called tip-links. If these tip-links are stretched, potassium channels open (starting the activation of the hair cell+ auditory nerve). An adaptation mechanism was proposed, which yields a downward movement of the plate of the tip link and thus a reduction in its stretching upon a constant force. This creates a region of negative stiffness in the otherwise linear force-displacement relationship [24]. Due to the nonlinear relationship, the hair bundle dynamics are nonlinear as well, including nonlinear dynamics effects like bifurcation to self-excited oscillations etc. The second discussed effect is somatic motility [40], which is a stretching and compressing movement of the outer hair cell, changing its length. For reviews, see [31, 39, 41]. Here, the influx of potassium (after deflection-induced channel opening) results in the dislocation of chloride ions from prestin molecule. Prestin is a molecule abundant in the OHC membrane, which changes its shape depending on the voltage due to the insertion or removal of chloride ions. If chloride ions leave the prestin molecule, it compresses, which yields a compression of the outer hair cell. This happens when the basilar membrane moves upwards (towards the tectorial membrane). Since the outer hair cell is connected to both (TM and BM), the compression of the OHC yields an amplification of the upward BM motion, i.e. an amplification of the sound input. For the other case, i.e., BM moves downwards (away from TM), potassium channels of OHC close, resulting in the returning of chloride ions into the prestin molecule, which in turn stretches. Thus, the OHC stretches and amplifies the downward motion of the BM.

Besides mammals, only birds, crocodiles, frogs and lizards have two types of hair cells, whereby only for mammals one might be capable of somatic motility [40, 42, 43]. Nevertheless, amphibians and reptiles like frogs, turtles, and lizards exhibit nonlinear dynamics in the hearing process as well [44–49]. Therefore, it was proposed [24, 43, 50, 51], that the hair bundle motility is the reason for the nonlinear amplification but has a low gain of a factor of ten, while somatic motility is intrinsically linear but offers an additional amplification (driven by hair bundle motility) to reach the amplification factor of 1000 observed in mammals.

Besides the frequency decomposition, nonlinear amplification and short-term adaptation described above, further pre-conditioning of the signals occurs in the inner ear, driven by an efferent feedback from subsequent processing stages [52–55]. This feedback is proposed to control OHC motility by a calcium driven motility. It is expected to be used to reduce unwanted/unattended (non-relevant) contributions in the sound signal by damping the amplification of OHCs. The efferent feedback controls thereby a larger group of OHCs simultaneously (as seen from innervation), but can address specific frequency ranges individually. The efferent feedback was shown to improve speech perception in particular in noisy environments. It is discussed that this improved perception is achieved by the active damping (or the lack of active amplification) for the noisy frequency bands in comparison to the amplification of the frequency bands associated with speech (or important sounds).

Concluding, in the hearing process up to transduction of the signal into electrical signals, a number of pre-processing steps are involved, which are thought to improve sound/speech perception and increase efficiency of the system. The most important pre-processing steps thereby are the frequency decomposition of the signal, the nonlinear (compressive) amplification and the various adaptation mechanisms.

3 Bio-inspired Acoustic Sensing

Does it provide advantages if bio-inspired pre-processing, as discussed above, is introduced into technological speech processing systems? Indeed, several researchers could demonstrate that integrating bio-inspired sensing (including a pre-processing) can strongly improve the performance of sound/speech processing systems. The first example is given by Araujo et al. [8], who demonstrated recently that the nonlinear filtering is an important factor for successful speech recognition. They used the task of spoken digit recognition to analyse the effect of the bio-inspired pre-processing onto the word success rate of their sound processing system. Therefore, they applied a frequency filtering to sound samples recorded with a microphone, and added various nonlinear amplification/filtering methods up to a cochlea-like nonlinear filtering. Then, the pre-processed signals were fed into a neural network for solving the recognition task. Araujo et al. [8] could show that the neural network itself had a success rate of 70–80% if combined with linear frequency filtering, whereas the combination of nonlinear filtering and neural network reached up to 96% recognition rate. They could determine that the contribution of the nonlinear filtering to this recognition rate was nearly 80%. Even for noise conditions (sound samples from subway, car etc.), high recognition rate of up to 86% were achieved, if nonlinear filtering was applied.

Another example for the advantage of bio-inspired sensing over conventional systems was given recently by Wang et al. [56]. They developed a sensor based on resonant operation, in contrast to the below resonance operation of microphones, and tested its effect on machine learning-based biometric authentication (speaker identification). It was shown that an exceptional error rate reduction in speaker identification with their bio-inspired sensor in comparison to MEMS microphone based system is achieved using only a small amount of training data.

The third example is the work of Kiselev et al. [18], who could demonstrate in their work the importance of adaptation in acoustic systems. They integrated an automatic gain control in their dynamic acoustic sensor, which changes an attenuation before and/or an amplification after the band-pass filtering, to keep the spike rate of their system in a pre-defined range. The application of this adapting sensing (pre-processing) system for differentiating speech from noise showed that the system performed much better for low signal-to-noise ratios (SNR), i.e. up to 15% increase in accuracy than the system without adaptation. Speech or sound in low SNR conditions is an acoustic environment with which speech processing systems still struggle.

Since it seems to be advantageous to integrate the bio-inspired pre-processing, we will discuss in this chapter several bio-inspired acoustic sensing systems, which

are capable of frequency decomposition, nonlinear amplification and/or adaptation. Sound/speech processing systems, incorporating or adapting mechanisms from the biological hearing, can be divided into two groups: (i) sensors with pre-processing properties (bio-inspired sensors) and (ii) systems incorporating models of pre-processing in the processing circuits. Before we will describe the bio-inspired sensors with integrated pre-processing in more detail, we will give a short introduction to systems with pre-processing in the stages after transduction in the next section.

3.1 Systems with Bio-inspired Pre-processing After the Sensor

Systems that incorporate pre-processing in circuits after the microphone are the standard case for speech processing systems like smartspeakers or in hearing aids etc. These apply mostly digital signal processing to reduce noise in signals, separate different speech signals from each other or amplify certain frequency bands. Despite the advantages due to the development of these software-based implementations, these have several drawbacks in terms of computation power, power consumption, latency and data-streaming and are typically not located near the sensor node but rather cloud-based. Here, hardware-based implementations can improve the performance in terms of the drawbacks listed above enabling a local and real-time performance. Neuromorphic cochleas have been designed starting from the late 1980s [57] and ongoing development resulted in a 64 channel binaural audition sensor, which is capable of speaker identification, source localization and was shown to be computationally less demanding [58]. Examples of neuromorphic sound processing platforms are the so-called silicon cochlea [57, 59], the Hopkins electronic ear [60], AER-EAR (also named dynamic auditory sensor) [9, 61], or the FPGA cochlea [62, 63]. These consist mainly of (i) a number of cascade or parallel filter banks, (ii) followed by nonlinear amplification stage and finally (iii) a spike generation stage. Furthermore, active coupling between filter stages [64] was introduced to improve roll-off of frequency response as well as the synchronization between two silicon cochleas for source localization [65]. Different adaptation mechanisms are successively included. Automatic gain control of the OHC model stage is included as a model describing feedback from medial olivocochlear nuclei [66].

3.2 Bio-inspired (Acoustic) Sensors

This section focuses on bio-inspired sensors, which integrate bio-inspired pre-processing in the sensor/sensor properties itself rather than adding it as an additional stage after the sensing stage. Thereby, we will review the different systems according to the integrated pre-processing, namely frequency decomposition, non-linear/compressive amplification and adaptation.

Frequency Decomposition

To incorporate frequency decomposition into the sensing system, resonant operation is applied. In contrast to this, microphones typically operate below resonance to guarantee a linear transfer characteristic over a large frequency range. Resonant operation offers two advantages, namely a higher sensitivity than non-resonant operation and a band-pass filtering functionality. Thereby, the increase in amplitude at resonance compared to out-of-resonance mode and the bandwidth of the filter are closely related and determined by the damping in the system. Quantitatively, the damping can be described the quality factor of the system. A higher quality factor indicates a lower damping, which results in larger amplitudes and a smaller bandwidth compared to lower quality factors. The resonance frequency in bio-inspired acoustic sensing systems is mainly determined by the geometrical dimensions of the sensor. Two geometric approaches are mainly applied: either beam structures or membranes, and examples for both are shown in Fig. 6.

Beam structures, sometimes termed artificial hair cells (AHC), can be either mounted single-sided or double-sided. To tune the frequency, typically the length or width of the beams is varied. The coverage of a larger frequency range is obtained

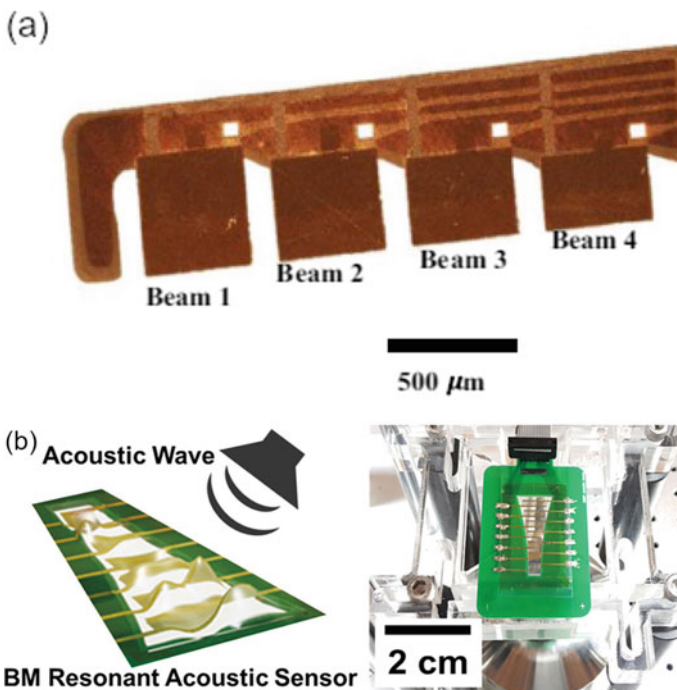


Fig. 6 Examples of **a** beam-based and **b** membrane based bio-inspired acoustic sensors (from [56, 68, 71])

using multiple beams with different resonance frequencies. However, a system of beams covering the complete auditory range was not yet realized, possibly due the small bandwidth and the thus large number of beams necessary. The read-out of the sensing signal is realized in various ways: piezo-electric, capacitive, piezo-resistive or optical.

Membrane structures, also termed artificial basilar membranes (ABM), typically have a triangular shape to model the frequency decomposition of the basilar membrane. In the case of optical readout, the analog coding of the frequency can be kept. In most cases, however, additional electrodes are added for easier readout/simpler calibration, which convert the membrane deflection into electrical signals. These are often based on the piezoelectric effect [56, 67, 68], but can be piezo-resistive or capacitive as well.

Operation was demonstrated in air or fluid for beam as well as membrane-based devices. Some of these devices were already tested for implantation to activate the inner ear [67, 68] in guinea pigs. Here, a brainstem response could be observed, demonstrating the activation of the auditory nerve by the implant.

Besides membrane and beam-based structures, also graded material properties using, e.g., acoustic meta-materials are applied for frequency-space coding [69, 70].

Nonlinear Dynamics

Diverse methods exist to tune sensors, not only acoustic ones, into a nonlinear regime to improve sensing properties like dynamic range, bandwidth, and (nonlinear) amplification and to exploit ensemble effects. Among these are feedback or feedforward loops [72–76], hydrodynamical coupling [77], electrostatic interaction [78], elastic properties, multi-mode coupling, and coupling of resonators [79].

Thereby, some of the feedback-based systems try to explicitly model the biological processes like hair bundle adaptation or somatic motility [73–76]. Unfortunately, their application to sound was either not yet achieved due to the size of the sensor [73, 74] or exhibited only a constant output level. The feedforward system by Crowley et al. requires a priori-knowledge on the input stimuli [72].

Besides the improvements for the sensing (like dynamic range, resolution etc), nonlinear, mechanical resonators offer the possibility to bring computation to the sensor domain, as was recently shown [80]. Thereby, single or networks of nonlinear, mechanical oscillators can be used to implement reservoir computing [81, 82]. The nonlinearity of the oscillators was applied furthermore to implement bit storage and bit flip operations [83, 84] and coupling of several mechanical oscillators can be used to realize logic circuits like binary comparators and XOR/NAND logic gates [85].

Adaptation

First attempts to realize the idea of adapting the sensor based on the computation of the sensed signal were undertaken only recently [86, 87]. Tsuji et al. [86] developed

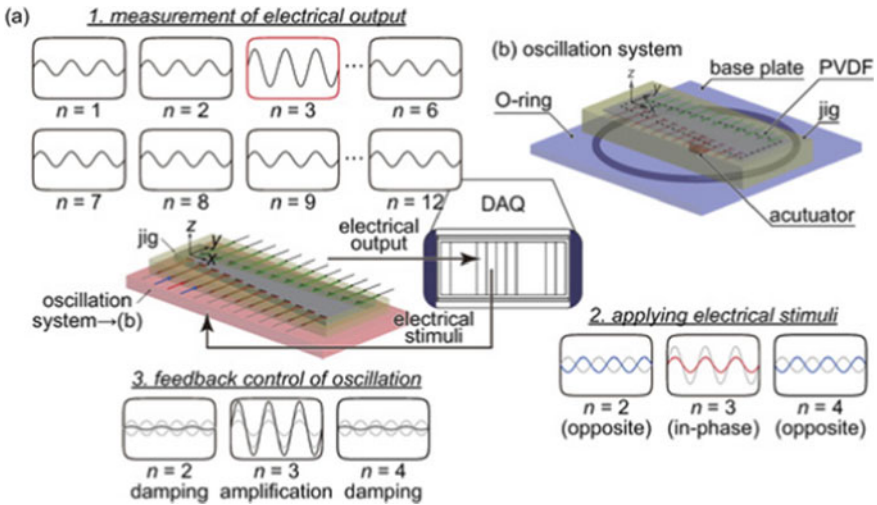


Fig. 7 Improved adaptation principle by Yamazaki et al. [88]

an artificial cochlea membrane, which resonates at different positions depending on the frequency. They used a two-step process for adaptation. In the first step the resonant position was detected and in the second step (control step) the oscillation at the neighboring positions was damped by applying a feedback signal. Further development [88] improved the speed of adaptation (Fig. 7).

Guerreiro et al. [87] used a single, mechanical resonator and fed its sensing signal to a leaky-integrate and fire neuron model, implemented in a microcontroller unit. The resulting pulses are used to apply an AC signal changing the Q-factor of the system to model short-term adaptation and a DC signal using a charge-pump circuit for changing the spring constant of the system to model long-term adaptation.

4 Recently Developed Adaptive, Acoustic Cantilever Sensor

The acoustic sensor, which we developed [20–22, 89], is based on a beam structure with appropriate feedback (see Fig. 8). Thereby, the beam is a multi-layer structure, build from silicon, silicon dioxide and aluminium with a typical length of 350 μm, a width of 150 μm and a thickness in the range of 1 – 5 μm. For sensing of the deflection, piezoresistive elements are integrated near the base of the beam. A thermo-mechanical actuation principle is integrated by the aluminium loop on top, which is heated upon voltage application. Due to the different thermal expansion coefficients of the different materials, the beam bends upon heating, i.e., due to the voltage signal applied at the aluminium loop. The beams, sometimes termed active cantilevers, were originally developed for application in atomic force microscopy [90, 91], but later

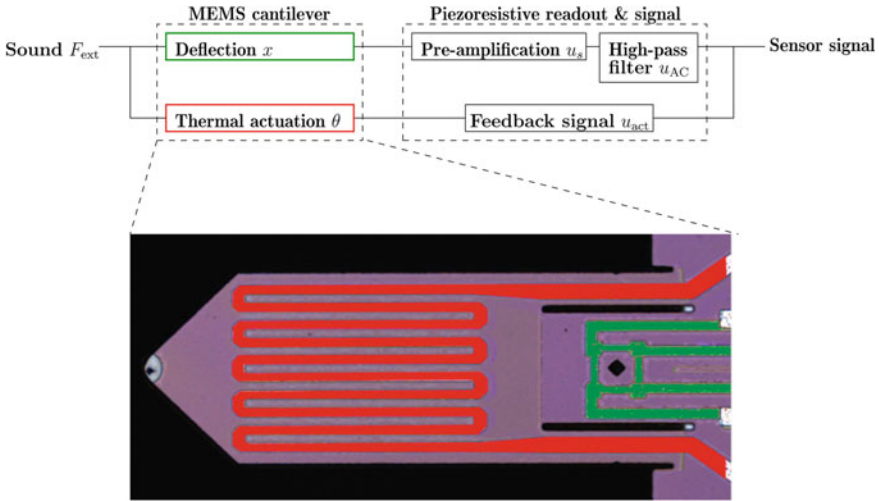


Fig. 8 Top: Schematic of the sensing system including the cantilever, feedback, and thermal actuator. All variables as in Eqs. (4)–(6). Bottom: Microscope image of the MEMS cantilever, consisting of silicon body, Al actuation loop (red) and piezo-resistive deflection sensing (green)

on successfully applied in scanning probe lithography, gas flow sensing and particle detection, IR sensing and other fields. Details on beam fabrication and applications can be found in several reviews. See, for instance, Refs. [92, 93].

The feedback loop, used to tune the dynamics of the beam, consists of (i) reading the sensor signal, (ii) calculating the feedback function and (iii) using the feedback signal to drive the actuator (see Fig. 8). The calculation of the feedback signal is done in an FPGA architecture on a STEMLab 125-14 board, which allows a near real-time feedback. Before the feedback is calculated, the sensor signal is amplified by a factor of 1000, high-pass filtered to use only the AC signal, and finally digitized by an analog-to-digital converter on the STEMLab-board (sample rate 125 MHz and 14 bit resolution). While the input range of the STEMLab board can be switched between 1 V and 20 V, the output range is limited to 1 V. Values of the feedback signal, outside of this range are mapped to the maximal value. After the calculation, the feedback signal is converted into an analog voltage signal by the digital-to-analog converter of the STEMLab board (sample rate 125 MHz), fed to a buffer board and finally applied to the actuator of the beam.

4.1 Frequency Decomposition

The intrinsic frequency $\omega_0 = 2\pi f$ of the cantilever sensor is determined by the length l_{Si} and the thickness d_{Si} of the sensor together with the elasticity module E_{Si} and density ρ_{Si} for Si as follows:

$$f = \frac{\omega_0}{2\pi} = \delta_n^2 \frac{d_{Si}}{2\pi l_{Si}^2} \sqrt{\frac{E_{Si}}{12\rho_{Si}}}, \quad (1)$$

where δ_n denotes the pre-factor for the n -th mode. Here, the first transverse mode is considered only. The material parameters of the cantilever are given in Table 2. The considered sensor dimensions yield a frequency of the beam of $f \approx 14$ kHz.

From the device parameter, we can calculate a quality factor Q_0 of the cantilever by means of the formula for an oscillating beam in air that was derived by Zoellner et al. [94]. It is mainly determined by the damping due to the surrounding fluid and given by

$$Q_0 = \frac{\frac{4\rho_{Si}d_{Si}}{\pi w_{Si}\rho_{gas}} + 1.05333 + \frac{3.7997}{\sqrt{2Re}}}{\frac{3.8019}{\sqrt{2Re}} + \frac{2.7364}{2Re}} \quad (2)$$

with the Reynolds number Re for this system given by

$$Re = \frac{2\pi f \rho_{gas} w_{Si}^2}{4\eta_{gas}}. \quad (3)$$

ρ_{gas} and η_{gas} denote the density and the dynamic viscosity of the surrounding media, i.e., air, respectively (cf. Table 2). For the investigated sensor, we use a quality factor of $Q_0 \approx 43.2$ in simulations ($Q \approx 40 - 50$ in experiments).

However, upon adding a feedback we see a decrease of bandwidth of the sensor (see Fig. 9). This effect can be described by a introducing an effective quality factor, which depends on the feedback strength. As was shown recently [22], the effective quality factor not only describes the change in bandwidth of the system but also the change in sensitivity (for the linear regime).

Table 2 Parameters of the sensor system and surrounding fluid (air)

Parameter	Value
Width of beam w_{Si}	150 μm
Length of beam l_{Si}	350 μm
Thickness of beam d_{Si}	1.25 μm
Density Silicon ρ_{Si}	2329 kg/m^3
Young's modulus Silicon E_{Si}	170 GPa
Mode factor δ_1 for frequency calculation	1.8751
Density air ρ_{gas}	1.189 kg/m^3
Dynamic viscosity air η_{gas}	18.232 $\cdot 10^{-6}$ Pa s

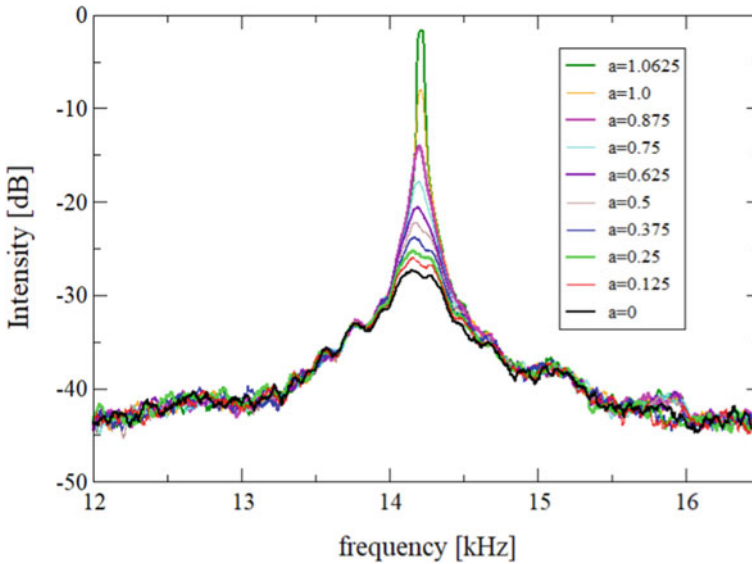


Fig. 9 Frequency response of bio-inspired sensor for different feedback strengths and constant sound pressure level

4.2 Nonlinear Dynamics

Human hearing system demonstrates nonlinear transfer characteristics, as was described in Sect. 2. In Fig. 10, the response of our developed sensor is shown in dependence of the sound pressure level for different values of the feedback strength a obtained from measurements. If the feedback strength is less than the critical feedback strength, linear transfer characteristics are observed. Thereby, the sensitivity strongly depends on the feedback strength (effective Q-factor). Near the critical feedback strength the transfer characteristics become nonlinear, whereas for larger feedback strengths autonomous oscillations are observed even without sound input. Thus, by tuning the feedback strength, nonlinear characteristics, comparable to the biological ones, are obtained.

To understand this transition from linear to nonlinear dynamics, we perform nonlinear dynamic analysis and simulations of the system, which are presented in the following. The modelling of the cantilever sensor is an extended version of the modal description proposed by Roeser et al. [96]:

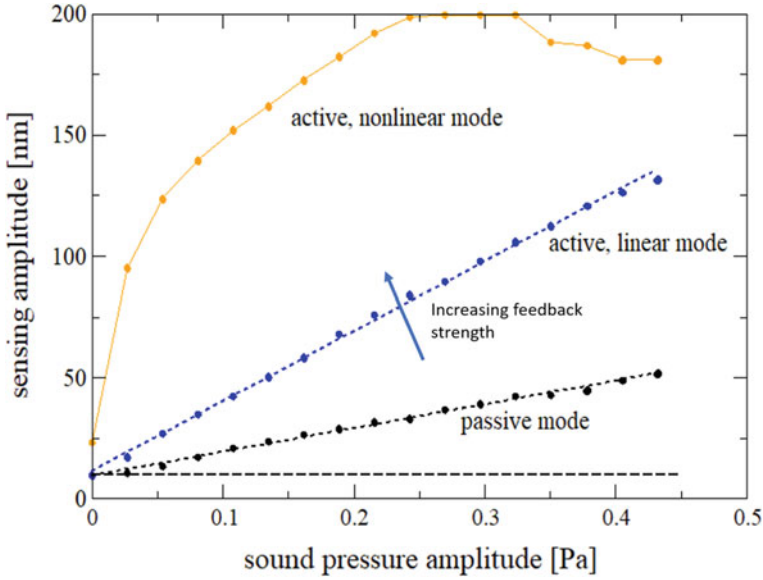


Fig. 10 Experimentally determined amplitude of oscillations in dependence of a for different sound pressure levels. From [95]

$$\ddot{x}(t) + \frac{\omega_0}{Q_0} \dot{x}(t) + \omega_0^2 x(t) = \alpha \theta(t) + \frac{F_{\text{ext}}(t)}{m} \tag{4a}$$

$$\dot{\theta}(t) + \beta \theta(t) = \gamma \left(\frac{\tanh u_{\text{act}}(t)}{R} \right)^2 \tag{4b}$$

$$\dot{u}_{\text{AC}}(t) = -\frac{u_{\text{AC}}(t)}{\tau} + \dot{u}_s(t). \tag{4c}$$

In Eq. (4), the dynamic variables x , θ , and u_{AC} denote the deflection of the cantilever (at the free end), the temperature difference between the beam structure and the surrounding, and the high-pass filtered version of the sensing voltage u_s , respectively. The latter is linearly related to the deflection with the scaling factor k :

$$u_s(t) = kx(t). \tag{5}$$

The actuation voltage $u_{\text{act}}(t)$ is given by

$$u_{\text{act}}(t) = au_{\text{AC}}(t) + u_{\text{DC}} \tag{6}$$

with a feedback strength a and a bias voltage u_{DC} . The model parameters in Eqs. (4) include the transfer factor α from temperature into deflection, the time constant

Table 3 Modelling parameters

Variable	Value
Frequency of beam $f = \omega_0/(2\pi)$	14 kHz
Quality factor Q_0	43.2
Transfer factor α (temperature to deflection)	$749.3702 \text{ m}/(\text{Ks})^2$
Time constant β (temperature change)	$1.0066 \cdot 10^3 \text{ 1/s}$
Transfer factor (voltage to temperature) γ	$4.2588 \cdot 10^7 \text{ K}\Omega^2/\text{sV}^2$
Heater resistance R	25Ω
Time constant of high-pass filter τ	10^{-3} s
Calibration factor k (piezoresistive sensing and pre-amplification)	10^6 V/m
Bias voltage u_{DC}	-200 mV

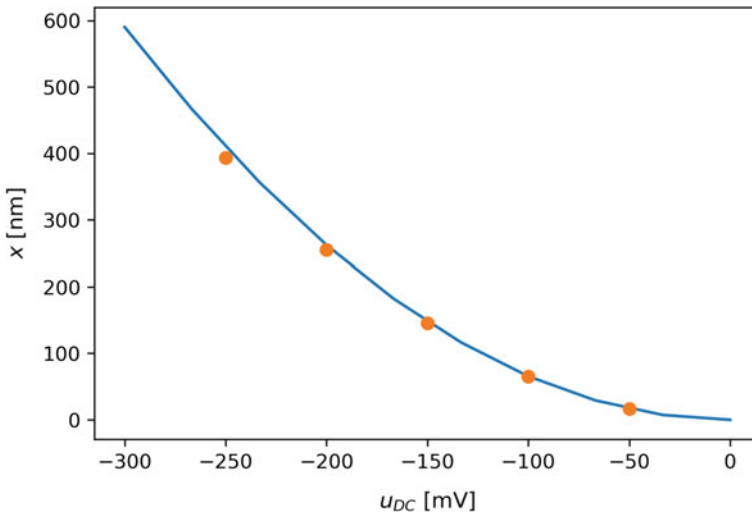


Fig. 11 Fixed point x^* in dependence on bias voltage u_{DC} as given in Eq. (7) (solid curve) and by direct simulation of Eqs. (4) (dots) for $a = 0.6$. Parameters as in Table 3

β for temperature changes, the transfer efficiency γ from actuation voltage into temperature changes, and the resistance R of the actuator. They are summarized in Table 3.

Setting all time derivatives in Eqs.(4) to zero, we find the fixed point of the cantilever sensor:

$$x^* = \frac{\alpha\gamma}{\omega_0^2\beta} \left(\frac{u_{\text{DC}}}{R} \right)^2. \tag{7}$$

A fixed point is a property of dynamical systems, which corresponds to the stationary state of the system. The dependence of the fixed point x^* (i.e. the steady state or DC

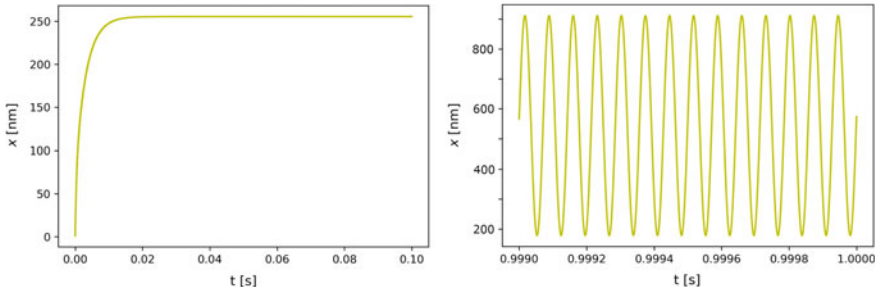


Fig. 12 Left (fixed point): Simulated time series of the deflection for $a = 0.6$. Right (oscillations): Simulated time series of the deflection for $a = 0.9$. Initial conditions $x(0) = 1$ nm. Other parameters as in Table 3

value of the deflection) on the bias voltage u_{DC} is depicted in Fig. 11. Depending on the feedback strength a , the dynamical behaviour of the system changes. Exemplary timeseries are shown in Fig. 12 for feedback strengths $a = 0.7$ (top) and $a = 0.9$ (bottom), i.e., below and above the bifurcation point, respectively. In the top panel, the timeseries quickly approaches the value of the fixed point starting from the initial condition $x(0) = 1$ nm (cf. Fig. 11). Above the bifurcation, the fixed point is unstable, and thus, the dynamics do not reach the steady state. Instead, an oscillation emerges with an amplitude of $0.35 \mu\text{m}$.

For further insight into the full dynamics for the case of oscillations (after bifurcation), Fig. 13 depicts all four variables of Eqs. (4) and (6): x , $y = \dot{x}$, θ , and u_{AC} .

The transition from the fixed-point behavior to oscillatory dynamics occurs via a Hopf bifurcation. Beyond the critical point, a square-root dependence of the oscillation amplitude on the bifurcation parameter (here: a) provides a good indication of this bifurcation. Indeed, this is shown in Fig. 14, which depicts the amplitude of oscillations in dependence in the feedback strength a . The amplitude is calculated as $\max[x(t)] - \min[x(t)]$ with $t \in [0.95 \text{ s}, 1 \text{ s}]$ starting from an initial condition $x(0) = 1$ nm. The bifurcation occurs for a critical feedback strength $a_{\text{crit}} \approx 0.8$.

4.3 Adaptation

Since the sensor dynamics can be easily tuned by the real-time feedback, adaptation can be implemented by changing dynamically the feedback parameters: feedback strength a and bias voltage u_{DC} . To implement an automatic adaptation of the sensor, the amplitude of the sensing signal is used for control of the adaptation. Conceptually, one of the feedback parameters (here: feedback strength a) switches to another value, if the amplitude is larger than a pre-defined threshold. Thereby, two adaptation variants were implemented, which differ in the reset of the feedback parameter to its initial value:

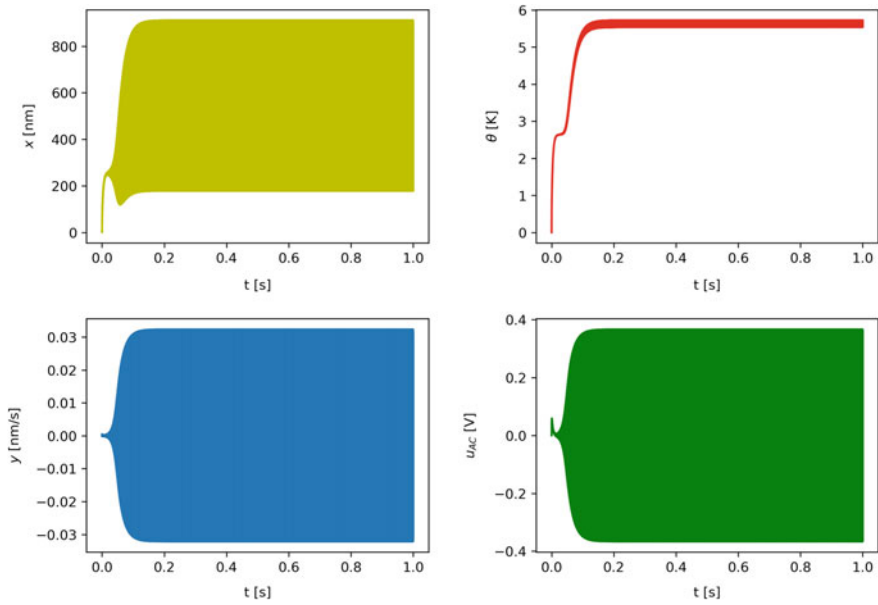


Fig. 13 Numerical results (Oscillations): Time series of the deflection for $a = 0.9$. The different panels show all four variables of Eqs. (4) and (6). Initial conditions $x(0) = 1$ nm. Other parameters as in Table 3

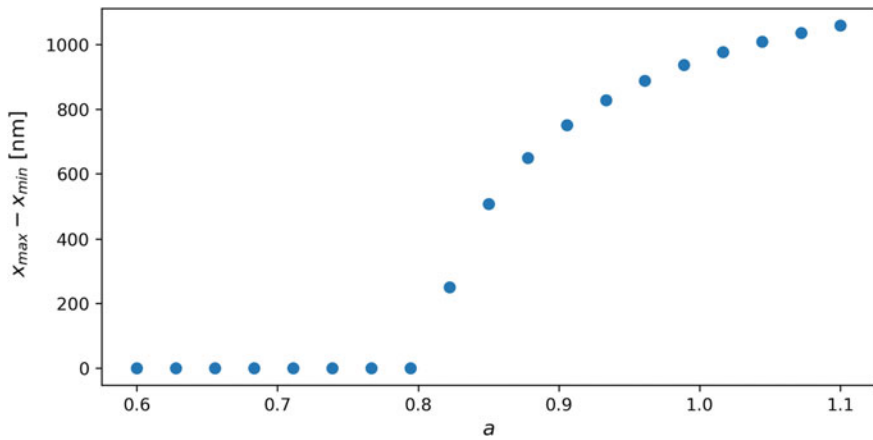


Fig. 14 Numerical results: Amplitude of oscillations in dependence on a for an initial condition $x(0) = 1$ nm. Other parameters as in Table 3

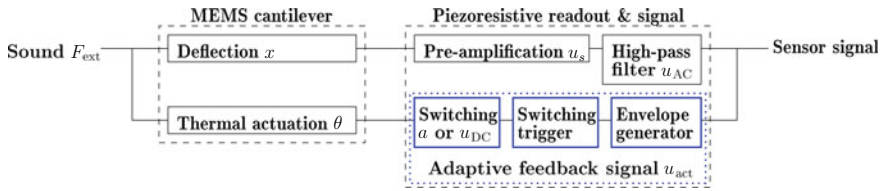


Fig. 15 Schematic of the sensing system including the cantilever, adaptive feedback (blue dashed box), and thermal actuator. The adaptation includes an envelope generator, a switching-trigger, and the switching of the feedback strength a or bias voltage u_{DC}

- *Variant 1:* The reset is done after a pre-defined time interval, independent of the signal amplitude. In this case, the response of the system resembles an event-based or spike-like response of the system to sound input. Thereby, with a fixed threshold, a spike-rate based encoding of the sound amplitude is possible.
- *Variant 2:* The feedback parameter is reset to its initial value, if the amplitude of the sensing signal becomes smaller than a second threshold. This variant enables an automatic gain control and dynamic range enlargement. For small sound pressure levels, the nonlinear response of the system has a larger gain than the linear one, whereas for larger sound pressure levels the gain for the linear response will be higher. Thus, operating the system in general in the nonlinear mode and switching to linear mode for larger sound pressure levels can ensure a sufficiently high gain for the complete input range. Note that the nonlinear operation increases the dynamic range for small SPL by approximately 6dB compared to the linear mode. Furthermore, the response of the system resembles the short-term adaptation at the hair cell-auditory nerve synapse for constant sound input (cf. Sect. 2).

Both variants enable a highlighting of the sound onset, which might improve the performance in processing tasks like localisation. These are typically based on the time difference of sound detection between two microphones. Thus, highlighting the sound onset might increase the efficiency of the processing for localisation.

A schematic of the implementation for both variants is shown in Fig. 15. It consists of (i) an envelope generator, which determines the envelope from the sensing signal, (ii) a switching-trigger stage that is responsible for comparison of the envelope with the pre-defined threshold(s), and (iii) a switching stage to set the feedback parameter to the respective value. Thereby, variant 1 was implemented in the FPGA, which is used for feedback calculation. Variant 2 was implemented using analog circuits, whereby the circuits incorporated additionally the feedback calculation [97]. For this implementation, the STEMLab board, including its ADC/DAC and FPGA as described in Sect. 4, was not used.

Figure 16 shows the measured time series of the sensing signal in case of variant 1 (left graph) and the envelope of the sensing signal for variant 2 (right graph). For variant 1, the feedback parameter a was varied between $a_0 = 0.75$ (nonlinear mode) and $a_1 = 0$ (linear (passive) mode) if the envelope of the sensing signal reached a

value of 200mV. In the linear mode, the sound input is not large enough to yield a sensing signal above noise level. Thus, a spike-like output is observed for this variant of adaptation.

The time interval between two spikes is composed of two components: first, the response time τ_1 of the sensor until switching, and second, the off-time τ_2 , i.e., the pre-defined time interval, for which the feedback parameter is kept at its lower value. Since the response time or transient behaviour of the sensor depends strongly on the input amplitude and the value of the feedback parameter a , a sound amplitude-dependent spike rate is observed, yielding an encoding of the sound amplitude by the spike rate, and the spike shape can be tuned by changing a . Such a spike-based output can improve the efficiency of the sensing system, because it reduces energy consumption, since feedback is turned off between spikes, and it reduces the amount of data, which needs to be sent to the processing unit, if only the spike times are transferred.

In the second variant [97], the feedback strength a is switched from a high value ($a = 0.6$, nonlinear mode) to a lower value ($a = 0.2$, linear mode), if the threshold is reached. The reset to the initial values occurs, if the envelope of the sensing signal decreases below a second threshold. If a constant sound input is applied, the sensor will react first with high sensitivity (indicated by the peak in the time series), which decreases after switching to a lower level (represented by the plateau in Fig. 16 right graph). Comparing the change of the amplitude at the peak and at the plateau for different sound pressures (i.e. different driving voltages for the loudspeaker), the change from nonlinear characteristics (at peak region) to linear characteristics (at plateau region) is clearly visible. This variant has three main purposes: first, it can

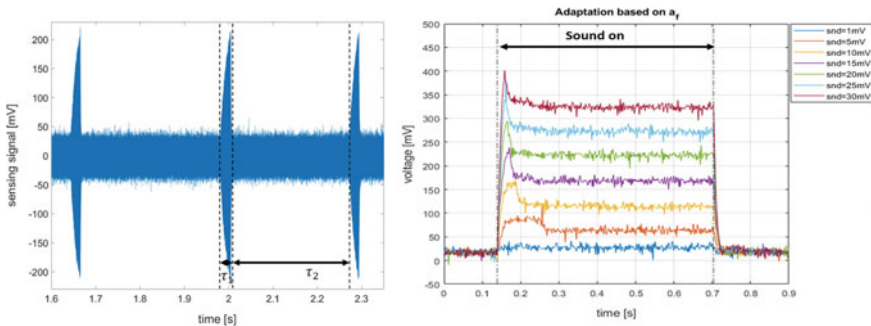


Fig. 16 (Left) Measured sensing signals for constant sound input (100mV driving voltage loudspeaker) obtained from experiments with FPGA-based implementation of adaptation variant 1, switching feedback strength a between $a_1 = 0.75$ and $a_0 = 0$ ($u_{DC} = -200mV$). Here, the feedback strength is kept at its lower value for a constant time interval τ_2 before resetting it to the high-sensitivity regime. The spike rate of this spike-like response of the sensor system depends on the sound-amplitude dependent part τ_1 and the fixed time interval τ_2 for reset. (Right) Envelope of sensing signal for adaptation variant 2 and constant sound input between 0.14 s and 0.7 s. Thereby, a was switched from 0.6 to 0.2 while u_{DC} was kept constant at -200 mV. Different sound pressures were applied by varying the amplitude snd for the loudspeaker (see legend). From [97]

increase the dynamic range, for which the gain (or sensitivity) is above a certain threshold. This is particularly important for large sound pressure amplitudes, since in this regime the nonlinear characteristics yield a small gain. Second, the overall shape highlights the onset of the sound, which might be important for processing tasks like localisation or recognition of speech. Third, the power consumption of the system might be decreased, since the feedback amplitude is reduced for the linear mode. Additionally, the shape of the response resembles the measured response at the synapse between hair cell and auditory nerve. However, it is until now, not clear if this adaptation is caused by biological restrictions (e.g. depletion of vesicles) or if it has a functional purpose for the subsequent processing.

5 Conclusions

The goal of this chapter has been two-fold: (1) to provide an overview of acoustic sensing from a biological point of view and (2) to present experimental and numerical results on an adaptable acoustic micro-electromechanical systems (MEMS)-based sensor. The mathematical model reflects the key experimental features. Besides the accessible sensor signal, the model also includes a variable for the thermal actuation of the electro-mechanical cantilever. Direct simulations indicate that the device undergoes a Hopf bifurcation. Dynamically, this corresponds to the transition from a fixed point to self-sustained oscillations. The response to different sound pressure levels was measured experimentally, where we identified passive, active linear, and active nonlinear modes. In the final parts of this chapter, we highlighted potential applications of the considered MEMS. We elaborated on experimental implementation of spiking functionality via feedback strength switching that was realized by an field programmable gate array (FPGA). We also demonstrated the possibility of sensory adaptation, which is based on dynamic switching of the feedback strengths. This will pave the way, for instance, for frequency-selective amplification of sound signals.

References

1. Zawawi, S.A., Hamzah, A.A., Majlis, B.Y., Mohd-Yasin, F.: *Micromachines*. **11**(5) (2020). <https://doi.org/10.3390/mi11050484>
2. Papastratis, I. <https://theaisummer.com/> (2021)
3. Adavanne, S., Politis, A., Nikunen, J., Virtanen, T.: *IEEE J. Sel. Top. Signal Process.* **13**, 34 (2018). <https://doi.org/10.1109/JSTSP.2018.2885636>
4. Alvarez, R., Park, H.J.: In: *IEEE International Conference on Acoustics, Speech and Signal Proc. (ICASSP)*, p. 6336 (2019). <https://doi.org/10.1109/ICASSP.2019.8683557>
5. Abeßer, J.A.: *Appl. Sci.* **10** (2020). <https://doi.org/10.3390/app10062020>
6. Bianco, M.J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M.A., Gannot, S., Deledalle, C.A.: *J. Acoust. Soc. Am.* **146**, 3590 (2019). <https://doi.org/10.1121/1.5133944>

7. Wu, J., Yilmaz, E., Zhang, M., Li, H., Tan, K.: *Front. Neurosci.* **14**, 199 (2020). <https://doi.org/10.3389/fnins.2020.00199>
8. Abreu Araujo, F., Riou, M., Torrejon, J., Tsunegi, S., Querlioz, D., Yakushiji, K., Fukushima, A., Kubota, H., Yuasa, S., Stiles, M.D., Grollier, J.: *Sci. Rep.* **10**(1), 1 (2020). <https://doi.org/10.1038/s41598-019-56991-x>
9. Liu, S.C., Delbruck, T., Indiveri, G., Whatley, A., Douglas, R.: *Event-Based Neuromorphic Systems*. John Wiley & Sons (2015). <https://doi.org/10.1002/9781118927601>
10. Schafer, P.B., Jin, D.Z.: *Neural Comput.* **26**(3), 523 (2014). https://doi.org/10.1162/NECO_a_00557
11. Barker, J., Vincent, E., Ma, N., Christensen, H., Green, P.: Special issue on speech separation and recognition in multisource environments. *Comput. Speech & Lang.* **27**(3), 621 (2013). <https://doi.org/10.1016/j.csl.2012.10.004>
12. Zai, A.T., Bhargava, S., Mesgarani, N., Liu, S.C.: *Front. Neurosci.* **9**, 347 (2015). <https://doi.org/10.3389/fnins.2015.00347>
13. Basinger, D.: Automatic gain control for implanted microphone. US Patent 8641595
14. Mortensen, R., Pedersen, B.D.: *J. Acoust. Soc. Am.* **120**(4) (2006). <https://doi.org/10.1121/1.2372363>
15. Kompis, M., Dillier, N.: *J. Acoust. Soc. Am.* **109**(3), 1123 (2001). <https://doi.org/10.1121/1.1338557>
16. Haila, O.: *J Acoust Soc Am.* **133**(3), 1844 (2013). <https://doi.org/10.1121/1.4795045>
17. Wen, B., Boahen, K.: In: 2006 IEEE International Solid State Circuits Conference-Digest of Technical Papers, pp. 2268–2277. IEEE (2006). <https://doi.org/10.1109/ISSCC.2006.1696289>
18. Kiselev, I., Liu, S.C.: In: 2021 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–5 (2021). <https://doi.org/10.1109/ISCAS51556.2021.9401742>
19. Josefsson, O.M.: MemS microphone with programmable sensitivity. US Patent 8,831,246 (2014)
20. Lenk, C., Seeber, L., Ziegler, M., Hövel, P., Gutschmidt, S.: In: 2020 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–4. IEEE (2020). <https://doi.org/10.1038/s41928-023-00957-5>
21. Lenk, C., Gutschmidt, S., Rangelow, I.W.: Vorrichtung und verfahren zur detektion von schall in gasen und flüssigkeiten. DE Patent 102018117481B8 (2019)
22. Lenk, C., Hövel, P., Ved, K., Durstewitz, S., Gutschmidt, S., Ivanov, T., Fritsch, T., Beer, D., Meurer, T., Ziegler, M.: *Nat. Electron.* **6**, 380 (2023). <https://doi.org/10.1038/s41928-023-00957-5>
23. Kandel, E., Schwartz, J., Jessell, T., Siegelbaum, S., Hudspeth, A.: *Principles of Neural Science*. McGraw-Hill (2012). ISBN 0-07-139011-1
24. Hudspeth, A.: *Nat Rev Neurosci.* **15**, 600 (2014). <https://doi.org/10.1038/nrn3786>
25. By Chittka I. Brockmann, cc by 2.5. https://upload.wikimedia.org/wikipedia/commons/5/58/10.1371_journal.pbio.0030137.g001-L-A.jpg
26. <http://www.cochlea.eu/en/ear/external-ear>
27. By original: Oarih, vector: Fred the oyster—own work based on: Cochlea-crosssection.png, cc by-sa 3.0. <https://commons.wikimedia.org/w/index.php?curid=9851471>
28. Spiegel, M.F., Watson, C.S.: *J. Acoust. Soc. Am.* **76**(6), 1690 (1984). <https://doi.org/10.1121/1.391605>
29. Sellon, J., Ghaffari, R., Freeman, D.: *Cold Spring Harb Perspect Med.* **9**(10), a033514 (2019). <https://doi.org/10.1101/cshperspect.a033514>
30. Boron, W.F., Boulpaep, E.: *Medical Physiology: a Cellular and Molecular Approach*, updated 2nd ed (2005). ISBN 1416023283
31. Dallos, P., Popper, A., Fay, R.: *The Cochlea*. Springer (1996)
32. Zilany, M.S.A., Bruce, I.C., Nelson, P.C., Carney, L.H.: *J. Acoust. Soc. Am.* **126**(5), 2390 (2009). <https://doi.org/10.1121/1.3238250>
33. Rhode, W.S., Smith, P.H.: *Hear. Res.* **18**(2), 159 (1985). [https://doi.org/10.1016/0378-5955\(85\)90008-5](https://doi.org/10.1016/0378-5955(85)90008-5)

34. Relkin, E.M., Turner, C.W.: *J. Acoust. Soc. Am.* **84**(2), 584 (1988). <https://doi.org/10.1121/1.396836>
35. Goutman, J.D., Glowatzki, E.: *Proc. Natl. Acad. Sci.* **104**(41), 16341 (2007). <https://doi.org/10.1073/pnas.0705756104>
36. Raman, I.M., Zhang, S., Trussell, L.O.: *J. Neurosci.* **14**(8), 4998 (1994). <https://doi.org/10.1523/JNEUROSCI.14-08-04998.1994>
37. De Vries, H.: *Phys.* **14**(1), 48 (1948). [https://doi.org/10.1016/0031-8914\(48\)90060-3](https://doi.org/10.1016/0031-8914(48)90060-3)
38. Sivian, L.J., White, S.D.: *J. Acoust. Soc. Am.* **4**(4), 288 (1933). <https://doi.org/10.1121/1.1915608>
39. Ashmore, J., Avan, P., Brownell, W.E., Dallos, P., Dierkes, K., Fettiplace, R., Grosh, K., Hackney, C.M., Hudspeth, A.J., Jülicher, F., Lindner, B., Martin, P., Meaud, J., Petit, C., Santos-Sacchi, J., Sacchi, J.R., Canlon, B.: *Hear. Res.* **226**(1), 1 (2010). <https://doi.org/10.1016/j.heares.2010.05.001>
40. Brownell, W.E., Bader, C.R., Bertrand, D., de Ribaupierre, Y.: *Sci.* **227**, 194 (1985). <https://doi.org/10.1126/science.3966153>
41. Fettiplace, R.: *Hair Cell Transduction, Tuning, and Synaptic Transmission in the Mammalian Cochlea*, pp. 1197–1227. John Wiley & Sons, Ltd (2017). <https://doi.org/10.1002/cphy.c160049>
42. Santos-Sacchi, J.: *J. Neurosci.* **11**, 3096 (1991). <https://doi.org/10.1523/JNEUROSCI.11-10-03096.1991>
43. Peng, A.W., Ricci, A.J.: *Hear. Res.* **273**(1), 109 (2011). <https://doi.org/10.1016/j.heares.2010.03.094>
44. Martin, P., Hudspeth, A.J.: *Proc. Natl. Acad. Sci.* **96**(25), 14306 (1999). <https://doi.org/10.1073/pnas.96.25.14306>
45. Crawford, A.C., Fettiplace, R.: *J. Physiol.* **364**(1), 359 (1985). <https://doi.org/10.1113/jphysiol.1985.sp015750>
46. Howard, J., Hudspeth, A.J.: *Proc. Natl. Acad. Sci.* **84**(9), 3064 (1987). <https://doi.org/10.1073/pnas.84.9.3064>
47. Denk, W., Webb, W.W.: *Hear. Res.* **60**(1), 89 (1992). [https://doi.org/10.1016/0378-5955\(92\)90062-R](https://doi.org/10.1016/0378-5955(92)90062-R)
48. Bensen, M.E., Marquis, R.E., Hudspeth, A.J.: *J. Neurosci.* **16**(18), 5629 (1996). <https://doi.org/10.1523/JNEUROSCI.16-18-05629.1996>
49. Martin, P., Mehta, A.D., Hudspeth, A.J.: *Proc. Natl. Acad. Sci.* **97**(22), 12026 (2000). <https://doi.org/10.1073/pnas.210389497>
50. Hudspeth, A.J.: *Neuron*. **59**, 530 (2008). <https://doi.org/10.1016/j.neuron.2008.07.012>
51. Maoiléidigh, D.Ó., Hudspeth, A.J.: *Proc. Natl. Acad. Sci.* **110**(14), 5474 (2013). <https://doi.org/10.1073/pnas.1302911110>
52. Guinan, J.J.J.: *Curr. Opin. Otolaryngol. & Head Neck Surg.* **18**(5), 447 (2010). <https://doi.org/10.1097/MOO.0b013e32833e05d6>
53. Rabbitt, R., Brownell, W.: *Curr. Opin. Otolaryngol. & Head Neck Surg.* **19**(5), 376 (2011). <https://doi.org/10.1097/MOO.0b013e32834a5be1>
54. Smith, D.W., Keil, A.: *Front. Syst. Neurosci.* **9**, 12 (2015). <https://doi.org/10.3389/fnsys.2015.00012>
55. Castellano-Munoz, M., Israel, S.H., Hudspeth, A.J.: *PLOS ONE*. **5**(10), 1 (2010). <https://doi.org/10.1371/journal.pone.0013777>
56. Wang, H.S., Hong, S.K., Han, J.H., Jung, Y.H., Jeong, H.K., Im, T.H., Jeong, C.K., Lee, B.Y., Kim, G., Yoo, C.D., Lee, K.J.: *Sci. Adv.* **7**(7), eabe5683 (2021). <https://doi.org/10.1126/sciadv.abe5683>
57. Lazzaro, J., Mead, C.: In: *Analog VLSI Implementation of Neural Systems*, p. 85 (1989). https://doi.org/10.1007/978-1-4613-1639-8_4
58. Li, C.H., Delbruck, T., Liu, S.C.: In: *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1159–1162 (2012). <https://doi.org/10.1109/ISCAS.2012.6271438>
59. Hamilton, T.J., Jin, C., van Schaik, A., Tapson, J.: *IEEE Trans. Biomed. Circuits Syst.* **2**(1), 30 (2008). <https://doi.org/10.1109/TBCAS.2008.921602>

60. Liu, W., Andreou, A., Goldstein, M.: *IEEE Trans. Neural Netw.* **3**(3), 477 (1992). <https://doi.org/10.1109/72.129420>
61. Liu, S.C., van Schaik, A., Minch, B.A., Delbruck, T.: *IEEE Trans. Biomed. Circuits Syst.* **8**(4), 453 (2014). <https://doi.org/10.1109/TBCAS.2013.2281834>
62. Xu, Y., Thakur, C.S., Singh, R.K., Hamilton, T.J., Wang, R., van Schaik, A.: *Front. Neurosci.* **12**, 213 (2018). <https://doi.org/10.3389/fnins.2018.00198>
63. Jiménez-Fernandez, A., Cerezuela-Escudero, E., Miro-Amarante, L., Domínguez-Morales, M.J., de Asís Gómez-Rodríguez, F., Linares-Barranco, A., Jiménez-Moreno, G.: *IEEE Trans. Neural Netw. Learn. Syst.* **28**, 804 (2017). <https://doi.org/10.1109/TNNLS.2016.2583223>
64. Wen, B., Boahen, K.A.: *IEEE Trans. Biomed. Circuits Syst.* **3**, 444 (2009). <https://doi.org/10.1109/TBCAS.2009.2027127>
65. Chan, V., Liu, S.C., van Schaik, A.: *IEEE Trans. Circuits Syst. I Regul. Pap.* **54**(1), 48 (2007). <https://doi.org/10.1109/TCSL.2006.887979>
66. Lyon, R.F.: *Human and Machine Hearing: Extracting Meaning from Sound*. Cambridge University Press (2017). <https://doi.org/10.1017/9781139051699>
67. Inaoka, T., Shintaku, H., Nakagawa, T., Kawano, S., Ogita, H., Sakamoto, T., Hamanishi, S., Wada, H., Ito, J.: *Proc. Natl. Acad. Sci.* **108**(45), 18390 (2011). <https://doi.org/10.1073/pnas.1110036108>
68. Zhao, C., Knisely, K.E., Colesa, D.J., Pflingst, B.E., Raphael, Y., Grosh, K.: *Sci. Rep.* **9**(1), 1 (2019). <https://doi.org/10.1038/s41598-019-39303-1>
69. Ammari, H., Davies, B.: *Proc. R. Soc. A.* **476**(2234), 20190870 (2020). <https://doi.org/10.1098/rspa.2019.0870>
70. Xie, Y., Tsai, T.H., Konneker, A., Popa, B.I., Brady, D.J., Cummer, S.A.: *Proc. Natl. Acad. Sci.* **112**(34), 10595 (2015). <https://doi.org/10.1073/pnas.150227611>
71. Han, J.H., Kwak, J.H., Joe, D.J., Hong, S.K., Wang, H.S., Park, J.H., Hur, S., Lee, K.J.: *Nano Energy.* **53**, 198 (2018). <https://doi.org/10.1016/j.nanoen.2018.08.053>
72. Crowley, K.M.: Thesis (2015). <http://hdl.handle.net/10919/51611>
73. Joyce, B.S., Tarazaga, P.A.: *J. Intell. Mater. Syst. Struct.* **28**(13), 1816 (2017). <https://doi.org/10.1177/1045389X16679289>
74. Joyce, B.S., Tarazaga, P.A.: *J. Intell. Mater. Syst. Struct.* **28**(6), 811 (2017). <https://doi.org/10.1177/1045389X16657425>
75. Kim, H., Song, T., Ahn, K.H.: *Appl. Phys. Lett.* **98**(1), 013704 (2011). <https://doi.org/10.1063/1.3533907>
76. Song, T., Park, H.C., Ahn, K.H.: *Appl. Phys. Lett.* **95**(1), 013702 (2009). <https://doi.org/10.1063/1.3167818>
77. Magar, K.S.T., Reich, G.W., Rickey, M., Smyers, B., Beblo, R.: *Aerodynamic Parameter Prediction on a Airfoil with Flap via Artificial Hair Sensors and Feedforward Neural Network*. American Institute of Aeronautics and Astronautics, p. 1540 (2016). <https://doi.org/10.2514/6.2016-1540>
78. Lee, C., Park, S.: *Bioinspiration & Biomim.* **7**, 046013 (2012). <http://hdl.handle.net/10919/51611>
79. Reichenbach, T., Hudspeth, A.J.: *Phys. Rev. Lett.* **106**, 158701 (2011). <https://doi.org/10.1103/PhysRevLett.106.158701>
80. Barazani, B., Dion, G., Morissette, J.F., Beaudoin, L., Sylvestre, J.: *J. Microelectromechanical Syst.* **29**(3), 338 (2020). <https://doi.org/10.1109/JMEMS.2020.2978467>
81. Dion, G., Mejaouri, S., Sylvestre, J.: *J. Appl. Phys.* **124**(15), 152132 (2018). <https://doi.org/10.1063/1.5038038>
82. Coulombe, J.C., Sylvestre, J.: *PLOS ONE.* **12**(6), e0178663 (2017). <https://doi.org/10.1371/journal.pone.0178663>
83. Mahboob, I., Yamaguchi, H.: *Nat. Nanotech.* **3**, 275 (2008). <https://doi.org/10.1038/nnano.2008.84>
84. Yao, A., Hikihara, T.: *Appl. Phys. Lett.* **105**(12), 123104 (2014). <https://doi.org/10.1063/1.4896272>

85. Hafiz, M.A.A., Kosuru, L., Younis, M.I.: *J. Appl. Phys.* **120**(7), 074501 (2016). <https://doi.org/10.1063/1.4961206>
86. Tsuji, T., Nakayama, A., Yamazaki, H., Kawano, S.: *Micromachines.* **9**(6) (2018). <https://doi.org/10.3390/mi9060273>
87. Guerreiro, J., Jackson, J.C., Windmill, J.F.: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing—Proceedings*, p. 1478 (2019). <https://doi.org/10.1109/ICASSP.2019.8682831>
88. Yamazaki, H., Yamanaka, D., Kawano, S.: *Micromachines.* **11**(7) (2020). <https://doi.org/10.3390/mi11070644>
89. Lenk, C., Seeber, L., Ziegler, M.: In: *Mikro-Nano-Integration; 8th GMM-Workshop*, pp. 1–3 (2020)
90. Pedrak, R., Ivanov, T., Ivanova, K., Gotszalk, T., Abedinov, N., Rangelow, I.W., Edinger, K., Tomerov, E., Schenkel, T., Hudek, P.: *J. Vac. Sci. & Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.* **21**(6), 3102 (2003). <https://doi.org/10.1116/1.1614252>
91. Ivanov, T.: Thesis (2003). urn:nbn:de:hebis:34-1153
92. Rangelow, I.W., Ivanov, T., Ahmad, A., Kaestner, M., Lenk, C., Bozchalooi, I.S., Xia, F., Youcef-Toumi, K., Holz, M., Reum, A.: *J. Vac. Sci. & Technol. B.* **35**(6), 06G101 (2017). <https://doi.org/10.1116/1.4992073>
93. Michels, T., Rangelow, I.W.: *Microelectron. Eng.* **126**, 191 (2014). <https://doi.org/10.1016/j.mee.2014.02.011>
94. Zöllner, J.P., Durstewitz, S., Stauffenberg, J., Ivanov, T., Holz, M., Ehrhardt, W., Riegel, W.U., Rangelow, I.W.: *MDPI Proc.* **2**(13), 846 (2018). <https://doi.org/10.3390/proceedings2130846>
95. Lenk, C., Ved, K., Gutschmidt, S., Ivanov, T., Hövel, P., Meurer, T., Ziegler, M.: *Dynamically Adaptable Acoustic Sensor with Nonlinear Filtering Functionality* (2021). <https://mne2021.exordo.com/programme/presentation/353>
96. Roesser, D., Gutschmidt, S., Sattel, T., Rangelow, I.: *J. Microelectromech. Syst.* **25**(1), 78 (2016). <https://doi.org/10.1109/JMEMS.2015.2482389>
97. Durstewitz, S., Lenk, C., Ziegler, M.: In: *2022 IEEE International Symposium on Circuits and Systems (ISCAS)* (2022). <https://doi.org/10.1109/ISCAS48785.2022.9937484>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Bio-inspired Perceptual Decision-Making Circuit Based on the Hassenstein-Reichardt Direction Detector



Tom Birkoben, Mirko Hansen, Marina Ignatov, Martin Ziegler,
and Hermann Kohlstedt

Abstract Decision-making belongs to one of the most prominent principles in the nerve system of living species. The crucial trade-off between time to decide (speed) and accuracy is essential for a creature to react to environmental stimuli and most importantly, relevant for survival. In-depth studies of coherent motion tasks in primates, based on two-choice decision experiments, led to a profound understanding of neuronal information processing and even bridged the gap to behavioural psychology. Three fundamental processing stages are needed to successfully perform a two-choice decision experiment based on a saccadic eye movement task: the neuronal representation of the sensory signal, the integration of the incoming stimuli and the comparison of the accumulated information to a threshold for a final decision. In accordance to the results of decision-making experiments with primates, an analogue electronic decision-making circuit was developed. Our concept study includes a LED-matrix showing moving light dots (i.e. the tasks screen in primate experiments), an array of photo diodes (i.e. retina of the primate, sensory evidence), a Hassenstein-Reichardt Detector (HRD) based motion detection (i.e. a correlation-type motion detector with a high biological relevance, based on an inhibitory circuit and delay lines) and finally a signal integration circuit. The latter stage represents the decision-making unit and is realised through an inhibitory mutual coupling of the accumulated signals. The entirely analogue circuit is based on integrate and fire relaxation oscillators to mimic firing neurons. We would like to emphasise that the presented real time decision-making system is closely related to biological information pathways in nerve systems. Neither a digital computer nor programming is necessary for decision-making, which might be an interesting approach for real time data processing in autonomous robots.

T. Birkoben (✉) · M. Hansen · M. Ignatov · H. Kohlstedt
Chair of Nanoelectronics, Faculty of Engineering, Institute for Electrical Engineering and Information Engineering, Kiel University, Kaiserstraße 2, 24143 Kiel, Germany
e-mail: tobi@tf.uni-kiel.de

H. Kohlstedt
e-mail: hko@tf.uni-kiel.de

M. Ziegler
Department of Electrical Engineering and Information Technology, Technische Universität Ilmenau, 98693 Ilmenau, Germany

Keywords Decision-making · Motion detector · Pulse oscillator · Visual system · Neuromorphic circuit

1 Introduction

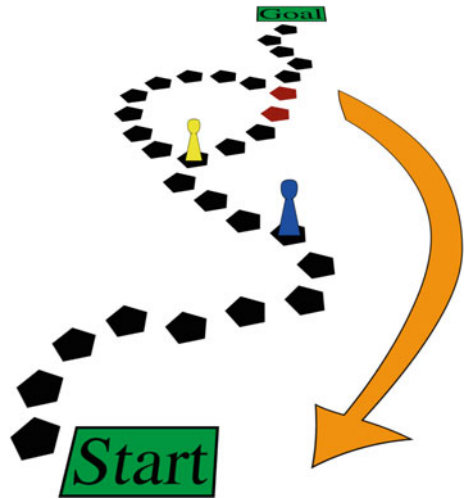
The remarkable interaction performance between biological nervous systems and their environment is a result of million years of evolution explained by Darwinism [1]. Indeed, humans, mammals and even simple life-forms as invertebrates are well adapted to permanently changing surroundings and only appropriate reactions ensure survival. Exploring the whole signal pathway from sensory transduction, decision-making neuron ensembles and the final response via motor neurons and muscles is the challenging task in neuroscience. Although a tremendous progress in the understanding of biological nervous systems can be noticed, there still is a considerable lack to understand all principles of information processing from the molecular length scale to the system size. In other words the “brain code” is not yet understood. Consequently, up to today it is impossible to engineer bio-inspired neuromorphic systems, which mimic the signal pathways of living species. The real benefits of biological information processing when it comes to cognitive tasks or pattern recognition together with extreme power efficiency, are still unattainable in their technical pendants. Instead of trying to solve all problems at once, it is much more efficient and pragmatic to mimic specific features of biological nerve systems by focusing on simpler and more visible goals. This strategy has been applied very successfully in the past [2]. We follow the same path. The presented work focuses on central aspects of biological two-alternatives decision-making and its neuromorphic realisation in a compact battery-powered analogue circuitry, which might be of relevance in the field of robotics.

The paper is organized in the following way: starting with an example of decision-making in daily life, models based on psychophysical experiments are shortly summarized. Hereafter, the most relevant model for this work, the leaky competing accumulator, is explained by means of a circuit block diagram. This scheme is transferred to an analogue circuit comprising a Hassenstein-Reichardt motion detector, relaxation type oscillators and delay-lines. Subsequently, experimental results based on moving dot tasks and the performance of the technical decision making system are presented. Finally, similarities and deviations between the bio-inspired electronics and the biological motivated, leaky competing accumulator are discussed.

Studies of decision-making encompass a broad range of research fields such as neuroscience, psychology, statistics, economics and robotics [3]. Essential ingredients of a typical decision making task are illustrated in the comprehensible cartoon in Fig. 1.

In a board game of dice you (represented by the yellow pawn in Fig. 1) may face two possible paths by reaching an intersection and a decision has to be made [4]. The right path is shorter (less steps) compared to the left one and you might be able to reach the goal faster but the path includes risky fields (labeled in red). With bad luck your pawn has to turn back to the start, with the consequence that you might

Fig. 1 Board game cartoon to illustrate a binary decision task. The yellow pawn could take the left path at the fork, i.e. many steps but a save track towards the goal or the right path, which means less steps but a higher risk because a hold on the red areas will lead to a reset. [adapted from [4]]



lose the game. Decision-making in our daily live, even under much more serious consequences, is omnipresent. How can a final choice, a decision, be explained in terms of neuronal signal pathways in our brains or in mammals in general? This question is at the heart of decision-making research and is discussed over the last decades [3, 5–10]. In addition we would like to emphasize that decision-making is often made under limited time and leads to a tradeoff between decision quality and decision-making time, also called accuracy—speed task [11–13].

Two major categories of decision-making models have been developed. The first category applies to complex decisions among valued alternatives [7, 14, 15]. Here the time for a final decision could take minutes or even much longer. The second category, the so-called sequential-sampling type of models, applies to binary or multiple perceptual and memory related tasks [3, 8, 16]. In contrast to the valued-based decisions, the response time is typically in the order of a few seconds. In Fig. 2 the taxonomy of the main decision-making models belonging to the second category (perceptual tasks) are summarized. The models are explained by the simplest decision-making process, i.e. to choose one of two-alternative evidences, let's say left, L, or right, R. Although neurobiological studies showed that these two model classes for decision-making (valid (not subject in Fig. 2) and perceptual) are linked to some extend [6, 9, 15], we will focus exclusively on two-alternatives perceptual decision-making models. Sequential-sampling models are divided in further subgroups, such as random-walk models and diffusion processes on the one hand and accumulator and counter models on the other hand [8]. For random-walk models as well as for accumulator/counter models it is assumed, that noisy input stimuli are accumulated (integrated) over time until a threshold criterion is reached, representing the final decision. Whereas, in random-walk models evidence is based on accumulating a single total, accumulator/counter models evidences are integrated over two processes separately, i.e. in case of two alternatives a kind of a distinct race towards both of the two thresholds.

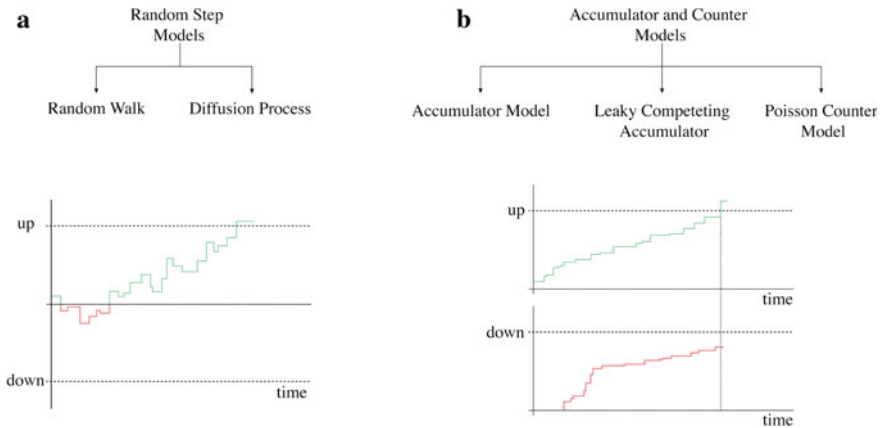


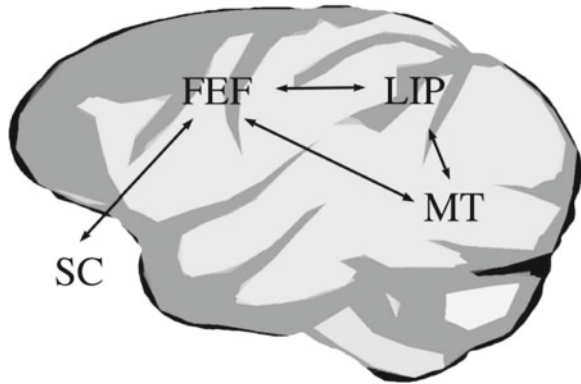
Fig. 2 Taxonomy of perceptual decision making models. **a** Random step models. Here the difference (e.g. left L signal evidence—right R signal evidence) between two alternatives is plotted versus time. An upper and lower threshold defines the final decision. Subgroups are the random-walk model and the diffusion type Wiener and Ornstein-Uhlenbeck models; **b** in accumulator-and counter-type models, the two evidences are separately integrated over time. Nonetheless an inhibitory, cross-wised coupling of the two sensory evidences assures a distinct final decision. Subgroups are the accumulator, the leaky competing accumulator and the Poisson type model [see [8] for more details]. In this work an analog circuit is realized in accordance to the leaky competing accumulator model

After this overview about theoretical models of decision-making, a short survey of neurobiological experiments, in the context of perceptual decision-making with a focus on dot motion discrimination tasks, is presented [8].

Since the first experimental attempt by Newsome et al., visual-saccadic decision-making became one of the most important methods to study visual signal pathways in primates [5, 17, 18]. For example, in so-called coherent motion (two-alternatives) sensory-motor tasks, the experimental procedure is as follows: A dynamic ensemble of dots is displayed on a screen. Some of these dots move randomly, whereas others move either coherently to the left, L, or to the right, R. A trained primate (mostly alert rhesus monkeys) observes the scenario on the screen and makes, after a response time (typically 2 s), a decision, i.e. a saccadic eye movement to the left or right, indicating its choice (decision) in which direction most of the coherent dots previously moved. Simultaneously to the monkey's decision-making, activities of visual cortical neurons are detected in vivo. A clear correlation between the monkey's behavior and neuronal activity was found. Later on, other psychophysical experiments e.g., the oddball task and the dot separation task were developed [8] and extended to non-invasive methods, such as MEG (magneto encephalography), fMRI (functional magnetic resonance imaging) and EEG (electro encephalography) [19].

As raised above, decision-making is a biological process, which translates noisy sensory information via perception and cognition into an action. So finally the motor system executes the decision made by higher brain functions and represents the weighting of the nerve system and takes the possible alternatives, set by the environment and its own goals, into account. We would like to emphasize, that reward,

Fig. 3 Areas of the oculomotor system involved in decision-making are exemplarily shown in a sketched brain of a primate. For decision making neuronal correlates such as, the middle temporal area (MT also called V5), lateral interparietal area (LIP), the frontal eye field (FEF) and the superior colliculus (SC) are involved



in form of stored information in the nerve system, i.e. the former experience of the creature, has also considerable influence in the decision process but is not considered in this work [15, 20].

Which neuronal correlates are responsible for decision-making? For two-choice discriminator (saccadic eye movement) tasks, it becomes clear, that oculomotor areas including the middle temporal area (MT also called V5), lateral interparietal area (LIP), the frontal eye field (FEF) and the superior colliculus (SC), are involved. More details about the interacting neuronal areas can be found in Refs. [8, 21–23]. In Fig. 3 these areas and their approximate locations in a non-human primate brain are depicted schematically.

It is believed that neurons in the area LIP mediate a simple decision process by integrating the sensory evidences over time towards a threshold, i.e. a criterion to commit one of the two possibilities. In other words, the decision process is a race between two competing alternatives. Binary, opposite movements in decision making are simultaneously elusive because a single motor response is required. Hence, models based on not fully independent sensory systems, including a mutual inhibitory component between the two signal pathways, were developed. The principles reflecting essential stages of the saccadic eye movement experiment are shown in Fig. 4a [24, 25]. A primate is facing a display to perform the coherent motion task. The behavioral task is illustrated in the same figure. Neuronal direction-sensitive motion detectors in the primate brain extract the visual strength of the moving dot ensembles. In more detail, the signals for dots, moving to the left and to the right are accumulated (integrated over time) by combined neurons. Both signal pathways are coupled via inhibitory neurons. Finally the output neurons indicate the decision and trigger motor neurons for the final saccadic eye movement (target T_1 —left or target T_2 —right).

In accordance to the behavioral task as shown in Fig. 4a, a neuromorphic circuitry was build and the corresponding circuit design is shown in Fig. 4b and will be explained hereinafter.

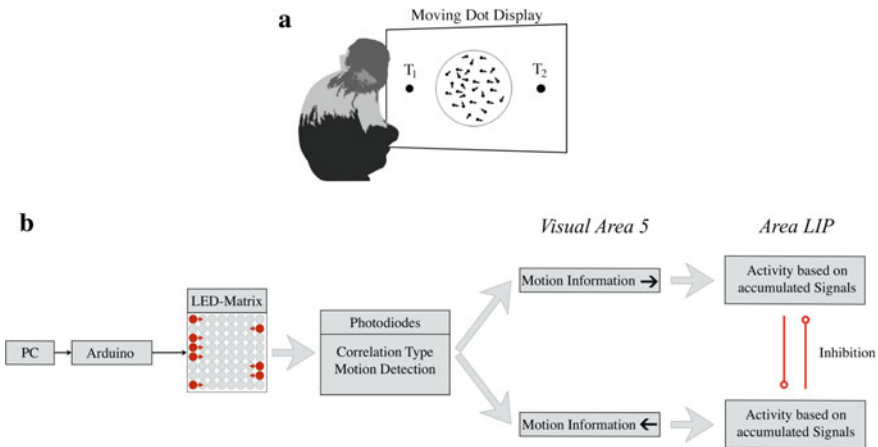


Fig. 4 **a** Sketch of the biological perceptual decision-making experiment. The primate facing a dot display performs a two-alternatives discrimination task defined by an ensemble of moving dots. The final saccadic eye movement reflects the decision whether the majority of the dots has been moving to the left or to the right; **b** block diagram of the biologically inspired perceptual decision making process of a two-alternatives discriminator task. From left to right and in correspondence to the biological pendant the following units are shown: LED (8×8) matrix to define computer-controlled moving dot tasks, a photo diode receptor unit as a retina pendant (a 5×5 matrix), a correlation-type Hassenstein-Reichardt motion detector (HRD), two integrating units for the two pathways, cross-wisely linked inhibitory neurons and two accumulation units. The corresponding biological units, the MT areal and the LIP-areal are marked [7, 26]

2 Materials and Methods

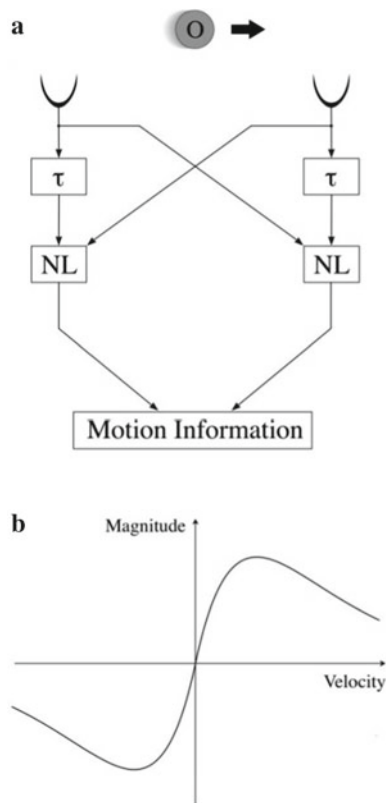
2.1 A Bio-inspired Block Diagram

The essential parts (left to right in Fig. 4b) are a computer controlled LED-matrix (representing the task screen in biological experiments), an array of photo diodes (i.e. retina of the primate, sensory evidence), motion detection (i.e. a correlation-type motion detector with a high biological relevance) based on an inhibitory circuit and delay lines) and finally a signal integration circuit for decision-making. The LED-display is connected to a digital computer via a microcontroller interface, in this way different computer-controlled tasks can be realized on the display. The motion detector is based on the Hassenstein-Reichardt motion detector (HRD) model, which was developed in the 1950's to explain signal pathways in insects, such as the fly [27–32]. The concept was further developed by Barlow and Levick [33]. Before continuing with the entire decision making system, the HRD will be described in more detail in sequel.

2.2 The Hassenstein-Reichardt Motion Detector (HRD)

The HRD is a correlation-type system and comprises in a mirror-like fashion two signal pathways including the photoreceptors $P_{1,2}$, two delay lines $\tau_{1,2}$ two (non-linear elements) $NL_{1,2}$ (e.g. multipliers or a vetogate) [27, 33–35]. Finally the signals of the two pathways are subtracted. To illustrate the function of the HRD, an object, O , moving with a positive velocity (i.e. from left to right) is shown in Fig. 5a. At first the photoreceptor P_1 detects a visual signal if the object passes by. The signal is split into two parts, the first serves as the input for the delay line τ_1 , while the second part of the signal represents the input of the non-linear element NL_2 . The qualitative steady-state response of a HRD is shown schematically in Fig. 5b. For a static object ($v = 0$ cm/degree) the output is zero. For objects moving either from left to right or vice versa, the HRD delivers a maximum/minimum response signal, respectively. This allows an insect (such as Chlorophanus or Drosophila) to distinguish between objects with different directions of motion. For larger absolute velocities beyond the maximum/minimum of the object, the signal response gradually vanishes for both directions.

Fig. 5 **a** Principle components of a HRD, comprising two photoreceptors, two delay lines ($\tau_{1,2}$ two non-linear elements $NL_{1,2}$). The information pathways are cross-linked; **b** schematic response curve of a HRD



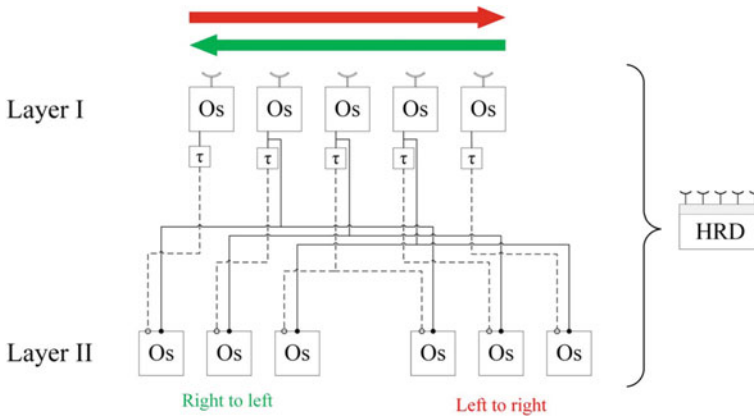


Fig. 6 View graph of the HRD circuit with an input stage comprising five photoreceptors, five input oscillators (layer I) and five delay lines (layer II). Six output oscillators (layer II) are used to indicate the decision between the two-alternatives (L or R). The red and green colored arrows correspond to object movements from left to right or right to left, respectively. On the right hand site a graphical symbol of the HRD is shown

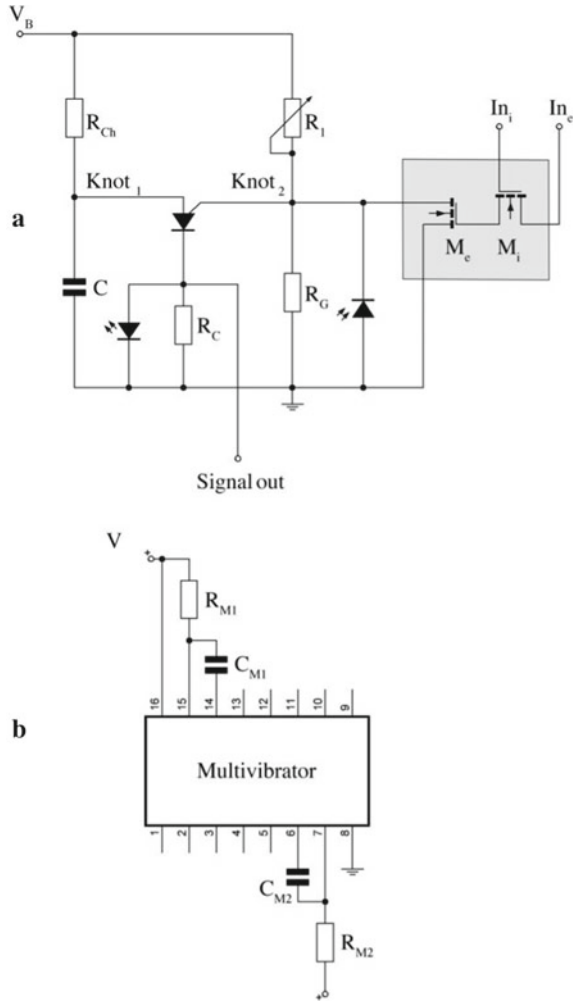
A more detailed layout of the developed HRD is shown in Fig. 6. The visual input received from the five photo receptors (photo diodes) is connected to five oscillator units Os (layer I). The first layer is connected via five delay lines (marked by τ) to the outputs which are further connected to another layer II comprising six oscillators. One part of layer II (the three oscillators marked in red) and the other part of layer II (marked in green) are indicators for motions from left to right and vice versa, respectively. It is worth to mention that one path in the middle is split into two subsequent oscillators on each site (one red and one green). The relaxation type oscillators offer the opportunity for excitatory and inhibitory inputs to realize pivotal neuronal functionalities and will be explained in more detail below [4, 36].

2.3 The Relaxation-Type Oscillator

A spiking neuron can be represented by an oscillator model [4, 37]. How have the oscillators and the delay-lines been realized electronically? The circuit diagram of the implemented relaxation-type oscillator including the excitatory (MOSFET M_e) and inhibitory (MOSFET M_i) coupling scheme and that of a delay-line are shown schematically in Fig. 7a and b, respectively.

The relaxation-type oscillators consist of a programmable unijunction transistor (PUT). (ON Semiconductor, Ebers, 1952, Floyd 1999). PUTs belong to the class of silicon rectifier devices and have four (npnp) alternating regions (see Fig. 8a). By applying a bias voltage between the anode and cathode, the PUT exhibits an $I - V$ curve with a negative differential resistance region as shown in Fig. 8b. Programmable means that the gate Voltage, V_G , applied to the gate terminal, modifies the oscillator

Fig. 7 a Circuit scheme of a programmable unijunction transistor (PUT) based oscillator. Two MOSFETs, M_e and M_i , are connected to the gate resistance R_G and mimic either the biological motivated excitatory (via M_e) or inhibitory (via M_i) coupling scheme of neurons. Both transistors are shown within the grey box. The LED between the cathode of the PUT and ground visualizes low-frequency oscillations (<20 Hz). A pulse at the cathode C of the PUT can be used to trigger other oscillators or delay-lines (signal out). The photo diode parallel to R_G represents one pixel of the receptor matrix (input oscillators in layer I in Fig. 6b). Delay-line based on a multivibrator [38]. The delay can be adjusted by the external resistor R_{M1} and the external capacitor C_{M1} of the multivibrator



frequency. The relaxation-type oscillator (see Fig. 7a), works as follows: The capacitor C is charged through the resistor R leading to the voltage V_C at knot 1. By reaching the voltage $V_{th} = V_G + 0.7$ V defined by the voltage divider R_G , R_1 and the bias voltage V_B , the PUT “fires”, i.e. the charge of the capacitor C is discharged via the cathode terminal to ground and hence the cycle perpetually repeats. Typical voltage traces of V_C of the capacitor C (at knot 1) and the gate terminal V_G (at knot 2) are depicted in Fig. 9. For low frequencies (<20 Hz) a LED between the cathode and ground (and parallel to the cathode resistor R_C) can be used to visualize the oscillations. An extended circuit was used to modify dynamically the gate resistance R_G and mimic excitatory and inhibitory properties of neurons. For those functions, two MOSFETs were integrated, one (M_e) which induces the oscillations (excitatory) and another (M_i) which suppresses the oscillations. The photo diode was implemented

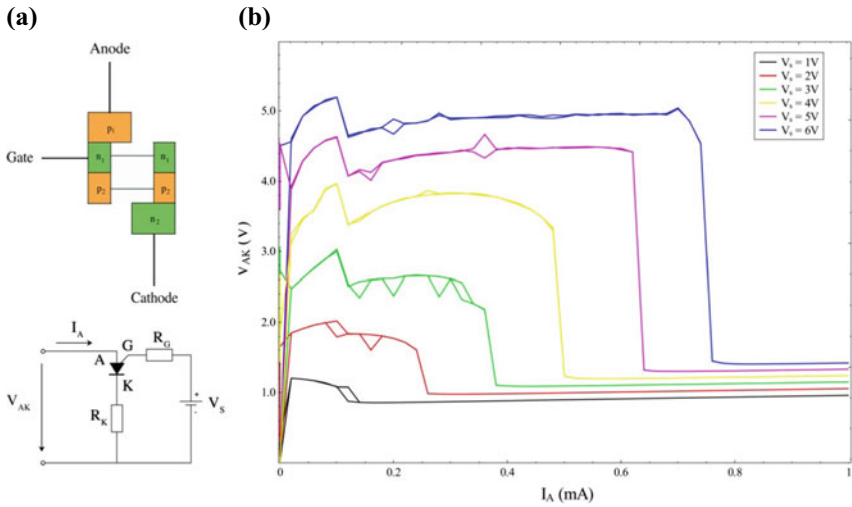


Fig. 8 **a** Schematic of the np regions and terminals of the PUT 2N 6027 [39]. Essential is the internal connection of the n_1 and p_2 regions to realize non-equilibrium charge carrier effects, necessary for a negative differential resistance region (NDR) in the $I - V$ curve [40]. **b** $I - V$ curves measured for positive bias voltages between the anode and cathode. The negative differential resistance regime is clearly visible

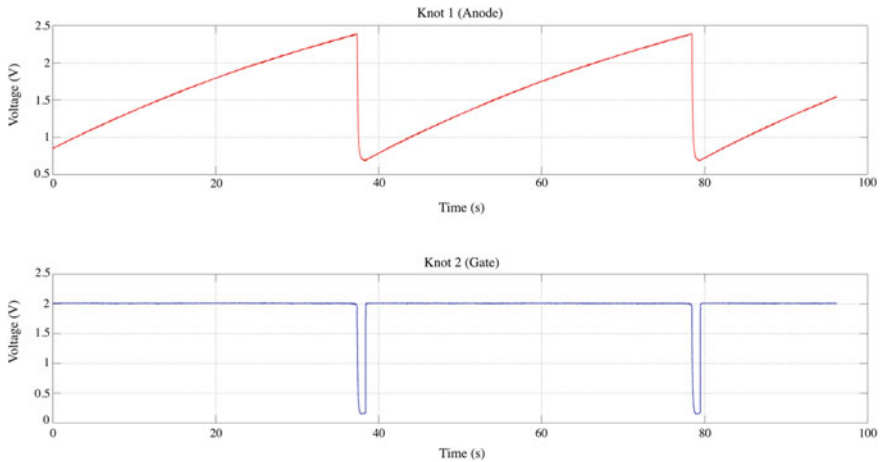


Fig. 9 Upper view graph: Characteristic relaxation-type voltage-time traces of the V_C (corresponds to knot 1 in Fig. 7) which represents the charging of the capacitor via the resistor R . Lower view graph: When the threshold voltage V_{th} (defined by the voltage divider R_1 , R_G and the bias voltage) is reached, the PUT fires (relaxes), which results in a short voltage drop at the gate terminal V_G (corresponds to knot 2 in Fig. 7). For this particular set of parameters the oscillator frequency was approximately 25 Hz

for the oscillators shown in layer I (Fig. 6). This input oscillators in layer I act as an opto-electrical interface and emulate the function of a the retina. The oscillators of layer I start to oscillate only, if the illuminated photo diodes receive a light stimuli whose intensity is high enough. We would like to emphasize, that for the basic version of the PUT (i.e. without the photo diode and the MOSFETs M_e and M_i at the gate terminal, Fig. 7a), the parameters were chosen in a way to suppress oscillations. This is the case when the voltage V_C does not reach the threshold voltage V_{th} of the PUT. Sole in the particular case where the external stimuli is large enough oscillations are observable. This stimuli is either optical for layer I or electrical for all other oscillators at the excitatory input (M_e) in the absence of any inhibitory signal (M_i). Then the relaxation pulses of the oscillators are easily visible via the LED between the PUT's cathode and ground.

2.4 LED-Matrix Chaser as a Dot-Task Display

In Fig. 10 a picture of the entire decision making circuit is shown. The LED-matrix was face-to-face with the 5×5 five photo receptor array. Since the matrix corresponds not one-to-one (8×8 LED to 5×5 photo receptors), only the inner area of it was used as an active part. For the microcontroller interface, an Arduino was used. To program the device it was connected to an laptop using a USB. For programming, tools of the

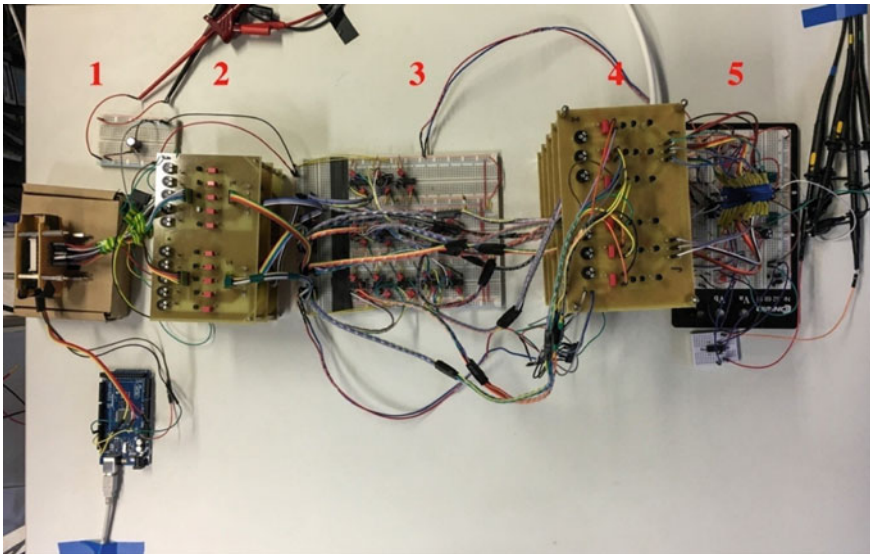


Fig. 10 Photograph of the entire decision making system. The LED display was face-to face with the 5×5 five photo receptor array. The circuit consists of five main subgroups: (1) LED-Matrix controlled via Arduino, (2) Layer I Oscillators, (3) HRD, (4) Layer II - Outcome HRD, (5) Accumulation Stage

Arduino software were used. For a typical dot-task presented to our decision-making system, the LED matrix was programmed as light chasers for each column. This means for example that for three adjacent LED columns the LEDs simultaneously chased from left to right and for the other two columns vice versa. The frequency, i.e. the repetition time for one cycle in a column was ca. 1 s. Under this imbalance condition of moving dots (more left to right then right to left) and similar to physiological motivated dot tasks, we tested the decision-making capability of the system.

3 Results and Discussion

3.1 Entire Electronic Decision Making System

The aforementioned components of a technical decision making system are summarized for one column of photo-detectors and sketched in Fig. 11. The two outputs of each HRD (representing L and R) are connected to the left and right accumulation stage. Each accumulation stage consists of one capacitor integrating the evidence for the right and left direction of motion, respectively. The diodes between the HRDs and the accumulation stages suppress coupling between the HRD oscillators and the discharging of the two capacitors. The voltages (V_L and V_R) of the two capacitors C_L and

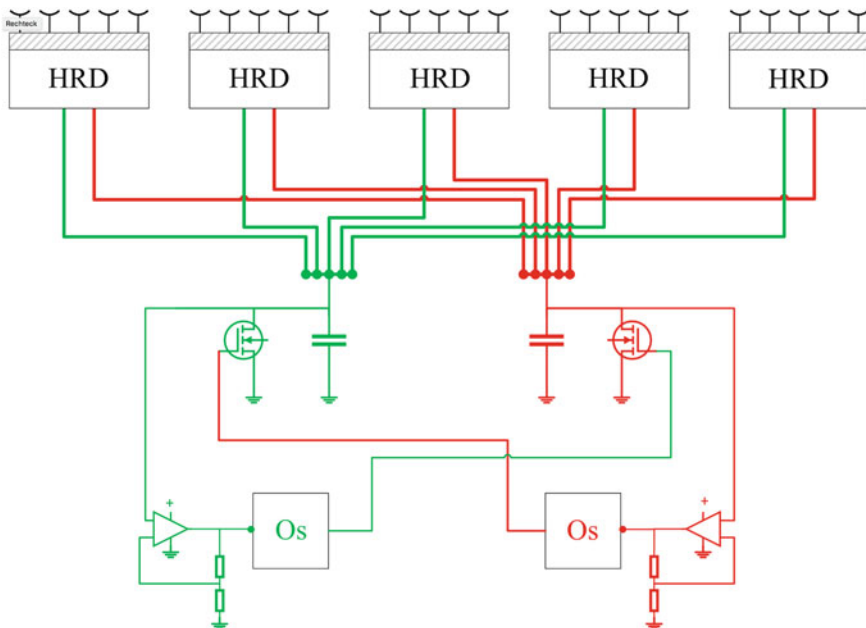


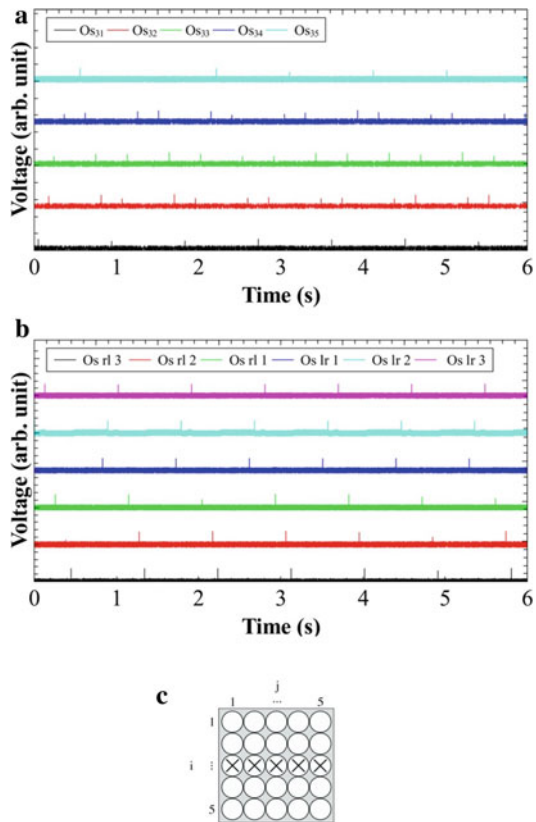
Fig. 11 Entire layout of the biologically-inspired perceptual decision making circuit. At the stages S_1 to S_4 signal traces were acquired

C_R are used to compare the two evidences. To indicate the final decision the voltages are encoded by two indication oscillators. To get the inhibitory coupling scheme as shown in Fig. 4, the encoded V_L and V_R serve as inhibitory inputs for the right and left indication oscillators, respectively. This assures the essential mutual influence of the two evidences as described by the leaky competing accumulator model (see Fig. 2).

3.2 Result of Decision-Making

To illustrate the function of our decision-making system, including the voltage traces at different stages, we defined a clear and simple task via the LED matrix. Only one column was used for the light chase and programmed in a way, that the direction of the light stimulus alternates between left to right and right to left establishing an symmetric input. The complete period was approximately 1 s. For this task the following voltage traces were measured at the stages S_1 to S_4 as labeled in Fig. 11. The results are shown in Fig. 12. In Fig. 12a the encoded light stimulus at the oscillation layer I

Fig. 12 **a** The five voltage versus time traces of layer I oscillators connected to one column of the photo diode array (see Fig. 6). The pattern represents the LED light chase in one row; **b** the six voltages traces of the Hassenstein-Reichardt detector (HRD) (see Fig. 6, layer II). **c** Layout of the photo-diode array of Layer I. The voltage traces of oscillators connected to the marked diodes are depicted in **a** (Os_{ij})



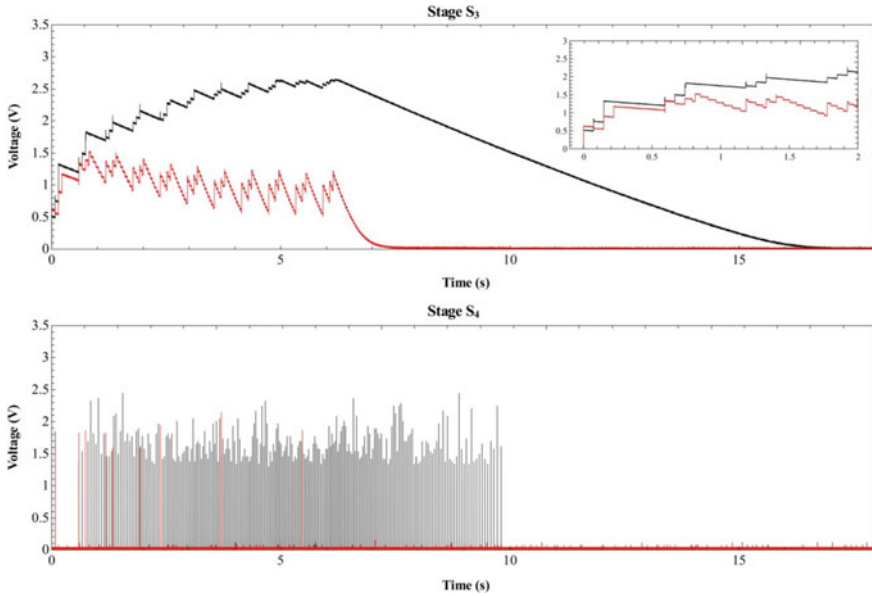


Fig. 13 Typical results of the decision making circuitry for a dot discrimination task. In the upper diagram the voltage versus time traces of stage S_3 , which correspond to the charge state of the two capacitors (see Fig. 11), are shown. Each trace represents an indicator for the strength, i.e. the number of dots moving to the left (movement R to L, black line) or right (movement L to R, red line) direction. The lower diagram depicts the encoded information in form of firing rates of the two output oscillators, used as a negative feedback

(stage S_1), i.e. the input for the HRDs is shown. The five spike sequences represent the sensory transformation of the optical pattern. In more detail, the sequential spike arrangement reflects the chronology of the programmed chasing light cascade which served as inputs for the correlation-type motion detectors. The output after the HRDs is depicted in Fig. 12b (at stages S_2) exhibits the two motion directions.

In Fig. 13 the voltage traces at S_3 and S_4 are depicted. S_3 represents the charge state of each capacitor. The time evolutions in the upper diagram reflect the accumulation of evidence during a decision process. At the beginning of the task ($t = 0$ s) the two signals increase simultaneously, because the system recognized evidence for the two alternatives. In the inset the time interval up to 2 s is enlarged. The race of the two signals is clearly visible. For times larger than 0.7 s, evidence 1 (black) increases, whereas evidence 2 (red) decreases. The latter is due to the negative feedback, established by the discharge MOSFET at the accumulation stage S_3 and reflects the implemented inhibitory function. We like to emphasize the similarity of these results to those, predicted in the biological motivated accumulator and counter models shown in Fig. 2b. The charge state is encoded in the firing rate detected at the cathodes of the two oscillators at stage S_4 . This ensures a counteracting response, i.e. one evidence increases at the expenses of the other evidence during decision-making.

4 Conclusion

In summary, we presented a bio-inspired analogue circuit for a two-alternatives perceptual moving dot task, known from physiological based discrimination experiments. Pivotal neuronal correlates of oculomotor areas in primates, such as the middle temporal area (MT also called V5), lateral intraparietal area (LIP), the frontal eye field (FEF) and the superior colliculus (SC) known to be involved in decision-making discriminators task, were mimicked by an analogue circuit. Based on the principle of a Hassenstein-Reichardt motion detector, relaxation-type oscillators, delay-lines, a LED-matrix and a photo-diode matrix, simple two-alternatives decision tasks were solved by our system. The circuit processes in real time and no computer is involved (we exclude here the laptop/microcontroller used to define the motion task). Rather than developing the system to break the world record in decision-making performed by a technical system, we focused on a closely bio-inspired version. Fundamental principles such as spiking neurons, inhibitory and excitatory coupling schemes and the Hassenstein-Reichardt motion detector were emulated.

In the future it might be possible to include a reward in form of memristive devices which change their state. Moreover studies as the system response-time versus the task difficulty can be performed [41, 42]. We hope that our approach makes a relevant contribution to the field of robotics based on neurobiological principles.

Acknowledgements This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors. Financial support by the German Research Foundation through FOR 2093 is gratefully acknowledged.

References

1. Shanahan, T.: *The Evolution of Darwinism: Selection, Adaptation and Progress in Evolutionary Biology*. Cambridge University Press (2004)
2. Ignatov, Marina, Ziegler, Martin, Hansen, Mirko, Petraru, Adrian, Kohlstedt, Hermann: A memristive spiking neuron with firing rate coding. *Front. Neurosci.* **9**, 376 (2015)
3. Gold, J.I., Shadlen, M.N., et al.: The neural basis of decision making. *Annu. Rev. Neurosci.* **30**(1), 535–574 (2007)
4. Gerstner, W., Kistler, W.M., Naud, R., Paninski, L.: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press (2014)
5. Newsome, W.T., Britten, K.H., Movshon, J.A.: Neuronal correlates of a perceptual decision. *Nat.* **341**(6237), 52–54 (1989)
6. Schall, J.D.: Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* **2**(1), 33–42 (2001)
7. Glimcher, P.W.: The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* **26**(1), 133–179 (2003)
8. Smith, P.L., Ratcliff, R.: Psychology and neurobiology of simple decisions. *Trends Neurosci.* **27**(3), 161–168 (2004)
9. Shadlen, M.N., Kiani, R.: Decision making as a window on cognition. *Neuron.* **80**(3), 791–806 (2013)

10. Cassey, P., Heathcote, A., Brown, S.D.: Brain and behavior in decision-making. *PLoS Comput. Biol.* **10**(7), e1003700 (2014)
11. Wickelgren, W.A.: Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* **41**(1), 67–85 (1977)
12. Uchida, N., Mainen, Z.F.: Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* **6**(11), 1224–1229 (2003)
13. Rinberg, Dmitry, Koulakov, Alexei, Gelperin, Alan: Speed-accuracy tradeoff in olfaction. *Neuron.* **51**(3), 351–358 (2006)
14. Luce, R.D., et al.: *Response Times: Their Role in Inferring Elementary Mental Organization*, no. 8. Oxford University Press on Demand (1986)
15. Sugrue, L.P., Corrado, G.S., Newsome, W.T.: Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**(5), 363–375 (2005)
16. Salzman, C.D., Newsome, W.T.: Neural mechanisms for forming a perceptual decision. *Sci.* **264**(5156), 231–237 (1994)
17. Zeki, S.M.: Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol.* **236**(3), 549–573 (1974)
18. Wurtz, R.H., Goldberg, M.E.: Activity of superior colliculus in behaving monkey. 3. cells discharging before eye movements. *J. Neurophysiol.* **35**(4), 575–586 (1972)
19. Heekeren, H.R., Marrett, S., Bandettini, P.A., Ungerleider, L.G.: A general mechanism for perceptual decision-making in the human brain. *Nat.* **431**(7010), 859–862 (2004)
20. Nichols, M.J., Newsome, W.T.: The neurobiology of cognition. *Nat.* **402**(6761), C35–C38 (1999)
21. Shadlen, M.N., Newsome, W.T.: Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *J. Neurophysiol.* **86**(4), 1916–1936 (2001)
22. Mazurek, M.E., Roitman, J.D., Ditterich, J., Shadlen, M.N.: A role for neural integrators in perceptual decision making. *Cereb. Cortex.* **13**(11), 1257–1269 (2003)
23. Sommer, M.A., Wurtz, R.H.: Brain circuits for the internal monitoring of movements. *Annu. Rev. Neurosci.* **31**, 317 (2008)
24. Shadlen, M.N., Britten, K.H., Newsome, W.T., Movshon, J.A.: A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**(4), 1486–1510 (1996)
25. Usher, M., McClelland, J.L.: The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**(3), 550 (2001)
26. Shadlen, M.N., Newsome, W.T.: Motion perception: seeing and deciding. *Proc. Natl. Acad. Sci.* **93**(2), 628–633 (1996)
27. Hassenstein, Bernhard, Reichardt, Werner: Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenbewertung bei der bewegungsperzeption des rüsselkäfers chlorophanus. *Zeitschrift für Naturforschung B* **11**(9–10), 513–524 (1956)
28. Reichardt, W.: Autocorrelation, a principle for evaluation of sensory information by the central nervous system. In: *Symposium on Principles of Sensory Communication 1959*, pp. 303–317. MIT press (1961)
29. Zanker, J.M., Srinivasan, M.V., Egelhaaf, M.: Speed tuning in elementary motion detectors of the correlation type. *Biol. Cybern.* **80**(2), 109–116 (1999)
30. Borst, Alexander: Correlation versus gradient type motion detectors: the pros and cons. *Philos. Trans. R. Soc. B Biol. Sci.* **362**(1479), 369–374 (2007)
31. Joesch, M., Schnell, B., Raghu, S.V., Reiff, D.F., Borst, A.: On and off pathways in drosophila motion vision. *Nat.* **468**(7321), 300–304 (2010)
32. Plett, Johannes, Bahl, Armin, Buss, Martin, Kühnlenz, Kolja, Borst, Alexander: Bio-inspired visual ego-rotation sensor for MAVs. *Biol. Cybern.* **106**(1), 51–63 (2012)
33. Barlow, H.B., Levick, W.R.: The mechanism of directionally selective units in rabbit's retina. *J. Physiol.* **178**(3), 477 (1965)
34. Hassenstein, Bernhard: Ommatidienraster und afferente bewegungsintegration. *Zeitschrift für vergleichende Physiologie* **33**(4), 301–326 (1951)

35. Borst, Alexander, Helmstaedter, Moritz: Common circuit design in fly and mammalian motion vision. *Nat. Neurosci.* **18**(8), 1067–1076 (2015)
36. Van Vreeswijk, C., Abbott, L.F., Bard Ermentrout, G.: When inhibition not excitation synchronizes neural firing. *J. Comput. Neurosci.* **1**(4), 313–321 (1994)
37. Izhikevich, E.M.: *Dynamical Systems in Neuroscience*. MIT press (2007)
38. Maini, A.K.: *Digital Electronics: Principles, Devices and Applications*. John Wiley & Sons (2007)
39. 2n6027, 2n6028—programmable unijunction transistor. <http://solarbotics.net/library/datasheets/2N602X.pdf>. Accessed 19 Sept 2022
40. Ebers, J.J.: Four-terminal pnpn transistors. *Proc. IRE.* **40**(11), 1361–1364 (1952)
41. Hanks, T., Kiani, R., Shadlen, M.N.: A neural mechanism of speed-accuracy tradeoff in macaque area lip. *Elife.* **3**, e02260 (2014)
42. Heitz, R.P., Schall, J.D.: Neural mechanisms of speed-accuracy tradeoff. *Neuron.* **76**(3), 616–628 (2012)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Pattern Recognition in the Box Jellyfish Rhopalial Nervous System Mimicked by an Ensemble of Pulsed Coupled Oscillators



Jan Bielecki, Tom Birkoben, Maximiliane Noll, Jan-Frederik Freiberg, Peer Wulff, Heinrich Terlau, and Hermann Kohlstedt

Abstract The rapidly expanding field of biomimetics emulates biological principles, distilled through evolutionary time, to solve problems in technology, medicine and material science. Information processing in neuronal circuitry of biological models may offer guidelines for future highly efficient computing systems exhibiting, as for example, low power consumption accompanied by excellent pattern recognition capabilities. Here, we consider the visual information processing in the nervous system of the box jellyfish *Tripedalia cystophora* as template for pattern recognition hardware to operate self-guiding robotic vehicles or automated driving assistants. LTSpice XVII simulations of coupled relaxation-type oscillators, based on programmable unijunction transistors (PUTs), enabled the emulation of basal visual functionalities of the *T. cystophora* central nervous system. The transfer of such simulations in real hardware circuits including possible limitations are discussed.

Keywords Biological model systems · Box jellyfish · Relaxation-type oscillators · Pulsed coupled oscillators · LTSpice XVII simulations

J. Bielecki (✉) · J.-F. Freiberg · P. Wulff · H. Terlau
Institute of Physiology, Kiel University, Hermann-Rodewald-Strasse 5, 24118 Kiel, Germany
e-mail: j.bielecki@physiologie.uni-kiel.de

J.-F. Freiberg
e-mail: j.freiberg@physiologie.uni-kiel.de

P. Wulff
e-mail: p.wulff@physiologie.uni-kiel.de

H. Terlau
e-mail: h.terlau@physiologie.uni-kiel.de

T. Birkoben · M. Noll · H. Kohlstedt
Nanoelectronics, Department of Electrical and Information Engineering, Kiel University,
Kaiserstraße 2, 24143 Kiel, Germany
e-mail: tobi@tf.uni-kiel.de

M. Noll
e-mail: mno@tf.uni-kiel.de

H. Kohlstedt
e-mail: hko@tf.uni-kiel.de

1 Introduction

Biological model systems are studied to understand distinctive biological processes and mechanisms. Along these lines various species are investigated because it is believed that explorations made in the model organism will render insight into the working principles of other organisms, including humans. Well known examples which have been extensively studied are mice, rats, zebrafish, the fruit fly *Drosophila melanogaster* and the nematode worm *Caenorhabditis elegans* [1–4]. Despite the immense differences in the details, it is believed that the most basal mechanisms of development are recapitulated in all animals and are evolved from the earliest animal ancestors. Hence, the elucidation of a developmental process in one animal is often of importance for understanding development in another [5]. Despite the significance of model systems for comparative studies in biology, they may also present novel guidelines for the field of bio-inspired engineering. The goal hereby is to extract evolutionary shaped basal biological mechanisms from the point of power efficiency and sensitivity and transfer those to technology. Visual information processing in the relatively simple *T. cystophora* nervous system offers a unique opportunity to reverse engineer pattern recognition electronic hardware, and thereby circumvent the need for processing computers to run pattern recognition software. Fast responding pattern recognition hardware could support self-guiding robotic vehicles or automated driving assistants. Here, we consider the visual system of the box jellyfish and realize their optical communication principles in electronic circuitry comprising a set of coupled relaxation-type oscillators [6, 7]. In the first stage we simulate the circuitry by LTSpice XVII [8]. In the present theoretical study, coupled relaxation-type oscillators with a fixed topology enabled essential box jellyfish motor effector activation used in coordinated swimming, hovering, and turning [9].

2 The Box Jellyfish: Anatomy, Dynamic and Behavior

The box jellyfish *Tripedalia cystophora* is emerging as a new model system for basic visual information processing. Photographs of the visual system and nervous system of *T. cystophora* are shown in Fig. 1 [10]. The *T. cystophora* model system is attractive due to its vertebrate like camera type eyes [11, 12], multiple distinct visual behaviors [13–16], and an experimentally tractable central nervous system comprising approximately 1000–1500 processing neurons [17, 18].

2.1 Visual system

The visual system of *T. cystophora*, and box jellyfish in general, comprises 24 eyes of four distinct different types. Eight of these are vertebrate like camera type lens eyes complete with ‘Matthiessen’s ratio’ lens with graded refractive index, retina with

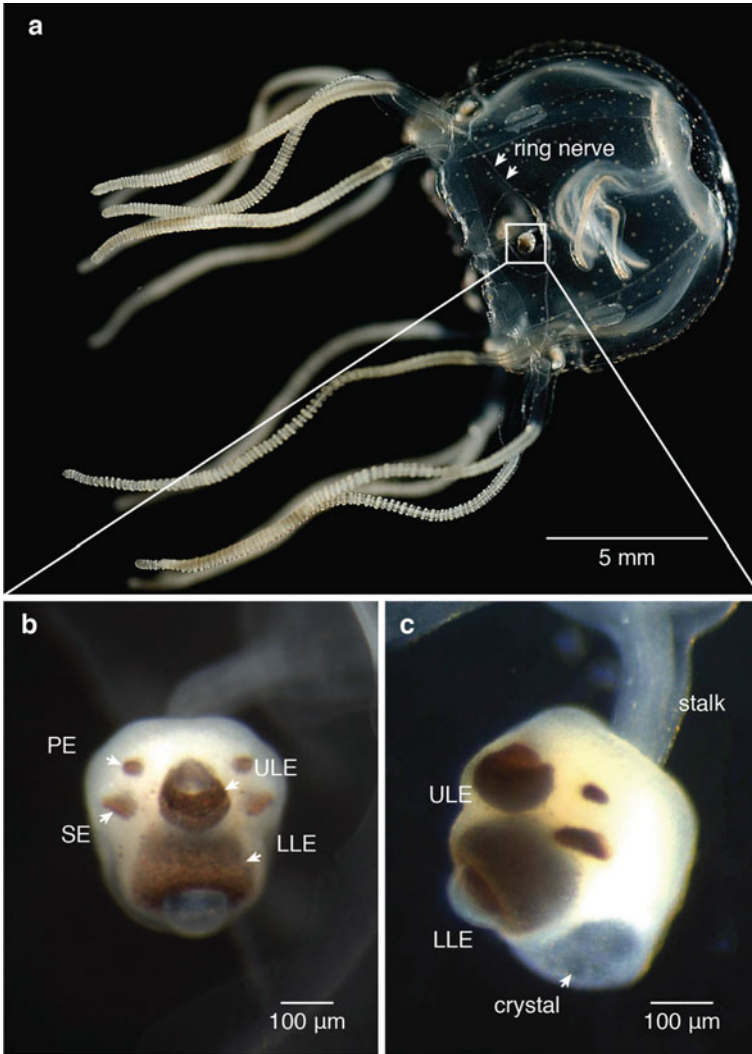


Fig. 1 The visual system and nervous system of *Tripedalia cystophora*. **a** An adult specimen of the mangrove dwelling box jellyfish *Tripedalia cystophora* has a bell diameter of about 10 mm. **b**, **c** The eye-brain complexes (rhopalia) each carries 6 eyes, of four morphological different types, and the processing neural circuitry. The four rhopalia are suspended from the sides of the bell by a flexible stalk and weighed down by a crystal on the distal end. The epidermal stalk nerve conveying the motor signal is located in the medial side of the stalk. A circumnavigating ring nerve enables inter-rhopalial communication (**a**). PE: pit eye, SE: slit eye, ULE: upper lens eye, LLE: lower lens eye. Figure adapted from [10]

opsin-based photoreceptors, and pigment screen to accommodate directional vision [12, 19, 20]. The eyes are distributed on four sensory structures termed rhopalia (see Fig. 1b and c), each of which is carrying two lens eyes and four ‘lesser’ eyes [12, 20]. The rhopalium also carries the processing neuronal circuitry used to process the visual information collected from the eyes. The rhopalia are suspended from the sides of the animal by a flexible stalk, which contains the efferent epidermal motor nerve, and are inter-connected through a ring nerve that circumnavigates the medusoid bell (Fig. 1a) [21]. It is thought that the ring nerve is involved in inter-rhopalial communication and, at least to some degree, coordinates the contraction of the bell [22]. A calcium sulphate hemihydrate crystal [23] at the distal end ensures a constant vertical orientation of the rhopalium and thereby constant field of view of the lens eyes [15, 24]. The lens eyes are thought to be special purpose eyes modulating specific visual behaviors. The upper lens eye is directed out of the water and into Snell’s window. Snell’s window is a physical phenomenon created by the difference in refractive indices between air and water that visually compresses the 180° hemisphere above water into an approximately 97° underwater cone. The visual field of just below 100° perfectly matches the upper lens eye to Snell’s window [12]. By adjusting the direction and angle of the visual field to accommodate visual information received from Snell’s window, the upper lens eye can detect the contrast line between the mangrove canopy and the open sky as a bright line on a dark background [24]. *T. cystophora* can thereby use terrestrial cues for long distance navigation [15]. Tidal currents constantly pose the risk of sweeping *T. cystophora* into the middle of the mangrove creek lagoon and subsequently to the open sea. Being swept away from the primary habitat would prove fatal to the animal since *T. cystophora* forage in light shafts created by sunlight shining through the mangrove tree foliage and between the prop roots [16], where their photo tactic copepod prey is abundant. Therefore, if the animal finds itself under the open sky, it will quickly turn and swim back under the canopy [15]. The lower lens eye is directed 60° into the water and with a visual field of approximately 170° scans the underwater environment for potentially dangerous objects [12, 13]. It has been shown that the lower lens eye overrides the upper lens eye, indicating that underwater input is regarded of higher importance than above water input [25]. Intuitively a sensible prioritization since collision with the prop roots of the mangrove trees poses an immediate danger of inflicting fatal damage to the fragile box jellyfish bell. As mentioned, the rhopalia are interconnected through the ring nerve and the overall behavior of the animal results from coordination between the rhopalia, prioritizing the rhopalium that detect changes in the visual environment [26–28].

2.2 Obstacle Avoidance Behavior

When *T. cystophora* encounters an underwater obstacle it will perform 4–5 fast bell contractions, turn 120–180° and swim away (Fig. 2) [13]. This distinct visual behavior is modulated by the lower lens eye and is based on true spatial vision. The animal

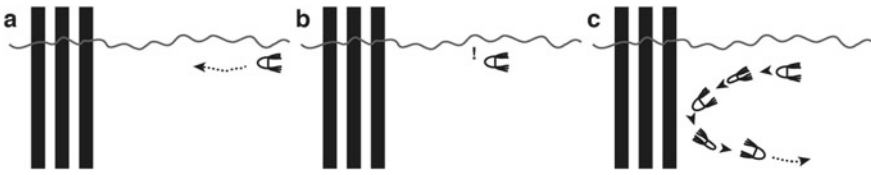


Fig. 2 Obstacle avoidance behavior. **a** *Tripedalia cystophora* navigating the mangrove habitat must avoid collision with underwater object to prevent damaging the fragile bell. **b** When the box jellyfish encounters an obstacle estimated close enough to pose an immediate danger, the animal will initiate obstacle avoidance behavior. **c** Obstacle avoidance behavior comprises 4–5 fast bell contractions, turning 120–180° and swimming away from the obstacle

shows a graded avoidance response to varying contrasts of the obstacles, higher contrast yielding greater response—presumably an inherent gauge of distance underwater [14]. Under the canopy, *T. cystophora* explores the mangrove habitat in search of foraging opportunities in the light shafts (Fig. 2a). When the animal encounters an underwater object, it will estimate the distance by the contrast of the object compared to the surrounding water. When the contrast surpasses the response threshold (Fig. 2b), the obstacle avoidance behavior will be initiated, which, as mentioned, consists of 4–5 fast bell contractions and directional control of the velarium (Fig. 2c). The velarium is a membrane-like constriction of the oral opening of the box jellyfish bell that can be asymmetrically constricted to form directionally controlled jet propulsion of the water flow created by the bell contraction [26]. This enables the animal to turn 180° within 2–3 bell contractions. The bell contracting motor output signal is generated by oscillatory, or pacemaker, cells located on the rhopalium at the base of the stalk [18, 29]. The pacemaker cells have never been unequivocally identified but are thought to be what has been described as ‘giant neurons’ in the literature [17, 18]. Regardless of location in the rhopalial nervous system (RNS), the pacemaker cells produce a swim pacemaker signal that contracts the box jellyfish bell in a 1:1 manner. The intrinsic pacemaker cell frequency is 3–4 Hz, which is suppressed by neuropeptidergic mediation [30] to approximately 0.5–1 Hz. The latter corresponds to leisurely swimming and foraging in the mangrove habitat. The duration of one bell contraction from fully expanded bell to fully expanded bell is approximately 250 ms, corresponding to an absolute physiological relevant maximum of 4 Hz. Usually the animals display swim contractions of less than 3 Hz even during obstacle avoidance events.

2.3 Pacemaker Activity Represents Behavior

The swim pacemaker signals can be recorded from the efferent epidermal stalk nerve (Fig. 3) by a suction electrode (extracellular electrophysiology). The pacemaker signals are 45 ms signals with a profile easily discernable from regular action potentials

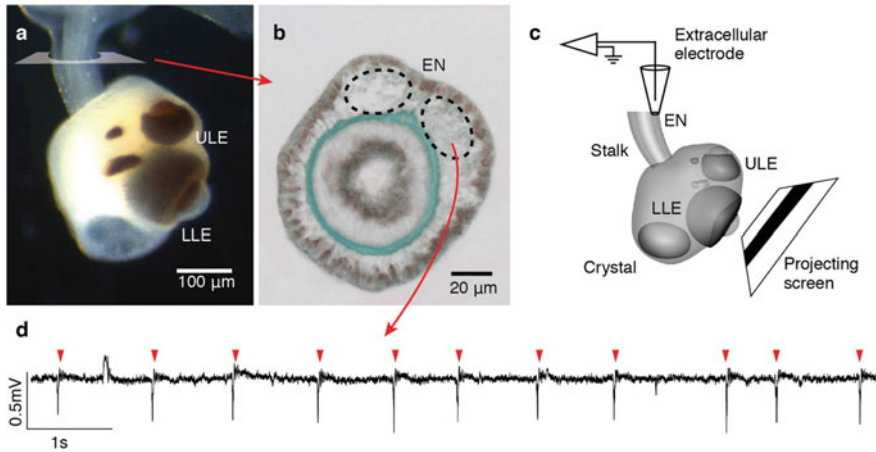


Fig. 3 Pacemaker signals predict the behavior of *Tripedalia cystophora*. **a** To record the motor neuron output signals (pacemaker signals) the stalk is transected at about 2/3 of the distance to the bell. **b** This section reveals the two medially located epidermal stalk nerves (EN) where an extracellular electrophysiological suction electrode can be attached. **c** With the electrode in place, visual stimulation of the lower lens eye (LLE) can be performed and the responding pacemaker activity recorded. **d** Intrinsic pacemaker signal activity recorded from the EN (red arrowheads). The pacemaker signals have significant longer duration (45 ms) than normal action potentials. ULE: upper lens eye. Figure adapted from [30]

(Fig. 3d). Since swim contractions are the only means of locomotion in *T. cystophora* the pacemaker signals characterize the behavior of the entire animal. This correlation is prominent when presenting light-ON and light-OFF stimuli to the lower lens eye. If a rhopalium is kept in the dark for a few minutes the pacemaker signal frequency will stabilize around 0.5–1 Hz. If subsequently the lower lens eye is presented with a light ON stimulus, the pacemaker signal frequency will suddenly decrease and sometimes completely cease. Conversely, a light-OFF stimulus to a light adapted lower lens eye will result in a sudden increase in pacemaker signal frequency. Whereas the light-ON response is long lasting (scale of minutes), the light-OFF response is transient and lasts about 10 s [25]. The pacemaker signal response is directly related to the foraging behavior of *T. cystophora*. When the box jellyfish enters a light shaft (light-ON), where phototactic copepod prey is abundant, the animal will cease swimming and start fishing by passively sinking with outstretched tentacles. If *T. cystophora* inadvertently swims out of the light shaft (light-OFF) it will perform a number of high frequency swim bell contractions, turn, and try to relocate the light shaft [16].

2.4 Lens Eye Morphology and Activity

Light entering the lens eye is first refracted by the lens and focused onto the retina. The lenses of the upper and lower lens eyes are of remarkable high quality and capable of producing images on the retina with high spatial resolution. However, the retinas of the lens eyes are displaced closer to the lens than the focal plane of the lens. This severe under-focus is thought to remove fine image details to avoid overloading the processing circuitry. With this under-focus the receptive fields of the photoreceptors are 15–20° depending on their location in the retina [12]. The upper and lower lens eyes comprise 400 and 600 cone-like photoreceptors respectively. The light sensitivity is accomplished through a c type opsin (cnidops), cAMP, and an opsin Gs pathway which hyperpolarizes the photoreceptor membrane [31, 32]. The photoreceptors are interconnected through invaginated chemical synapses [33] and, since the box jellyfish retina is everted, articulate directly on putative second order neurons. The everted retina means that the light strikes the outer photoreceptive signals directly and that the animals do not have blind spots as in vertebrate eyes, where the afferent optic nerve exits the retina (optic disk).

2.5 Rhopalial Nervous System Organisation

The organization of the RNS is not yet understood in detail. However, visual information, detected by the lens eyes, is processed in putative second order neurons and subsequently applied to modulate swim pacemaker signal frequency. The swim pacemaker cell cluster is a central convergent and translatory hub for processed sensory information. The visual input in turn modulates the pacemaker cell cluster activity and is here translated into the motor signal that controls the motor effectors of the box jellyfish bell. The swim bell contractions are the only means of locomotion for the box jellyfish so the swim pacemaker signals can be used as indication of the behaviour of the animal. The advantage is, as mentioned, that we do not need to examine the entire animal to determine a behavioural reaction to a given visual stimulus [25, 34].

2.6 Retinal Organization, Bipolar Cells Connectivity to ON- and OFF- Ganglion Cells

Vertebrate photoreceptors signal through a type c-opsin and a G-protein cGMP transduction cascade [35] and the photoreceptors thereby hyperpolarize in response to light ON stimuli [36]. In mammalian retinas, bipolar cells respond selectively to the photoreceptor activity. In the dark, the constant release of glutamate from the cone cells keep ON-bipolar cell membranes hyperpolarized, but at light-ON hyper-

polarization of the photoreceptor membrane reduces glutamate release, inducing depolarization of the bipolar cell membrane. In contrast OFF-bipolar cells remain hyperpolarized in light-ON conditions and depolarizes in response to light-OFF events. Retinal ganglion cells subsequently integrate neuronal information from several bipolar cells to produce firing rates corresponding to the received information. Ganglion cells have circular receptive fields with specialized center and antagonistic surround regions (eg. ON-center-OFF-surround). Depending on the type of ganglion cell the firing rate is modulated by light illumination of the center- or surround region. ON-center ganglion cells have low rates of firing under dim illumination and rapidly increase firing in response to light-ON stimulus in the center of their receptive field. In contrast, OFF-center ganglion cells discharge at low rates under light conditions and rapidly increase firing rates by light-OFF stimulus to their receptive field center [36]. The corresponding classes of bipolar and ganglion cells have excitatory connections so that an ON-center bipolar cell depolarization increases ON-center ganglion cell firing rate. The same is true for the OFF-center bipolar and OFF-center ganglions cell connection. There are species differences across the mammalian class but bipolar cells commonly integrate visual information from several light sensitive photoreceptors and, similar to the ganglion cells, have antagonistic center-surround receptive field organization [36].

2.7 Retinal Pre-processing of Basic Shapes (Bars, Contrast Lines, etc.)

The firing rate of ganglion cells provides a measure of the difference in the intensities of light illuminating the center and surround regions. Information about small differences in intensities and the center-surround organization of the ganglion cells receptive fields are therefore designed to report principally on contrast lines rather than absolute intensity [36]. Retinas adapted to optimally scan the appropriate environments and pattern recognition are inherent functions in visual systems across the animal kingdom [36, 37]. Unlike the human general-purpose eyes, that guide all our visual behavior, box jellyfish have special-purpose eyes, which scan the visual field for specific visual information (cf. matched filters) [37]. A visual strategy based on multiple special purpose eyes may seem energetically expensive, but requires far fewer computational circuits in the central nervous system. Based on immunohistochemical staining, it is estimated that *T. cystophora* has approximately one thousand neurons available to process visual information from all six eyes on the rhopalium and to modulate several visual behaviors [17]. Accordingly, the rhopalial nervous system (RNS) is presumed highly compartmentalized, indicating that the actual neuronal network serving one particular eye is considerably smaller [24, 30]. For the transfer into electronics, the pacemakers can be viewed as oscillators with a certain intrinsic frequency that is modulated by visual input. For *Tripedalia cystophora* it would be advantageous to have this type of matched filter in the lower

lens eye to scan the retina for contrast lines that would imply an oncoming underwater obstacle. This could initiate evasive actions and steer the animal clear of collision. Presently, we have little knowledge of the cellular mechanisms involved in obstacle avoidance behavior but from previous work we can form an idea of how obstacle detection – or pattern recognition – can be accomplished in the box jellyfish visual system, and how the motor signal is generated. In the everted retina of *T. cystophora* the photoreceptors directly articulate on second order neurons, which could then be similar in function to the bipolar cells in the human retina. The second order neurons could translate the hyperpolarizing photoreceptor signal into ON- and OFF-center excitatory signals to the pacemaker cells. Three sets of giant neurons, presumably pacemaker cell clusters, have been identified in the *T. cystophora* RNS in the proximal part of the stalk region. Each set has positive immunoreactivity (ir+) to a specific neuropeptide [18]. Since ir+ has been discovered for three specific neuropeptides in three distinct different pacemaker clusters and swim pacemaker signals (Fig. 3d) can be recorded from the efferent epidermal stalk nerve (Fig. 3c), it is attractive to suggest that two subordinate pacemaker cell clusters converge into a terminal pacemaker cell cluster, which produces the motor signal for the effectors. Additionally, it stands to reason that, due to general morphology of photoreceptor cells and the distance from the lens eyes to the putative pacemaker cell clusters, at least one level of interneurons is present between the two cell types. This again supports the presumption of bipolar cell organizational analogy. Assuming two major control units in the rhopalium; the upper and the lower lens eyes, second order neurons from each eye could converge onto one giant neurons/pacemaker cell cluster (sub-cluster) analogous to retinal ganglion cells in mammals. The sub clusters would converge onto the terminal pacemaker cell cluster that would integrate the signal from the entire RNS and produce the appropriate behavioral modulatory motor signal of the epidermal stalk nerve. The convergence of signaling from the two lens eyes and the bias of the two potential cluster outputs in the terminal cluster is beyond the scope of this chapter, but there is evidence that sensory input from the lower lens eye has priority over the other eyes on the rhopalium [25]. Theoretically, this bias could be integrated at the level of the sub-cluster–terminal cluster junction.

2.8 Modeling RNS Visual Information Processing to the Mammalian Retina

Here we will consider the modulation of obstacle avoidance behavior by the lower lens eye. Putatively this could be based on pattern recognition in the retina and oscillator frequency synchrony in the pacemaker cell clusters. Consider the photoreceptors of the lower lens eye, the second order neurons, and the corresponding pacemaker cell cluster in *T. cystophora* to be three levels of matrices analogues to the photoreceptors, bipolar cells, and ganglion cells of the mammalian retina.

Matrix level I: photoreceptors

Approximately 600 photoreceptors make up the retina of the lower lens eye in the box jellyfish, each with an acceptance angle of about 20° [12]. As mentioned, photoreceptors with c type opsins hyperpolarize in response to a light-ON stimulus and assumedly this is the case for box jellyfish as well [31, 32]. A light-ON response implies that the photoreceptors only register a change in light intensity, whereas constant light causes adaptation and image fading or blindness [38]. Adaptation occurs in photoreceptors across the Metazoa regardless of opsin type, and animals must have strategies in place to counteract adaptation. In mammals, oculomotor generated fixational eye movements constantly refresh the retinal image to avoid immobility blindness [39]. The box jellyfish utilizes the bell contractions as fixational eye movements and can in this way detect immobile objects in their visual field [24].

Matrix level II: second order neurons

The photoreceptors articulate directly on second order neurons in the *T. cystophora* everted retina. Potentially, and similar to bipolar cells, OFF-center second order neurons respond to depolarized photoreceptors (light-OFF) by remaining in a depolarized state, and hyperpolarize in response to hyperpolarized photoreceptors (light-ON). Conversely, ON-center second order neurons depolarize in response to hyperpolarized photoreceptors (light-ON) and hyperpolarize with depolarized photoreceptors (light-OFF) [36]. In short, ON-center second order neurons signal in response to light-ON, OFF-center neurons signal in response to light-OFF. Applied to a matrix, we would now have an array of second order neurons activated when a suitable visual stimulus (light or dark) is applied to their respective field of view (acceptance angle). When in the second order neurons a ‘pre-programmed’ visual image fills the field of view of the lower lens eye, all second order neurons could potentially depolarize simultaneously. So, if *T. cystophora* would encounter a speed limit sign (Fig. 4), the image would dovetail the matrix of ON- and OFF-center second order neurons and entirely excitatory signals would be relayed to the respective ON- and OFF-center sub-cluster pacemaker cells. One important aspect of the pattern detection is that since the photoreceptors are binary in nature (either responding to a stimulus or not) the response threshold is important. If the threshold is too low, too much redundant information will be relayed to the processing circuitry—if the threshold is too high, important information could be lost. Figure 4c displays a threshold of sufficient sensitivity to detect a speed limit sign in a cluttered visual environment (numbers within a circle). The *T. cystophora* response to the speed limit sign is an approximation since their photoreceptors do not respond to red light, but this problem will not be encountered in a CCD or CMOS monochrome sensor.

Matrix level III: pacemaker cells

Hypothetically, and similar to bipolar and ganglion cell connectivity in the mammalian retina, several second order neurons converge on the sub-cluster pacemaker cells. This neuronal circuit architecture seems logical in that there are considerably more neurons in the neuropil than there are pacemaker cells. This morphology

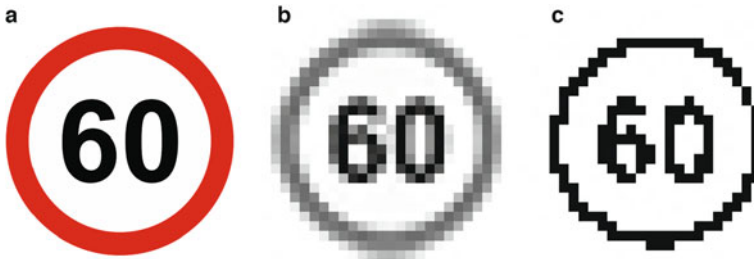


Fig. 4 Representation of the image quality perceived by the *Tripadalia cystophora* retina. **a** Speed limit sign 60 km/h. **b** The same image as in (a) monochrome (single opsin), and adjusted for the approximately 600 pixels/photoreceptors resolution in the lower lens eye of *T. cystophora*. **c** Same image as in (b) with adjusted response threshold

suggests a matrix of i.e. ON-center second order neurons to converge on one ON-center sub-cluster pacemaker cell. Likewise the OFF-center second order neurons converge on OFF-center sub-cluster pacemaker cells. The ON- and OFF-center second order neurons form excitatory synapses with ON- and OFF-center sub-cluster pacemaker cells respectively. This results in sub-cluster pacemaker cells (oscillators) high frequency firing in response to excitatory presynaptic potentials from the second order neurons [36]. Following this line of reasoning when recognizing a pre-encoded pattern, such as the 60 km/h speed limit sign (Fig. 4), under water, all sub-cluster pacemaker cells synchronize to high frequency firing, terminal pacemaker cells initiates obstacle avoidance motor signal. In our automated driving assistant analogy, a potential vehicle would be slowed (or accelerated) to 60 km/h.

3 An Engineered Box Jellyfish by an Ensemble of Pulsed-Coupled Oscillators

As described previously, the box jellyfish’s visual system and rhopalial nervous system (RNS) can be viewed as an actuator coupled vertebrate-like retina. The level of visual information processing in the RNS is comparable to the pre-processing in a human retina in regard to pattern recognition, contrast line detection, and contrast estimation. The difference is that where human retinal information is sent to higher processing centers in the brain to actuate behavioral actions, in box jellyfish the motor effector signal is created directly within the RNS. In the present work, we modeled the described pattern recognition and oscillator synchrony to achieve the terminal pacemaker cluster output necessary for obstacle avoidance behavior. To simplify, we used the humblest pattern possible: one white pixel, one black. The light intensity of each pixel was monitored by a sensory oscillator, which would increase firing frequency when an internal criterion was met. In our experiment, the pattern of the two pixels could be varied in four ways: ON-OFF, ON-ON, OFF-ON, OFF-OFF. By

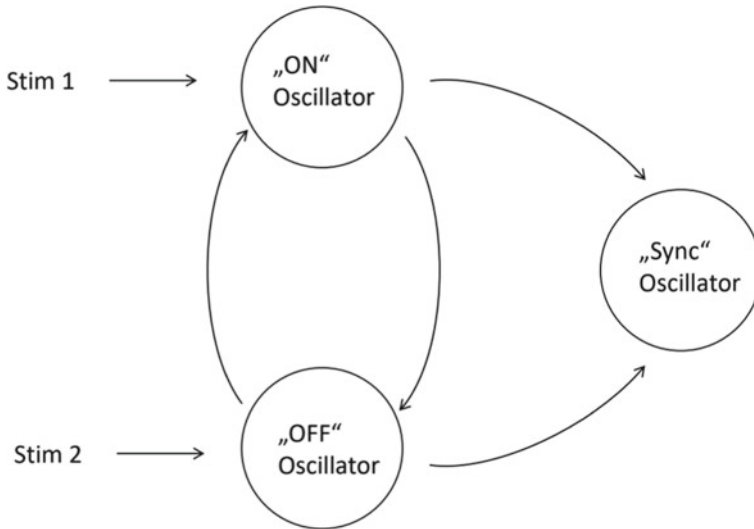


Fig. 5 Abstract functional diagram to mimic visual information processing in *T. cystophora*. The two input oscillators respond to the light intensity on the photodiode (sensor input). When both input oscillators are activated in response to the appropriate visual stimulus, the sync oscillator initiates an activation sequence of 2 Hz spikes. See also Fig. 10

using one ON-center and one OFF-center sensory oscillator (sub-cluster pacemaker cell equivalents) to monitor the presented patterns, the pattern meeting the criteria of both oscillators simultaneously would be the ON-OFF combination. In Fig. 5 this functionality is portrayed in an abstract view graph comprising two interacting input oscillators for the stimulus 1 (ON-oscillator) and stimulus 2 (OFF-oscillator). These oscillators were coupled and connected to a “Sync” oscillator unit (terminal cluster pacemaker cell equivalent). When the two sensory oscillators recognized the ON-OFF pattern simultaneously, the oscillators fired in phase and produced increased signal amplitude. The sync oscillator responded to the increased amplitude by commencing firing and producing the motor signal needed to initiate obstacle avoidance behavior. When the oscillators were uncoupled no increased amplitude could be observed, and the sync oscillator remained silent. The technical realization of this process is discussed below.

This sensory oscillatory experiment is a proof of principle that the rhopalial nervous system can accomplish pattern recognition and respond with evasive action by utilizing interconnected sensory modulated sub-cluster pacemaker cells. The terminal cluster pacemaker cell firing activity is in turn modulated by the synchronized activity (and thereby signals amplitude) of the sub-cluster pacemaker cells. In order to partly mimic the complex functionalities of the box jellyfish neuronal network in electronic hardware, an ensemble of relaxation-type oscillators were applied and

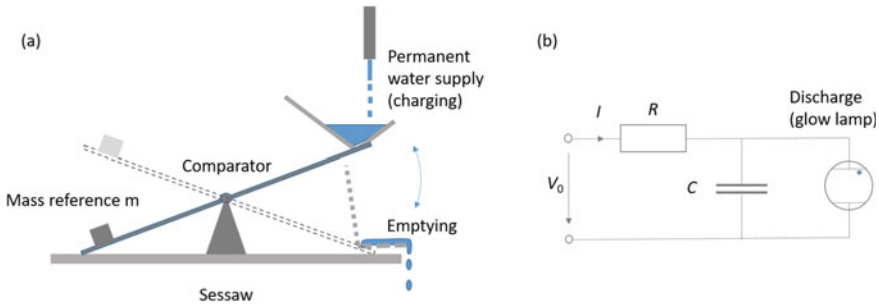


Fig. 6 A mechanical (a) and an electronic version (b) of relaxation-type oscillators

simulated by LTspice XVII (Linear Technology). Relaxation-type oscillators exhibit the leaky integrate and firing (LIF) mechanism of a neuron and therefore were applied to electronically mimic basic neuronal functions [40].

To explain the basic function of such oscillators a mechanical and an electrical version of relaxation-type oscillators are sketched in Fig. 6a and b, respectively. The mechanical model (Fig. 6a) is based on a seesaw comprising a mass on one wing (left) and a permanent charging (here water) on the right wing. In case the mass in the right container becomes larger than that on the left wing, the seesaw suddenly, seeps and the water container empties. The seesaw turns back to its original position and the procedure repeats again in a self-sustained way. The electronic version (Fig. 6b) consists of a constant voltage source V_0 , a resistor R , a capacitor C and a glow discharge lamp. The capacitor is charged by the current I via the resistor. If the threshold voltage of the glow lamp is reached, the lamp discharges the capacitor, accompanied by a short visible light spark and the entire process starts again [41]. The period of the pulse are set by the resistor, the capacitor and the threshold voltage of the lamp. The similarity to the leaky-integrate and firing (LIF) model for neurons is obvious. In this work semiconductor based relaxation-type van der Pol oscillators (programmable unijunction transistors) were applied [42]. The programmable unijunction transistor (PUT) 2N6027 functions as an electrical switch [43]. The PUT is part of a small circuitry comprising passive devices. The basic circuit is sketched in Fig. 7a. The three terminals of a PUT are the gate G, the cathode C and the anode A. With a constant supply voltage V , a current flows through $R1$. This current charges the capacitor $C1$ and the voltage at the anode of the PUT increases. If it reaches the threshold voltage of the PUT, the PUT will switch into the conducting state. Hence, the capacitor gets discharged by a current flow from anode to cathode. At the resistor $R4$ this discharge can be measured in form of a voltage spike. If the voltage from anode to cathode is sufficiently high, a current can flow. The threshold voltage depends on the voltage divider consisting of the resistors $R2$ and $R3$ and hence can be adjusted. The spiking frequency of the oscillator varies in dependence of its circuit parameters. The time constant $\tau = R_1 * C_1$ as well as the voltage divider $R2$ and $R3$ define the pulse frequency f_{osc} of the oscillator. The

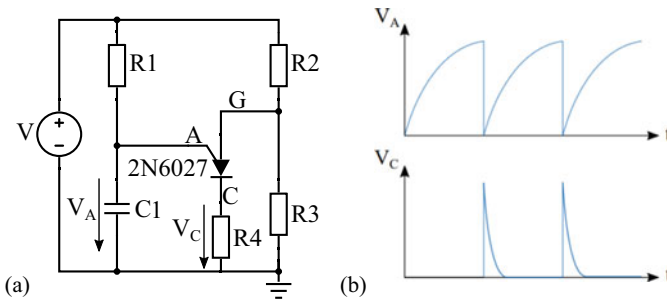


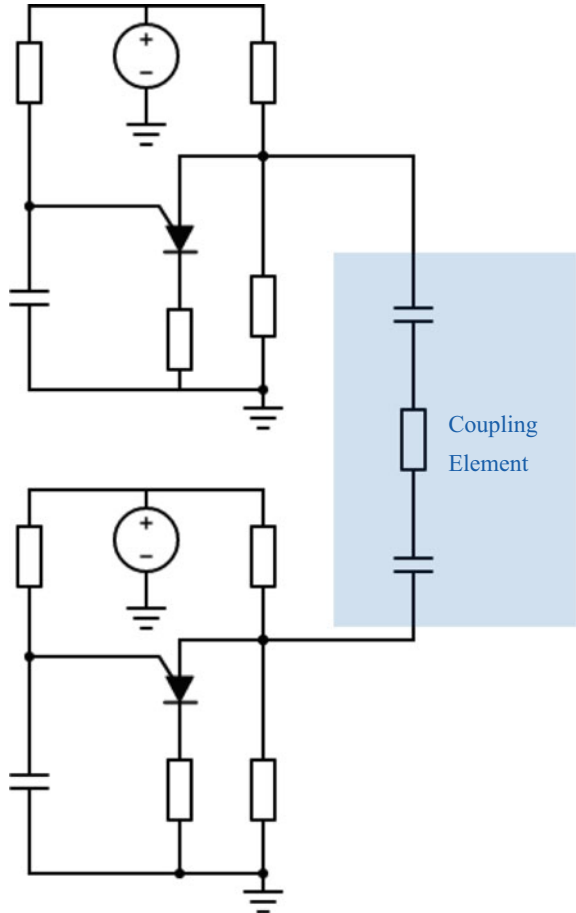
Fig. 7 Electrical circuit of a relaxation-type van der Pol oscillator (a), and the corresponding voltage curves (b). V_A shows the charging curve of the capacitor $C1$ and the instant discharge when the threshold voltage of the PUT is reached. The voltage V_C shows the discharge spikes of the capacitor at the cathode of the PUT

voltage traces at the anode terminal V_A and at the cathode terminal V_C are depicted qualitatively in Fig. 7b. Three of such PUT-based oscillators were coupled to mimic basal mechanisms of the box jellyfish's visual system.

As coupling mechanism between the oscillators, a gate to gate coupling is used (Fig. 8). The voltage at the gate is constant when the threshold voltage of the PUT is not reached. When the PUT switches, a current can flow and the voltage breaks down for a short moment in form of the reversed voltage spike at the cathode. The capacitors in the coupling are used to filter low frequencies, especially the DC part of the voltage. In the moment of a spike the capacitors are equivalent to short circuits and a current can flow. This means that the voltage divider changes its value and it is more likely for the second oscillator to fire, too. Therefore the two oscillators are pulsed-coupled [44–47]. If the frequency difference between both oscillators is small enough a synchronization will occur. Synchronization can only be achieved with oscillators that are similar enough to each other in regards of their frequencies. In this work the coupled oscillator scheme as shown in Fig. 5 has been transferred to an electronic circuitry design and in the first stage simulated correspondently by LTspice XVII [8]. The pattern recognition functions with three oscillators and persists of two stages. The visual stimuli are encoded in the voltages $V1$ and $V2$. A high voltage is equal to an ON pattern and a voltage of $0V$ equals an OFF pattern. Both voltage sources are chosen in a way that all possible combinations of two binary inputs (OFF-OFF, OFF-ON, ON-OFF, ON-ON) occur over time (Fig. 9).

For the charging resistors $R2$ and $R6$ of the two oscillators the values $750\text{ k}\Omega$ and $420\text{ k}\Omega$) are chosen in a way that no synchronization can occur without the right stimuli from $V1$ and $V2$. The 2N7002 are n-mosfets and function as voltage-controlled switches. If a positive voltage is applied by $V1$ or $V2$, the corresponding n-MOSFET switches and a current can flow through the device. The oscillator influenced by $V1$ is called the ON-oscillator. If a positive voltage is applied to the n-MOSFET, it switches and a second resistor is parallel connected to the charging resistor. The new value of

Fig. 8 Electrical circuit of two coupled relaxation-type van der Pol oscillators. The two oscillators are coupled via a RC circuit. The capacitors serve as a DC-de-coupling between the oscillators. A firing of one oscillator leads to a phase shift of the other oscillator and may lead to synchrony in case the coupling resistance is finite



the overall charging resistor can be calculated with $R_{ges} = R_1 * R_2 / R_1 + R_2$. The second resistor is chosen in a way that the overall resistance R_{ges} is almost equal to the charging resistor of the OFF-oscillator. Therefore, in the moment of a stimulus synchronization of the two oscillators becomes possible. The second oscillator is called OFF-oscillator because synchronization is only possible in absence of a stimulus from V2. When a positive voltage is applied as stimulus by V2, the n-MOSFET switches into the on state and now a resistor is parallel connected to the capacitor. This leads to a leaky-current through the resistor and less current arrives at the capacitor. Therefore, the charging process of the capacitor takes longer than before and the frequency of the oscillator drops. This drop is so high, that again no synchronization is possible between the two oscillators. The at the cathodes generated voltage spikes are the input for the second stage. The second stage consists of only one oscillator: It only spikes when the oscillators of the first stage are firing in synchrony. Two optocouplers (PC817A) are applied as voltage dependent switches and likewise to

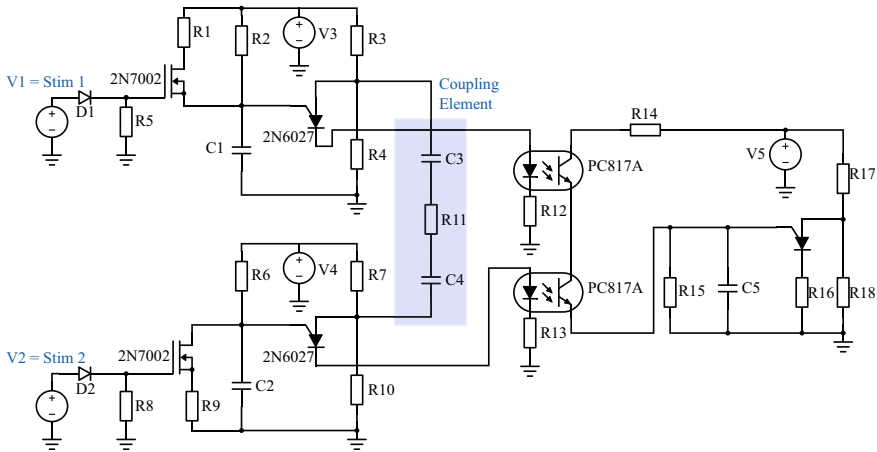


Fig. 9 Electrical circuit of the complete setup. The ON- and OFF-oscillator are on the left and receive their input from V1 and V2. With the coupling element in place and the right input pattern synchronization can occur and the synchronization oscillator on the right starts to spike. Chosen Values: $R1 = 1\text{ M}\Omega$, $R2 = 750\text{ k}\Omega$, $R3 = R7 = R17 = 100\text{ k}\Omega$, $R4 = R10 = R18 = 200\text{ k}\Omega$, $R5 = R8 = R12 = R13 = 1\text{ k}\Omega$, $R6 = 420\text{ k}\Omega$, $R9 = 1.6\text{ M}\Omega$, $R11 = 2\text{ M}\Omega$, $R14 = R16 = 10\text{ k}\Omega$, $R15 = 9\text{ M}\Omega$, $C1 = C2 = C5 = 0.1\text{ }\mu\text{F}$, $C3 = C4 = 1\text{ }\mu\text{F}$, $V3 = V4 = V5 = 10\text{ V}$

realize a potential decoupling from the input circuit. Both optocouplers are connected with a cathode of one of the PUTs. Hence, the generated voltage spikes of the oscillators are the inputs of the optocouplers. Only when both optocouplers are switched on simultaneously, a current charges the capacitor of the sync oscillator. Therefore, only when the optocouplers receive a voltage spike at the same time from their input-oscillators can the synchronization oscillator generate spikes—which is equivalent for synchronization of the input oscillators.

4 Conclusion

Biological model systems offer attractive guidelines to develop novel computing architectures, which may exhibit enhanced pattern recognition capabilities, by equally low power dissipation. In this work, we demonstrate how two input oscillators and a sync oscillator can mimic basal visual information processing in the *T. cystophora* rhopalial nervous system. The sync oscillator fires only when the two input oscillators achieve the correct input pattern and subsequently initiates the activation sequence (2 Hz spiking signal) (Fig. 10). This validates not only that pattern recognition functionality in the box jellyfish nervous system can be accomplished by sequentially connected oscillators, but also that this functionality can be reversed engineered into hardware electronics. The present work considers a very humble pattern of just two pixels but with a bit of engineering effort, hardware components

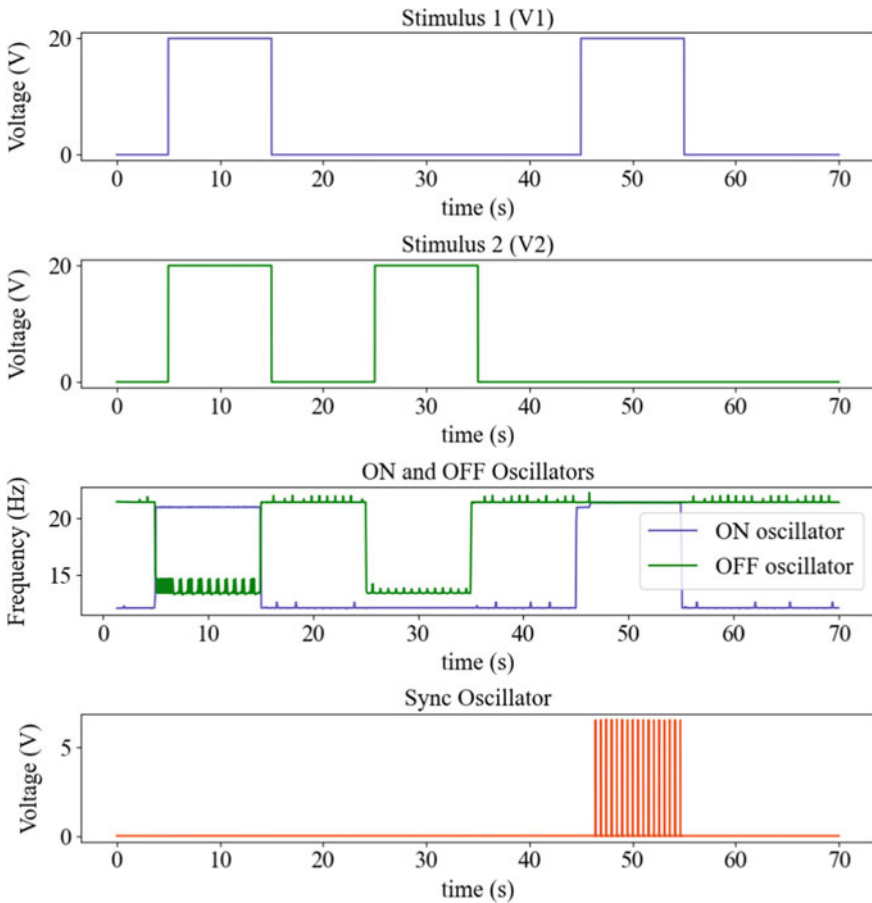


Fig. 10 The first two panels show the input variation over time (V_1 and V_2 in Fig. 9). Panel 3 shows the frequency of the ON and OFF oscillator in dependence of the input stimulus. Panel 4 shows the spiking behavior of the synchronization oscillator that only spikes when the right input pattern occurs and induces a synchronization of the ON and OFF oscillator

can be extended to support the approximately 600 pixels in the box jellyfish retinal sensor. This would enable extraction of specific visual information (such as speed limits) from a complex visual environment (Fig. 4). Pattern recognition filters for each possible speed limit inserted in parallel would then instantaneously report on the specific speed allowed, since only the sync oscillator recognizing the accurate visual cue would initiate the activation sequence. Subsequently, an autonomously driving vehicle would then be able to maintain correct speed using hardware units rather than occupying processing capacity on an on-board computer. Integrated in the *T. cystophora* model system is an override mechanism favoring sensory information of the most present danger. The special purpose lens eyes, upper and lower lens

eyes, each has pattern recognition function but for vastly different visual input. The upper lens eye is directed out of the water column and concerned about the keeping the contrast line between the mangrove canopy and the open sky in the periphery of the visual field, indicating that the animal is nested well within the habitat [15]. The lower lens eye is directing obstacle avoidance behavior and is responsible for keeping a safe distance to underwater obstacles that could damage the fragile bell of the animal [13]. Collision with obstacles present greater danger and avoidance is more time critical than to adjust the position of the animal in respect to the habitat. It is then logical that the lower lens eye has override privileges compared to the upper lens eye [25]. By adding pattern recognition filters with overriding privileges, of e.g. pedestrians or approaching vehicles, our hardware model would then have built in safety measures for clear and present danger – if the sync oscillators monitoring the override filters would initiate an activation sequence, the vehicle could be brought to an abrupt stop regardless of the input from the speed limit sync oscillators. The *T. cystophora* neuronal network integrates the bias of the sub cluster pacemaker cells into the terminal cluster pacemaker cells, but for the technical modeling we need another layer of sync oscillators to accomplish the decision making functionality (Fig. 11).

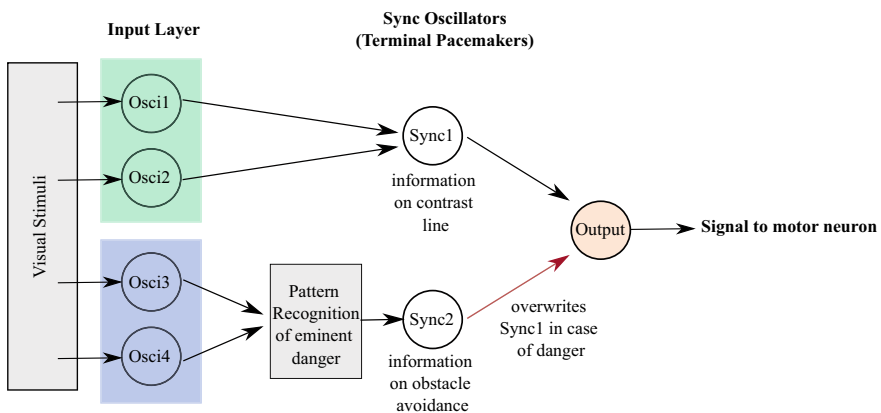


Fig. 11 The *T. cystophora* nervous system contains an intrinsic terminal pacemaker override mechanism to avoid immediate danger. For transference into electronic circuitry an extra decision making oscillator level must be added to accomplish this intrinsic override function. In this figure, “visual stimuli” represent the photoreceptor and second order neuron layers. The input layer comprises the sub-cluster pacemaker cells which respond to the appropriate visual input by high frequency firing. The sync oscillators in turn produce the 2 Hz activation sequence in response to aligned spiking from the sync oscillators. Here Sync2 has override privileges in case Sync1 and Sync2 are simultaneously activated, and an added decision making oscillatory circuit would evaluate this bias and produce the activation sequence appropriate for Sync2

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Project-ID 434434223–SFB 1461 and within the Research Unit FOR 2093.

References

1. Hales, K.G., Korey, C.A., Larracuenta, A.M., Roberts, D.M.: Genetics on the fly: a primer on the drosophila model system. *Genetics* **201**(3), 815–842 (2015)
2. Phifer-Rixey, M., Nachman, M.W.: Insights into mammalian biology from the wild house mouse *Mus musculus*. *ELife* **4**(April), e05959 (2015)
3. Meyers, Jason R.: Zebrafish: Development of a vertebrate model organism: zebrafish: development of a vertebrate model organism. In: *Current Protocols Essential Laboratory Techniques*, vol. 16, no. 1, p. e19 (2018)
4. Rankin, C.H., Beck, C.D.O., Chiba, C.M.: *Caenorhabditis elegans*: a new model system for the study of learning and memory. *Behav. Brain Res.* **37**(1), 89–92 (1990)
5. Blair Hedges, S.: The origin and evolution of model organisms. *Nat. Rev. Genet.* **3**, 838–849 (2002)
6. Pikovsky, A., Rosenblum, M., Kurths, J.: *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge Nonlinear Science Series 12. Cambridge University Press, New York (2003)
7. Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd edn. Westview Press, A Member of the Perseus Books Group, Boulder, CO (2015)
8. Brocard, G.: *The LTSpice XVII Simulator: Commands and Applications: Manual, Methods and Applications*, 1st edn. Würth Elektronik. Künzelsau: Swiridoff Verlag (2021)
9. Low, L.A., Reinhall, P.G., Storti, D.W., Goldman, E.B.: Coupled van Der Pol Oscillators as a simplified model for generation of neural patterns for jellyfish locomotion. *Struct. Control Health Monit.* **13**(1), 417–429 (2006)
10. Bielecki, J., Garm, A.: Vision made easy: cubzoans can advance our understanding of systems-level visual information processing. In: *Results and Problems in Cell Differentiation*, vol. 65, pp. 599–624 (2018)
11. Kozmik, Z., Ruzickova, J., Jonasova, K., Matsumoto, Y., Vopalensky, P., Kozmikova, I., Strnad, H., et al.: Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc. Natl. Acad. Sci.* **105**(26), 8989–8993 (2008)
12. Nilsson, D.-E., Gislén, L., Coates, M.M., Skogh, C., Garm, A.: Advanced optics in a jellyfish eye. *Nature* **435**(7039), 201–205 (2005)
13. Garm, A., O’Connor, M., Parkefelt, L., Nilsson, D.-E.: Visually guided obstacle avoidance in the box jellyfish *Tripedalia Cystophora* and *Chiropsella Bronzie*. *J. Exp. Biol.* **210**(20), 3616–3623 (2007)
14. Garm, A., Heddal, I., Islin, M., Gurska, D.: Pattern- and contrast-dependent visual response in the box jellyfish *Tripedalia Cystophora*. *J. Exp. Biol.* **216**(24), 4520–4529 (2013)
15. Garm, A., Oskarsson, M., Nilsson, D.-E.: Box jellyfish use terrestrial visual cues for navigation. *Curr. Biol.* **21**(9), 798–803 (2011)
16. Buskey, E.: Behavioral adaptations of the Cubozoan Medusa *Tripedalia Cystophora* for feeding on copepod (*Dioithona Oculata*) Swarms. *Mar. Biol.* **142**(2), 225–232 (2003)
17. Skogh, C., Garm, A., Nilsson, D.-E., Ekström, P.: Bilaterally symmetrical Rhopalial nervous system of the box jellyfish *Tripedalia Cystophora*. *J. Morphol.* **267**(12), 1391–1405 (2006)
18. Nielsen, S.K.D., Koch, T.L., Wiisbye, S.H., Grimmelikhuijzen, C.J.P., Garm, A.: Neuropeptide expression in the box jellyfish *Tripedalia Cystophora*—new insights into the complexity of a ‘Simple’ nervous system. *J. Comp. Neurol.* **529**(11), 2865–2882 (2021)

19. Matthiessen, L.: Untersuchungen über den Aplanatismus und die Periscopie der Krystallinsen in den Augen der Fische. *Pflüger, Archiv für die Gesamte Physiologie des Menschen und der Thiere* **21**(1), 287–307 (1880)
20. Garm, A., Andersson, F., Nilsson, D.-E.: Unique structure and optics of the lesser eyes of the box jellyfish *Tripedalia Cystophora*. *Vis. Res.* **48**(8), 1061–1073 (2008)
21. Garm, A., Poussart, Y., Parkefeld, L., Ekström, P., Nilsson, D.-E.: The ring nerve of the box jellyfish *Tripedalia Cystophora*. *Cell Tissue Res.* **329**(1), 147–157 (2007)
22. Satterlie, R.A.: Neuronal control of swimming in jellyfish: a comparative story. *Can. J. Zool.* **80**(10), 1654–1669 (2002)
23. Boelmann, F., Epple, M., Stje, I., Tiemann, H.: Statoliths of calcium sulfate hemihydrate are used for gravity sensing in Rhopaliophoran Medusae (Cnidaria). In: Buerlein, E. (ed.) *Handbook of Biomineralization*, pp. 261–272. Wiley-VCH Verlag GmbH, Weinheim, Germany (2007)
24. Bielecki, J., Høeg, J.T., Garm, A.: Fixational eye movements in the earliest stage of metazoan evolution. *PLoS ONE* **8**(6), e66442 (2013). (Edited by Eric James Warrant)
25. Garm, A., Mori, S.: Multiple photoreceptor systems control the swim pacemaker activity in box jellyfish. *J. Exp. Biol.* **212**(24), 3951–3960 (2009)
26. Petie, R., Garm, A., Nilsson, D.-E.: Visual control of steering in the box jellyfish *Tripedalia Cystophora*. *J. Exp. Biol.* **214**(17), 2809–2815 (2011)
27. Petie, R., Garm, A., Nilsson, D.-E.: Contrast and rate of light intensity decrease control directional swimming in the box jellyfish *Tripedalia Cystophora* (Cnidaria, Cubomedusae). *Hydrobiologia* **703**(1), 69–77 (2013)
28. Stöckl, A.L., Petie, R., Nilsson, D.-E.: Setting the pace: new insights into central pattern generator interactions in box jellyfish swimming. *PLOS ONE* **6**(11), 1–11 (2011)
29. Yatsu, N.: Notes on the physiology of *Charybdea Rastonii*. *J. Coll. Sci. Imp. Univ. Tokyo*, **40**(December), 1–12 (1917)
30. Bielecki, J., Nachman, G., Garm, A.: Swim pacemaker response to bath applied neurotransmitters in the cubozoan *Tripedalia Cystophora*. *J. Comp. Phys. A* **199**(9), 785–797 (2013)
31. Plachetzki, D.C., Degnan, B.M., Oakley, T.H.: The origins of novel protein interactions during animal opsin evolution. *PLoS ONE* **2**(10), e1054 (2007). (Edited by Berend Snel)
32. Koyanagi, M., Takano, K., Tsukamoto, H., Ohtsu, K., Tokunaga, F., Terakita, A.: Jellyfish vision starts with CAMP signaling mediated by Opsin-Gs cascade. *Proc. Natl. Acad. Sci.* **105**(40), 15576–15580 (2008)
33. Clark Gray, G., Martin, V.J., Satterlie, R.A.: Ultrastructure of the retinal synapses in cubozoans. *Biol. Bull.* **217**(1), 35–49 (2009)
34. Garm, A., Bielecki, J.: Swim pacemakers in box jellyfish are modulated by the visual input. *J. Comp. Physiol. A* **194**(7), 641 (2008)
35. Stryer, L.: Cyclic GMP cascade of vision. *Ann. Rev. Neurosci.* **9**(1), 87–119 (1986)
36. Kandel, E.R., Schwartz, J.H. (eds.): *Principles of Neural Science*. Internat, 4 edn. McGraw-Hill, New York, NY (2000)
37. Wehner, R.: Matched Filters’-neural models of the external world. *J. Comp. Physiol. A* **161**(4), 511–531 (1987)
38. Fain, G.L., Matthews, H.R., Carter Cornwall, M., Koutalos, Y.: Adaptation in vertebrate photoreceptors. *Physiol. Rev.* **81**(1), 117–151 (2001)
39. Kowler, E.: Eye movements: the past 25 years. *Vis. Res.* **51**(13), 1457–1483 (2011)
40. Lim, H., Ahn, H.-W., Kornijcuk, V., Kim, G., Seok, J.Y., Kim, I., Hwang, C.S., Jeong, D.S.: Relaxation oscillator-realized artificial electronic neurons, their responses, and noise. *Nanoscale* **8**(18), 9629–9640 (2016)
41. Pearson, S.O., St. H., Anson, G.: The neon tube as a means of producing intermittent currents. *Proc. Phys. Soc. Lond.* **34**(1), 204–212 (1921)
42. van der Pol, B.: LXXXVIII. On ‘Relaxation-Oscillations.’ *Lond. Edinburgh Dublin Philos. Mag. J. Sci.* **2**(11), 978–992 (1926)
43. Floyd, T.L.: *Electronic Devices*, 5th edn. Prentice Hall, Upper Saddle River, N.J (1999)

44. Mirollo, R.E., Strogatz, S.H.: Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math.* **50**(6), 1645–1662 (1990)
45. Bard Ermentrout, G., Bard, G., Chow, C.C.: Modeling neural oscillations. *Physiol. Behav.* **77**(4), 629–633 (2002)
46. Christensen, A.L., O’Grady, R., Dorigo, M.: From fireflies to fault-tolerant swarms of robots. *IEEE Trans. Evol. Comput.* **13**(September), 754–766 (2009)
47. Ignatov, M., Hansen, M., Ziegler, M., Kohlstedt, H.: Synchronization of two memristively coupled van Der Pol oscillators. *Appl. Phys. Lett.* **108**(8), 084105 (2016)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Biologically Inspired and Energy-Efficient Neurons



Xiaoyan Cheng, Sebastian Simmich, Finn Zahari, Tom Birkoben, Maximiliane Noll, Tobias Wolfer, Eckhard Hennig, Robert Rieger, Hermann Kohlstedt, and Andreas Bahr

Abstract Silicon neurons represent different levels of biological details and accuracies as a trade-off between complexity and power consumption. With respect to this trade-off and high similarity to neuron behaviour models, relaxation-type oscillator circuits often yield a good compromise to emulate neurons. In this chapter, two exemplified relaxation-type silicon neurons are presented that emulate neural behaviour with energy consumption under the scale of nJ/spike. The first proposed fully CMOS relaxation SiN is based on mathematical Izhikevich model and can mimic a broad range of physiologically observable spike patterns. The results of kinds of biologically plausible output patterns and coupling process of two SiNs are presented in 0.35 μm CMOS technology. The second type is a novel ultra-low-frequency hybrid CMOS-memristive SiN based on relaxation oscillators and analog memristive devices. The hybrid SiN directly emulates neuron behaviour in the range of physiological spiking frequencies (less than 100 Hz). The relaxation oscillator is

X. Cheng · A. Bahr

Sensor System Electronics, Department of Electrical and Information Engineering,
Kiel University, Kaiserstraße 2, 24143 Kiel, Germany

e-mail: xc@tf.uni-kiel.de

A. Bahr

e-mail: ab@tf.uni-kiel.de

S. Simmich · R. Rieger (✉)

Networked Electronic Systems, Department of Electrical and Information Engineering,
Kiel University, Kaiserstraße 2, 24143 Kiel, Germany

e-mail: rri@tf.uni-kiel.de

F. Zahari · T. Birkoben · M. Noll · H. Kohlstedt

Nanoelectronics, Department of Electrical and Information Engineering, Kiel University,
Kaiserstraße 2, 24143 Kiel, Germany

T. Wolfer · E. Hennig

Electronics & Drives, Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany

A. Bahr

Institute for Integrated Circuits, School of Electrical Engineering, Computer Science and
Mathematics, Hamburg University of Technology, Eissendorfer Straße 38, 21073 Hamburg,
Germany

implemented and fabricated in 0.13 μm CMOS technology. An autonomous neuronal synchronization process is demonstrated with two relaxation oscillators coupled by an analog memristive device in the measurement to emulate the synchronous behaviour between spiking neurons.

Keywords Silicon neuron · Izhikevich model · Hybrid CMOS-memristive · Integrate-and-fire · Ultra-low-frequency relaxation oscillator · Artificial synapse · Analog memristive device · Neuronal synchronization

1 Introduction

Living species are well adapted to their environments, a result of a hundred million years of evolution on earth. Due to constraints in space, time and energy, biological information processing in nervous systems of creatures are shaped during evolution towards an optimum between capabilities and resource consumption [1–3]. It provides a benchmark for technical systems particularly when it comes to elaborate brain functions such as conscious awareness or decision-making which consume an incredible small amount of energy within a limited volume (i.e. a space of a brain) [4]. Although, there are continuous remarkable progresses in very large scale integration (VLSI) technology, the gap still remains between digital processors and biological computing systems [5, 6].

Neuromorphic engineering emerges and endeavors to develop intelligent machines with comparable biological computation and energy efficiency [6–12]. More recently, bio-inspired silicon neurons (SiNs) and artificial neural networks have been vastly investigated and developed to imitate the biological computing scheme [13–15]. The nervous system is a fine-grained parallel processing and highly linked neural network [5, 6, 16, 17]. Therefore, the large superiority of biological computing systems for certain tasks like sensory processing or pattern recognition comes from its real-time analog computation in which data processing (e.g., computing and learning) and data storage (i.e., memory) are inseparably linked. However, digital processors based on von Neumann architecture execute binary computing with a strict separation of data processing and storing [6, 18, 19]. Hence, a large part of power consumption is distributed on the data transfer between processing units and memory. As a consequence, the artificial neurons capable of real-time analog computing have recently received increasing interest in silicon neuron design [20]. Furthermore, due to high similarity to neuron behaviour models, the relaxation-type oscillators draw more attentions in recent researches [21, 22].

In this chapter, two exemplified high energy efficient relaxation-type silicon neurons are presented. These two analog silicon neurons represent two promising development strategies in artificial neuron design. One strategy is fully CMOS technology spiking neurons based on mathematical neuron models (here: the Izhikevich-model) and generating a broad range of physiologically observable spike patterns. Spiking frequency, i.e. operating speed, of these artificial neurons is considerably higher than

that of biological neurons (e.g. typical mean firing frequencies of biological neurons are in the range of up to 100 Hz). Another strategy is hybrid silicon neurons, which use CMOS analog oscillators and analog memristive devices to implement neuron and synapse circuits. These hybrid SiNs directly emulate the behaviour of neurons at the biological frequency. Both SiNs introduced here allow real-time analog computation rather than binary data processing and enable the realization of large scale monolithic analog neural network in the future.

Before introducing details of two exemplified SiNs, the analog memristive device used in the hybrid SiNs is introduced below firstly. Then Izhikevich-model based low-power SiN is introduced in Sect. 2 including the theories, topology, simulation and measurement results. In Sect. 3 the hybrid SiN based on low frequency relaxation oscillator is described. In order to understand the synchronization process of memristive coupling, the design and measurement results of single oscillator is introduced at first. Subsequently, the autonomous synchronization processes of two pulse-coupled oscillators via a RC network and an analog memristive device are experimentally verified, respectively. Finally, a brief conclusion in Sect. 4 completes the chapter.

1.1 Memristive Devices

The mutual memristive coupling of two self-sustained relaxation oscillators has been successfully realized experimentally [22], through the use of an Ag-doped-TiO_{2-x}-Al memristive device with digital switching behavior. Compared with Ag-doped-TiO_{2-x}-Al memristive devices, interface-based memristive devices like the double barrier memristive devices (DBMDs) [23] with analog switching behavior (i.e., a continuous change in resistance) behave more similarly to synapses in nervous systems. Consequently, to come closer to emulate the process of synchronization in neurons, in Sect. 3.3.3 an interface-based device will be adopted as an artificial synapse in our experiments. A brief overview of DBMDs is provided in the following. Interested readers are referred to the literature for more details [23–30].

The schematic of a double barrier memristive device is shown in Fig. 1a. It consists of a Nb/NbO₂/Al₂O₃/Nb_xO_y/Au material stack [23, 28], in which Nb and Au are the bottom and top electrodes, respectively, while Al₂O₃ acts as tunnel barrier and the Nb_xO_y/Au interface forms a Schottky-like contact [23, 24, 29]. The devices are produced by DC magnetron sputtering of all materials on 100 mm wafer without breaking the vacuum and subsequent structuring by standard photolithography, lift-off and etching [23, 24]. Figure 1b shows a typical absolute value of the current density versus voltage ($|J| - V$) hysteresis curve of DBMDs [23]. A gradual resistive switching behavior is present [23] rather than abrupt resistance jumps observed by digital memristive devices [22]. In the $|J| - V$ measurement, the voltage has been ramped from 0 to 3 V to set the device from its initial high resistance state (HRS) to a low resistance state (LRS). Afterwards, the voltage has been ramped down to -2 V and back to 0 V to reset the device. Voltage was applied to the top electrode while the bottom electrode was grounded. Using low voltages (e.g., 0.5 V) allows

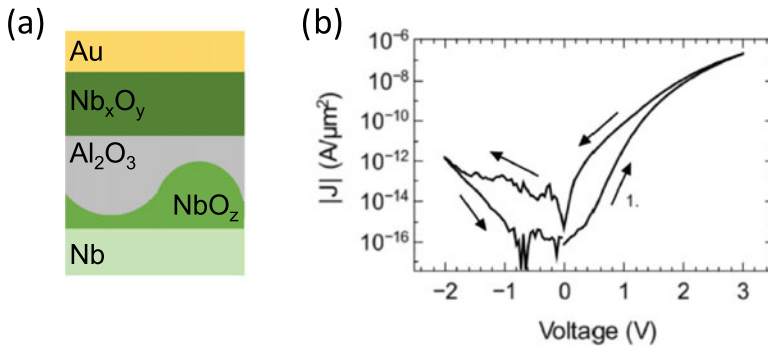


Fig. 1 A schematic structure **a** and current density versus voltage characteristics of a double barrier memristive device plotted as absolute value on a semi-logarithmic scale (**b**). **b** is reproduced from [23] (licensed under CC BY 4.0—<https://creativecommons.org/licenses/by/4.0/>)

for non-destructive read-out of the device state. The diode-like characteristic with a high J - V non-linearity obtained by the Schottky barrier facilitates integration into passive crossbar arrays [27], since the current at negative bias voltage is negligible compared to positive bias. The devices can further be gradually switched by using voltage pulses with different amplitudes and widths [26]. Switching occurs for pulses with a width in the millisecond regime or beyond. The non-linear switching process is further crucial dependent on the voltage amplitude. Moreover, the memristive state relaxes towards HRS with time [23]. These effects are considered in the coupling experiments shown in Sect. 3.3.3. Due to the diode-like characteristics, the main interaction between two oscillators happens when a positive voltage is applied to the memristive device, i.e. higher output voltage from an oscillator connected to the top electrode and lower output voltage from an oscillator connected to the bottom electrode, as it is shown in Sect. 3.3.3.

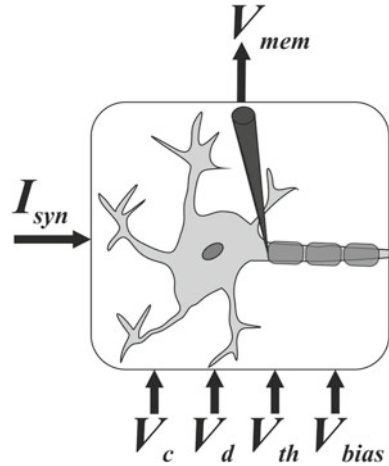
2 Izhikevich-Model Based Low-Power Neuron

In 2003, Izhikevich presented a mathematical model [31] that describes the spiking and bursting behavior of cortical neurons. This model combines the biological plausibility of the Hodgkin-Huxley type [32] and the computational efficiency of the integrate-and-fire model [33]. It has been described, in [31], as a two-dimensional system:

$$v' = 0.04v^2 + 5v + 140 - u + I \quad (1)$$

$$u' = a(bv - u) \quad (2)$$

Fig. 2 Block diagram of the Izhikevich-model based low-power neuron



where v' and u' are the derivatives of v and u with respect to time, respectively: v represents the membrane voltage; and u is a membrane recovery variable that provides a negative feedback to the membrane voltage [31]. The variable I adds incoming synaptic currents to the system. All variables in (1) and (2) and also the parameters a and b are dimensionless.

In [13], the Izhikevich model was implemented in a 0.35 μm CMOS technology, consisting of 14 transistors for a single neuron. This neuron’s output, which represents the membrane potential measured at the axon, is tunable by the five different inputs. Figure 2 shows the different inputs and one output of the neuron. The output V_{mem} stays at a resting potential, as long as no currents at the postsynaptic input I_{syn} exceed the threshold. By changing the inputs V_c , V_d and V_{th} , different biologically plausible spike patterns can be generated [13]. The neuron uses a silicon area of $70 \times 40 \mu\text{m}^2$ and has a power consumption of 8 pJ/spike.

2.1 Topology

The schematic of this neuron is shown in Fig. 3. We adapted the circuit presented in [13] for the usage in a low-power analog neural network. As shown in Fig. 3, the current input I_{syn} is directly connected to the membrane voltage output. In a neural network built without additional current output, the input current of the first neuron will couple directly to the second neuron. In the biological neuron as in Fig. 2, however, the current input I_{syn} is separated from the axon voltage V_{mem} . Therefore, a dedicated current output node delivering the current I_{axon} is added to the circuit. This output can be taken to represent an action potential propagating towards the synapse. Furthermore, we modified the biasing circuit of the original circuit. Instead

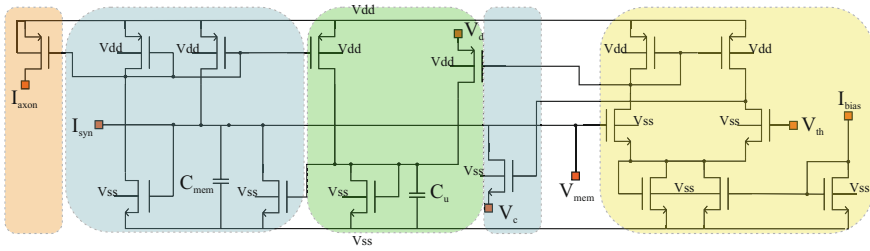


Fig. 3 Schematic of the neuron circuit, divided in four parts: the membrane voltage (blue), the slow variable (green), the comparator (yellow) and the axon output (orange)

of a voltage bias via V_{bias} , we employ current-biasing using a mirrored input current I_{bias} . This enables more sensitive bias control in an integrated circuit.

The neuron circuit can be structured into four parts: a circuit representing the cell membrane and, specifically, the membrane voltage (Fig. 3, blue), the slow variable (Fig. 3, green), a comparator (Fig. 3, yellow), and the axon output (Fig. 3, orange). The membrane circuit integrates the I_{syn} current on the capacitor C_{mem} . The comparator controls a transistor which discharges the membrane capacitance to the resting voltage between spikes. This voltage level is defined by the applied input voltage V_c . The slow variable circuit represents Eq. (2), and slows the depolarization of the membrane voltage. The input voltage V_d controls the amount of charge stored on the capacitor C_u .

A possible configuration for the neural network is presented in Fig. 4. The first layer neuron output current is connected to the synaptic inputs of the second layer. The coupling strength is set via the axon output current by programmable weighted current mirrors. Although the neurons are interconnected, the individual membrane voltages of each neuron can be monitored through the output V_{mem} . In doing so, the output of the neural network can be combined using the information of the different membrane voltages.

2.2 Simulation Results

The circuit shown in Fig. 3 is implemented in a 0.35 μm CMOS technology and simulated in the Cadence design environment. It is able to generate multiple biologically plausible spike patterns when a stimulation current is applied to the input. By changing the control parameters presented in the previous section, the output of the neuron varies.

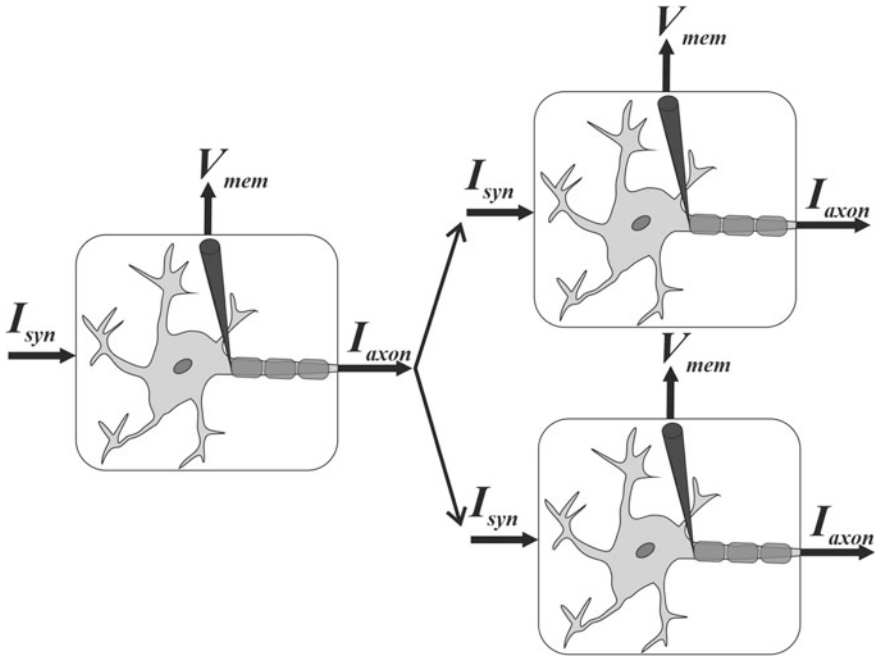


Fig. 4 Configuration of multiple neurons in a neural network

2.2.1 Individual Izhikevich-Model Based Low-Power Neuron

The circuit can either be supplied with the nominal operating voltage for this technology of 3 V or with reduced power consumption with 1.5 V. The different operating modes are capable of generating different output patterns. The simplest pattern—single spikes—is shown in Fig. 5. The input current is integrated on the membrane capacitance until the membrane voltage reaches the threshold voltage and a spike is generated. After the spike, the slow variable keeps the circuit reset for a short time. This pattern is generated with a supply voltage of $V_{DD} = 3\text{ V}$, $V_d = V_c = 0$, $V_{th} = 700\text{ mV}$, $I_{bias} = 1\text{ }\mu\text{A}$, and $I_{syn} = 200\text{ nA}$. All voltages refer to V_{ss} ($-1.5\text{ V}/-750\text{ mV}$). Even though the spike frequency is considerably higher than that of a biological neuron, the spike shape is realistic.

With $V_c = 100\text{ mV}$ and $I_{syn} = 1\text{ }\mu\text{A}$, fast spikes with a reset offset of 100 mV (shown in Fig. 6a) are generated. The larger the value of I_{syn} , the faster C_{mem} is charged; thus, the membrane voltage reaches the threshold voltage of the comparator faster. The state trajectory presented in Fig. 6b shows that the circuit reaches a limit cycle. This happens very quickly compared to the circuit in [13]. Additionally, there is only very small jitter in this state.

In the low-power configuration the supply voltage is reduced to 1.5 V, I_{bias} is set to 20 nA, and I_{syn} is set between 100 nA and 1 μA . For a fair comparison of

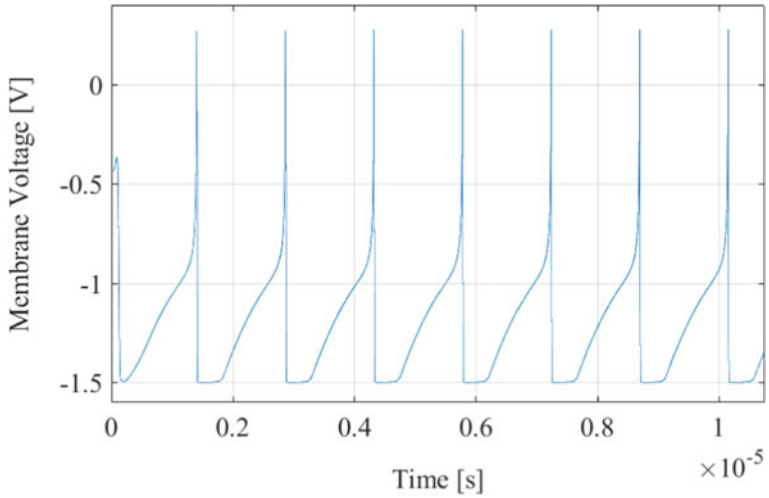
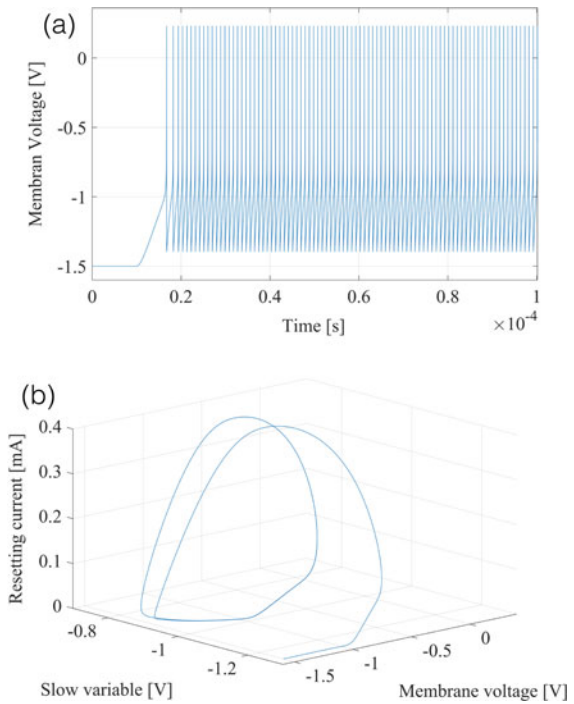


Fig. 5 Simulated transient waveform of the circuit's spiking membrane voltage

Fig. 6 a Transient waveform of Membrane Spiking Voltage. b State trajectory of (a)



different circuit implementations and spike patterns, the energy consumption per spike is frequently reported, referring to the spike as the smallest computational unit. Thus, the total current consumed by the circuit is integrated over the time of a single spike event and multiplied by the supply voltage. Under low-power operation, an energy consumption of 8–10 pJ per spike is simulated, which is in good agreement with what was achieved in [13].

2.2.2 Coupled Izhikevich-Model Based Low-Power Neuron

Two neurons are connected as shown in Fig. 4 to form a simple network. The current output I_{axon} of neuron 1 connects to the current input I_{syn} of neuron 2. The neurons are both operated in their low-power configurations, and the variables are set to $V_c = 130$ mV, $V_d = 0$ V, $V_{th} = 540$ mV, $I_{syn1} = 300$ nA, and $I_{syn2} = 200$ nA, where I_{syn1} is applied to neuron 1 and I_{syn2} to neuron 2. Thereby, both neurons generate the same spike pattern with different spike frequencies. The state of each neuron can be seen by assessing the different V_{mem} outputs. The membrane voltages of neurons 1 and 2 are shown in Fig. 7a. After about 30 μ s, both neurons become locked and oscillate in synchrony. Figure 7b shows the change in amplitude, while Fig. 7c shows the change in the frequencies during the settling process. As expected, the initially faster neuron 1 remains unchanged, while neuron 2 changes its output to the same frequency as neuron 1. Thereby, the frequency first rises close to that of neuron 1, but then drops to 830 Hz before beginning to rise again. Additionally, the amplitude changes during the settling process. While the frequency of neuron 2 is slow—between 15 and 20 μ s—its amplitude is high, as more energy can be stored in C_{mem} . After both neurons settle to the same frequency, the amplitude of neuron 1 remains larger than that of neuron 2, as I_{syn} differs between the neurons.

2.3 Measurement Results

The circuit was implemented and fabricated in 0.35 μ m CMOS technology. Figure 8a shows a microphotograph of the die, where the rectangular shape indicates the location of the circuit. Figure 8b presents the related layout of this neuron, with an area of $33.4 \times 55.9 \mu\text{m}^2$

For all measurements, the neuron was operated in low-power mode (supply voltage 1.5 V). All variable inputs were generated externally, using a laptop-controlled multifunctional input/output card. The output waveforms measured during a constant I_{syn} input are presented in Fig. 9a–d. Figure 9a shows simple slow spikes ($V_c = V_d = 0$ V, $V_{th} = 0.9$ V, $I_{bias} = 20$ nA, $I_{syn} = 200$ nA). The frequency was about 900 Hz and, therefore, close to biological spike frequencies. Additionally, the spike shape was in good agreement with that of biological spikes. By changing the input voltage to $V_c = 100$ mV and setting $V_{th} = 0.96$ V, the neuron generated fast spikes, as presented in Fig. 9b. The spikes were fired with a frequency of 30 kHz. Due to the

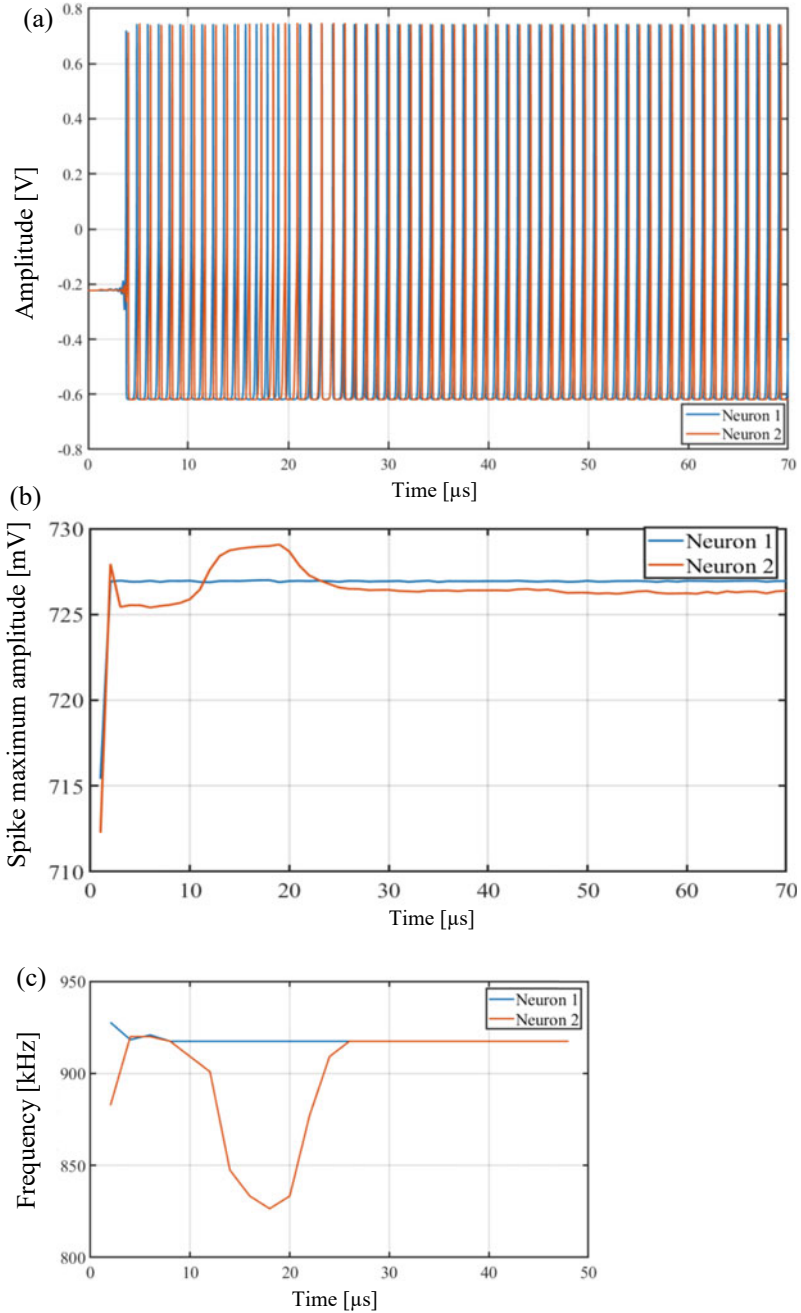


Fig. 7 **a** V_{mem} of the coupled neurons during the settling process. **b** Change in amplitude during the settling process. **c** Change in frequency during the settling process

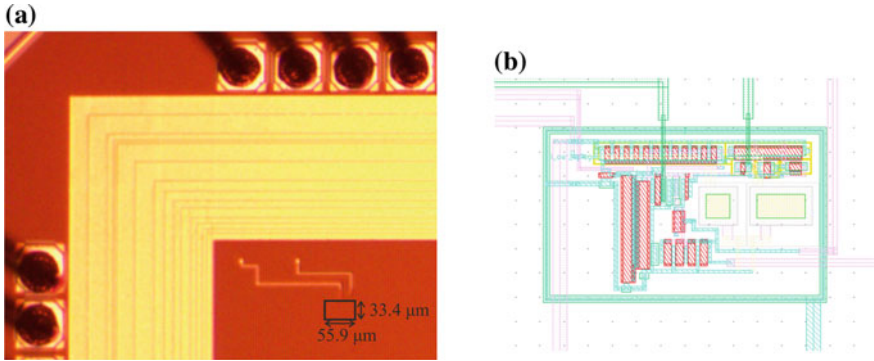


Fig. 8 **a** Photograph of the fabricated die with marked dimensions of $33.4 \times 55.9 \mu\text{m}^2$. **b** Layout of the presented neuron

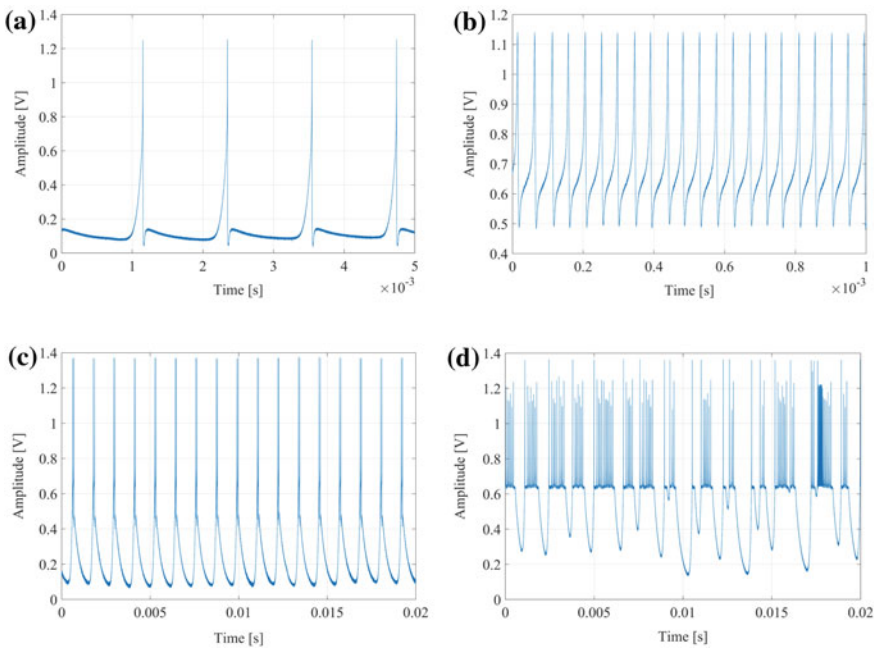


Fig. 9 Measurement of different output pattern during a constant current input: **a** Slow simple spikes with a frequency of 900 Hz. **b** Fast simple spikes with a frequency of 30 kHz. **c** Periodical bursting behavior. **d** Aperiodic bursting behavior

high frequency, the neuron was not able to fully discharge to V_c . Figure 9c shows periodic bursting ($V_c = 0.1 \text{ V}$, $V_d = 0.1 \text{ V}$, $V_{th} = 0.85 \text{ V}$, and $I_{syn} = 1 \mu\text{A}$). The neuron generated three very fast spikes and then discharged to V_c periodically. The opposite -aperiodic bursting- is presented in Fig. 9d. The setup was the same as that for periodic bursting, except that V_d was set to V_{dd} .

3 Ultra-Low-Frequency Hybrid CMOS-Memristive Silicon Neuron

Synchronization and memory of spiking neurons are vastly accepted and explained as underlying mechanisms of neuronal signal processing, such as recognition, perception and awareness, in brains of living creatures [17, 34–40]. In neuromorphic engineering, to emulate neuron activities, a variety of ways to transfer these biological mechanisms to electronic circuits are realized [10, 41–43]. Researches in recent years demonstrate that the neuromorphic system built by hybrid CMOS-memristive silicon neurons consisting of relaxation oscillators and memristive devices is a promising candidate for neuromorphic computing, since they allow one to emulate neuronal synchronization and synaptic functionalities in a detailed way with energy efficiency and a high packing density [11, 19, 22, 44–49]. One example is given in Fig. 10. It shows an experiment based on two discrete self-sustained relaxation oscillators coupled with an Ag-doped- TiO_{2-x} -Al memristive device to emulate basal coupling and an autonomous synchronization scheme for neuronal ensembles [22]. More specifically, two relaxation oscillators with intrinsic oscillation frequencies of 543 Hz and 414 Hz, respectively, are pulse-coupled through a resistor–capacitor network comprising a single memristive device. A synchronization process is observed as anticipated, i.e. the left ‘slow’ oscillator eventually follows the right ‘fast’ oscillator.

The experiment successfully demonstrates that, with the memristive pulse-coupling of two relaxation oscillators, two relevant dynamic aspects of biology (memory and synchronization) can be transferred to electronic circuits in a small-scale compact system. In the future, the construction of a large scale pulse-coupled memristive oscillator network is expected to facilitate the emulation of higher cognitive functions and perceptual processes. Very large-scale integrated (VLSI) circuits

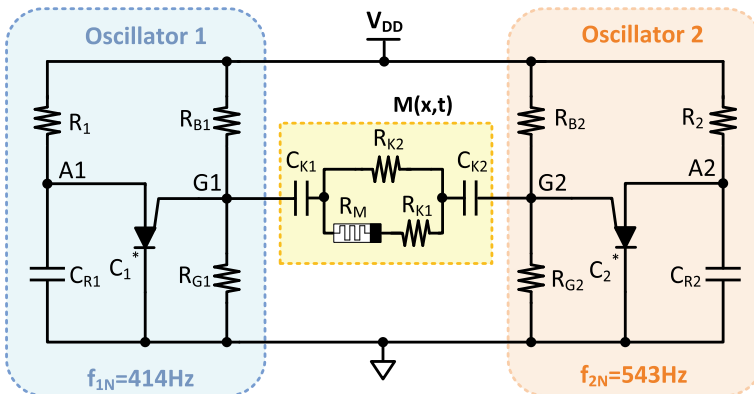


Fig. 10 Block diagram of discrete setup of two mutually coupled self-sustained relaxation oscillators with intrinsic frequencies of $f_{1N} = 414$ Hz and $f_{2N} = 543$ Hz, respectively. $M(x,t)$ (middle) is a memristive resistor network, where x is the state variable of the memory process

have been shown as a feasible solution owing to their variety of advantages: highly complex neural network connections, highly efficient parallel computing, and real-time processing [7, 12, 19, 20, 50]. However, to directly transfer the design in the experiment given in Fig. 10 to neuromorphic VLSI network, the structure of the oscillator from the experiment has two constraints: at first, the oscillator is based on a programmable unijunction transistor (PUT; 2N2067), discrete resistors and capacitors. It is necessary to find a new circuit structure to replace PUT by using basic elements (i.e. CMOS transistors) in VLSI; In addition, the frequency of oscillation and static power consumption of the oscillator from the above experiment are limited by discrete resistors and capacitors. Especially for the biological frequency applications—at which the frequency should be less than one hundred hertz and below—it is infeasible to realize ultra-large resistors or capacitors in the custom on-chip design to satisfy the ultra-low frequency requirements. As a result, a new integrated relaxation-type oscillator is highly demanded.

In this section, we present a study of a hybrid CMOS-memristive silicon neuron consisting of a monolithic ultra low frequency relaxation oscillator which is designed and realized in CMOS technology and the analog memristive device introduced before. The circuit design, simulation and measurement results of the proposed relaxation oscillator are introduced at first. Then, the principle of pulse coupled oscillators is explained and examined by experiments of resistive coupling system. At the end, an autonomous synchronization process is demonstrated by experiments of memristive coupling system. This hybrid SiN is strongly biologically oriented and paves the way for large neuromorphic VLSI system.

3.1 Ultra-Low-Frequency Relaxation Oscillator

This section introduces the design, the oscillation mechanism, the simulation and measurement results of the relaxation-type oscillator which is fabricated with 130 nm IHP technology. The supply voltage 3.3 V is adopted to meet the state transition requirements of the analog memristive devices used to realize a coupling between on-chip oscillators. To make oscillator self sustained, negative differential resistance (NDR) circuit is adopted here.

3.1.1 Circuit Design and Oscillation Mechanism

The architecture of the proposed integrated self-sustained relaxation oscillator [9] is shown in Fig. 11. To overcome the obstacle of integration of ultra-large resistors or capacitors induced from the structure of oscillator in previous experiment, it adopts a pA-scale charging current to realize ms-scale charging time and biological oscillation frequency. The output current of a MOS-only current reference I_{REF} (1 nA) [51] is mirrored through two pairs of NMOS current mirrors (N3 and N4, N3 and N5) and two pairs of PMOS current mirrors (P1 and P3, P2 and P4), in order to supply

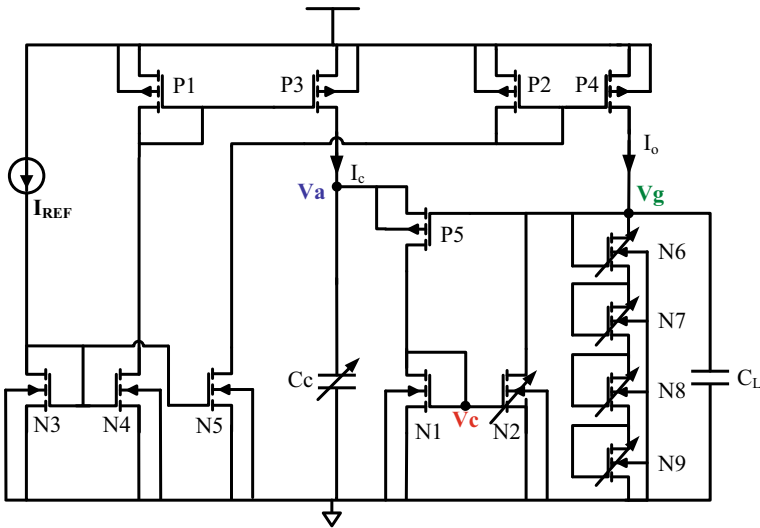


Fig. 11 Schematic of proposed oscillator, which is designed to represent the discrete oscillator in Fig. 10. The output voltage port V_g is an interface to outside

the charging current I_c (250 pA) and output stage current I_o (3 nA), respectively. To improve the accuracies of the current mirrors, cascode current mirror structures are adopted (but not shown here). Furthermore, the output stage current I_o flows into configurable diode-connected stacked NMOS transistors (N6, N7, N8, and N9) to generate the output voltage V_g , which adjusts the threshold voltage V_{thG} of negative differential resistance (NDR) circuit [52]. It avoids the difficult tradeoff of size of resistors or power consumption in the structure designed by the experiment shown in Fig. 10. The NDR circuit consists of PMOS transistor P5, a pair of NMOS transistors N1 and N2, and the configurable diode-connected stacked NMOS transistors.

The oscillation frequency is tunable by applying a digital configuration to the programmable capacitor C_c . It consists of binary weighted capacitance with the size of 1, 2, 4, 8 and 16 C_0 . By considering chip area and reducing parasitic effects, the unit capacitance C_0 is designed to be 0.94 pF with the size of $25 \mu\text{m} \times 25 \mu\text{m}$. This enables a programming range of C_c from 0.94 pF to 30 pF. In addition, the tunable transistor N2 can adjust the NDR region. Another configurable module is the diode-connected stacked NMOS transistors which control the output voltage V_g in the range of 2.6–2.8 V in order to ensure the state transition of the memristive device.

A timing diagram for the voltages V_a , V_g , and V_c under the conditions of $f= 19.7$ Hz and $C_c= 4.7$ pF is illustrated in Fig. 12. When the circuit is powered on, the potential of V_a and V_g is zero and P5 is turned off. The constant charging current I_c from P3 begins to flow into the programmable capacitor C_c and potential V_a rises linearly to V_{aH} . At the same time, the output stage current I_o from P4 flows into the diode-connected stacked NMOS transistors and generates output voltage V_g . To

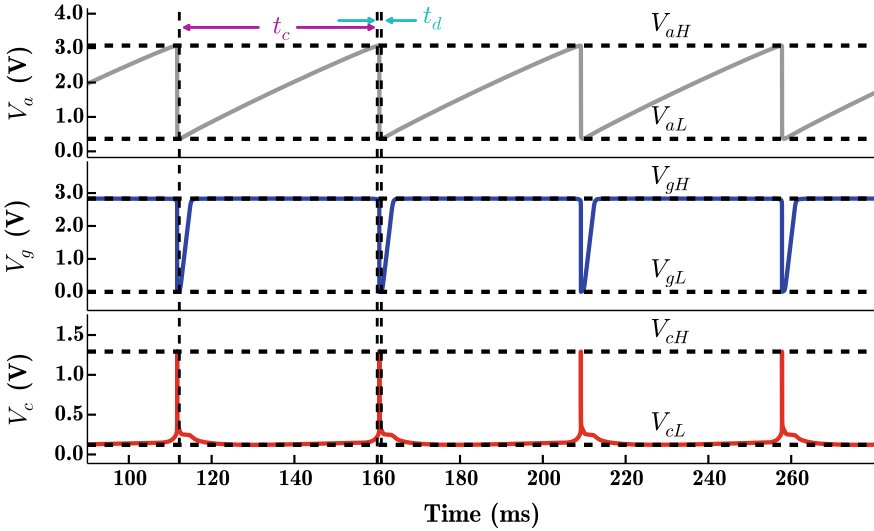


Fig. 12 Timing diagram of oscillator output voltage V_g and internal voltage potential V_a and V_c ; t_c indicates the time required for charging C_c , while t_d is its discharge time ($C_c = 4.7$ pF, $f = 19.7$ Hz)

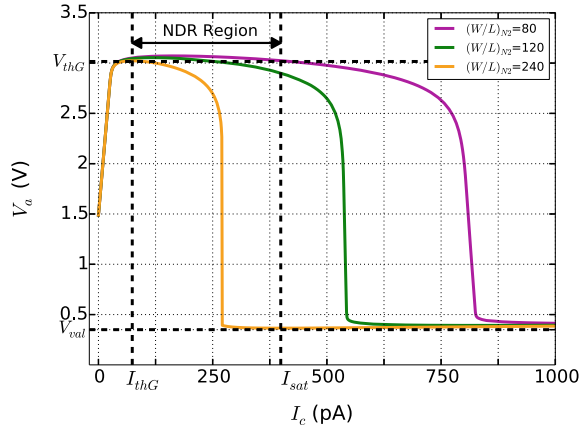
make sure P5 keeps turned off before V_a arrives at V_{aH} , output voltage V_g should reach V_{gH} earlier than V_a reaches V_{aH} . The time of V_g from 0 to V_{gH} depends on the output stage current I_o and C_L which is parasitic capacitance at V_g port. Considering parasitic capacitance from I/O Pad and additional test point on PCB, to make sure V_g arriving V_{gH} earlier, I_o (3nA) is chosen 12 times bigger than I_c (250pA) here. When the voltage V_a reaches the turn-on voltage V_{thG} (i.e., V_{aH} in Fig. 12) of the NDR, the transistor P5 turns on, C_c begins discharging and I_c current flows into N1. It generates potential V_c and turns on N2. In the design, the size of N2 is large enough to swallow all the current from I_o when its V_{gs} equals V_c with small drain-source voltage V_{ds} (i.e. V_g). The drain-source voltage of N2 is then immediately pulled down (i.e., output voltage V_g is pulled down to be V_{gL}). This speeds up the discharging of C_c and generates spike signal V_c .

The capacitor C_c discharges until the voltage V_a reaches V_{aL} and the gate voltage V_c of N1 and N2 also decreases, which reduces the drain current of N2. I_o flows back to the diode-connected stacked NMOS transistors. Thus, V_g becomes high and switches off P5. Then, the next charging phase starts. From simulated waveforms in Fig. 12, we can see that V_a presents a saw-tooth-type signal, V_g is a rectangular pulse-type signal, and V_c is a neuron spiking-type signal.

3.1.2 Negative Differential Resistance Regime

The V_a-I_c DC characteristics of the proposed oscillator, with different sizes of N2, are shown in Fig. 13. The width of the NDR region changes with the size of N2; for

Fig. 13 V_a - I_c DC characteristic for different values of N2, showing the negative differential resistance area. For increasing size (W/L) of N2, the NDR area decreases



example, with an aspect ratio of 240 for N2, when V_a is less than V_{thG} , P5 is cut off. When $V_a = V_{thG}$, then P5 turns on, and the oscillator enters an NDR region in which the voltage V_a decreases as I_c increases.

The negative-resistance region extends until the valley point is reached, where the current is defined by I_{sat} . Beyond the valley point, further increases in V_a produce increases in I_c . This region is the so-called saturation region, which should be avoided in the circuit designed for oscillation-based applications. An appropriate size for N2 is chosen, in order to make sure that the charging current I_c should be greater than I_{thG} , thus guaranteeing the turning on of the oscillator and also not exceeding I_{sat} . For this purpose, to make sure I_c (i.e. 250pA) is in the NDR region, the W/L of N2 should be less than 240. In the design, the W/L of N2 is decided 220 and a 4-bit transistor bank is added to the oscillator to calibrate the size of N2, considering the deviation of the process. The unit size of the transistor bank is $W/L = 1\mu\text{m}/1\mu\text{m}$.

3.2 Simulation and Measurement Results of Relaxation Oscillator

As shown in Fig. 14, the total area of chip layout is 0.05546 mm^2 , including the current reference, the programmable capacitor bank, and the core of the oscillator. The simulated static power consumption of a single oscillator is 24 nW excluding the power consumption of the current reference, as the current reference will be shared with other oscillators in future neuromorphic networks. The power consumption of the current reference is 255 μW . Figure 15 summarizes the performance of the proposed oscillator for different configurations, based on post-layout simulations and measurement results. The energy per spike is defined as the power consumption of a single oscillator core (active power) integrated over the period of spiking. Post-layout simulations indicate that it varies from 0.8 nJ/spike to 7.12 nJ/spike, and the

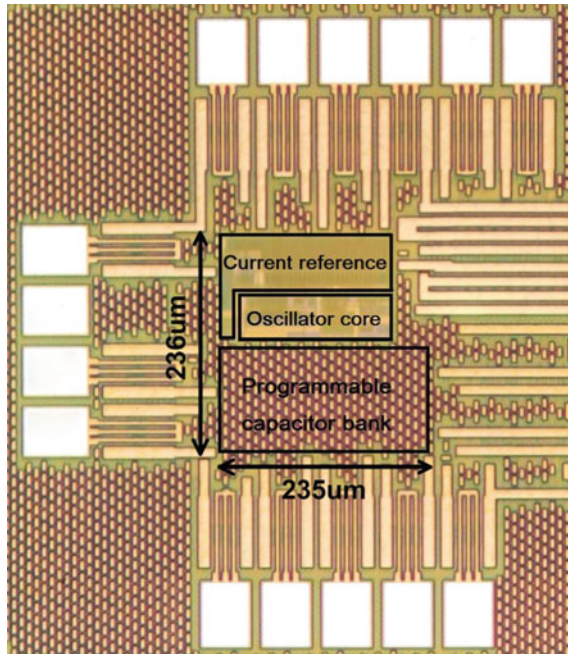


Fig. 14 The photograph of die with core area of 0.05546 mm^2 . The three main parts of this work, i.e. current reference (top), oscillator core (middle) and programmable capacitor bank (bottom) are shown in it

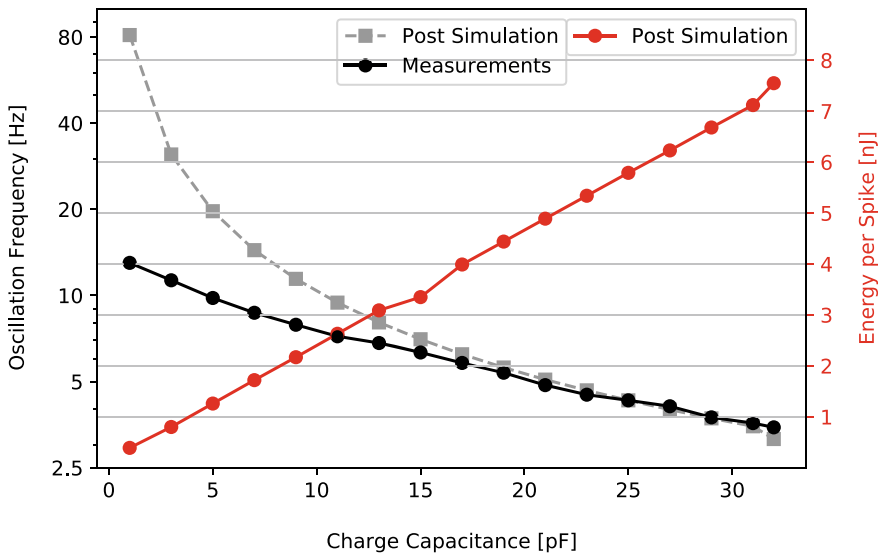


Fig. 15 Post-layout simulation and measurements of oscillation frequency and calculated energy per spike versus programmable capacitance

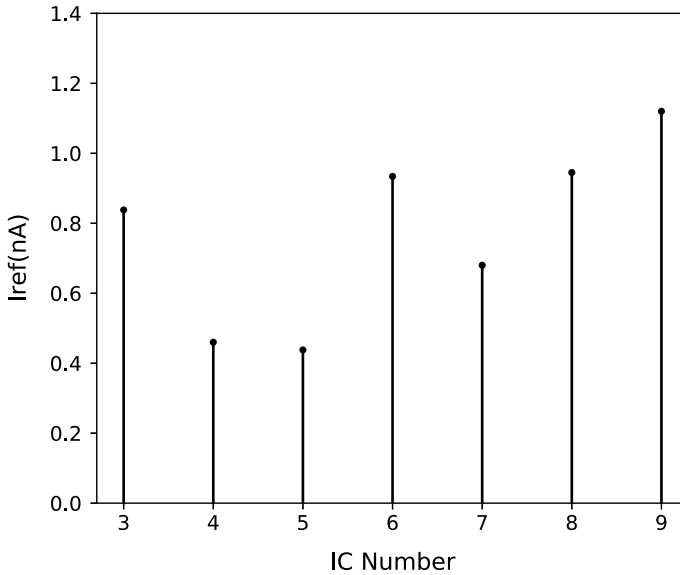


Fig. 16 Measured current versus IC number

oscillation frequency is tunable from 3.15 to 81.30 Hz by programming the on-chip capacitors. A total of 10 chips were bonded and tested. The output of the current reference was tested, as shown in Fig. 16. Seven chips were within the acceptable limits for low-frequency neuromorphic applications. Compared with the simulation results, the oscillation frequency of chip 9 in the test results was programmable from 3.58 to 13.01 Hz. The oscillation frequency was relatively lower when the charging capacitor was small (from 1 to 15 C_0), due to the parasitic capacitor at the input of oscillators from PCB and chip package being comparably large, with respect to the small charging capacitor C_c . The parasitic effects can be reduced in the integrated on-chip design for neuromorphic network applications.

3.3 Experiments of Coupling Systems

In a previous work the coupling of two relaxation-type oscillators built by discrete Programmable Unijunction Transistors (PUTs) and by a digital type of memristive device was investigated by Ignatov et al. In this section two types of coupling experiments between both on-chip oscillators were discussed. The first experiment uses a RC network as intermedia, which aims to check the functionality of synchronization of the two oscillators. In the second experiment, the two oscillators are coupled by an analog memristive device. Due to its memristive characteristic, the test allows us to further observe the autonomous transition process from unsynchronization to

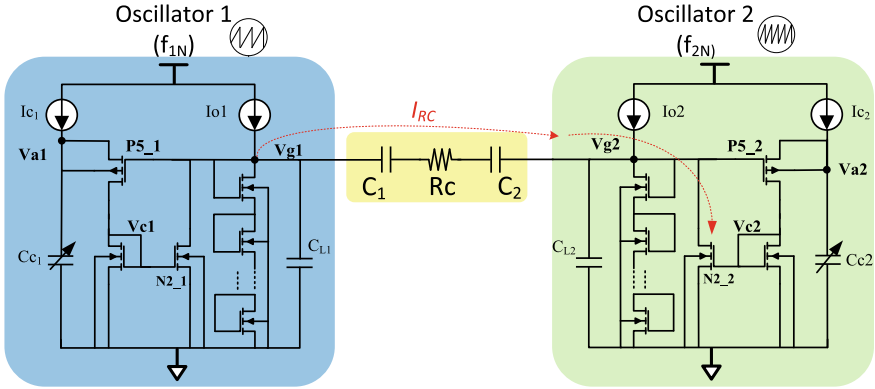


Fig. 17 Coupling test with RC network

synchronization. This transition phase mimics the synchronization process of neuron signals [45, 46].

3.3.1 Coupling System with RC Network

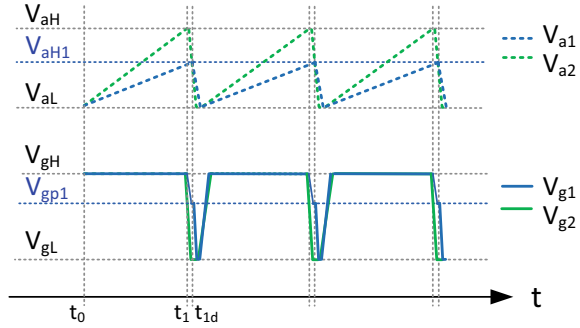
The coupling system of two oscillators via a RC network is shown in Fig. 17. Two capacitors $C_1 = C_2$ within the coupling network form the DC potential decoupling between both oscillators. The coupling network which consists of R_C series connected with C_1 and C_2 is a passive high-pass filter with a cut-off frequency f_c determined by the following equation [22]:

$$f_c = \frac{1}{\pi R_c C_1} \tag{3}$$

In the uncoupled state (i.e. without RC network), the intrinsic frequency f_{1N} of oscillator 1 differs from the intrinsic frequency f_{2N} of oscillator 2. Hereby, $f_{1N} < f_{2N}$ and the frequency difference is $\Delta f_N = f_{2N} - f_{1N}$.

After the coupling system with RC network is powered on and at t_0 , the gate voltage V_{g1} and gate voltage V_{g2} are charged to be V_{gH} (the blue and green curve in the lower graph in Fig. 18). At this moment, there is no current flowing through RC network due to $V_{g1} = V_{g2}$. The voltage V_{a1} from oscillator 1 (green dotted curve in the upper graph in Fig. 18) and V_{a2} from oscillator 2 (blue dotted curve in the upper graph in Fig. 18) rise linearly towards V_{aH} as the capacitors C_{c1} and C_{c2} are charged with constant charging currents I_{c1} and I_{c2} , respectively, as shown in Fig. 18. Induced by the difference in the charging capacitor selected for C_{c1} and C_{c2} , V_{a2} reaches the turn-on voltage V_{aH} ahead of V_{a1} , as $C_{c2} < C_{c1}$, under the condition of the same charging current (i.e., $I_{c1} = I_{c2} = I_c$). At the same time, V_{g2} falls at the moment t_1 ahead of V_{g1} . Therefore, there is a low-resistance signal path existing from V_{g2}

Fig. 18 Waveform of coupled process in synchronous phase



to ground and the decreasing of V_{g2} cause voltage difference between V_{g1} and V_{g2} . There is current flowing from V_{g1} to V_{g2} through RC network. When V_{g2} falls down to be V_{gL} ($V_{gL} \approx 0$) at t_{1d} , so V_{g1} is

$$V_{gp1} = R_c \cdot I_{RC} \tag{4}$$

where, I_{RC} is the current flowing through the coupled resistor at the moment of t_{1d} . I_c comes mainly from the I_{O1} and the discharging current from parasitic capacitance at the node of V_{g1} . To ensure that oscillator 1 follows the faster oscillator 2 (i.e., that the two oscillators synchronize), V_{aH1} (the value of V_{a1} at the moment of t_{1d}) should be a threshold voltage V_{on} bigger than V_{gp1} in order to turn on pmos transistor $P5_1$ in oscillator 1 and lead C_{c1} into the self discharging state. Then, V_{g1} is pulled down to be V_{gL} due to the turn-on of $P5_1$ oscillator 1. After that, the two oscillators independently go into self-sustained charging state again, and the synchronization process is repeated. As a result, the two self-sustained oscillators synchronize with the same frequency of f_{2N} .

Therefore, R_c needed for synchronization of RC coupled two oscillators can be calculated as

$$V_{gp1} = R_c \cdot I_{RC} \leq (V_{aH1} - V_{on}) \tag{5}$$

Therefore,

$$R_c \leq \frac{(V_{aH1} - V_{on})}{I_{RC}}, \tag{6}$$

where V_{aH1} is determined by

$$I_c \cdot t_{1d} = V_{aH1} \cdot C_{c1}$$

hence,

$$V_{aH1} = \frac{I_c \cdot t_{1d}}{C_{c1}} \approx \frac{I_c}{f_{2N} \cdot C_{c1}} \tag{7}$$

where, $t_{1d} \approx T_2$. From $I_c \cdot T_{1N} = I_c/f_{1N} = V_{aH} \cdot C_{c1}$, where, T_{1N} is the intrinsic oscillation period of oscillator 1, we can have

$$C_{c1} = \frac{I_c}{f_{1N} \cdot V_{aH}}. \quad (8)$$

Substituting Eq. 8 into Eq. 7, we obtain

$$V_{aH1} = V_{aH} \cdot f_{1N}/f_{2N}. \quad (9)$$

Therefore,

$$R_c \leq \frac{(V_{aH} \cdot \frac{f_{1N}}{f_{2N}} - V_{on})}{I_{RC}} = \frac{[V_{aH} \cdot (1 - \frac{\Delta f_N}{f_{2N}}) - V_{on}]}{I_{RC}} \quad (10)$$

From Eq. 10, it can be easily concluded that, when the frequency ratio of the two oscillators f_{1N}/f_{2N} gets smaller (i.e., $\Delta f_N/f_{2N}$ is larger), R_c should be smaller, such that the two oscillators can synchronize. Therefore, we can obtain the following conclusions: The coupling resistance is proportional to the frequency ratio and inversely proportional to the frequency difference under the same value of f_{2N} .

3.3.2 Measurement Results of Resistive Coupling

To verify the above conclusions, five experiments have been executed and the corresponding results are summarized in Table 1. When the frequency ratio (f_{1N}/f_{2N}) varies from 0.207 to 0.889, relative frequency difference ($\Delta f_N/f_{2N}$) identically decreasing from 0.793 to 0.111, the maximal coupling resistance R_c required for synchronization increases from 564 k Ω to 251 M Ω . The variations of R_c matches our expectation, since Eq. 10 is derived by ignoring parasitic factors.

3.3.3 Coupling System with Analog Memristive Device

In this section, an analog memristive device was chosen as coupling element in the coupling system. Here, devices showing similar electrical characteristics as the DBMDs presented above are used. These devices incorporate HfO₂ instead of Nb_xO_y. While a typical $|J|$ - V curve is shown below, a detailed analysis of the device performance will be published elsewhere. The circuit demonstrates an autonomous phase-locking and frequency synchronization process due to resistance changes in an analog memristive device.

To test the performance of the memristive coupling system, two oscillators were assembled with an analog memristive device in a single test PCB as shown in Fig. 19a. The structure of test PCB is depicted in Fig. 19b. Oscillator 1 with lower oscillation frequency f_{1N} and oscillator 2 with higher oscillation frequency f_{2N} are connected to the top and bottom electrode (BE) of the memristive device, respectively, through

Table 1 Measurement results of maximal coupled resistor for synchronization

Oscillators	Charging capacitor(C_0)	Intrinsic frequency of oscillators(Hz)	Freq. Ratio (f_{1N}/f_{2N})	Freq. Diff. ($\Delta f_N/f_{2N}$)	Coupling resistor ($M\Omega$)
f1	32	2.8	0.207	0.793	0.564
f2	2	13.5			
f1	16	4.5	0.333	0.667	0.599
f2	2	13.5			
f1	9	6.4	0.475	0.526	0.701
f2	2	13.5			
f1	32	2.8	0.622	0.378	11
f2	29	4.5			
f1	20	4.0	0.889	0.111	251
f2	29	4.5			

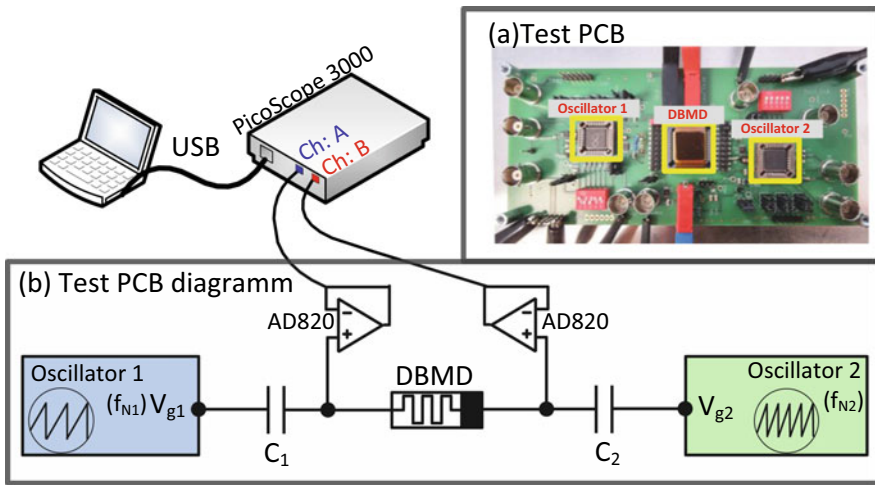


Fig. 19 Test system for the memristive coupling

series connected DC decoupled capacitor C_1 and C_2 . The gate voltages (V_{g1} and V_{g2}) of both oscillators were recorded using a PicoScope 3000 Series mixed-signal oscilloscope after buffered AD820 amplifiers with high input impedance low input bias current to reduce the loss of signal current.

The layout of a $5 \times 5 \text{ mm}^2$ chip containing analog memristive device is shown in Fig. 20a [48]. It contains 25 sub-cells arranged in a 5x5 array. Except for two sub-cells (Tcells, gray boxes) containing test structures, the other 23 sub-cells can be used in the experiments. Six devices with area sizes increasing from $100 \mu\text{m}^2$ to $2500 \mu\text{m}^2$ [numbered with “1” to “6” in Fig. 20b] are located in a single sub-cell [23, 48]. To connect the memristive devices with the two oscillators, the chip was

glued onto a JLCC 44 chip carrier shown in Fig. 20c and some devices with an area of $400 \mu\text{m}^2$ (size 3) are wire-bonded for experiments. A typical $|J| - V$ hysteresis of the HfO_2 -based devices is given in Fig. 20d. The JLCC 44 chip carrier is then placed into a PLCC 44 socket on the test PCB to connect two oscillators, as shown in Fig. 19 for experiments.

3.3.4 Measurement Results of Memristive Coupling

As discussed above, the bigger the frequency ratio (f_{1N}/f_{2N}) is, the larger coupling resistance is needed for synchronization. Hence, the two oscillators can immediately synchronize without a small coupling resistance, when the frequency ratio (f_{1N}/f_{2N}) is big. The initial resistance value of the memristive device is small enough for their mutual coupling and synchronization. However, for small frequency ratios (f_{1N}/f_{2N}), the two oscillators require a small coupling resistance for synchronization. When the initial resistance value is not small enough, as long as the resistance of the memristive device decreases gradually to R_c defined by the Eq. 10, the two oscillators can synchronize.

A representative synchronization process of two oscillators coupled with a memristive device is shown in Fig. 21. A device with area size of $400 \mu\text{m}^2$ numbered with “3” in Fig. 20b was used in this experiment. By considering the switch dynamic and retention characteristics of the devices, the intrinsic frequencies of the two oscillators 1 and 2 were $f_{1N} = 2.24$ Hz and $f_{2N} = 5.90$ Hz, respectively. Initially, due to the high resistance of the memristive device, the two self-sustained oscillators can not synchronize (i.e. they are in the desynchronous state (DS)). In each discharging period of oscillator 2, the V_{g1} was pulled down to be $V_{gp1} = R_M \cdot I_M$ when V_{g2} reaches V_{gL} (about 0 V), where R_M was the resistance of memristive device, I_M was the current flowing through the memristive device during each discharging from oscillator 2. As a consequence, every discharge of the faster oscillator (in this case, oscillator 2) would trigger a discharge of V_{g1} to be V_{gp1} . During this period, voltage V_{gp1} exerted on the top electrode of the memristive device and the bottom electrode of the memristive was V_{gL} . The voltage difference happened at each discharge period of the faster oscillator and gradually changed the resistance value of the memristive device. As long as the resistance of coupled memristive device decreased less than R_c calculated by Eq. 10 and then $V_{gp1} \leq V_{aH1} - V_{on}$, oscillator 1 can follow the rhythm of oscillator 2. In this experiment, after about 125 s, the two oscillators synchronized due to the state transition of the memristive device from a high to low resistance state.

In upper graph of Fig. 21, the transient frequency variations of oscillator 1 and 2 are given. In the desynchronous state phase (DS phase), the frequency of oscillator 1 was equal to its intrinsic frequency ($f_1 = f_{1N} = 2.24$ Hz which is calculated by 0.3 V threshold voltage). For the first 66 s, the oscillator network remained desynchronous until the intermediate phase (I phase) was reached. The intermediate state is characterized by the fact that the frequency of oscillator 1 jumps between the frequency f_{2N} and f_{1N} at irregular time intervals. The intermediate phase ended till synchronous

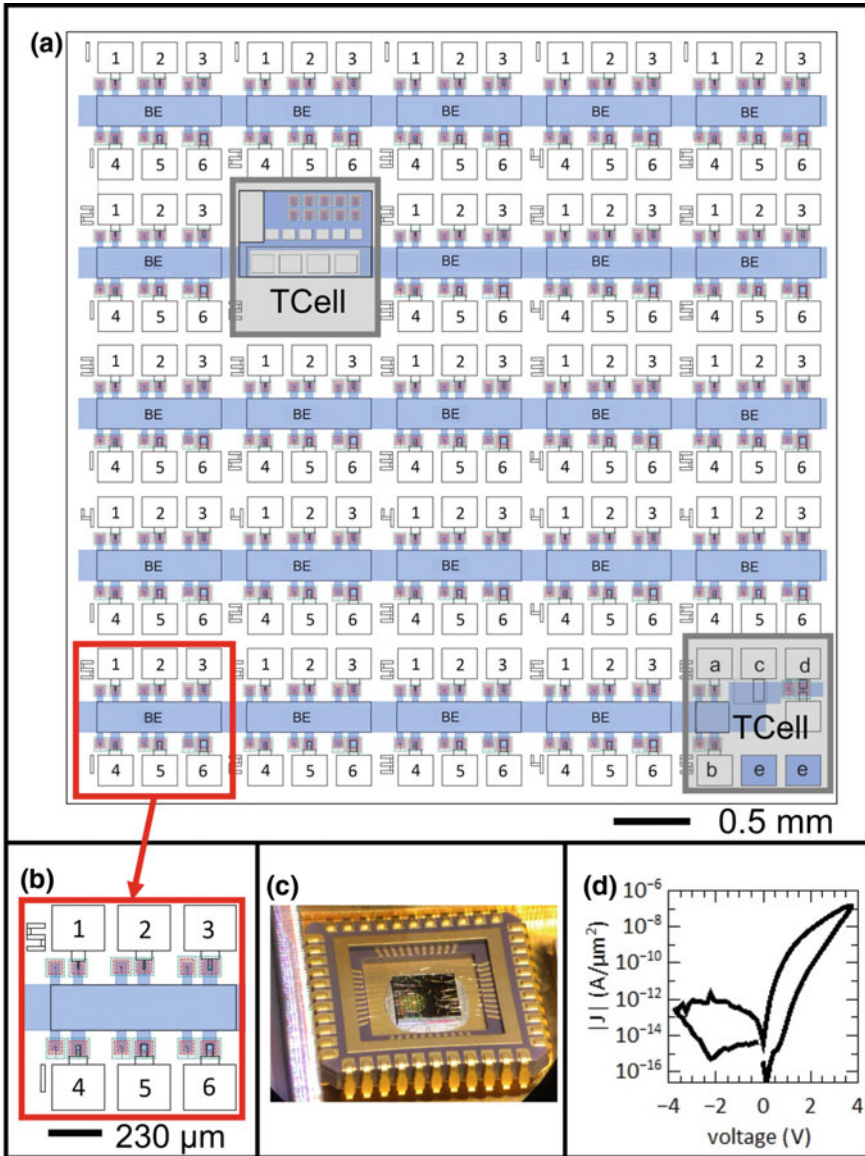


Fig. 20 **a** The layout of a 5×5 chip containing analog memristive device. **b** Layout of a sub-cell. **c** Chip glued on JLCC 44 chip carrier and wire-bonded. **d** Typical $|J| - V$ hysteresis of the HfO₂-based devices

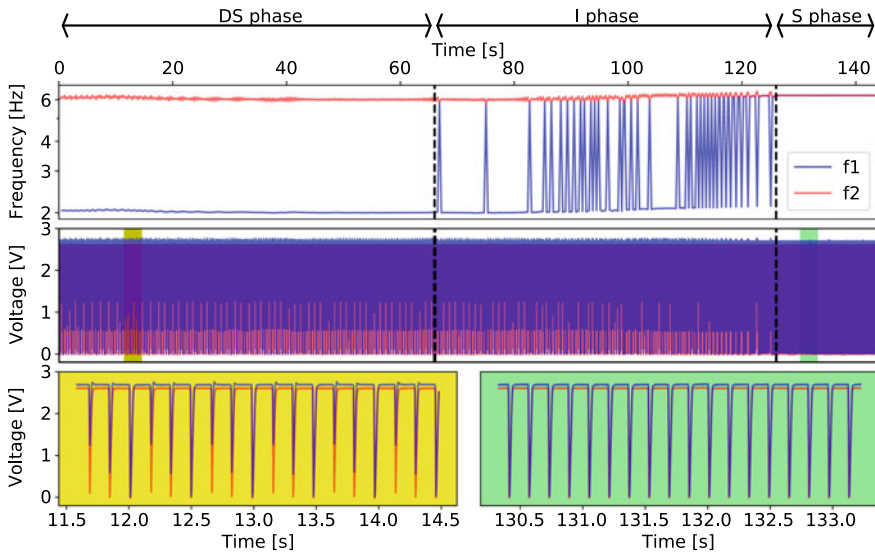


Fig. 21 Synchronization process of memristive coupling: Upper graph: f_1 and f_2 were the transient frequency variations of both self-sustained oscillators. The continuous resistance change of the memristive device caused a desynchronous phase (DS phase), an intermediate phase (I phase), and a synchronous phase (S phase). Middle graph: Transient waveform of V_{g1} and V_{g2} signal. Lower graph: Zoomed in transient waveforms of V_{g1} and V_{g2} signal in desynchronous phase (in yellow background) and synchronous state (in green background), respectively, corresponding time slots in the whole synchronization process are marked by the yellow bar and the green bar in the middle graph

phase (S phase) was reached at 125 s when phase and frequency of both oscillators synchronized. In the synchronous state phase, each V_{g2} pulse of oscillator 2 triggered oscillator 1 to discharging fully and generated a full gate pulse (i.e. amplitude of V_{g1} is from V_{gH} to ground). Oscillator 2 oscillated at its intrinsic frequency f_{2N} over the entire period of time.

4 Conclusion

In this work, two types of biologically plausible silicon neurons have been realised in integrated circuit technology. The first low-power spiking silicon neuron is inspired by mathematical Izhikevich model and built by relaxation oscillator implemented in 0.35 μm CMOS technology. It achieves an energy consumption of 8~10 pJ per spike under low-power operation mode and synchronous processes of two coupled neurons is demonstrated in the simulation results. The measurement results show that it can mimic a broad range of physiologically observable spike patterns. This area and energy efficient fully CMOS silicon neuron could be used as a universal

neuron circuit integrated in large scale analogue VLSI systems. The second ultra-low-frequency hybrid SiN achieves the biological spiking frequencies below 100 Hz with energy consumption in the range of 0.8–7.12 nJ/spike. To emulate a basic neural network, two integrated ultra-low-frequency relaxation oscillators coupled by an external analog memristive device were proposed. The autonomous neuronal synchronization processes of this basic neural network are presented and analysed. The realisation of biologically plausible oscillators in CMOS technology and an analog memristive device fabricated on silicon wafer pave the way towards large scale hybrid analog integrated neural network realisation in future.

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Project-ID 434434223–SFB 1461 and within the Research Unit FOR 2093.

References

1. Darwin, C.: On the origin of species: a facsimile of the first edition. Harvard University Press (1964)
2. Brodal, P.: The Central Nervous System: Structure and Function. Oxford University Press (2004)
3. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009)
4. Poon, C.-S., Zhou, K.: Neuromorphic silicon neurons and large-scale neural networks. *Front. Neurosci.* **5** (2011)
5. Chicca, E., Indiveri, G.: A recipe for creating ideal hybrid memristive-CMOS neuromorphic processing systems. *Appl. Phys. Lett.* **116**, 120501 (2020). <https://doi.org/10.1063/1.5142089>
6. Beyond von Neumann. *Nat. Nanotechnol.* **15**, 507 (2020). <https://doi.org/10.1038/s41565-020-0738-x>
7. Mead, C.: Analog VLSI and Neural Systems. Addison-Wesley, Reading, MA (1989)
8. Mead, C.: Neuromorphic electronic systems. *Proc. IEEE* **78**(10), 1629–1636 (1990). <https://doi.org/10.1109/5.58356>
9. Cheng, X., Birkoben, T., Kohlstedt, H., Bahr, A.: A CMOS integrated low-power, ultra-low-frequency relaxation oscillator for neuromorphic applications. In: 2021 IEEE 64th International Midwest Symposium on Circuits and Systems (MWSCAS), 978-1-6654-2461-5/21 (2021)
10. Torrejon, J., Riou, M., Araujo, F.A., Tsunegi, S., Khalsa, G., Querlioz, D., Bortolotti, P., Cros, V., Yakushiji, K., Fukushima, A., Kubota, H., Yuasa, S., Stiles, M.D., Grollier, J.: Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017)
11. Mahadevaiah, M.K., Perez, E., Wenger, Ch., Grossi, A., Zambelli, C., Olivo, P., Zahari, F., Kohlstedt, H., Ziegler, M.: Reliability of CMOS Integrated Memristive HfO₂ Arrays with Respect to Neuromorphic. In: Computing IEEE International Reliability Physics Symposium (IRPS) (2019)
12. Chicca, E., Stefanini, F., Bartolozzi, C., Indiveri, G.: Neuromorphic electronic circuits for building autonomous cognitive systems. *Proc. IEEE* **102**, 1367–1388 (2014)
13. Wijekoon, J.H.B., Dudek, P.: Compact silicon neuron circuit with spiking and bursting behaviour. *Neural Netw.* **21**(2), 524–534 (2008). (The official journal of the International Neural Network Society)
14. Oh, S., Shi, Y., del Valle, J. et al.: Energy-efficient Mott activation neuron for full-hardware implementation of neural networks. *Nat. Nanotechnol.* **16**, 680–687 (2021). <https://doi.org/10.1038/s41565-021-00874-8>

15. Jiao, X., Akhlaghi, V., Jiang, Y., Gupta, R. K.: Energy-efficient neural networks using approximate computation reuse. In: 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 1223–1228 (2018). <https://doi.org/10.23919/DATE.2018.8342202>
16. Pfeiffe, M., Pfeil, T.: Deep learning with spiking neurons: opportunities and challenges. *Front. Neurosci.* **12** (2018)
17. Von der Malsburg, C.: The Correlation Theory of Brain Function. *Models of Neural Networks*, pp. 95–119. Springer New York (1994)
18. Kasabov, N., Sengupta, N., Scott, N.: From von Neumann, John Atanasoff and ABC to Neuromorphic computation and the NeuCube spatio-temporal data machine. In: 2016 IEEE 8th International Conference on Intelligent Systems (IS), pp. 15–21 (2016). <https://doi.org/10.1109/IS.2016.7737434>
19. Ziegler, M.: Memristive Devices for Neuronal Systems. Postdoctoral thesis. Christian-Albrechts-Universität zu Kiel, Faculty of Engineering (2015)
20. Douglas, R., Mahowald, M., Mead, C.: Neuromorphic analogue VLSI. *Ann. Rev. Neurosci.* **18**, 255–281 (1995)
21. Gerstner, W., Kistler, W.M.: *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press (2002)
22. Ignatov, M., Hansen, M., Ziegler, M., Kohlstedt, H.: Synchronization of two memristively coupled van der pol oscillators. *Appl. Phys. Lett.* **108**(8), 084–105 (2016)
23. Hansen, M., Ziegler, M., Kolberg, L., Soni, R., Dirkmann, S., Mussenbrock, T., Kohlstedt, H.: A double barrier memristive device. *Sci. Rep.* (2015)
24. Hansen, M., Ziegler, M., Kohlstedt, H.: Double barrier memristive devices for neuromorphic computing. In: IEEE International Conference on Rebooting Computing (ICRC) (2016)
25. Solan, E., Dirkmann, S., Hansen, M., Schroeder, D., Kohlstedt, H., Ziegler, M., Mussenbrock, T., Ochs, K.: An enhanced lumped element electrical model of a double barrier memristive device. *J. Phys. D Appl. Phys.* **50** (2017)
26. Hansen, M., Zahari, F., Ziegler, M., Kohlstedt, H.: Double-barrier memristive devices for unsupervised learning and pattern recognition. *Front. Neurosci.* **11** (2017)
27. Hansen, M., Zahari, F., Kohlstedt, H., Ziegler, M.: Unsupervised Hebbian learning experimentally realized with analogue memristive crossbar arrays. *Sci. Rep.* **8**, 8914 (2017)
28. Strobel, J., Hansen, M., Dirkmann, S., Neelisetty, K.K., Ziegler, M., Haberfehlner, G., Popescu, R., Kothleitner, G., Chakravadhanula, V.S.K., Kübel, C., Kohlstedt, H., Mussenbrock, T., Kienle, L.: In depth nano spectroscopic analysis on homogeneously switching double barrier memristive devices. *J. Appl. Phys.* **121**(24), 245307 (2017)
29. Dirkmann, S., Hansen, M., Ziegler, M., Kohlstedt, H., Mussenbrock, T.: The role of ion transport phenomena in memristive double barrier devices. *Sci. Rep.* **6** srep35686 (2016)
30. Zahari, F., Schlichting, F., Strobe, J., Dirkmann, S., Cipo, J., Gauter, S., Trieschmann, J., Marquardt, R., Haberfehlner, G., Kothleitner, G., Kienle, L., Mussenbrock, T., Ziegler, M., Kersten, H. and Kohlstedt, H.: Correlation between sputter deposition parameters and I-V characteristics in double-barrier memristive devices. *J. Vac. Sci. Technol. B* **37**, 061203 (2019)
31. Izhikevich, E.M.: Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**(6), 1569–1572 (2003)
32. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952)
33. Knight, B.W.: Dynamics of encoding in a population of neurons. *The J. General Physiol.* **59**, 734–766 (1972)
34. Engel, A.K., Fries, P., Singer, W.: Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001)
35. Gerstner, W., Kistler, W. M., Naud, R., Paninski, L.: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press (2014)
36. Buzsáki, G.: *Rhythms of the Brain*. Oxford University Press (2006)
37. Singer, W.: Neuronal synchrony: a versatile code for the definition of relations. *Neuron* **24**, 49–65 (1999)

38. Singer, W., Gray, C.M.: Visual feature integration and the temporal correlation hypothesis. *Ann. Rev. Neurosci.* **18**, 555–586 (1995)
39. Von der Malsburg, C.: The what and why of binding. *Neuron* **24**, 95–104 (1999)
40. Fries, P.: A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**(10), 474–480 (2005)
41. Nikonov, D.E., Csaba, G., Porod, W., Shibata, T., Voils, D., Hammerstrom, D., Young, I.A., Bourianoff, G.I.: Coupled-oscillator associative memory array operation for pattern recognition. *IEEE J. Explor. Solid-State Comput. Dev. Circ.* **1**, 85–93 (2015)
42. Raychowdhury, A., Parihar, A., Smith, G.H., Narayanan, V., Csaba, G., Jerry, M., Porod, W., Datta, S.: Computing with networks of oscillatory dynamical systems. *Proc. IEEE* **107**(1) (2019)
43. Csaba, G., Porod, W.: Coupled oscillators for computing: a review and perspective. *Appl. Phys. Rev.* (2020)
44. Zahari, F., Hansen, M., Mussenbrock, T., Ziegler, M., Kohlstedt, H.: Pattern recognition with TiOx-based memristive devices. *AIMS Mater. Sci.* (2015)
45. Ignatov, M.: Emulation of Neural Dynamics in Neuromorphic Circuits Based on Memristive Devices. Doctoral thesis, Christian-Albrechts-Universität zu Kiel, Faculty of Engineering (2017)
46. Ignatov, M., Ziegler, M., Hansen, M., Kohlstedt, H.: Memristive stochastic plasticity enables mimicking of neural synchrony: memristive circuit emulates an optical illusion. *Sci. Adv.* **3**, e1700849 (2017)
47. Hansen, M., Zahari, F., Kohlstedt, H., Ziegler, M.: Unsupervised Hebbian learning experimentally realized with analogue memristive crossbar arrays. *Sci. Rep.* (2018)
48. Hansen, M.: On the Development of Memristive Devices for Electroforming-free and Analog Memristive Crossbar Arrays. Doctoral thesis, Christian-Albrechts-Universität zu Kiel, Faculty of Engineering (2017)
49. Wenger, C., Zahari, F., Mahadevaiah, M. K., Perez, E., Beckers, I., Kohlstedt, H., Ziegler, M.: Inherent stochastic learning in CMOS integrated HfO2 arrays for neuromorphic computing. *IEEE Electron Device Lett.* (2019)
50. Liu, S., Kramer, J., Indiveri, G., Delbrück T., Douglas, R.: *Analog VLSI*. MIT Press (2002)
51. Sayed, G.M., Mendoza-Ponce, P., Krautschneider, W., Kuhl, M.: Ultra-low-power self-biased 1 nA current reference circuit for medical monitoring devices in 350 nm and 180 nm cmos technology. In: *ANALOG 2018; 16th GMM/ITG-Symposium*, pp. 1–4 (2018)
52. Gonzalez, G.: *Foundations of Oscillator Circuit Design*. Artech House (2007)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Synchronization Phenomena in Oscillator Networks: From Kuramoto and Chua to Chemical Oscillators



Petro Feketa, Alexander Schaum, and Thomas Meurer

Abstract This chapter addresses the problems of synchronization analysis in various types of oscillator networks. In particular, we derive sufficient conditions for emergence of multi-cluster formations in Kuramoto networks with dynamic coupling, prove the output-feedback synchronization of chaotic behavior in networks of Chua oscillators with nonlinear static coupling, and study the synchronization of complex spatiotemporal patterns in coupled infinite-dimensional reaction-diffusion models of chemical oscillators. The obtained results contribute towards a deeper understanding of the internal organization of oscillator networks, explain the prerequisites for the emergence of patterns of synchrony and justify their stability properties in terms of the dynamical characteristics of oscillators, parameters of couplings, and the interconnection topology of the network. The interplay of these three ingredients is required for the complex and dynamically rich behavior of the network. Theoretical results of the chapter are supplemented with numerical case studies.

Keywords Oscillator network · Synchronization · Multi-clustering · Adaptive coupling · Spatiotemporal patterns

P. Feketa (✉)

School of Mathematics and Statistics, Victoria University of Wellington, PO Box 600, 6140 Wellington, New Zealand
e-mail: petro.feketa@vuw.ac.nz

A. Schaum

Chair of Automation and Control, Kiel University, Kaiserstraße 2, 24143 Kiel, Germany
e-mail: alsc@tf.uni-kiel.de

T. Meurer

Digital Process Engineering Group, Institute of Mechanical Process Engineering and Mechanics, Karlsruhe Institute of Technology (KIT), Hertzstr. 16, 76187 Karlsruhe, Germany
e-mail: thomas.meurer@kit.edu

© The Author(s) 2024

M. Ziegler et al. (eds.), *Bio-Inspired Information Pathways*, Springer Series on Bio- and Neurosystems 16, https://doi.org/10.1007/978-3-031-36705-2_16

385

1 Introduction

Oscillator networks proved to be natural mathematical models for a variety of complex processes in various areas ranging from biology and social sciences to physics and modern technology [36]. Due to a broad spectrum of application domains, oscillator networks as mathematical objects constitute a wide range of dynamical systems: from paradigmatic phase oscillators to multi-dimensional neuro-inspired models, chaotic systems, and infinite-dimensional chemical oscillators. This together with the possibility of complex dynamical interactions between oscillators and complex interconnection topologies make the proper analysis of such systems challenging and complicates the understanding of the emergence of collective behavior therein. One of the fundamental types of oscillator networks behavior is synchronization that is known to be essential for many purposes. For example, the synchronization of oscillatory signals between different brain regions supports the interaction between working and long-term memory [16]. The phenomena of partial synchronization and multi-clustering are of practical importance for neurophysiological systems [35] and distributed power generation [3, 6].

Motivated by the above discussion, this chapter focuses on three essentially different types of oscillator networks and provides their synchronization analysis. Thus, sufficient conditions for the emergence of multi-cluster formations in Kuramoto networks with dynamic coupling are provided in Sect. 2. This is made by proposing an alternative characterization of multi-cluster behavior in terms of the existence of invariant manifolds for the corresponding error-system [14]. These manifolds are of a special topological structure and their dimensions coincide with the number of clusters in the network. The conditions for the existence and stability of the invariant manifolds interrelate the natural frequencies of oscillators, plasticity characteristics, and the interconnection topology of the network. The proofs are based on the perturbation theory of invariant tori of dynamical systems [37, 38].

Section 3 addresses the problem of the output-feedback synchronization of chaotic behavior generated by the ensembles of Chua oscillators and synchronization of spatiotemporal patterns in the master-slave configuration of two Gray–Scott reaction-diffusion models. The latter one is a prototypic model of complex isothermal autocatalytic reactions that is governed by the pair of coupled partial differential equations. Synchronization analysis of this infinite-dimensional model is carried out employing late-lumping observer design techniques [12] thus relating the synchronization property with the convergence of the observer error dynamics. It is worth noting that every considered setup requires a distinct method for its formal analysis. The usage of the presented theorems is demonstrated in the numerical case studies. A short conclusion in Sect. 4 completes the chapter.

Notation

The following notation will be used throughout the chapter. Let \mathbb{N} , \mathbb{R} , $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, and \mathbb{C} denote the sets of natural, real, positive real, non-negative real, and complex numbers, respectively. For given $n, m \in \mathbb{N}$ let \mathbb{R}^n and \mathcal{T}_m denote the n -dimensional

Euclidean space and m -dimensional torus, respectively. The one-dimensional torus \mathcal{T}_1 is the one-sphere (circle \mathcal{S}_1). Let $f : \mathcal{T}_m \rightarrow \mathbb{R}^n$ be a function of the variable $\varphi = (\varphi_1, \dots, \varphi_m)^\top \in \mathcal{T}_m$ which is continuous and 2π -periodic with respect to each $\varphi_s, s = 1, m$. Finally, $C(\mathcal{T}_m)$ denotes the space of all such functions f equipped with the norm $\|f\|_0 = \max_{\varphi \in \mathcal{T}_m} \|f(\varphi)\|$, where $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^n , i.e., $\|f(\varphi)\|^2 = \sum_{i=1}^n |f_i(\varphi)|^2$, $|f_i(\varphi)|$ stands for the absolute value of the i th component of f evaluated at φ . By $C^1(\mathcal{T}_m)$ we denote the subspace of $C(\mathcal{T}_m)$ with every $f \in C^1(\mathcal{T}_m)$ having a continuous partial derivative with respect to each $\varphi_s, s = 1, m$ and $\|f\|_1 = \max\{\|f\|_0, |\frac{\partial f}{\partial \varphi_1}|_0, \dots, |\frac{\partial f}{\partial \varphi_m}|_0\}$. For a given set $\mathcal{V}, |\mathcal{V}|$ denotes the number of elements in \mathcal{V} . $\text{Re } \lambda(A)$ denotes the set of real parts of all eigenvalues of square matrix A and any set $B < 0$ if and only if for any $b \in B$ it holds that $b < 0$.

2 Multi-clustering in Networks of Phase Oscillators with Dynamic Coupling

In this section, the phenomenon of multi-clustering in networks of phase oscillators with dynamic coupling is studied. To this end, a general nonlinear model is introduced and the multi-cluster behavior of the network is then characterized in terms of the existence and stability of the corresponding invariant toroidal manifolds. These manifolds have a particular topological structure and their dimension coincides with the number of clusters in the network. Finally, the proposed approach is applied to a Kuramoto network with adaptive coupling and sufficient conditions are derived for the emergence of multi-cluster behavior. These conditions interrelate the dynamic properties of oscillators, the plasticity parameters of the adaptive couplings, and the interconnection topology of the network. A numerical example concludes this section.

2.1 General Nonlinear Model

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the directed graph representing the network of nonlinear oscillators, where $\mathcal{V} = \{1, \dots, N\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represent the oscillators and their interconnection edges, respectively. Let $A = [a_{ij}]_{(i,j) \in \mathcal{V} \times \mathcal{V}}$ be the adjacency matrix of \mathcal{G} , where $a_{ij} = 1$ if the edge $(i, j) \in \mathcal{E}$, and $a_{ij} = 0$ when $(i, j) \notin \mathcal{E}$. Contrary to many existing results, we do not impose any connectivity assumptions on the interconnection graph. The dynamics of the network is given by

$$\dot{\theta}_i = f_i(\theta_i) + \sum_{j \in \mathcal{V}} a_{ij} k_{ij} g_{ij}(\theta_i, \theta_j), \quad i \in \mathcal{V}, \tag{1a}$$

$$\dot{k}_{ij} = \Gamma_{ij}(k_{ij}, \theta_i, \theta_j), \quad (i, j) \in \mathcal{E}, \tag{1b}$$

where $\theta_i(t) \in \mathcal{S}_1$ and $k_{ij}(t) \in \mathbb{R}$ denote the phase of the oscillator $i \in \mathcal{V}$ and the coupling strength of the link $(i, j) \in \mathcal{E}$ at time $t \in \mathbb{R}_{\geq 0}$, respectively. The function $f_i \in C(\mathcal{T}_1)$ defines the intrinsic dynamics of the i th oscillator, $i \in \mathcal{V}$, and $g_{ij} \in C(\mathcal{T}_2)$, $(i, j) \in \mathcal{E}$ defines the influence of oscillator j on oscillator i depending on their phases. The dynamics of the coupling strength k_{ij} , $(i, j) \in \mathcal{E}$ is defined by the function $\Gamma_{ij} : \mathbb{R} \times \mathcal{T}_2 \rightarrow \mathbb{R}$ such that $\Gamma_{ij}(k, \cdot) \in C(\mathcal{T}_2)$ for every fixed $k \in \mathbb{R}$. For example, if $f_i(s) \equiv w_i$, $g_{ij}(s_1, s_2) = \sin(s_2 - s_1)$, and $\Gamma_{ij}(k, s_1, s_2) = -k + \cos(s_2 - s_1)$, then (4) is a network of Kuramoto oscillators with dynamic Hebbian-type coupling [5, 14, 18]. If $\Gamma_{ij} \equiv 0$ for all $(i, j) \in \mathcal{E}$, then (1) reduces to (1a) with constant coupling strengths $k_{ij} \in \mathbb{R}$, $(i, j) \in \mathcal{E}$, and it defines the network of phase oscillators with static coupling (see e.g., [25] and [26], for the synthesis of electrical circuits for both adaptive and non-adaptive Kuramoto models, respectively).

Definition 1 Let $\mathcal{P} = \{\mathcal{P}_1, \dots, \mathcal{P}_m\}$ with $m \in \mathbb{N}$, $1 < m \leq |\mathcal{V}|$ be a partition of \mathcal{V} , where $\cup_{i=1}^m \mathcal{P}_i = \mathcal{V}$ and $\mathcal{P}_i \cap \mathcal{P}_j = \emptyset$ if $i \neq j$ for all $i, j \in \mathcal{V}$. The network exhibits cluster synchronization when the oscillators can be partitioned so that the phases of the oscillators in each cluster evolve identically.

Contrary to the notion of multi-clustering given in Definition 1, the term multi-clustering may also refer to the network's behavior that is characterized by a partition of nodes into subsets so that the frequencies of oscillators in each subset coincide, see, e.g., [4, 5]. This type of behavior also admits an alternative characterization in terms of the existence of a suitable invariant low-dimensional manifold (see Remark 1 for details).

2.2 Synchronization Invariant Manifolds

In sequel, an alternative characterization of multi-clustering given in Definition 1 will be provided in terms of the existence of synchronization invariant manifolds of the corresponding error-system. To this end, for every cluster \mathcal{P}_s , $s = \overline{1, m}$ we pick an arbitrary oscillator $i_s \in \mathcal{P}_s$ within the cluster, denote its phase by $\varphi_s := \theta_{i_s}$, and collect all φ_s , $s = \overline{1, m}$ into the vector of reference phases $\varphi = (\varphi_1, \dots, \varphi_m)^\top \in \mathcal{T}_m$. For every oscillator $i \in \mathcal{P}_s \setminus \{i_s\}$, let $e_i = \theta_i - \varphi_s$ denote the relative phase difference within the cluster \mathcal{P}_s , $s = \overline{1, m}$. All relative phase differences e_i , $i \in \mathcal{V} \setminus \cup_{s=1}^m \{i_s\}$ and all coupling strengths k_{ij} , $(i, j) \in \mathcal{E}$ are collected into the vectors e and k of dimensions $|\mathcal{V}| - m$ and $|\mathcal{E}|$, respectively.

In many cases, following the introduced notation, system (1) can be rewritten in the form

$$\dot{\varphi} = F_1(\varphi, e, k), \quad (2a)$$

$$\dot{e} = F_2(\varphi, e, k), \quad (2b)$$

$$\dot{k} = F_3(\varphi, e, k), \quad (2c)$$

with suitably defined functions F_1, F_2, F_3 such that $F_i(\cdot, e, k) \in C(\mathcal{T}_m), i \in \{1, 2, 3\}$ for every fixed (e, k) . For instance, such a transformation is always possible when the interaction between nodes are of a diffusive type, i.e., $g_{ij}(\theta_1, \theta_2) \equiv \tilde{g}_{ij}(\theta_j - \theta_i), \Gamma_{ij}(k_{ij}, \theta_i, \theta_j) \equiv \tilde{\Gamma}_{ij}(k_{ij}, \theta_j - \theta_i)$. System (2) has the same number of equations as (1): Eq.(2a) governs the dynamics of a single selected node within every set $\mathcal{P}_s, s = \overline{1, m}$, Eq.(2b) governs the dynamics of the relative phase differences within every set $\mathcal{P}_s, s = \overline{1, m}$, and Eq.(2c) governs the dynamics of coupling strengths.

We are now in position to introduce an alternative characterization of multi-clustering in terms of the existence of invariant manifolds of (2) of a special topological structure.

Definition 2 We say that system (1) admits multi-clustering defined by the partition \mathcal{P} if (1) can be rewritten in form of (2) and there exists an m -dimensional invariant toroidal manifold for (2) given by

$$\mathcal{M} = \{(e, k, \varphi) \in \mathbb{R}^{|\mathcal{V}|-m} \times \mathbb{R}^{|\mathcal{E}|} \times \mathcal{T}_m : e = 0, k = u(\varphi), \varphi \in \mathcal{T}_m\} \quad (3)$$

for some $u = (u_1, \dots, u_{|\mathcal{E}|}) \in C(\mathcal{T}_m)$.

The invariant manifold \mathcal{M} corresponds to the oscillating behavior of coupling strengths k preserving zero phase $e = 0$ difference within clusters. Although the function u is periodic with respect to every $\varphi_i, i = \overline{1, m}$, the oscillations of the coupling strengths are not necessarily periodic, since they are generated by trajectories on the m -dimensional torus \mathcal{T}_m , which can be, for example, quasi-periodic in time.

The existence of the invariant manifold (3) does not automatically imply the emergence of multi-cluster behavior in (1). It only indicates the possibility of such behavior if the trajectory of the error-system (2) reaches the manifold. The invariance property guarantees that the multi-clustering will be maintained for all times since the moment when the trajectory of (2) reaches the manifold. The emergence of the multi-cluster behavior can be associated to the asymptotic stability of the manifold (3): If the manifold is asymptotically stable then any point from a vicinity of the manifold will eventually converge towards it. In Sect.2.3, sufficient conditions for the existence and local asymptotic stability of invariant toroidal manifolds of type (3) will be derived for a particular subclass of networks (1), namely, for Kuramoto networks with adaptive coupling. The conditions will provide a trade-off between the natural frequencies of oscillators, plasticity parameters of adaptive couplings, and the interconnection topology of the network.

Remark 1 In the spirit of Definition 2, the frequency multi-clustering can be also characterized in terms of the existence of an invariant toroidal manifold

$$\mathcal{M}_{\text{freq}} = \{(e, k, \varphi) \in \mathbb{R}^{|\mathcal{V}|-m} \times \mathbb{R}^{|\mathcal{E}|} \times \mathcal{T}_m : e = d, k = u(\varphi), \varphi \in \mathcal{T}_m\}$$

for some $u = (u_1, \dots, u_{|\mathcal{E}|}) \in C(\mathcal{T}_m)$ and constant $d \in \mathbb{R}^{|\mathcal{V}|-m}$. Constant phase difference between any two oscillators within a cluster will guarantee the same frequency for all oscillators that belong to the same cluster. Further extensions and

generalizations of manifolds \mathcal{M} and $\mathcal{M}_{\text{freq}}$ are possible to cover many other types of collective behavior like, generalized multi-clustering or practical synchronization. These are, however, out of the scope of the current chapter.

2.3 Application to Adaptive Kuramoto Networks

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the directed graph defined in Sect. 2.1 that represents the network of oscillators with the adjacency matrix $A = [a_{ij}]_{i,j=\overline{1,N}}$. Additionally, it is assumed that the graph does not have self-loops, i.e., $a_{ii} = 0$ for all $i = \overline{1, N}$. The dynamics of the network is given by [14]

$$\begin{aligned} \dot{\theta}_i &= w_i + \sum_{j=1}^N a_{ij} k_{ij} \sin(\theta_j - \theta_i), \quad i = \overline{1, N}, \\ \dot{k}_{ij} &= -\gamma k_{ij} + \mu_{ij} \Gamma(\theta_j - \theta_i), \quad i, j = \overline{1, N}, \end{aligned} \quad (4)$$

where $w_i \in \mathbb{R}$ and $\theta_i(t) \in \mathcal{T}_1$ denote the natural frequency and the phase of the i th oscillator. The dynamics of the coupling strength $k_{ij}(t) \in \mathbb{R}$ is defined by parameters $\mu_{ij}, \gamma \in \mathbb{R}_{>0}$ and $\Gamma \in C^1(\mathcal{T}_1)$ with $|\Gamma|_1 = \delta \in \mathbb{R}_{>0}$.

We are interested in establishing the relationships between the emergence of multi-cluster behavior in (4) (in the sense of Definition 1) and the dynamical properties of oscillators, the adaptive couplings, and the interconnection topology. To this end, for a given partition \mathcal{P} , let \mathcal{E}_{in} and \mathcal{E}_{out} be the subsets of \mathcal{E} that correspond to the intra-cluster links and inter-cluster links, respectively. The cardinalities of these sets $c_{in} = |\mathcal{E}_{in}|$ and $c_{out} = |\mathcal{E}_{out}|$ characterize the interconnection structure of \mathcal{G} with respect to the partition \mathcal{P} . Additionally, let $w_{min} = \min_{i=\overline{1,N}} |w_i|$ and $w_{max} = \max_{i=\overline{1,N}} |w_i|$ denote the minimal and maximal absolute value of the natural frequencies.

Sufficient conditions for the existence and construction procedure of the invariant toroidal manifolds that correspond to the m -cluster behavior of the network (4) have been proposed in [14] for the case of identical plasticity parameters $\mu_{ij} \equiv \mu \in \mathbb{R}_{>0}$. However, the mentioned result does not answer the question whether the constructed invariant manifold is (asymptotically) stable. Later, this question has been answered in [13] by proposing sufficient conditions for the asymptotic stability of the invariant manifold for the case when the plasticity parameters μ_{ij} are different for the intra-cluster and inter-cluster links. Namely, let $\mu_{ij} \equiv \tilde{\mu} \in \mathbb{R}_{>0}$ if the link (i, j) connects nodes within some cluster $\mathcal{P}_s, s = \overline{1, m}$, and $\mu_{ij} \equiv \mu \in \mathbb{R}_{>0}$ otherwise. For convenience, system (4) with the chosen set of plasticity parameters will be denoted as $\Sigma(\tilde{\mu}, \mu)$ from now on, where the first argument stands for the plasticity parameter μ_{ij} of the adaptive link connecting nodes within the same cluster, and the second argument stands for the plasticity parameter μ_{ij} of the adaptive link connecting nodes belonging to different clusters.

The derivation of the error-system follows the steps introduced in Sect. 2.2. Let the partition \mathcal{P} be given. For every cluster $\mathcal{P}_s, s = \overline{1, m}$ pick an arbitrary oscillator $i_s \in \mathcal{P}_s$

and denote its phase and natural frequency by $\varphi_s := \theta_{i_s}$ and $\bar{w}_s := w_{i_s}$, respectively. For every oscillator $i \in \mathcal{P}_s$, let $e_i = \theta_i - \varphi_s$ define the relative phase-error within a given cluster \mathcal{P}_s , $s = \overline{1, m}$. Then, $\Sigma(\tilde{\mu}, \mu)$ can be rewritten in the following form:

$$\dot{\varphi}_s = \bar{w}_s + \sum_{j \in \mathcal{P}_s} a_{i_s j} k_{i_s j} \sin e_j + \sum_{r \neq s} \sum_{j \in \mathcal{P}_r} a_{i_s j} k_{i_s j} \sin(e_j + \varphi_r - \varphi_s), \quad s = \overline{1, m}, \quad (5a)$$

$$\begin{aligned} \dot{e}_i &= w_i - \bar{w}_s + \sum_{j \in \mathcal{P}_s} [a_{ij} k_{ij} \sin(e_j - e_i) - a_{i_s j} k_{i_s j} \sin e_j] \\ &+ \sum_{r \neq s} \sum_{j \in \mathcal{P}_r} [a_{ij} k_{ij} \sin(e_j - e_i + \varphi_r - \varphi_s) - a_{i_s j} k_{i_s j} \sin(e_j + \varphi_r - \varphi_s)] \end{aligned} \quad \forall i \in \mathcal{P}_s \setminus \{i_s\}, \quad s = \overline{1, m}, \quad (5b)$$

$$\dot{k}_{ij} = -\gamma k_{ij} + \mu \Gamma(e_j - e_i + \varphi_r - \varphi_s) \quad \forall i \in \mathcal{P}_s, \quad \forall j \in \mathcal{P}_r, \quad s \neq r, \quad s, r = \overline{1, m}, \quad (5c)$$

$$\dot{k}_{ij} = -\gamma k_{ij} + \tilde{\mu} \Gamma(e_j - e_i) \quad \forall i, j \in \mathcal{P}_s, \quad i \neq j, \quad s = \overline{1, m}. \quad (5d)$$

System (5) is a counterpart of system (2) for the case of Kuramoto networks with adaptive coupling and it has the same number of equations as system (4). Equations (5a) describe the dynamics of m arbitrarily selected oscillators (one from every cluster). Equations (5b) describe the error dynamics within each cluster. Equations (5c) describe the dynamics of the coupling strengths between nodes of different clusters. Finally, (5d) describe the dynamics of the intra-cluster coupling strengths. Let $\varphi = (\varphi_1, \dots, \varphi_m)^\top \in \mathcal{T}_m$ and $e = (e_{i_1}, \dots, e_{i_m})^\top \in \mathbb{R}^{N-m}$ be the vectors collecting all cluster phases φ_i , $i = \overline{1, m}$ and all intra-cluster relative phase errors e_i , $i \in \mathcal{P}_s \setminus \{i_s\}$, $s = \overline{1, m}$, respectively. Similarly, all inter- and intra-cluster coupling strengths are collected into the vectors $k^{inter} \in \mathbb{R}^{c_{out}}$ and $k^{intra} \in \mathbb{R}^{c_{in}}$, respectively, and $k = (k^{inter^\top}, k^{intra^\top})^\top$. Following Definition 2, the multi-cluster behavior in network $\Sigma(\tilde{\mu}, \mu)$ is possible if system (5) possesses an invariant toroidal manifold

$$\mathcal{M} = \{(e, k, \varphi) \in \mathbb{R}^{N-m} \times \mathbb{R}^{c_{in}+c_{out}} \times \mathcal{T}_m : e = 0, k = u(\varphi), \varphi \in \mathcal{T}_m\} \quad (6)$$

for some $u \in C(\mathcal{T}_m)$. This invariant manifold corresponds to the oscillating behavior of the coupling strengths k preserving zero phase error e within clusters. In [14], it has been shown that the inter-cluster coupling strengths cannot converge to some constant value say d simultaneously guaranteeing the convergence of the phase errors to zero, i.e., $\{(e, k, \varphi) \in \mathbb{R}^{N-m} \times \mathbb{R}^{c_{in}+c_{out}} \times \mathcal{T}_m : e = 0, k = d, \varphi \in \mathcal{T}_m\}$ is *not* an invariant set of (5) for any constant $d \in \mathbb{R}^{c_{in}+c_{out}}$. Hence, the oscillating behavior of the inter-cluster coupling strengths is necessary for the emergence of multi-cluster formations in (4).

Following the steps of the proof of Theorem 3 from [14], the following result can be obtained.

Theorem 1 (Adapted from [14], Theorem 3) *Let the following conditions hold true for system $\Sigma(\tilde{\mu}, \mu)$ and a given partition \mathcal{P} :*

- (A1) for any $s = \overline{1, m}$ and for any $i, j \in \mathcal{P}_s$ it holds that $w_i = w_j$,
 (A2) for any $s, r = \overline{1, m}$, $s \neq r$ there exist constants $c_{sr} \in \mathbb{N}$ such that for any $i \in \mathcal{P}_s$

$$\sum_{j \in \mathcal{P}_r} a_{ij} = c_{sr},$$

- (A3) given $c_{max} := \max_{s=1, m} \sum_{r \neq s} c_{sr}$ it holds that

$$w_{min} - \mu\gamma^{-1}\delta c_{max} > 0 \tag{7}$$

and

$$4 \frac{\mu}{\gamma^2} \delta \sqrt{c_{out}} \sum_{\substack{s, r = \overline{1, m} \\ s \neq r}} c_{sr} \frac{w_{max} + \mu\gamma^{-1}\delta c_{max}}{w_{min} - \mu\gamma^{-1}\delta c_{max}} < 1. \tag{8}$$

Then, system (5) has an invariant toroidal manifold \mathcal{M} that corresponds to the m -cluster behavior of $\Sigma(\tilde{\mu}, \mu)$ defined by the partition \mathcal{P} .

Proof The proof is based on the perturbation theory of invariant tori for nonlinear extensions of dynamical systems on torus. We refer the interested readers to the monographs [22, 38], which provide the fundamentals of the mathematical theory of multi-frequency oscillations and to the papers [11, 33, 37], which address the persistence of the Green–Samoilenko function of the invariant tori problem under the perturbations of the right-hand side of the corresponding ordinary differential equations. A relation of these concepts to the synchronization analysis of oscillator networks is discussed in [14]. □

Conditions (A1)–(A3) allow for the following interpretation:

- (A1) requires the natural frequencies to be equal within every cluster.
- (A2) requires that the number of incoming links to every node within a given cluster \mathcal{P}_s from a different cluster \mathcal{P}_r , $r \neq s$ is the same. Condition (A2) restricts only the number of links and does not require any symmetry of the corresponding adjacency matrix. It is worth to highlight that the intra-cluster couplings are generally not required for the emergence of multi-cluster behavior in the network since (A2) restricts only the structure of the inter-cluster connections.
- (A3) establishes the relations between the natural frequencies of the oscillators, plasticity parameters μ, γ, δ and the inter-cluster interconnection topology. The procedure for verifying this conditions will be demonstrated in Example 1.

In order to derive sufficient conditions for the asymptotic stability of the manifold \mathcal{M} , the following auxiliary notation is introduced : For a given cluster \mathcal{P}_s , $s = \overline{1, m}$, let

- n_s be the number of elements in the set \mathcal{P}_s ;
- $\mathcal{G}_s \subset \mathcal{G}$ be a subgraph that correspond to the nodes from \mathcal{P}_s and intra-cluster connections, i.e.,

$$\mathcal{G}_s = \{(\mathcal{P}_s, \mathcal{E}_s) : \mathcal{E}_s = \mathcal{P}_s \times \mathcal{P}_s \cap \mathcal{E}\};$$

- A_s be the adjacency matrix of \mathcal{G}_s .

To define the residual connectivity of \mathcal{G}_s with respect to the node i_s , the nodes inside each cluster are enumerated according to the rule $\mathcal{P}_s = \{i_1^s, \dots, i_{k_s}^s, \dots, i_{n_s}^s\}$, where $i_{k_s}^s = i_s$, i.e., the selected node i_s has a sequential number k_s in the cluster \mathcal{P}_s . Then, let

- A_s^- be an $(n_s - 1) \times (n_s - 1)$ -dimensional matrix constructed from A_s by removing its k_s th row and column;
- \tilde{A}_s be the residual adjacency matrix w.r.t. the node i_s , i.e.,

$$\tilde{A}_s = A_s^- - \begin{pmatrix} a_{i_s i_1^s} & \dots & a_{i_s i_{k_s-1}^s} & a_{i_s i_{k_s+1}^s} & \dots & a_{i_s i_{n_s}^s} \\ a_{i_s i_1^s} & \dots & a_{i_s i_{k_s-1}^s} & a_{i_s i_{k_s+1}^s} & \dots & a_{i_s i_{n_s}^s} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i_s i_1^s} & \dots & a_{i_s i_{k_s-1}^s} & a_{i_s i_{k_s+1}^s} & \dots & a_{i_s i_{n_s}^s} \end{pmatrix} \quad (9)$$

- D_s be the degree matrix of A_s , i.e., the diagonal matrix with diagonal elements equal to the sum of all elements in the corresponding row of A_s , and D_s^- be an $(n_s - 1) \times (n_s - 1)$ -dimensional matrix constructed from D_s by removing its k_s th row and column.

Theorem 2 (Adapted from [13], Theorem 2) *Let the network $\Sigma(\tilde{\mu}, \mu)$ satisfy conditions (A1), (A2), and (A3) of Theorem 1, i.e., there exist the synchronization invariant toroidal manifold \mathcal{M} that corresponds to the multi-cluster behavior of the network given by the partition \mathcal{P} . If*

(A4) *for every $s = \overline{1, m}$*

$$\text{sign } \Gamma(0) \text{Re } \lambda \left(\tilde{A}_s - D_s^- \right) < 0, \quad (10)$$

then there exist $\mu_0 \leq \mu$ such that for all $v < \mu_0$ the invariant toroidal manifold that corresponds to the multi-clustering of $\Sigma(\tilde{\mu}, v)$ is locally asymptotically stable.

The 'stability-condition' (A4) depends on the value of the plasticity function Γ evaluated at zero (sometimes called learning rule) and the eigenvalues of the matrix that is constructed of those elements of the adjacency matrix A , which correspond to the *intra-cluster* connections. This makes a crucial contrast to the 'existence-condition' (A2) that restricts the *inter-cluster* connectivity of the interconnection graph \mathcal{G} .

The usage of both Theorems 1 and 2 is demonstrated in the following example:

Example 1 [13]

Consider a network of $N = 7$ Kuramoto oscillators (4) with adjacency matrix

$$A = \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{array} \right),$$

natural frequencies $w = (0.5, 0.5, 0.5, \sqrt{0.8}, \sqrt{0.8}, \sqrt{0.8}, \sqrt{0.8})^\top$, plasticity parameters $\gamma = 0.2, \tilde{\mu} = \mu = 0.001$, Hebbian learning rule $\Gamma(s) = \cos(s)$, and the desired two-cluster partition $\mathcal{P} = \{1, 2, 3\} \cup \{4, 5, 6, 7\}$.

Condition (A1) is satisfied thanks to the choice of w . Every node in cluster \mathcal{P}_1 has exactly one incoming link from the nodes of cluster \mathcal{P}_2 , and vice versa. Therefore, the resulting interconnection topology satisfies the condition (A2) and the network satisfies (A3) with characteristics $c_{out} = 7, c_{max} = c_{12} = c_{21} = 1, \delta = 1$. Indeed,

$$w_{min} - \mu\gamma^{-1}\delta c_{max} = \frac{1}{2} - \frac{0.001}{0.2} = 0.495 > 0$$

and

$$4 \frac{\mu}{\gamma^2} \delta \sqrt{c_{out}} \sum_{\substack{s,r=1,m \\ s \neq r}} c_{sr} \frac{w_{max} + \mu\gamma^{-1}\delta c_{max}}{w_{min} - \mu\gamma^{-1}\delta c_{max}} \approx 0.9615 < 1.$$

All conditions of Theorem 1 are satisfied. Asymptotic stability of the two-cluster formation can be concluded from the condition (A4) of Theorem 2. For this purpose, the nodes 3 and 7 are chosen as k_1 and k_2 for clusters \mathcal{P}_1 and \mathcal{P}_2 , respectively. Since $\text{sign } \Gamma(0) = \text{sign } \cos(0) = 1 > 0$, (A4) for the cluster \mathcal{P}_1 reads

$$A_1 = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{array} \right), \quad \tilde{A}_1 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \quad D_1^- = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that

$$\text{Re } \lambda(\tilde{A}_1 - D_1^-) = \text{Re } \lambda \begin{pmatrix} -2 & 1 \\ -1 & -1 \end{pmatrix} = \text{Re} \left(-\frac{3}{2} \pm \frac{\sqrt{3}}{2}i \right) < 0.$$

For the cluster \mathcal{P}_2 , condition (A4) reads

$$A_2 = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \end{array} \right), \quad \tilde{A}_2 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}, \quad D_2^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

thus

$$\operatorname{Re} \lambda(\tilde{A}_2 - D_2^-) = \operatorname{Re} \lambda \begin{pmatrix} -2 & 1 & 0 \\ -1 & -1 & 1 \\ -1 & 0 & -1 \end{pmatrix} = \begin{bmatrix} \operatorname{Re}(-1 \pm i) \\ \operatorname{Re}(-2) \end{bmatrix} < 0.$$

All conditions (A1)–(A4) of Theorems 1 and 2 are satisfied. Simulation results are presented in Figs. 1 and 2. In particular, the numerical simulations in Fig. 1 demonstrate the conclusions of Theorem 2 and show that starting from a vicinity of the invariant toroidal manifold, which corresponds to the desired multi-cluster behavior of oscillators, the coupling strengths and phase-errors of the considered Kuramoto network converge to this manifold. Figure 2 depicts two snapshots of the graph \mathcal{G} at the beginning ($t = 0$) and at the end ($t = 500$) of simulation.

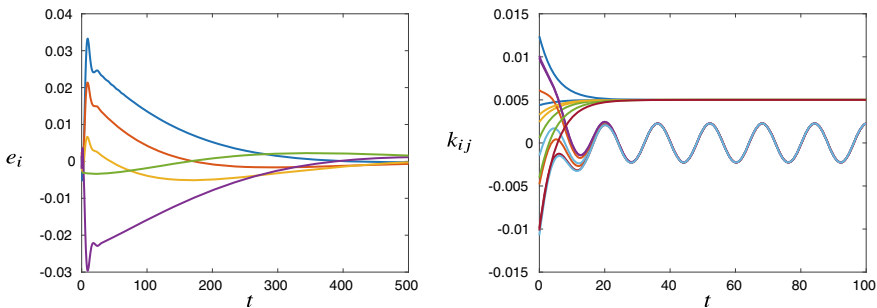


Fig. 1 *Left figure:* Evolution of absolute values of the phase-errors $e_i, i = \overline{1, N}$ within clusters. All $e_i(t) \xrightarrow{t \rightarrow \infty} 0, i = \overline{1, N}$ that correspond to the asymptotic stability of the invariant toroidal manifold \mathcal{M} and the emergence of two-cluster formation given by partition \mathcal{P} . *Right figure:* Evolution of coupling strengths k_{ij} . The intra-cluster coupling strengths converge to the constant value, and the inter-cluster couplings exhibit quasiperiodic oscillations. Due to the choice of rationally independent natural frequencies, these oscillations are not periodic in time

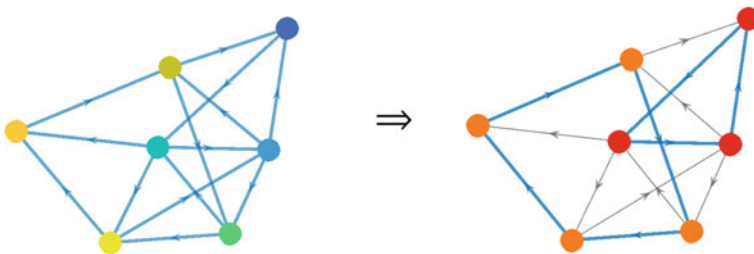


Fig. 2 Snapshots of the graph \mathcal{G} at the beginning (left figure) and at the end of simulation (right figure). Colors of the nodes represent their phases. In the right figure, the blue connections denote intra-cluster links whose coupling strengths converge to a constant value. Light-grey links correspond to the oscillating inter-cluster couplings (see also Fig. 1). Red and orange nodes in the right figure form to two different clusters

3 Synchronization of Complex Dynamics

This section is devoted to the synchronization analysis of complex dynamics. In particular, we address the problem of the synchronization of chaotic behavior generated by ensembles of Chua oscillators and the problem of the master-slave synchronization of spatiotemporal patterns generated by the Gray–Scott reaction-diffusion PDE model. Besides theoretical results on the synchronization, both problems are illustrated in numerical simulations.

3.1 Synchronization of Chaotic Behavior

Synchronization of chaotic systems has been intensively studied during the last decades [2, 27, 31, 32] due to their numerous applications, e.g., in secure communication [1, 41, 48], laser physics [28], and biomedical engineering [40]. In this context, the Chua circuit appeared to be one of the most interesting objects of exploration since it exhibits extremely rich dynamical behavior and variety of bifurcation phenomena despite its structural simplicity. As pointed out in [46], the investigation of synchronizability of coupled Chua oscillators may help to understand complex dynamical phenomena arising in networks of chaotic systems of more general types.

Synchronization of two linearly coupled Chua oscillators has been studied in [7–9, 44, 46, 51]. A graph-spectral approach for the synchronization of networks of resistively coupled nonlinear oscillators has been proposed in [47], and the upper bound on the coupling conductance required for synchronization has been obtained therein. In this chapter, we present sufficient conditions for the synchronization of the network of $N \in \mathbb{N}$ Chua oscillators interconnected with the static nonlinear coupling via the first state coordinate only. Our conditions provide a trade-off between the interconnection topology of the network, properties of nonlinear coupling function, and parameters of the Chua circuits in order to achieve synchronization. In Sect. 3.1.1, the network under consideration is defined and the main problem of synchronization is formulated. Then, we propose sufficient conditions for the synchronization of $N \in \mathbb{N}$ coupled Chua oscillators with static nonlinear coupling satisfying a so-called sector condition (see Eq. (15)), and provide a numerical example to illustrate the usage of the derived conditions.

3.1.1 Chua Oscillator

We consider the extended Chua circuit

$$\dot{\xi}_1 = \alpha[-\xi_1 + \xi_2 - \phi(\xi_1)] + u, \quad \xi_1(0) = \xi_{10} \quad (11a)$$

$$\dot{\xi}_2 = \xi_1 - \xi_2 + \xi_3, \quad \xi_2(0) = \xi_{20} \quad (11b)$$

$$\dot{\xi}_3 = -\beta\xi_2 - \gamma\xi_3, \quad \xi_3(0) = \xi_{30} \quad (11c)$$

$$y = \xi_1 \quad (11d)$$

with scalar piecewise linear function

$$\phi(s) = as + \frac{1}{2}[b - a][|s + 1| - |s - 1|] \quad \forall s \in \mathbb{R}$$

and parameters $\alpha, \beta > 0, \gamma \geq 0, a < b < 0$, or in a matrix form as

$$\dot{x} = Px + bu + f(x), \quad x(0) = x_0 \quad (12a)$$

$$y = c^T x \quad (12b)$$

with the state $x(t) = [\xi_1(t) \ \xi_2(t) \ \xi_3(t)]^T \in \mathbb{R}^3$, $x_0 = [\xi_{10} \ \xi_{20} \ \xi_{30}]^T \in \mathbb{R}^3$, external input $u(t) \in \mathbb{R}$, which will be later used to interconnect Chua oscillators, and the matrix P and vectors f, b, c given by

$$P = \begin{bmatrix} -\alpha & \alpha & 0 \\ 1 & -1 & 1 \\ 0 & -\beta & -\gamma \end{bmatrix}, \quad f(x) = \begin{bmatrix} -\alpha\phi(\xi_1) \\ 0 \\ 0 \end{bmatrix}, \quad b = c = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Now, consider a network of $N \in \mathbb{N}$ nodes described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with node set \mathcal{V} and edge set \mathcal{E} , so that $|\mathcal{V}| = N$. Let the associated adjacency matrix be given by $A = \{a_{ij}\}_{i,j=1,\dots,N}$ with zero main diagonal. Then, applying the output feedback

$$u_i = - \sum_{j=1}^N a_{ij} k(y_i - y_j), \quad i = \overline{1, N}$$

with an arbitrary nonlinear locally Lipschitz continuous coupling function $k : \mathbb{R} \rightarrow \mathbb{R}$, the dynamics of $N \in \mathbb{N}$ coupled Chua oscillators can be written as

$$\dot{x}_i = Ax_i - b \sum_{j=1}^N a_{ij} k(y_i - y_j) + f(x_i), \quad x_i(0) = x_{i0} \quad (13)$$

$$y_i = c^T x_i,$$

$i = 1, \dots, N$. For any given $\eta \in \mathbb{R}^{3N}$ let $x = (x_1, \dots, x_N) : \mathbb{R} \rightarrow \mathbb{R}^{3N}$ denote a solution to (13) satisfying initial condition $x(0) = \eta$. The Lipschitz continuity of the right-hand side of (3) guarantees the existence and uniqueness of the solution for any initial value $\eta \in \mathbb{R}^{3N}$.

Associated to this network consider the relative synchronization errors with respect to the node 1

$$e_i = x_i - x_1, \quad i = \overline{1, N}. \tag{14}$$

The relative errors e_{ij} between arbitrary nodes i and j can be expressed using the relative errors e_i and e_j

$$e_{ij} = x_i - x_j = x_i - x_1 - (x_j - x_1) = e_i - e_j.$$

Accordingly, instead of analyzing $\frac{N(N-1)}{2}$ relative errors e_{ij} between connected nodes, it is sufficient to consider the behavior of the $N-1$ errors $e_i, i = \overline{2, N}$.

The problem addressed in the sequel consists in providing sufficient conditions on the system parameters, the nonlinear coupling function k and the network topology which ensure the synchronization of $N \in \mathbb{N}$ coupled Chua oscillators, i.e., the global convergence of the norms of errors e_i to zero:

$$\lim_{t \rightarrow \infty} \|e_i(t)\| = 0, \quad i = \overline{2, N}.$$

3.1.2 Nonlinear Couplings Satisfying a Sector Condition

We consider a special case of Lipschitz continuous odd coupling $k : \mathbb{R} \rightarrow \mathbb{R}$ satisfying a sector condition: Let there exist two constants $k_1, k_2 \geq 0$ with $k_2 \geq k_1 \geq 0$ such that for all $s \in \mathbb{R}$:

$$k(s)s \geq 0, \quad k(-s) = -k(s), \quad k_1|s| \leq |k(s)| \leq k_2|s|. \tag{15}$$

Additionally, define the following auxiliary matrices

$$\mathcal{A}_1 = \begin{bmatrix} 0 & |a_{23}| & \cdots & |a_{2N}| \\ |a_{32}| & 0 & \ddots & |a_{3N}| \\ \vdots & \ddots & \ddots & \vdots \\ |a_{N2}| & |a_{N3}| & \cdots & 0 \end{bmatrix}, \quad \mathcal{A}_2 = \begin{bmatrix} 0 & |a_{23} - a_{13}| & \cdots & |a_{2N} - a_{1N}| \\ |a_{32} - a_{12}| & 0 & \ddots & |a_{3N} - a_{1N}| \\ \vdots & \ddots & \ddots & \vdots \\ |a_{N2} - a_{12}| & |a_{N3} - a_{13}| & \cdots & 0 \end{bmatrix}$$

with zero main diagonals, $(N - 1) \times (N - 1)$ -dimensional identity matrix I , and $K = \text{diag}\{\kappa_2, \dots, \kappa_N\}$, where $\kappa_i = \sum_{j=1}^N a_{ij}$ denotes the degree of node $i = \overline{1, N}$. Sufficient conditions for the synchronization of the entire network can be formulated in terms of eigenvalues of a matrix M defined below that interrelates the parameters of oscillators, coupling function, and the interconnection topology.

Theorem 3 (Adapted from [15], Theorem 3) *Let matrix*

$$M = \begin{bmatrix} -(\alpha - \alpha|a|)I - k_1K + (k_2 - k_1)\mathcal{A}_1 + k_2\mathcal{A}_2 & \alpha I \\ I & -\mu_0 I \end{bmatrix}$$

be Hurwitz, where

$$\mu_0 = \frac{1 + \gamma}{2} - \operatorname{Re} \left(\sqrt{\frac{(1+\gamma)^2}{4} - (\gamma + \beta)} \right).$$

Then, the synchronization errors (14) between the states of Chua oscillators (13) with nonlinear coupling k satisfying (15) converge exponentially to zero.

Example 2 [15]

Consider a complete graph hosting $N = 20$ identical Chua oscillators (13) in its nodes. The oscillators are connected via the nonlinear coupling

$$k(s) = 3s + \arctan s \quad \text{for all } s \in \mathbb{R}, \tag{16}$$

and the parameters of oscillators are $\alpha = 15.61$, $\beta = 25.581$, $\gamma = 0$, $a = -1.142$, $b = -0.715$. These parameters correspond to the chaotic behavior of each oscillator with the double scroll attractor [34] (see Fig. 3 (left)). The considered nonlinear coupling strength (16) satisfies the sector condition (15) with constants $k_1 = 3$ and $k_2 = 4$. For the chosen parameters of the network matrix M from Theorem 3 reads as

$$M = \begin{bmatrix} -54.7834I & 15.61I \\ I & -0.5I \end{bmatrix} + \begin{bmatrix} \mathbb{1}\mathbb{1}^\top - I & \mathbb{0} \\ \mathbb{0} & \mathbb{0} \end{bmatrix},$$

where $\mathbb{0}$ denotes zero $(N - 1) \times (N - 1)$ -matrix, and $\mathbb{1}$ denotes $(N - 1)$ -dimensional vector $(1 \ 1 \ \dots \ 1)^\top$. The eigenvalues of M lie in the open left half-plane of the complex plane \mathbb{C} so that $\operatorname{Re}(\lambda(M)) \in [-56.0643, -0.0748]$. Hence, from Theorem 3 we conclude that oscillators achieve synchronization. The state evolution of the oscillators is shown in Fig. 3 (right).

3.2 Synchronization of Spatiotemporal Patterns

3.2.1 Gray–Scott Reaction-Diffusion Model

The Gray–Scott model [17, 21], that is a simple prototype for the models of complex isothermal autocatalytic reactions, is governed by the pair of coupled reaction-diffusion equations

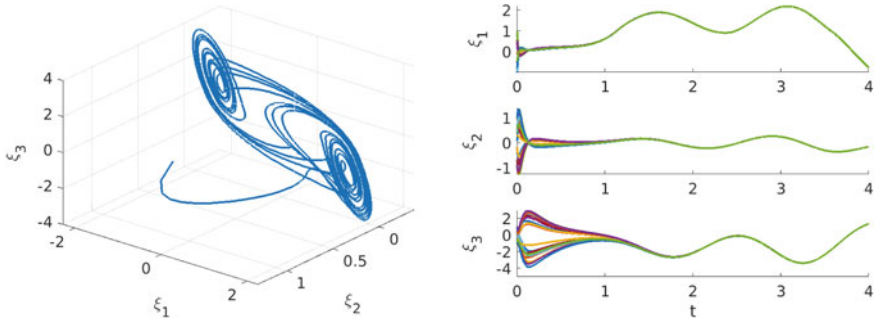


Fig. 3 *Left figure:* Double scroll attractor for the chaotic Chua oscillator. *Right figure:* Time evolution of $N = 20$ coupled Chua oscillators from the Example 2

$$\partial_t a = D_a \partial_z^2 a - ab^2 + \alpha(1 - a) \tag{17a}$$

$$\partial_t b = D_b \partial_z^2 b + ab^2 - (\alpha + \beta)b \tag{17b}$$

with homogeneous Neumann (non-flux) boundary conditions

$$\partial_z a(t, 0) = \partial_z a(t, L) = 0, \quad t \geq 0 \tag{17c}$$

$$\partial_z b(t, 0) = \partial_z b(t, L) = 0, \quad t \geq 0 \tag{17d}$$

and initial conditions given by

$$a(0, z) = a_0(z), \quad z \in \Omega \tag{17e}$$

$$b(0, z) = b_0(z), \quad z \in \Omega, \tag{17f}$$

where $a(t, z) \in \mathbb{R}_{\geq 0}$ and $b(t, z) \in \mathbb{R}_{\geq 0}$ denote the concentrations of two chemical species at time $t \in [0, \infty)$ and at position $z \in \Omega := (0, L)$, $L > 0$, parameters D_a, D_b, α , and β are positive scalars, initial conditions $a_0, b_0 \in L^2(\Omega)$. The space of square integrable functions $f : \Omega \rightarrow \mathbb{R}^n$, $n \in \mathbb{N}$ is denoted by $L^2(\Omega)$ and it is equipped with the norm $\|f\| = (\int_{\Omega} |f(z)|^2 dz)^{\frac{1}{2}}$, where $|\cdot|$ denotes either the absolute value or the \mathbb{R}^n -distance, depending on whether its argument is scalar or a vector.

System (17) may exhibit a variety of irregular spatiotemporal patterns in response to finite-amplitude perturbations of the steady state $(a, b) = (1, 0)$ [10, 21] (see also Fig. 4). These include a diversity of complex dynamical regimes ranging from steady states and stationary periodic solutions to traveling waves, pulse splittings, and spatiotemporal chaotic behavior. Pattern formation capabilities and the underlying mechanisms of the Gray–Scott model in one- and two-dimensional spatial domains have been reported and analyzed by mathematical analysis methods [23, 24, 45], by computer simulations [29, 30, 42, 43], and by experiments [19, 20].

The problem of synchronization of spatiotemporal chaotic behavior generated by the Gray–Scott model have been addressed in [49, 50] by means of the impulsive control techniques. In particular, a class of pinning impulsive controllers has been

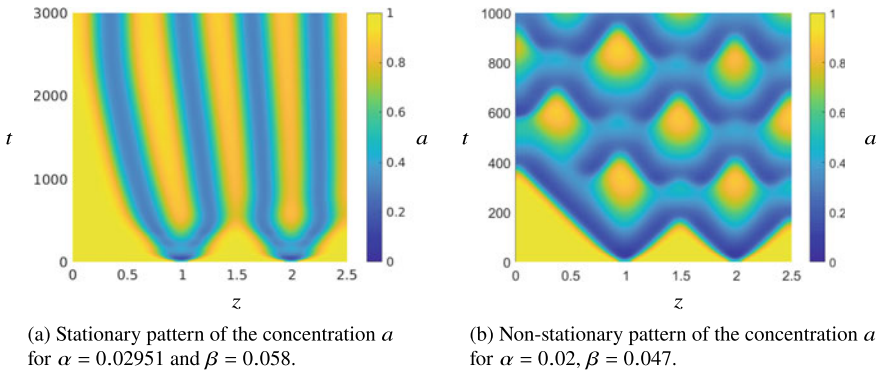


Fig. 4 Typical patterns generated by the Eq. (17) in the interval $(0, 2.5)$ with diffusion coefficients $D_a = 2 \cdot 10^{-4}$ and $D_b = 10^{-4}$. Both patterns **a** and **b** emerge from the same initial conditions, but for different parameters α, β

designed to stabilize and synchronize the spatiotemporal chaotic behavior. The current chapter addresses the synchronization problem of spatiotemporal patterns in the master-slave configuration in which the master system is given by the Gray–Scott model (17), and the slave Gray–Scott system is coupled to the original system via a finite number of spatial locations. This coupling enters the slave system dynamics as the in-domain injections of the master system state variables at the respective locations (see Eq. (19) for details). Following the late-lumping approach that relies directly on the original PDE description, we examine the minimal number of coupling connections sufficient for the synchronization of the master-slave configuration and additionally analyze the results in numerical simulations. In the considered setup, the slave system can be seen as a special kind of the point-wise innovation (PWI) observer [39] for the master Gray–Scott model. The synchronization between the master and the slave systems will correspond to the convergence of the observation error of the PWI-estimator to zero and sufficient conditions for this convergence can be concluded using analysis techniques from [12].

3.2.2 Master-Slave Setup and Main Results

Let the master system be given by (17) and be coupled to the slave system via $m + 1$ spatial locations at positions $z = \zeta_i \in \bar{\Omega} := [0, L], i = \bar{0}, m, m \in \mathbb{N}$ so that $\zeta_0 = 0, \zeta_m = L$, and $z_i - z_{i-1} =: d_i > 0$ for all $i = 1, \dots, m$. System (17) is then equipped with the outputs

$$y_i^a(t) = a(t, \zeta_i), \quad y_i^b(t) = b(t, \zeta_i), \quad t \geq 0, i = 0, \dots, m. \quad (18)$$

A particular case of $m = 1$ with $d_1 = L$ corresponds to the availability of the boundary coupling only, i.e., no in-domain coupling points. The slave system is defined as

a copy of (17) with the in-domain injection of the concentrations y_i^a, y_i^b at respective locations $\zeta_i, i = 1, \dots, m$:

$$\partial_t \hat{a} = D_a \partial_z^2 \hat{a} - \hat{a} \hat{b}^2 + \alpha(1 - \hat{a}) \tag{19a}$$

$$\partial_t \hat{b} = D_b \partial_z^2 \hat{b} + \hat{a} \hat{b}^2 - (\alpha + \beta) \hat{b} \tag{19b}$$

$$\hat{a}(t, \zeta_i) = y_i^a(t), \quad t \geq 0, \quad i = \overline{0, m} \tag{19c}$$

$$\hat{b}(t, \zeta_i) = y_i^b(t), \quad t \geq 0, \quad i = \overline{0, m} \tag{19d}$$

$$\hat{a}(0, z) = \hat{a}_0(z), \quad z \in \Omega \tag{19e}$$

$$\hat{b}(0, z) = \hat{b}_0(z), \quad z \in \Omega. \tag{19f}$$

The main result is given in the following theorem.

Theorem 4 (Adapted from [12], Theorem 1) *Let the initial conditions $a_0, b_0 \in L^2(\Omega)$ and $\hat{a}_0, \hat{b}_0 \in L^2(\Omega)$ for both master and slave systems satisfy $0 \leq a_0(z), b_0(z), \hat{a}_0(z), \hat{b}_0(z) \leq 1$ for all $z \in \Omega$. Then, there exist constants $\sigma_i > 0, i = \overline{1, m}$ such that the states of the master and slave systems asymptotically synchronize in the L_2 -sense, i.e.,*

$$\lim_{t \rightarrow \infty} \|a(t, \cdot) - \hat{a}(t, \cdot)\| = \lim_{t \rightarrow \infty} \|b(t, \cdot) - \hat{b}(t, \cdot)\| = 0,$$

provided that $d_i \leq \sigma_i, i = \overline{1, m}$.

Proof A constructive proof is based on the observer design technique proposed in [12]. The resulting constants σ_i provide the largest distance between the spatial coupling locations that enable the global synchronization of two Gray–Scott models, i.e., synchronization for any the initial conditions satisfying Theorem 4. \square

Example 3

The master system (17) is considered in the domain $\Omega = (0, 2.5)$ with parameters $D_a = 2 \cdot 10^{-4}, D_b = 1 \cdot 10^{-4}, \alpha = 0.02, \beta = 0.047$ and the initial conditions $(a_0, b_0) = (1, 0)$ which are perturbed to $(a_0(z), b_0(z)) = (0.5, 0.25)$ at locations $z \in (0.925, 1.05)$ and $z \in (1.925, 2.05)$ and to $(a_0(z), b_0(z)) = (0.25, 0.75)$ at locations $z \in (0.25, 0.375)$. The slave system (19) is a copy of (17) with 6 uniformly distributed coupling locations. The initial conditions for the slave system are selected at the steady state $(\hat{a}_0, \hat{b}_0) = (1, 0)$. A comparison of the behavior of the master system and the slave one is given in Fig. 5, and the corresponding synchronization errors are provided in Fig. 6. The norms of these errors converge to zero as $t \rightarrow \infty$. Additionally, the synchronization errors for the cases of 3, 6, 11, 51, and 126 uniformly distributed coupling locations are depicted in Fig. 7. These numerical simulations demonstrate that the smaller gaps d_i between the coupling locations lead to smaller observation errors and their faster convergence to 0.

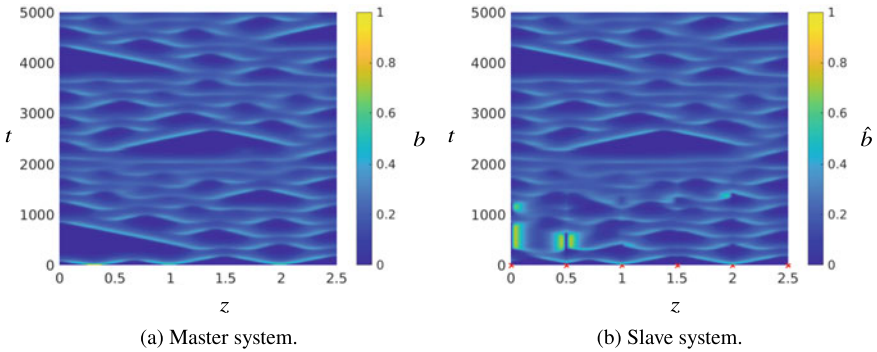


Fig. 5 A comparison of the evolution of concentrations b and \hat{b} . Six equidistantly located coupling points are marked with red crosses (right figure)

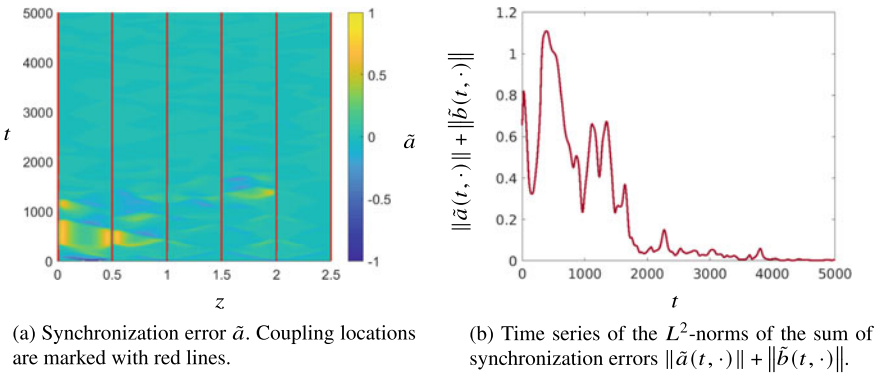


Fig. 6 Synchronization errors in case of 6 equidistantly located coupling points

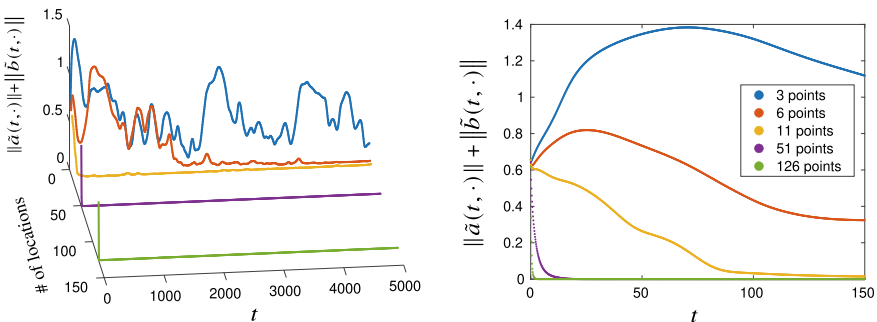


Fig. 7 Evolution of the sum of the L^2 -norms of synchronization errors depending on the number of equidistantly located coupling points for $t \in [0, 5000]$ (left figure) and $t \in [0, 150]$ (right figure). The slave system driven by 3 coupling points does not converge to the master system, whereas all other setups lead to the convergence of the synchronization error to 0 with the convergence rate increasing with the number of coupling points

4 Conclusion

The chapter provides synchronization analysis of three essentially different types of oscillator networks, namely, (i) Kuramoto networks with adaptive coupling, (ii) networks of Chua oscillators connected via nonlinear static output-feedback term, and (iii) the master-slave configuration of two Gray–Scott reaction-diffusion models. Sufficient conditions have been presented for the emergence of synchronization phenomena in these networks in terms of the dynamical properties of oscillators, characteristics of the couplings, and the interconnection topology of the networks. Theoretical results are supplemented with numerical case studies.

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 434434223—SFB 1461 and within the Research Unit FOR 2093. The authors are grateful to Karlheinz Ochs and Dennis Michaelis for fruitful discussions on the synchronization analysis of networks of Chua oscillators.

References

1. Argyris, A., Syvridis, D., Larger, L., Annovazzi-Lodi, V., Colet, P., Fischer, I., Garcia-Ojalvo, J., Mirasso, C.R., Pesquera, L., Shore, K.A.: Chaos-based communications at high bit rates using commercial fibre-optic links. *Nature* **438**(7066), 343 (2005)
2. Azar, A.T., Vaidyanathan, S., Ouannas, A.: *Fractional Order Control and Synchronization of Chaotic Systems*, vol. 688. Springer (2017)
3. Balaguer, I.J., Lei, Q., Yang, S., Supatti, U., Peng, F.Z.: Control for grid-connected and intentional islanding operations of distributed power generation. *IEEE Trans. Industr. Electron.* **58**(1), 147–157 (2010)
4. Berner, R.: *Patterns of Synchrony in Complex Networks of Adaptively Coupled Oscillators*. Springer Nature (2021)
5. Berner, R., Scholl, E., Yanchuk, S.: Multiclusters in networks of adaptively coupled phase oscillators. *SIAM J. Appl. Dyn. Syst.* **18**(4), 2227–2266 (2019)
6. Berner, R., Yanchuk, S., Schöll, E.: What adaptive neuronal networks teach us about power grids. *Phys. Rev. E* **103**(4), 042315 (2021)
7. Bowong, S., Tewa, J.J.: Practical adaptive synchronization of a class of uncertain chaotic systems. *Nonlinear Dyn.* **56**(1–2), 57 (2009)
8. Chen, F., Ji, G., Zhai, S., Wang, S., Zhou, S., Zhang, T.: Uncertain Chua system chaos synchronization using single variable feedback based on adaptive technique. In: *Proceedings of the 2012 IEEE International Conference on Information and Automation (ICIA-2012)*, pp. 196–199 (2012)
9. Chen, F., Zhang, C., Ji, G., Zhai, S., Zhou, S.: Chua system chaos synchronization using single variable feedback based on LaSalle invariance principal. In: *Proceedings of the 2010 IEEE International Conference on Information and Automation (ICIA-2010)*, pp. 301–304 (2010)
10. Doelman, A., Kaper, T.J., Zegeling, P.A.: Pattern formation in the one-dimensional Gray-Scott model. *Nonlinearity* **10**(2), 523 (1997)
11. Feketa, P., Perestyuk, Y.: Perturbation theorems for a multifrequency system with pulses. *J. Math. Sci.* **217**(4) (2016). <https://doi.org/10.1007/s10958-016-2988-6>
12. Feketa, P., Schaum, A., Meurer, T.: Distributed parameter state estimation for the Gray-Scott reaction-diffusion model. *Systems* **9**(4) (2021). <https://doi.org/10.3390/systems9040071>

13. Feketa, P., Schaum, A., Meurer, T.: Stability of cluster formations in adaptive Kuramoto networks. *IFAC-PapersOnLine* **54**(9), 14–19 (2021)
14. Feketa, P., Schaum, A., Meurer, T.: Synchronization and multicluster capabilities of oscillatory networks with adaptive coupling. *IEEE Trans. Autom. Control* **66**(7), 3084–3096 (2021). <https://doi.org/10.1109/TAC.2020.3012528>
15. Feketa, P., Schaum, A., Meurer, T., Michaelis, D., Ochs, K.: Synchronization of nonlinearly coupled networks of Chua oscillators. *IFAC-PapersOnLine* **52**(16), 628–633 (2019). <https://doi.org/10.1016/j.ifacol.2019.12.032>
16. Fell, J., Axmacher, N.: The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* **12**(2), 105 (2011)
17. Gray, P., Scott, S.: Autocatalytic reactions in the isothermal, continuous stirred tank reactor: Oscillations and instabilities in the system $A + 2B \rightarrow 3B$; $B \rightarrow C$. *Chem. Eng. Sci.* **39**(6), 1087–1097 (1984)
18. Ha, S.Y., Noh, S.E., Park, J.: Synchronization of Kuramoto oscillators with adaptive couplings. *SIAM J. Appl. Dyn. Syst.* **15**(1), 162–194 (2016)
19. Lee, K.J., McCormick, W., Ouyang, Q., Swinney, H.L.: Pattern formation by interacting chemical fronts. *Science* **261**(5118), 192–194 (1993)
20. Lee, K.J., McCormick, W.D., Pearson, J.E., Swinney, H.L.: Experimental observation of self-replicating spots in a reaction-diffusion system. *Nature* **369**(6477), 215–218 (1994)
21. McGough, J.S., Riley, K.: Pattern formation in the Gray-Scott model. *Nonlinear Anal. Real World Appl.* **5**(1), 105–121 (2004)
22. Mitropolsky, Y.A., Samoilenko, A.M., Kulik, V.L.: *Dichotomies and Stability in Nonautonomous Linear Systems, Stability and Control: Theory, Methods and Applications*, vol. 14. Taylor & Francis, London (2003)
23. Morgan, D.S., Kaper, T.J.: Axisymmetric ring solutions of the 2D Gray-Scott model and their destabilization into spots. *Phys. D* **192**(1–2), 33–62 (2004)
24. Muratov, C., Osipov, V.V.: Static spike autosolitons in the Gray-Scott model. *J. Phys. A* **33**(48), 8893 (2000)
25. Ochs, K., Michaelis, D., Jenderny, S., Szymendera, M.K.: Kuramoto model with Hebbian learning mimics spatial correlations causing an optical illusion. In: 2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 36–39. IEEE (2021)
26. Ochs, K., Michaelis, D., Roggendorf, J.: Circuit synthesis and electrical interpretation of synchronization in the Kuramoto model. In: 2019 30th Irish Signals and Systems Conference (ISSC), pp. 1–5. IEEE (2019)
27. Ochs, K., Michaelis, D., Solan, E., Feketa, P., Schaum, A., Meurer, T.: Synthesis, design, and synchronization analysis of coupled linear electrical networks. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**(12), 4521–4532 (2020)
28. Ohtsubo, J.: Chaos synchronization and chaotic signal masking in semiconductor lasers with optical feedback. *IEEE J. Quantum Electron.* **38**(9), 1141–1154 (2002)
29. Ouyang, Q., Swinney, H.L.: Transition from a uniform state to hexagonal and striped Turing patterns. *Nature* **352**(6336), 610–612 (1991)
30. Pearson, J.E.: Complex patterns in a simple system. *Science* **261**(5118), 189–192 (1993)
31. Pecora, L.M., Carroll, T.L.: Synchronization of chaotic systems. *Chaos: Interdiscip. J. Nonlinear Sci.* **25**(9), 097611 (2015)
32. Pecora, L.M., Carroll, T.L., Johnson, G.A., Mar, D.J., Heagy, J.F.: *Fundamentals of synchronization in chaotic systems, concepts, and applications*. *Chaos: Interdiscip. J. Nonlinear Sci.* **7**(4), 520–543 (1997)
33. Perestyuk, M., Feketa, P.: On preservation of the invariant torus for multifrequency systems. *Ukr. Math. J.* **65**(11), 1661–1669 (2014). <https://doi.org/10.1007/s11253-014-0887-x>
34. Pivka, L., Wu, C.W., Huang, A.: Chua's oscillator: a compendium of chaotic phenomena. *J. Frankl. Inst.* **331**(6), 705–741 (1994)
35. Rattenborg, N.C., Amlaner, C., Lima, S.: Behavioral, neurophysiological and evolutionary perspectives on unihemispheric sleep. *Neurosci. & Biobehav. Rev.* **24**(8), 817–842 (2000)

36. Rodrigues, F.A., Peron, T.K.D., Ji, P., Kurths, J.: The Kuramoto model in complex networks. *Phys. Rep.* **610**, 1–98 (2016)
37. Samoilenko, A.: Perturbation theory of smooth invariant tori of dynamical systems. *Nonlinear Anal. Theory Methods Appl.* **30**(5), 3121–3133 (1997)
38. Samoilenko, A.M.: *Elements of the Mathematical Theory of Multi-frequency Oscillations, Mathematics and Its Applications (Soviet Series)*, vol. 71. Kluwer Academic Publishers Group, Dordrecht (1991)
39. Schaum, A., Alvarez, J., Meurer, T., Moreno, J.: State-estimation for a class of tubular reactors using a pointwise innovation scheme. *J. Process Control* **60**, 104–114 (2017)
40. Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press (2018)
41. Tse, C., Lau, F.: *Chaos-Based Digital Communication Systems. Operating Principles, Analysis Methods and Performance Evaluation* (Springer, Berlin, 2004) (2003)
42. Vigelius, M., Meyer, B.: Stochastic simulations of pattern formation in excitable media. *PLoS ONE* **7**(8), e42508 (2012)
43. Wang, W., Lin, Y., Yang, F., Zhang, L., Tan, Y.: Numerical study of pattern formation in an extended Gray-Scott model. *Commun. Nonlinear Sci. Numer. Simul.* **16**(4), 2016–2026 (2011)
44. Wang, X.F., Wang, Z.Q., Chen, G.: A new criterion for synchronization of coupled chaotic oscillators with application to Chua’s circuits. *Int. J. Bifurc. Chaos* **9**(06), 1169–1174 (1999)
45. Wei, J., Winter, M.: Asymmetric spotty patterns for the Gray-Scott model in \mathbb{R}^2 . *Stud. Appl. Math.* **110**(1), 63–102 (2003)
46. Wu, C.W., Chua, L.O.: A unified framework for synchronization and control of dynamical systems. *Int. J. Bifur. Chaos* **4**(04), 979–998 (1994)
47. Wu, C.W., Chua, L.O.: Application of graph theory to the synchronization in an array of coupled nonlinear oscillators. *IEEE Trans. Circuits Syst. I: Fund. Theory Appl.* **42**(8), 494–497 (1995)
48. Yang, T., Chua, L.O.: Impulsive stabilization for control and synchronization of chaotic systems: theory and application to secure communication. *IEEE Trans. Circuits Syst. I: Fund. Theory Appl.* **44**(10), 976–988 (1997)
49. Zhang, K.: *Impulsive control of dynamical networks*. Ph.D. thesis, University of Waterloo (2017)
50. Zhang, K., Liu, X., Xie, W.C.: Impulsive control and synchronization of spatiotemporal chaos in the Gray-Scott model. In: *Interdisciplinary Topics in Applied Mathematics, Modeling and Computational Science*, pp. 549–555. Springer (2015)
51. Zheng, Y., Liu, Z., Zhou, J.: A new synchronization principle and application to Chua’s circuits. *Int. J. Bifurc. Chaos* **12**(4), 815 (2002)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Emulation of Learning Behavior in the Hippocampus: From Memristive Learning to Behavioral Tests



Christian Kaernbach, Thorsten Bartsch, Maximilian Brütt, Annika Hanert, Nick Diederich, and Martin Ziegler

Abstract Neuromorphic engineering is concerned with the emulation of biological learning and memory processes in hardware. The use of memristive devices, i.e., non-volatile memory devices, has given this field a significant boost in the last decade. However, most of today's efforts are aimed at the hardware implementation of artificial intelligence computational methods, while the emulation of biological computational methods is less pursued. In the latter, however, there is enormous potential for information technology. For this, however, network-dependent cognitive functionalities from biology must be identified and transferred to technical systems. In this chapter, we will show a possible approach. Using the hippocampus, which is the central structure of the mammalian brain responsible for learning new information, as an example, it is shown how elementary cognitive functions can be investigated by behavioral tests in humans and how their functionality can be broken down to the network dependent functionalities. Furthermore, it is shown how these functionalities can be technically reproduced in a memristive network model.

Keywords Neuromorphic computing · Memristive devices · Learning · Information processing · Cognitive systems

1 Introduction

Neuromorphic engineering goes back to Carver Mead, who in the 1980s used the then new silicon technology to emulate biological circuits of nervous systems [1]. His work was motivated by the hope to better understand the functions of the brain and to partly reproduce complex processes in technical systems. In addition to Mead, the physicist and Nobel laureate Richard Feynman and the mathematician John

C. Kaernbach · T. Bartsch · M. Brütt · A. Hanert
Christian-Albrechts-Universität Zu Kiel, 24143 Kiel, Germany

N. Diederich · M. Ziegler (✉)
Micro- and Nanoelectronic Systems, Technische Universität Ilmenau, 98684 Ilmenau, Germany
e-mail: martin.ziegler@tu-ilmenau.de

Hopfield were among the fathers of neuromorphic engineering. With the phrase on his blackboard at time of his death "what I cannot create I do not understand", Richard Feynman provided the motivation for generations of scientists in the field. In recent years, neuromorphic engineering has experienced an enormous technological interest, as it also enables applications for the field of artificial intelligence (AI) by providing hardware that can be better tailored to the needs of AI [2]. In this context, neuromorphic systems are now seen as hardware realizations of biologically-inspired computational architectures, and the field has expanded significantly, especially in the last decade [2, 5, 56].

Important components for the realization of biological computational architectures in hardware are memristive devices (also called memristor), which allow to emulate local biological learning paradigms in hardware, in particular to emulate synaptic plasticity [2–4]. Memristive devices are two-terminal electronic devices that change their resistance by applying electrical signals and retain the change even after the electrical signals are switched off [5]. Thus, this class of devices belongs to non-volatile memory devices and, as we will show, are ideal for use as artificial synapses in neuromorphic systems [6].

The change in coupling strength between neurons, i.e., the synaptic plasticity, is the essential building block for learning and memory processes in biological nervous systems and forms the cellular correlate of dynamic biological information processing [7], as we will consider in more detail in Sect. 2.1. Thus, the emulation of cellular learning forms by means of memristive devices is central to the development of neuromorphic systems and requires the emulation of critical neural information processes within memristive devices. However, this directly raises the question of what the requirements for the devices are and which functions are essential for the network-level processes of biological information processing and memory formation. Therefore, it is indispensable to acknowledge the neuronal network architecture and learning rules to design suitable memristive devices. This poses, however, further questions to biology as to how information processes operate in our brain. Here, it is the global network level that allows higher forms of learning and determines our behavior and actions [64].

Even though scientists have been researching for centuries how our brain works and what the secrets of information processing are, the functioning of the brain is only incompletely understood and we are far from a true comprehension of biological information processing [64]. However, in neurobiology, significant progress has been made in the last decades in exploring information processing and specifically memory formation at the mesoscopic level of different networks in our brain [8]. The progress made in the field of neurobiology has not only provided a deeper understanding of the computational processes underlying cognitive functions but also revealed the structural foundations of neural circuits and network architectures responsible for these processes. These developments offer an excellent opportunity to investigate the relationship between the structure and function of cognitive processes, with a particular focus on the memory function, as a model for technological emulation and modeling. In this context, the hippocampus, which plays a critical role in memory formation (as described in Sect. 2.2), has emerged as an especially promising candidate for the

emulation of neurobiological learning pattern with memristive devices. In addition to a large number of anatomical findings, hippocampus-associated and network-dependent memory functions have been identified (e.g. spatial navigation, pattern completion and pattern separation) that can be attributed to distinct neuronal networks in the hippocampus [9]. This allows the investigation of learning and memory functions on a global level in the form of behavioral tests in humans, which is subject of Sect. 3. However, this poses the challenge of using suitable tests to study forms of learning and to obtain as detailed information as possible about the performance of the hippocampus [9]. Starting from those functional relationships of global network-based learning forms, we will show in Sect. 4 how these can be reproduced within neuromorphic networks based on memristive devices. We will address particularly the question which device properties are important to mimic synaptic plasticity for the emulation of network-based learning forms.

Thus, in this book chapter, we want to address the original motivation of neuromorphic engineering and show how biological paradigms of network-dependent learning forms can be adapted for emulation with memristive devices. Furthermore, we demonstrate how global functionalities of memory formation can be extracted and reconstructed in a way that they can be described by simple network structures. The aim is to show how a conceptual bridge that can be built from the multi-dimensional global network level to the cellular level. For this purpose, we follow the systematics shown in Fig. 1 and demonstrate how cellular learning forms can be emulated so that it can be transferred within a multidimensional network level, i.e., how one can move from the microstructure of synapses and neurons to a mesoscopic structure of only a few connected neurons to a macroscopic structure of a nervous system.

2 Neurobiological Learning Principles

2.1 Cellular Learning Paradigms

Basis of learning and memory are activity-dependent changes of the connections between individual neurons. These changes are achieved by a temporary strengthening or diminution of synaptic connections and are referred to as **synaptic plasticity** [7]. Synaptic plasticity describes an increase in the efficiency of a synaptic transmission through repeated or persistent activity of the connected input cells.

Long-term potentiation (LTP) is the classic paradigm of synaptic plasticity and is regarded as the cellular basis for memory formation. This cellular phenomenon was first described by Bliss and Lømo in 1973 in the CA3¹ area of the hippocampus [10]. However, most experiments on LTP have been performed at the junction between

¹ CA: Cornu Ammonis.

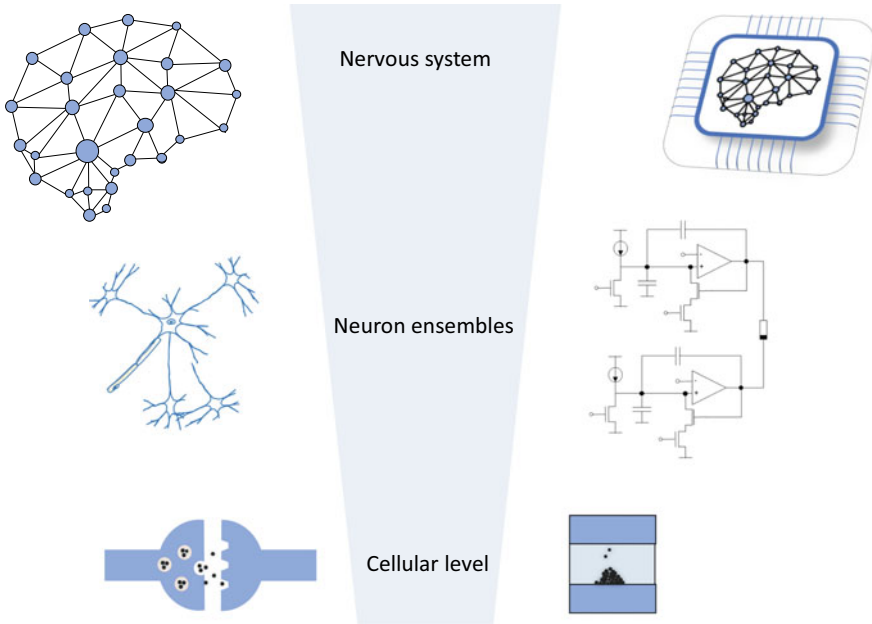


Fig. 1 Modelling of memristive learning behavior: emulation of cellular learning forms and their integration into a neuronal circuit architecture to mimic cognitive behavior. For this purpose, a bridge must be built from the cellular level to the multidimensional network level

CA3 and CA1 areas of the hippocampus [65]. Thereby, a short, high-frequency stimulation of axons² connecting areas CA3 and CA1 can elicit a sustained enhancement of the excitatory postsynaptic response potential (EPSP) (see Fig. 2). The transmission of these potentials from CA3 to CA1 occurs via the neurotransmitter glutamate to voltage-dependent NMDA (N-methyl-D-aspartate) receptors, whereby the glutamate influx triggers the EPSP at the synapse between Schaffer collaterals and CA1 pyramidal cells. This high-frequency stimulation triggering an EPSP leads to a persistent strengthening of CA1 synapses and thus producing a long-lasting increase in signal transmission between two neurons. This elicited synaptic plasticity makes a crucial contribution to memory formation regarding LTP acting as a surrogate of information storage in the central nervous system (CNS) [11]. In this process of input-dependent increased cellular excitability, the synaptic transmission efficiency is increased over hours to days via a functional amplification and weighting of synaptic connections. In contrast to LTP, a **long-term depression (LTD)**, in the sense of a reduction in the strength of synaptic transmission, can be evoked by low-frequency stimulation. Four principal properties underlie LTP/LTD (see Fig. 2): (i) **input specificity**, i.e., LTP only occurs at this synaptic connection, (ii) **associativity**, i.e., simultaneous activity at a stronger synaptic connection enables LTP at an associatively linked weaker

² Schaffer collaterals.

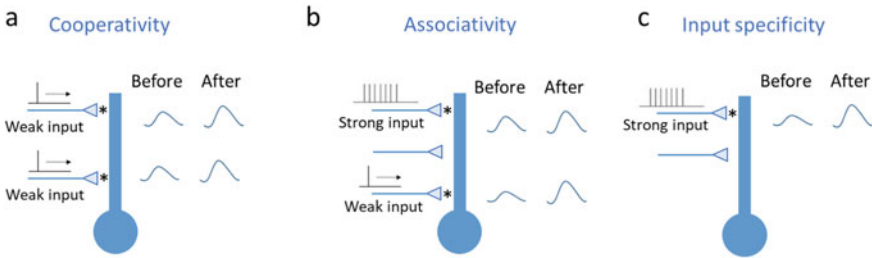


Fig. 2 Properties of long-term potentiation (LTP) in CA1 of the hippocampus exemplified by a single pyramidal cell receiving strong and weak stimuli: **a** Cooperativity, **b** Associativity, and **c** Specificity. The postsynaptic response potential (EPSP: Excitatory Postsynaptic Potential) before after long-term potentiation is outlined on the right of each figure. Adapted from [7]

synapse, (iii) **cooperativity**, i.e., LTP can be evoked by cooperative activity of input signals, and (iv) **persistency**, synaptic transmission is increased over hours to weeks.

NMDA receptors associated with LTP are also thought to be closely linked to the formation of higher-order network activity i.e., hippocampal place cell representations leading to cognitive maps, and theta rhythm [12, 13]. It has been shown that learning indeed directly induces LTP processes in CA1 [11]; also, an impediment to the maintenance of LTP results in a disruption of spatial memory of already stored information [14]. Importantly, these fundamental mechanisms of memory formation in the CA1 region are also putatively disrupted in human neurological memory disorders, such as transient global amnesia and Alzheimer’s disease [15].

Another correlate of synaptic plasticity is **spike-time-dependent plasticity** (STDP) and **paired-pulse facilitation** (PPF). In these plastic processes, augmentation of synaptic transmission is achieved by coupling input signals within a critical time window of a few to several hundred milliseconds. A detailed description of STDP is given in Sect. 4.

2.2 Network Dependent Learning Paradigms

The previously discussed mechanisms of synaptic plasticity are the basis for memory and learning in the brain. However, memory and learning, i.e., memories and events within a temporal-local framework, require more complex network structures. The biological substrates of various types of memory can be assigned to different areas of the brain [16]. In particular, episodic memory content of our personal experiences is critically reliant on the hippocampus [17, 18]. An effective memory system (i.e. minimal interference and maximal capacity) must provide at least two cognitive functions: first, the rapid storage of experiences as individual events (thereby avoiding the ‘overwriting’ of similar information—‘catastrophic interference’), and second, the retrieval of those memories, when similar events are encountered [19].

In this context, two functions of hippocampus, namely pattern separation and pattern completion are highlighted here, as they are essential cognitive processes for the encoding and retrieval of episodes [19, 20]. Theories regarding pattern separation and completion processes that derived from computational approaches have been consistently supported by studies in rodents (see 21 for a review). Recent data recorded in humans using modern imaging techniques show that pattern completion and separation play a critical role in human (and other mammalian) learning and are subject to aging or degenerative processes, such as in Alzheimer's disease [21, 22].

A schematic representation of pattern separation and pattern completion is shown in Fig. 3a. Pattern completion allows incomplete representations to be 'completed' by previously stored representations. Pattern separation is the process of representing and storing similar representations uniquely and non-overlapping (orthogonalized). The biological significance of this information discrimination is that new information does not overwrite similar, previously stored information. Overwriting would lead to catastrophic interference and ultimately to no new learning.

A schematic representation of the network topology of the individual hippocampal fields is shown in Fig. 3b. The hippocampal dentate gyrus (DG) and the CA3 field have been attributed the function of pattern separation and completion, respectively. Recurrent axon collaterals serve as the basis of pattern completion, which in turn enable auto-association networks [23]. In detail, information is routed from the entorhinal

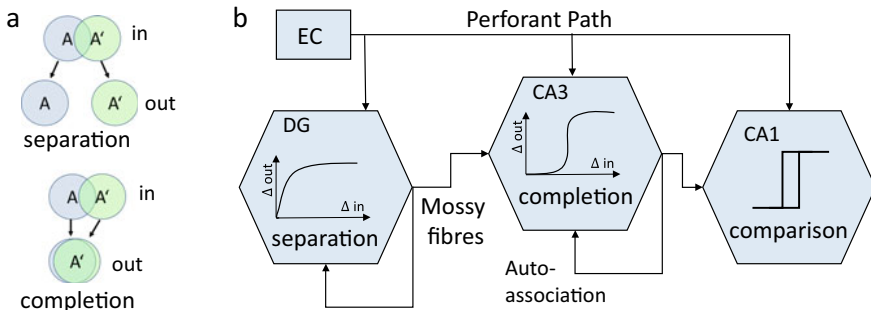


Fig. 3 Schematic representation of pattern separation and pattern completion in the hippocampus. Pattern separation provides a distinct representation of similar and overlapping information, whereas pattern completion unifies similar and overlapping representations. **b** Formalized circuit of hippocampal function. The input via the tractus perforans is the input pathway into the hippocampus. DG: Emergence and self-organization of patterns of EC mediated activity by means of a heteroassociative network (completion of one piece of information by presentation of information from another category). CA3: Excitatory recurrent connections (12,000 collaterals/CA3 cell) mediate autoassociative encoding and recollection of episodic memory content in the sense of completing a memory content upon partial presentation of information. In addition, direct projections from EC to CA3 (2600 connections/CA3 cell) provide presentation of known information. Schaffer collaterals encode and mediate associations between CA3 activity and CA1 activity (directly from EC). In addition, recurrent connections from CA3 to DG exist. CA1: direct input from EC forms new representations, which are compared with the predicted representations from CA3 (comparator function). Adapted from [21]

cortex (EC) directly to the subnetworks in the DG and CA3 via the perforant path. Additionally, these subnetworks receive further projections via their own recurrent collaterals. The network in CA3 additionally receives information about the mossy fibers from the granule cells of the DG. In particular, the distinct recurrent network in CA3 serves as an auto-associative network that enables the completion of previously stored information against the background of the presentation of incomplete stimuli. The mossy pathway from the DG to CA3 serves to establish separate pattern representation in the context of new learning and to reduce interference, whereas the direct input from EC enables the retrieval or presentation of known information. Animal experimental data provide evidence that dentate gyrus networks are necessary for pattern separation, whereas CA3 networks are critical for pattern completion [21]. Finally, the CA1 compares new representations from EC with the predicted representations from CA3 and has therewith a comparator function (cf. Fig. 3b [39]).

Theories based on computational models of pattern separation processing and experimental rodent studies that measure the behavioural outcome on the basis of hippocampal place cell remapping both corroborate that the DG/CA3 network is critically involved in pattern separation [24]. Studies in humans support this finding by means of fMRI investigations that measured the activity of hippocampal areas during behavioural paradigms that tax pattern separation [22, 25].

This network function is further characterized by the fact that the represented network circuits are arranged multilayered in three-dimensional space to cause an increase in computational capacity. The pattern separation function is favored by adhering to the principle of sparse connectivity from the DG to CA3 cells (46 inputs per CA3 cell) [9]. Furthermore, LTP for augmentation or deaugmentation of synaptic transmission is required for the storage of information in these networks (see Fig. 3b). Also, these networks show a significant fault tolerance (graceful degradation), in the sense that errors due to lost components can be compensated by the distributed representation.

3 Investigating Hippocampal Functions in Animals and Humans

High demands are placed on our memory system in daily life due to constantly changing environmental conditions that require a continuous mnemonic processing. As shown in Sect. 2.2, this functionality is mapped in the hippocampus by the cognitive network-dependent functions of pattern completion and pattern separation. These functions reduce interference between memories and generalizations about similar events, thus contributing to memory formation [21]. The challenge, however, is the study of these functions and, in particular, the identification of network-dependent mechanisms with respect to memory performance, as sketched in Fig. 1. In particular, it is challenging to elucidate the role of hippocampal subfield processing in pattern separation and -completion in humans.

In this section, we will show to what extent this can be optimized by applying behavioural tests addressing the human non-semantic memory. Here, we will first discuss classic memory tests and especially the Mnemonic Similarity Task (MST, [27]). We further show how behaviour in a pattern separation task can be studied in two human hippocampal lesion models: first, the selective CA1 subfield lesions in amnesia and, second, preferential neurodegeneration in DG/CA3 subfields in a patient cohort with a rare inflammation of the brain (LGI1 encephalitis). These natural hippocampal lesion models help to examine and understand a causal relationship between anatomical structures and pattern separation performance. The development of an alternative to the MST using sensory stimuli instead of depictions of objects, the Visual Sensory Memory Task (VSMT), then concludes this section.

3.1 Mnemonic Similarity Task

Classic standard recognition tests feature only two types of stimuli, old and new stimuli (see, e.g., 26). These two types of stimuli are typically called “Repeats” (for the old stimuli) and “Foil” (for the new stimuli). Inspired by concepts from computational neuroscience, new concepts of memory subfunctions came up, including pattern separation and completion. Behavioural pattern separation in humans is commonly measured by means of specific match-to-sample tasks that include a third type of stimulus, so-called “Lures” [21, 22]. Lures are stimuli that are similar to but not identical with old stimuli. In due course participants are now given three response possibilities, namely, “old”, “new”, and “similar”.

An established test comprising lures is the Mnemonic Similarity Task (MST, [27], see Fig. 3.1). This memory test comprises an encoding phase of items of everyday objects and a retrieval phase (see Ref. [27] for examples). During encoding, the participant is asked to classify these stimuli as either indoor or outdoor objects. At that point in time, the participant is not aware that they will have to remember the stimuli at a later point in time. During recall, the participant is presented with old, new, and similar stimuli and given the corresponding three response possibilities. Supposedly, those similar lures tax hippocampal pattern separation so that correctly identifying lures as similar suggests successful pattern separation abilities, whereas confusing similar lures with their corresponding targets would indicate a bias towards pattern completion. The advantage of the MST is the usage of everyday objects that allows the application in patients with neurological disease and aging participants.

3.2 Human Hippocampal Lesion Models

Memory impairment is commonly caused by an impairment of hippocampal functions due to neurological disorders or aging [28, 29]. In the following section, the aims are to show the mechanistic contribution of the human hippocampus to pattern

separation and to demonstrate the neurobiological processes within the hippocampus during consolidation of mnemonic information. Consolidation refers to the (time-dependent) stabilization of a memory after initial acquisition (encoding) into a long-lasting form. To study these subjects, we investigated natural hippocampal lesion models in memory impaired patients with selective hippocampal damage.

Although regional neural activity of the hippocampus can be tested using functional MRI (see Ref. 25 for an example), those studies lack information regarding mechanistic aspects of causality about the subfield-specific computational processes and the causal role of hippocampal structure and its function. Therefore, we examined two hippocampal lesion models, where specific hippocampal subfields are impaired due to neurological diseases, as shown in Fig. 5.

In Study I [38], the transient global amnesia (TGA) served as a model for a selective disruption of hippocampal CA1 neurons. A TGA is characterized by an abrupt cognitive deficit limited to an anterograde amnesia in the acute phase that resolves within 24 h. Typically, focal lesions restricted to area CA1 can be detected in MR-Imaging [30, 31]. Hence, it was suggested that a selective impairment of CA1 during TGA causes a deficit in pattern separation.

Study II [43] aimed at further elucidating the causal role of hippocampal subfield contributions to pattern separation. Here, an extremely rare patient cohort was studied. Patients with an inflammatory brain disorder positive for LGI1 antibodies were examined. These patients develop limbic (-hippocampal) encephalitis with persisting memory deficits (so-called LGI1 encephalitis; 32, 33) and have structural damage to the hippocampal system [34, 35]. The aim of this study was to investigate the pattern separation performance from DG and CA3 that are predominantly affected by neuroinflammatory changes due to LGI1 encephalitis [36, 37]. The hypothesis suggested that inflammatory lesions within the DG and CA3 subfields correlate to hippocampal pattern separation.

The role of the hippocampal CA1 networks in pattern separation and recognition memory.

Here, we investigated the critical relay function of CA1 neurons in pattern separation performance using TGA as natural lesion model of a CA1 deficit [38]. Information processing within CA1 is characterized by the comparison of dual afferent projections—from EC via the perforant path and from CA3 via the Schaffer collaterals [19, 23, 39]. The integration of those two projections within CA1 is assumed to facilitate an immediate retrieval and consolidation in neocortical long-term stores [9]. With regard to the contribution of CA1 neurons to pattern separation processes in humans, our results complement the current concept developed in computational models and experimental animal models [40]. Pattern separation function is relayed and facilitated by the DG that is in turn assumed to decorrelate overlapping memories by sparse coding of neural signals from EC to CA3 [9]. For the transfer of mnemonic information to extra-hippocampal areas, CA3 projects to area CA1, the main hippocampal output area [39, 41]. The selective CA1 dysfunction caused an impairment in the transmission of the then-separated information from the DG/CA3 network to the neocortex resulting in ineffective pattern separation performance on

the behavioral level [38]. This suggests that CA1 does not perform pattern separation on the neural level per se, but relays memory information from DG/CA3 networks to neocortical areas. Complementing the results of our studies, this assumption applies to the result of the dependence of pattern separation performance on hippocampal DG volume, but a weaker association to the volume of CA1 [38, 43]. The role of recognition memory by the CA1 volume can be conceptualized by the central position of CA1 as the functional readout of hippocampal circuit projections. By measuring hippocampal volume in LGI1 encephalitis, we showed that the volume of CA1 was the best predictor of recognition memory [43]. The integration of the dual afferent projections from EC and CA3 facilitates the restoration of a memory trace and thus recognition of an environmental cue [42]. Together, these results corroborate the view that CA1 is involved in both pattern separation and recognition memory processes. The functional readout of the hippocampal circuit to neocortical areas involved in hippocampus-dependent memory formation is thus highly dependent on the dynamics within the subnetworks.

Dentate Gyrus Networks in Pattern Separation

Theoretical models state that the DG performs pattern separation by the transformation of overlapping input patterns into distinct, non-overlapping representations [17, 20]. Hanert et al. [43] showed evidence that the DG volume was the best predictor of behavioral pattern separation compared to the volume of regions CA2/3 and CA1. This functional model of the DG in pattern separation processes has been confirmed by electrophysiological recordings in rodents [24]. Evidence for a separation-like activity within the DG in humans has been provided by high-resolution fMRI during a mnemonic similarity recognition paradigm [25]. These results complement our findings by presenting a structure–function relationship between pattern separation and the DG. These findings are also in accordance with previous studies that demonstrated a greater volume of the DG to be associated with a better discrimination of overlapping items [44, 45].

In summary, our results in neurological patients clarified the role of the human hippocampus and its specific subfield contributions to pattern separation and memory consolidation. It was found that the hippocampal DG as well as intact CA1 neurons are essential for pattern separation in humans. We also demonstrated that pattern separation was best predicted by the volume of the DG, whereas recognition memory was stronger associated with the volume of CA1. However, we also found that an impairment due to a lesion restricted to CA1 neurons compromised pattern separation performance. These results emphasize and refine the current view on hippocampus-dependent memory processing within the hippocampal DG as a critical ‘pattern separator’, and CA1 essentially involved in transferring the mnemonic output to neocortical long-term stores.

3.3 The Visual Sensory Memory Task (VSMT)

The hippocampus plays an important role when information is transferred from short-term to long-term memory. In principle, that could be any sort of information. Here we will focus on the difference between categorical-semantic and acategorical-sensory, non-semantic information.

During information processing, the original sensory information is possibly enriched with semantic content: The letters of the words of a poem are initially nothing but a pattern of black and white areas. This sensory black-and-white pattern is processed to be interpreted as letters, the letters form words, and the words transfer meaning. Likewise, the sensory colored patterns of the stimuli of the MST (see Fig. 4.) gain meaning as they are processed, interpreted, and recognized. The four-leaf clover, for instance, could be seen as a sign of good luck by some persons. Semantic content is organized in hierarchical categories: A four-leaf clover is part of the category clover, which is part of the category plants.

In the presence of such semantic content, the original sensory form of the input loses its importance; memorizing this information is based more on the higher representations rather than on the lower ones [46]. The exact amount and content of the meaning associated with such semantic stimuli might, however, vary enormously between persons; a florist might associate quite different things when seeing a picture of a four-leaf clover than a rabbit breeder. This introduces a massive variation of retention performance which is not under the control of the experimenter.

Here we opted for non-semantic sensory information. This type of information shows behavioral characteristics identical to those known for semantic information [47]. It can thus serve as an example for the type of information typically treated by the

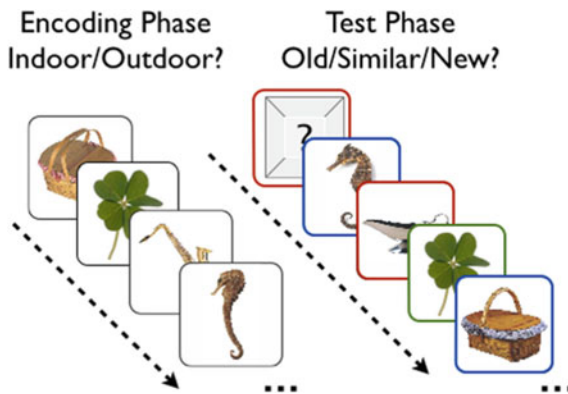


Fig. 4 A schematic presentation of the Mnemonic Similarity Task. The four-leaf clover is present in both the encoding and the testing phase; it is an example of a “repeat” stimulus. The oilcan is not present in the depicted encoding stimuli; it might be a new stimulus, a “foil”. The seahorse shown in the test phase is similar to but not identical with the seahorse shown in the encoding phase; it represents a “lure” stimulus. The same holds true for the picnic basket

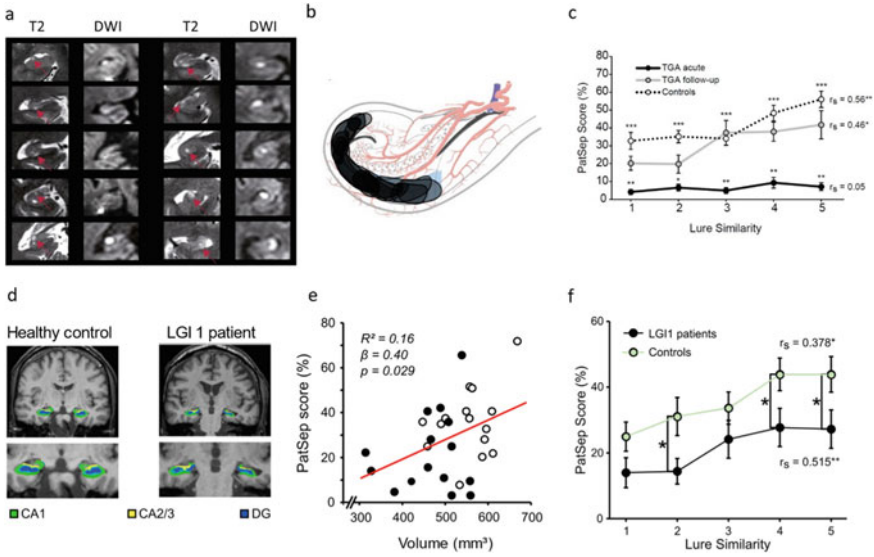


Fig. 5 Human hippocampal lesion models. **a–c** Study 1. **a** T2-weighted and diffusion-weighted magnetic resonance imaging (DWI) show representative lesions confined to the CA1 area of the hippocampus in TGA patients. **b** Overview of all DWI/T2 lesions of TGA patients participating in Study 1 transferred to an anatomical template of the cornu ammonis demonstrating selective CA1 lesions [38]. **c** TGA patients showed a strongly impaired pattern separation performance measured by the MST regarding all lure similarity levels (black). After recovery from TGA, the same patients’ performance (grey) was equal to healthy control participants (white). **d–f** Study 2 [43]. **d** T1-weighted magnetic resonance imaging shows hippocampal subfield segmentation in a representative healthy control participant compared with a patient with volume loss in DG and CA3 regions due to LGI1 encephalitis. **e**) Higher volume of hippocampal DG predicts better pattern separation performance measured by the MST in patients (full circle) versus healthy controls. **f**) Patients with reduced hippocampal DG and CA3 volume show significant gradual increases in pattern separation performance from high to low similarity. However, the healthy control group exhibited significantly better performance in pattern separation from highly similar lures to lures with low similarity to targets. * $p < 0.05$, ** $p < 0.01$

hippocampus. Void of semantic information organized in varying hierarchical categorical structures, it eliminates the variance in retention performance due to varying amount, type, and linkage of associations. Furthermore, with categorical stimuli, also similarity judgements are subject to a multitude of uncontrollable influences. A zoologist might judge the similarity of two seahorse depictions differently from a layman. A first tentative neuromorphic model of basic hippocampus function should not have to deal with high-level representations acting on the perceived similarity of two stimuli. This was a major reason to consider sensory stimuli.

A second and just as important reason was the parametrical control of similarity by mixing sensory stimuli without loss of validity. Sensory stimuli can be mixed to create intermediate stimuli with any desired degree of similarity. This is of specific importance for studies sizing pattern completion and separation, and cannot be done

with categorical stimuli as provided by the MST. One cannot mix the picture of a four-leaf clover with the picture of an oil can to get a picture of something that is somewhere between a four-leaf clover and an oilcan and reliably identifiable as such.

Sensory memory is usually reported in two variants. For instance, Cowan [48] describes short and long auditory stores, with short auditory stores keeping information for up to 300 ms, and long auditory stores retaining auditory information for at least several seconds. Please note that the lifetime of the so-called long stores corresponds to the lifetime of classic short-term memory. There is, however, ample evidence that sensory information may be stored for even longer periods, paralleling the classic findings for long-term memory (for a review see Ref. 49). In other words, there is no reason not to use sensory stimuli in classic memory experiments; as to the new paradigms that came along with the concepts of pattern separation and completion, sensory stimuli offer the advantage of parametric control of similarity.

The Visual Sensory Memory Task (VSMT) assesses pattern separation and completion in a way similar to the MST, using sensory stimuli instead of depictions of objects. In addition, it comes with an analysis different from that presented by the authors of the MST: Instead of subtracting certain entries of the 3×3 response matrix, it is analyzed in terms of classic Gaussian Signal Detection Theory [50]. Gaussian SDT assumes that a monitored quantity—in our case familiarity—is distributed normally on a decision axis, with equal standard deviations for the different stimulus classes (foils, lures, repeats). Only the means differ, with repeats having the highest familiarity, and foils the lowest. Decisions for a specific response—in our case “old”, “similar”, or “new”—are made on the basis of criteria on this decision axis. The differences between the means of the distributions of stimulus classes with a certain familiarity (lures, repeats) against the stimulus class with the lowest familiarity (foils) are measures of the so-called sensitivity (d'). The possibility to quantify the physical correlation of lures with their respective targets allows to establish a psychophysical relationship between physical correlation on the one hand and psychological similarity expressed in d' on the other hand.

Figure 6a shows two examples of the stimuli used in the VSMT, with the second example being a lure to the first one. The stimuli are composed from grayscale visual pink noise, i.e., visual noise that has a $1/f$ distribution of spatial frequencies. The use of pink noise was inspired by the finding that natural images show a $1/f$ distribution of spatial frequencies, and that cortical cells are tuned to exactly this distribution of spatial frequencies [51]. This pink noise is then subjected to a Gaussian envelope, smoothening out the grayscale variations of the noise at the borders of the stimuli and preserving them most prominently in the center of the stimuli. This measure should prevent memorizing strategies based on visual artifacts occurring at the borders of the stimuli.

Analyzing the lure and target sensitivities expressed as d' values as a function of the physical lure-target correlation we found a psychophysical law which links the sensitivity to the quartic correlation of lure and target grayscale values: $d' \propto r_{LT}^4$, where r_{LT} is the correlation of lure and target. Figure 6b illustrates this result for two different mixing algorithms, called blending and pixel substitution. Figure 6c shows exemplary behavioral data, with the lure correlations already chosen such that

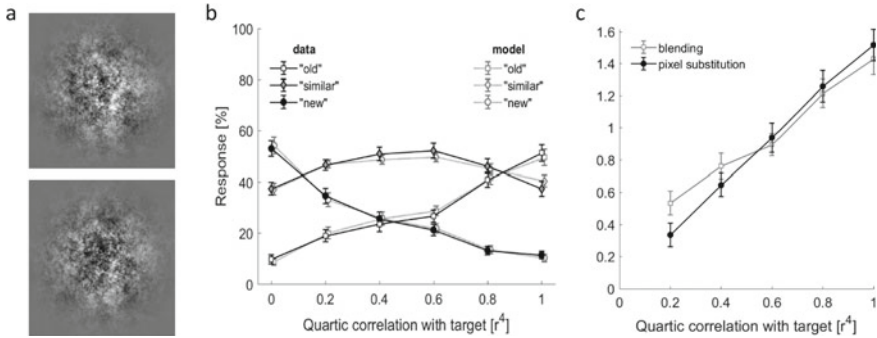


Fig. 6 The Visual Sensory Memory Task (VSMT): **a** Two example stimuli of the VSMT. The lower stimulus is a lure to the upper stimulus: the physical correlation of the grayscale values of the two stimuli is $r^2 = 0.69$, i.e., 69% of the variance of the grayscale values of one of the two stimuli is due to the variance of the grayscale values of the other stimulus. **b** Sensitivity of lures ($r < 1$) and targets ($r = 1$) as a function of lure-target correlation. If the latter is expressed as quartic correlation (r^4), the measured sensitivities approach a linear relation. **c** Exemplary behavioral results of the VSMT. With increasing correlation between lures and targets, the frequency of the response “new” decreases, and the frequency of the response “old” increases. The frequency of the response “similar” follows a nonmonotonic course, increasing up to a maximum at about $r^4 = 0.5$ and decreasing thereafter. The symmetry of these results demonstrates that the quartic correlation of the physical grayscale values of the pink noise stimuli is a good predictor of the psychological similarity, expressed in sensitivity values d'

they are distributed equally-spaced on a quartic correlation axis in order to obtain equally-spaced sensitivities.

The process of *pattern separation*, i.e., the differentiation between similar stimuli, can be examined quantitatively when using sensory stimuli, based on familiarity judgements. However, in order to test *pattern completion* one needs an indicator of identification. This can be realized by having the participants name the stimuli: Completing a fragmentary pattern can be verified if the pattern is identified by telling its name.

So, we tested the ability of participants to learn names of the sensory stimuli and to correctly identify them immediately after learning, and after one week. Figure 7 shows the naming performance directly after the learning, and one week later. Naming performance is calculated for “hits”, i.e., trials, where participants recognized the stimuli correctly as targets, as well as for “misses”, i.e., trials, where participants claimed the (old) stimuli to be new but were nevertheless requested to guess a name. Obviously, the naming performance is better for hits than for misses, but it is still well above chance level for misses. Most importantly, naming performance does not decrease strongly after a week.

The latter data illustrate that sensory stimuli can be used to test the transition of short-term memory to long-term memory, which is the principal role of hippocampus in human memory. Devoid of the influence of higher-level semantic representations and with the additional advantage of the possibility to construct stimuli with any

Fig. 7 Naming performance for sensory stimuli directly after learning (Test 1) and after one week (Test 2). Naming performance is established independently for hits and misses



desired degree of similarity, they provide an excellent opportunity to directly compare neuromorphic models of the hippocampus with human behavioral performance.

4 Neuromorphic Investigation Pathways

In the last subsection we have shown the special importance of pattern completion and pattern separation for memory formation and their significance for higher cognitive functionalities of the brain. For the emulation of these functionalities within neuromorphic systems, the bridge to cellular learning paradigms is important. In particular, this requires reconstructing cellular learning and network architectures in a way that ensures global functionality through local building blocks. One possible approach will be presented in this subsection, with special use of memristive devices. Based on the Hebbian learning theory, we will show how this theory can be applied to memristive devices and how it can be used to construct network architectures that enable the emulation of hippocampal learning forums.

4.1 Hebbian Learning

As early as 1949, the psychologist Donald Hebb suspected a temporal relationship between local neuronal activity and the change in synaptic connectivity there. Today, this relationship is known as Hebb's learning rule and it reads as follows [52]: "When an axon of cell A is near enough to excite cell B or repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." Now often and somewhat more simply stated this sentence says: "neurons that fire together wire together." Thus, the temporal correlation of signals determines the strengthening of

synaptic connections. In the original sense, this only includes learning processes, whereby a similar rule also applies to unlearning. An extension made to this comes from Stent in 1973 and reads as: “Neural connections weaken when they are inactive at the same time that the postsynaptic neuron is active, or they will weaken if they are active, but consistently fail to cause the postsynaptic cell to fire an action potential” [53].

While Hebb’s learning rule gives an intuitive approach to cellular learning mechanisms, a mathematical approach is needed for a formal quantitative model description. Following Ref. [54], this can be obtained via the characteristics of synaptic plasticity. As mentioned above, a fundamental feature of synaptic plasticity is **input specificity**. This means that the change in synaptic weight ω_{ij} between pre-neuron j and post-neuron i depends only on local variables and thus only on information available at the synapse site. This can be formulated mathematically as follows [54]:

$$\frac{d\omega_{ij}}{dt} = F(\omega_{ij}, A_j, A_i)$$

Here the pre- and post-synaptic activities A_j or A_i are given by the voltage dependent function $A_{i(j)} = \mathbf{g}(u_{i(j)})$, where $F(\omega_{ij}, A_j, A_i)$ is a function dependent on the learning process, which we will specify in more detail next. An important property of cellular learning is **cooperativity**. This means that the synaptic weight changes when neurons are active simultaneously. This allows a simple ansatz for the function $F(\omega_{ij}, A_j, A_i)$:

$$F(\omega_{ij}, A_j, A_i) = \alpha A_j \cdot A_i$$

where α is a positive constant ($\alpha > 0$) named as learning rate. However, it is useful to make α weight-dependent, so that one excludes unlimited weight growth:

$$\alpha(\omega_{ij}) = \gamma \cdot (1 - \omega_{ij})$$

Since γ is a positive constant and if ω_{ij} is normalized between zero and one, a weight saturation is therewith obtained. The third important property of cellular learning is **associativity**. This allows local learning rules to be transferred from cellular level to a multidimensional network level [54]. Furthermore, competition considers the limitation of shared synaptic resources, which leads to the fact that weights can only grow at the expense of other synaptic weights. One commonly used and simple possibility to implement competition is an adaptive threshold voltage (or sometimes referred to as a sliding threshold) for post-neuron activation [54, 55].

4.2 Memristive Hebbian Learning

Neuromorphic electronics is essentially concerned with the implementation of biological information processing within electrical circuits and systems [1]. Therefore, the emulation of synaptic plasticity via electronic devices is an important aspect for the hardware realization of neural circuits. In this context, memristive devices have shown their potential in recent years and are increasingly used as substitutes for synapses in artificial neural networks [2]. In the following, we will show how they can be described using Hebbian learning theory [56]. First, however, a short formal presentation of the concept of memristive device is given.

In their simplest form, memristive devices (also known as memristors) consist of a metal insulator metal structure as sketched in Fig. 8a. The insulator is the memristive layer which changes its resistance when an electrical voltage is applied. Thus, this device has a memory effect for electrical signals, which also explains the name as memory resistor. The idea of memristive devices goes back to Leon Chua, who first formally described the class of electronic devices within a mathematical theory [57]. Following Chua, the current–voltage characteristics (I - V curve) of memristive devices can be described by the following set of differential equations:

$$I = G(x, V, t) \cdot V \text{ with } \frac{dx}{dt} = f(x, V, t),$$

where x is called a state variable and ranges between zero and one, while $f(x, V, t)$ is a function describing the state dynamics under external voltage stimuli. The conductance $G(x, V, t)$ is named as memductance and can be linked to the state variable via [58]

$$G = x \cdot G_{on} + (1 - x) \cdot G_{off}.$$

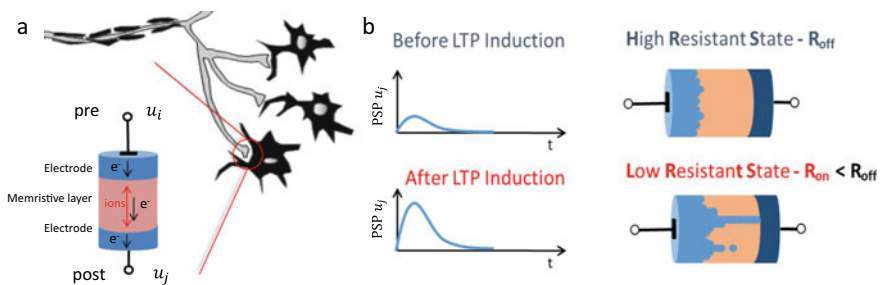


Fig. 8 Emulating synaptic plasticity with memristive devices: **a** schematic representation of a memristive device for emulating synaptic functionality. The change in atomic structure in the memristive layer results in a change in the resistance value of the device. **b** Emulation of long-term potentiation (LTP): after induction of LTP, there is a long-lasting increased depolarization of the post synaptic potential (PSP) upon activity of the pre neuron. This can be reproduced by the bifurcation of filamentary conducting structures within a memristive device. (Adapted from [56])

Here G_{on} and G_{off} are the maximum and minimum conductance of the device, respectively. In 2007, scientists at HP labs made a connection between this theory, developed by Chua in the 1970s, and the resistive switching properties of resistive materials [59]. In this class of materials, a change in the local atomic configuration results in a change in the resistance of the devices [2]. Thus, the function $f(x, V, t)$ describes the voltage-driven atomic reorganization within the memristive layer [59].

The emulation of synaptic plasticity using memristive devices is shown in Fig. 8b. When a high electrical voltage is applied, the atomic configuration within the memristive layer changes. Shown here is the injection of metal atoms into the insulating memristive layer building a metallic bridge between the two electrodes of the device. This leads to a permanent change in the resistance of the device and can emulate long-term potentiation (LTP). Furthermore, by applying a sufficiently negative voltage, the metallic bridge can be broken. This increases the resistance of the memristive device again, which can be identified as a long-term depression (LTD). A special property of memristive devices is that they also have gradual resistance changes, which can be used to simulate Hebb's learning. Thus, using the Hebbian learning theory and the Memristor equation, the following relationship can be identified:

$$\frac{d\omega_{ij}}{dt} \rightarrow \frac{dx_{ij}}{dt} = \frac{1}{G_{on} - G_{off}} \frac{dG_{ij}}{dt} = f(x_{ij}, g(u_{ij}), t)$$

Thus, synaptic plasticity can be simulated via the voltage-driven resistance switching mechanism in memristive devices. Here, $f(x_{ij}, g(u_{ij}), t)$ represents a voltage-dependent function that describes the voltage curve across the memristive device. It was shown that for memristive devices this equation can be identified with the logistic equation [60]:

$$f(x_{ij}, g(u_{ij}), t) = \beta(x_{ij}, g(u_{ij}), t)x_{ij}(t)(1 - x_{ij})$$

It is important to note that this equation has a state and voltage dependent learning rate β and therefore differs from the normal logistic regression. This takes account of the special properties of memristive devices, whose functional mechanisms are usually based on ionic processes and have threshold properties. Thus, a change of the resistance of the device depends on the duration and amplitude of a voltage stimulus, as well as on the current resistance state [60]. Although the specific form of β depends on the type of the memristive device and its physical properties, some important statements can already be made at an abstract (descriptive) level: (i) in the simplest case, the voltage function $g(u_{ij})$ can be represented by the voltage drop across the memristive device, i.e. at constant potential at the one electrode via the voltage u_j at the opposite electrode. (ii) Following reference [60], the learning rate is linearly dependent on the state variable ($\beta = \gamma x_{ij} u_j$), so that the memristive state change for the emulation of synaptic learning processes in the framework of the Hebbian learning theory can be summarized as.

$$\frac{dx_{ij}}{dt} = \gamma(1 - x_{ij})x_{ij}^2 \cdot u_j$$

Thus, we have a compact model for a memristive device within the framework of Hebbian learning theory. Figure 9a shows the I-V curve obtained for sweeping the voltage u_j between 1 V and -1 V (see arrows) with that model. Therein, a state-dependent learning rate $\beta(u_j)$ (red curve) is compared with a state-independent learning rate $\beta(x_{ij}, u_j)$ (black curve). What can be seen is that by considering the state variables in the learning rate, an asymmetry between the set process (at positive voltage) and the reset process (at negative voltage) is obtained. In addition, the state variable dependent characteristic leads to a stronger non-linearity. For better visualization, the change of the state variable x is depicted in Fig. 9b and the function f is plotted in Fig. 9c. One sees a clearly stronger non-linear behavior of a state dependent learning rate (black curves) over the time course of the voltage change, which agrees with the threshold behavior of real memristive devices. It is shown in [4] how this model can be adapted to real device characteristics by choosing suitable representations for $\beta(x_{ij}, g(u_{ij}), t)$. In the following, however, we will discuss the biological plausibility of the model in more detail and show how it can be used to reproduce synaptic plasticity.

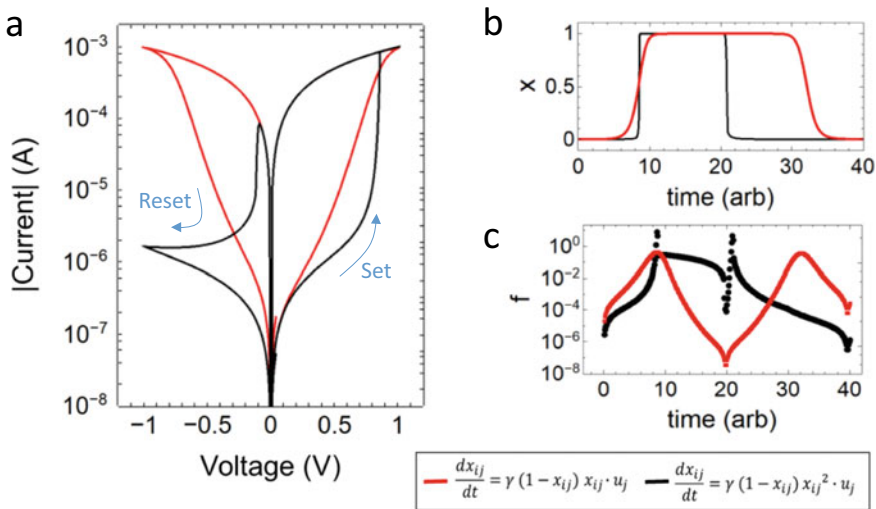


Fig. 9 Phenomenological model of a memristive device with a state-dependent learning rate $\beta(u_j)$ (red curve) and a state-independent learning rate $\beta(x_{ij}, u_j)$ (black curve). **a** I-V curve for a voltage sweep between 1 V and -1 V . **b** Change of the corresponding state variables. **(c)** Comparison of the two functions used

4.3 Emulation of Synaptic Learning

To show the biological plausibility of the memristive Hebbian learning model, we will follow reference [6] and first consider the emulation of LTP and LTD. For this purpose, the **voltage clamp method** was emulated first and compared with physiological data obtained from hippocampal CA1 neurons. The results obtained by the memristive learning model are shown in Fig. 10. For this purpose, the post-synaptic potential was used as a constant voltage offset ($u_j = V_{post}$), while a voltage pulse train was applied to the pre-neuron side (cf. Figure 10a). As sketched in Fig. 10b, this causes the resistance state of the memristive device to be set or reset depending on the post-synaptic offset V_{post} . Here, u_{set} is to symbolize the critical voltage for changing the device state. Therewith, the post-synaptic potential determines whether the resistance of the memristive cell is increased (emulation of LTD) or decreased (emulation of LTP) in a good agreement with physiological data of [61], as shown in Fig. 9c.

An important property of synaptic learning is the temporal correlation between neuronal activities on the pre-synaptic side in relation to the post-synaptic side. In this context, spiking neural networks (SNNs) have been shown to mimic information processing in biological neural networks very well. A central component of information coding in SNN is provided by the neurons of the network, whose activity is coded in spike patterns (voltage pulse trains). The strength of a stimulus is translated into a number of spikes. Therefore, the all-or-nothing principle is used, which means that if the input stimulus (which can be represented by a current I_{in}) depolarizes the post-synaptic potential u_j to such an extent that it exceeds a critical threshold θ_{thres} ,

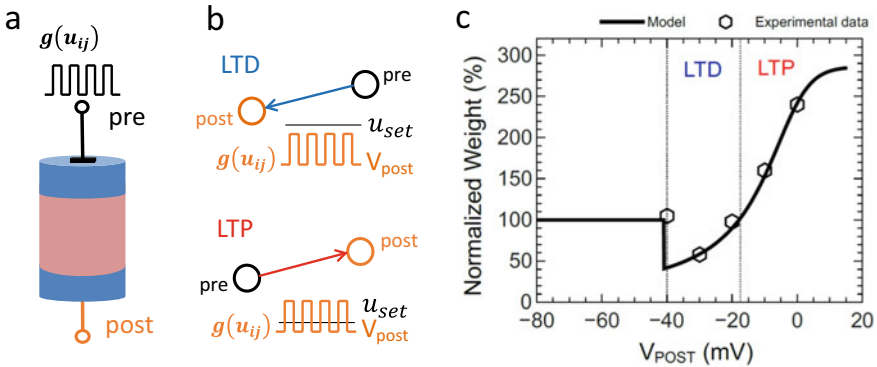


Fig. 10 Emulation of the voltage clamp method for measuring long-term potentiation and long-term depression with the memristive learning model. **a** Schematic representation of the voltage function over the memristive device. **b** For the simulation, the post-synaptic contact was used as the voltage offset, while a voltage was applied to the pre-synaptic connection. u_{set} describes the required set voltage of the memristive device. **c** The change in resistance value obtained with the model as a function of post-synaptic voltage (solid line) compared to data from hippocampal neurons from the CA1 region. (Adapted from [6])

a short voltage pulse (spike) is generated. Depending on the stimulus strength, more or fewer spikes are generated. Thus, a high activity of the neuron can be identified with a high number of spikes, i.e. with a higher frequency than in comparison to a weak input current, which generates only a few spikes. Mathematically, this can be represented quite elegantly within the framework of the leaky-integrate-and-fire neuron model which is generally given by the following equation:

$$\frac{du_j}{dt} = F(u_j) + \delta \cdot I_{in}(t)$$

Here $F(u_j)$ is a function that describes the voltage integration and must be defined in more detail and δ is a constant. For the determination of $F(u_j)$ there are several possibilities, whereby [54] gives a good overview. In the context of the consideration made here we use the quadratic leaky-integrate-and-fire (QIF) neuron model specified by the following expression:

$$C \frac{du_j}{dt} = g_L(u_j - u_r)(u_j - u_c) + I_{in}(t)$$

where g_L is a constant that has the dimension conductivity per voltage, C is the membrane capacitance, u_r is the resting potential, and u_c is the threshold potential for self-induced spiking of the neuron. A schematic representation of the generated voltage curve $u_j(t)$ of the QIF neuron model is shown in Fig. 11a. Here, a spike signal from neuron A or B (orange or black) is always transmitted to the network as soon as u_j exceeds the threshold value θ_{thres} . In the model used in reference 6, the neuron model was also extended in such a way that when one of the neurons in the small network of two neurons sketched in Fig. 11a spikes, the current value of the respective other neuron is used for the offset potential V_{post} . Then the synaptic weight can be increased (LTP) or decreased (LTD) from the respective value $u_j(t)$ of the other neuron.

As already mentioned, the temporal order of the spikes is of relevance for SNN. An important learning schema in this context is **spike-timing-dependent plasticity (STDP)**, which is sketched in Fig. 11b. STDP is an asymmetric form of Hebbian learning and allows causality in undirected neuronal networks. It specifies the strengthening of the synaptic connection (potentiation) if the pre-neuron is active before the post-neuron and a reduction of the synaptic connection strength (depression) if the post-neuron was active before the pre-neuron (cf. Fig. 11b). If neuron A is chosen as the pre-neuron and neuron B as the post-neuron, SDTP can be simulated in this network using the memristive Hebbian learning rule as shown in Fig. 11b. Here, the parameters of the model were adjusted to show good agreement with experimental data of Bi and Poo [61]. However, the temporal ordering of STDP, which leads to unidirectional connectivity in networks, is not valid for the complete frequency range. For spike patterns at frequencies above 40 Hz, the temporal order of spikes from pre- and post-synaptic neurons is cancelled as shown in Fig. 11c

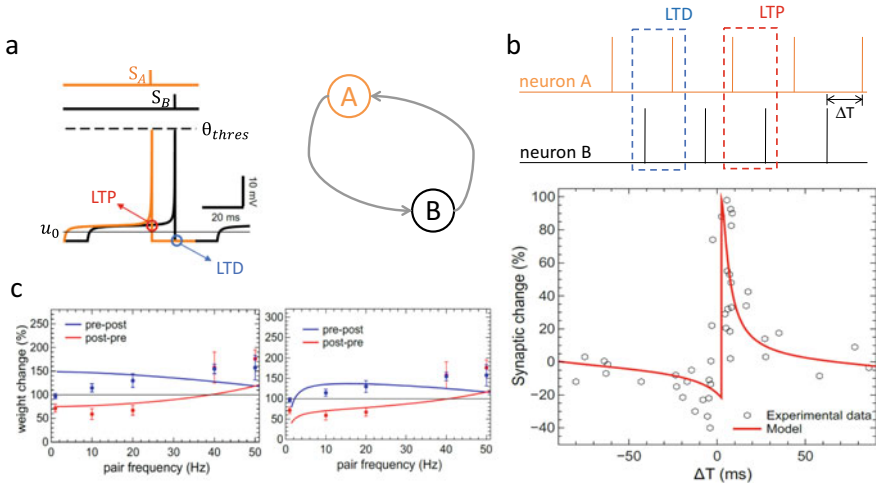


Fig. 11 Emulation of spike-timing-dependent plasticity (STDP). **a** Schematic and time course of a network of two leaky-integrate and fire neurons for emulation of STDP. When one neuron fires, the respective voltage value of the other neuron is used to change the resistance state of the memristive devices coupling the two neurons (represented by the gray arrows). Here u_i describes the threshold between setting and resetting the memristive device **b** schematic representation of the spike order described by the STDP and the STDP curve obtained with the model (red line) compared to measured data (dots). **c** Dependence of the spike order on the frequency of the spike pairs. The solid lines are results from the simulation, while the dots are measured values of real neurons. Here, in the right image, the model of the memristive devices was extended by a leakage rate, which leads to a better match at low frequencies. (The figure was adapted from [6])

for the example of cortical pyramidal neurons (red and blue data points) [62]. To reproduce this behavior in the model, a fixed time delay of ± 10 ms between the spike trains (A and B) was chosen, so that the individual spike pairings are constant with respect to each other. Hereby two cases were considered in reference [6]: (i) the resistance state is constant between two following voltage pulses and (ii) the resistance value increases between those pulses as a result of a continuous reset mechanism typically found in ionic memristive devices [4]. As a result, Fig. 11c (red and blue curves) shows that within the model the transition between asymmetric and symmetric behavior can be well emulated. The latter is especially relevant for the formation of bidirectional connections in networks as we will see in the next chapter. Furthermore, it can be stated that a low and autonomous reset of the device resistant state leads to an improved match at low frequencies of less than 10 Hz.

4.4 Emulation of Network Dependent Learning Schemes

Learning and memory formation in biological networks is closely linked to network topology [63]. Thus, it is the connectivity between individual neurons that determines

how a network responds to external stimuli and thus how it processes the information. While in machine learning the topology is fixed in the form of neuronal layers and the connectivity between the layers is determined by training processes, in biological neural networks a large part of the topology is formed by the learning process itself. In this process, connections are formed between networks that did not exist before and thus the network grows with its task.

Emulating growing networks within electronic circuits is difficult to realize in hardware, but all-to-all connecting networks, that are, networks in which each neuron is connected to every other neuron provide a simple opportunity to come a step closer to study the development of the topology of biological networks. A good example of this are the DG and CA3 regions of the hippocampus, which act as biological role models of the mentioned essential tasks pattern completion and pattern separation. This is considered to be the basis of associative memory in particular and is characterized by incoming information being stored sequentially and independently of each other. Thus, stored information can be retrieved and restored even if it was only partially presented. In the network topology of the hippocampus, recurrent connections in the DG and CA3 region could be identified for this purpose, forming an auto-associative network. This can be well modeled by an all-to-all network [6].

The memristive network model of Ref. [6] which emulates those two key functionalities, i.e. sequential learning and pattern completion, in form of an all-to-all network is shown in Fig. 12. Therefore, the memristive learning model was used and external information was presented to the network in two different ways: (i) by applying the individual pixels of visual patterns in parallel and (ii) by applying the individual pixels of a visual pattern sequentially. A schematic representation for better explanation is shown in Fig. 12a. The obtained connectivity matrices for these two cases are shown for pattern completion and sequential learning in Fig. 12b and c, respectively. For pattern completion, it was shown that using the memristive learning model, there is an increase in bidirectional connectivity at higher frequencies, which can be explained by the behavior shown in Fig. 12c. A unidirectional connectivity pattern is the result in the case of a sequential pattern presentation during the learning process (see Fig. 12c).

This allows the topology of the network to be shaped depending on the presentation of the input data. The latter is particularly important for learning in biological networks, since it allows for experience-dependent learning, i.e., depending on how information is presented (experiences have been made), the topology of the network and thus the functionality of the neural network changes. Thus, memories can shape the performance of the system. This is an important feature related to episodic memory, in which events are linked by their temporal sequence.

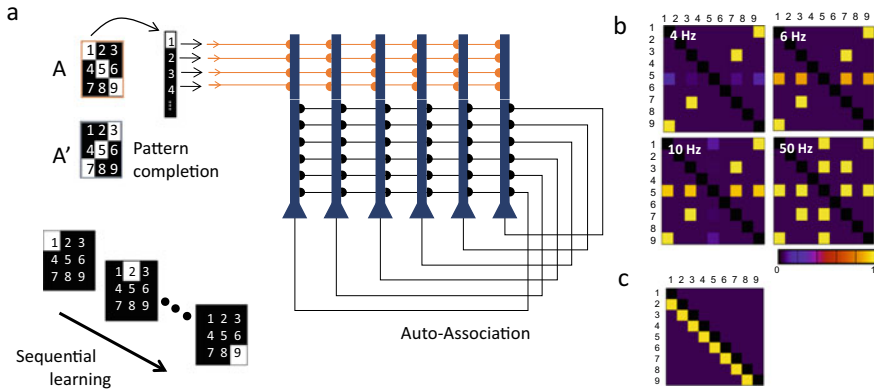


Fig. 12 Memristive network model of hippocampal structure to emulate pattern completion and sequential learning in the context of an all-to-all network. **a** Test pattern and network topology. The way a pattern is presented to the network gives rise to different network topologies. **b** Connectivity matrix of the network after learning for different repetition frequencies in the case of pattern completion (pattern A and A'). **c** Preserving connectivity matrix for the sequential creation of pixels of the test pattern during learning. (Adapted from [6])

Acknowledgements The authors acknowledge financial support via the Deutsche Forschungsgemeinschaft (DFG) by the Research Unit 2093: memristive devices for neuronal systems.

References

1. Mead, C., Ismail, M.: (eds.): Analog VLSI Implementation of Neural Systems, vol. 80. Springer Science & Business Media (1989)
2. Ielmini, D., Ambrogio, S.: Emerging neuromorphic devices. *Nanotechnology* **31**(9), 092001 (2019)
3. Adam, G.C., Khiat, A., Prodromakis, T.: Challenges hindering memristive neuromorphic hardware from going mainstream. *Nat. Commun.* **9**, 5267 (2018)
4. Hansen, M., Zahari, F., Ziegler, M., Kohlstedt, H.: Double-Barrier memristive devices for unsupervised learning and pattern recognition. *Front. Neurosci.* **11** (2017)
5. Waser, R., Wuttig, M.: Memristive Phenomena - from fundamental physics to neuromorphic computing: Lecture Notes: Spring School organized by Peter Grünberg Institute, Forschungszentrum Jülich and Physics Institutes, RWTH Aachen University, Jülich Aachen Research Alliance, Section Fundamentals of Future Information Technology (JARA-FIT): in collaboration with universities, research institutes and industry (Forschungszentrum Jülich, Zentralbibliothek) (2016)
6. Diederich, N., Bartsch, T., Kohlstedt, H., Ziegler, M.: A memristive plasticity model of voltage-based STDP suitable for recurrent bidirectional neural networks in the hippocampus. *Sci. Rep.* **8**, 9367 (2018)
7. Kandel, E.C., Schwartz J.H., Jessell T.M.: Principles of Neuronal Science, 3rd edn. Elsevier Science Publishing, Amsterdam (1991)
8. Andersen, P., Morris, R., Amaral, D., Bliss, T., O'Keefe, J., (eds.): The Hippocampus Book. Oxford University Press (2006)

9. Rolls, E.T.: The mechanisms for pattern completion and pattern separation in the hippocampus. *Front Syst. Neurosci.* **7**, 74 (2013)
10. Bliss, T., Schoepfer, R.: Neuroscience. Controlling the ups and downs of synaptic strength, *Science* **304**, 973 (2004)
11. Whitlock, J.R., Heynen, A.J., Shuler, M.G., Bear, M.F.: Learning induces long-term potentiation in the hippocampus. *Science* **313**, 1093 (2006)
12. Bliss, T., Collingridge, G., Morris, R.: *Long-Term Potentiation Enhancing Neuroscience for 30 Years*. Oxford University Press, New York (2004)
13. Larson, J., Lynch, G.: Role of N-methyl-D-aspartate receptors in the induction of synaptic potentiation by burst stimulation patterned after the hippocampal θ -rhythm. *Brain Res.* **441**, 111 (1988)
14. Pastalkova, E., Serrano, P., Pinkhasova, D., Wallace, E., Fenton, A.A., Sacktor, T.C.: Storage of spatial information by the maintenance mechanism of LTP. *Science* **313**, 1141 (2006)
15. Bartsch, T., Schonfeld, R., Muller, F.J., Alfke, K., Leplow, B., Aldenhoff, J., Deuschl, G., Koch, J.M.: Focal lesions of human hippocampal CA1 neurons in transient global amnesia impair place memory. *Science* **328**, 1412–1415 (2010)
16. Squire, L.R., Zola, S.M.: Structure and function of declarative and nondeclarative memory systems. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13515–13522 (1996)
17. McClelland, J.L., McNaughton, B.L., O'Reilly, R.C.: Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995)
18. Scoville, W.B., Milner, B.: Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20**, 11–21 (1957)
19. O'Reilly, R.C., McClelland, J.L.: Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* **4**, 661–682 (1994)
20. Marr, D.: Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. B Biol. Sci.* **262**, 23 (1971)
21. Yassa, M.A., Stark, C.E.: Pattern separation in the hippocampus. *Trends Neurosci* **34**(10), 515–525 (2011)
22. Bakker, C.B., Kirwan, M.M., Stark, C.E.: Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* **319**, 1640 (2008)
23. Lisman, J.E.: Relating hippocampal circuitry to function: recall of memory sequences by reciprocal dentate-CA3 interactions. *Neuron* **22**, 233 (1999)
24. Leutgeb, J.K., Leutgeb, S., Moser, M.-B., Moser, E.I.: Pattern separation in the dentate gyrus and CA3 of the Hippocampus. *Science* **315**, 961 (2007)
25. Berron, D., Schütze, H., Maass, A., Cardenas-Blanco, A., Kuijff, H.J., Kumaran, D., Düzel, E.: Strong evidence for pattern separation in human dentate gyrus. *J. Neurosci.* **36**, 7569–7579 (2016)
26. Strong, M.H., Strong, E.K.: The nature of recognition memory and of the localization of recognitions. *Am. J. Psychol.* **27**(3), 341–362 (1916)
27. Stark, S.M., Stevenson, R., Wu, C., Rutledge, S., Stark, C.E.L.: Stability of age-related deficits in the mnemonic similarity task across task variations. *Behav. Neurosci.* **129**(3), 257–68 (2015)
28. Bartsch, T., Wulff, P.: The hippocampus in aging and disease: from plasticity to vulnerability. *Neuroscience* **309**, 1–16 (2015)
29. Small, S.A., Schobel, S.A., Buxton, R.B., Witter, M.P., Barnes, C.A.: A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nat Rev Neurosci* **12**, 585–601 (2011)
30. Bartsch, T., Alfke, K., Deuschl, G., Jansen, O.: Evolution of hippocampal CA-1 diffusion lesions in transient global amnesia. *Ann. Neurol.* **62**, 475–480 (2007)
31. Bartsch, T., Alfke, K., Stingele, R., Rohr, A., Freitag-Wolf, S., Jansen, O., Deuschl, G.: Selective affection of hippocampal CA-1 neurons in patients with transient global amnesia without long-term sequelae. *Brain* **129**, 2874–2884 (2006)
32. Bettcher, B.M., et al.: More than memory impairment in voltage-gated potassium channel complex encephalopathy. *Eur. J. Neurol.* **21**, 1301–1310 (2014)

33. Butler, C.R., Miller, T.D., Kaur, M., Baker, I.W.S., Boothroyd, G.D., Illman, N.A., Rosenthal, C.R., Vincent, A., Buckley, C.J.: Persistent anterograde amnesia following limbic encephalitis associated with antibodies to the voltage-gated potassium channel complex. *J. Neurol. Neurosurg. Psychiatry* **85**, 387–391 (2014)
34. Irani, S.R., Michell, A.W., Lang, B., Pettingill, P., Waters, P., Johnson, M.R., Schott, J.M., Armstrong, R.J.E., S Zagami, A., Bleasel, A.F., Somerville, E.R., Smith, S.M.J., Vincent, A.: Faciobrachial dystonic seizures precede Lgi1 antibody limbic encephalitis. *Ann. Neurol.* **69** (2011)
35. Malter, M.P., Frisch, C., Schoene-Bake, J.-C., Helmstaedter, C., Wandinger, K.-P., Stoecker, W., Urbach, H., Surges, R., Elger, C.E., Vincent, A., Bien, C.G.: Outcome of limbic encephalitis with VGKC-complex antibodies. Relation to antigenic specificity. *J. Neurol.* **261**, 1695–1705 (2014)
36. Finke, C., Prüss, H., Heine, J., Reuter, S., Kopp, U.A., Wegner, F., Then Bergh, F., Koch, S., Jansen, O., Münte, T., Deuschl, G., Ruprecht, K., Stöcker, W., Wandinger, K.-P., Paul, F., Bartsch, T.: Evaluation of cognitive deficits and structural hippocampal damage in encephalitis with leucine-rich, glioma-inactivated 1 antibodies. *JAMA Neurol.* **74**, 5059 (2017)
37. Miller, T.D., Chong, T.T., Aimola Davies, A.M., Ng, T.W.C., Johnson, M.R., Irani, S.R., Vincent, A., Husain, M., Jacob, S., Maddison, P., Kennard, C., Gowland, P.A., Rosenthal, C.R.: Focal CA3 hippocampal subfield atrophy following LGI1 VGKC-complex antibody limbic encephalitis. *Brain* **140**, 1212–1219 (2017)
38. Hanert, A., Pedersen, A., Bartsch, T.: Transient hippocampal CA1 lesions in humans impair pattern separation performance. *Hippocampus* (2019)
39. Knierim, J.J., Neunuebel, J.P.: Tracking the flow of hippocampal computation: pattern separation, pattern completion, and attractor dynamics. *Neurobiol. Learn. Mem.* **129**, 38–49 (2016)
40. O'Reilly, R.C., Rudy, J.W.: Computational principles of learning in the neocortex and hippocampus. *Hippocampus* **10**(4), 389–397 (2000)
41. Insausti, R., Amaral, D.G.: Hippocampal formation. In: Paxinos, G., (ed.), *The Human Nervous System*, pp. 871–914. Elsevier (2004)
42. Hasselmo, M.E., Eichenbaum, H.: Hippocampal mechanisms for the context-dependent retrieval of episodes. *Neural Netw.* **18**, 1172–1190 (2005)
43. Hanert, A., Rave, J., Granert, O., Ziegler, M., Pedersen, A., Born, J., Finke, C., Bartsch, T.: Hippocampal dentate gyrus atrophy predicts pattern separation impairment in patients with LGI1 encephalitis. *Neuroscience* **400**, 120–131 (2019)
44. Dillon, S., Tsivos, D., Knight, M.J., McCann, B., Pennington, C.M., Shiel, A.I., Conway, M.E., Newson, M.A., Kauppinen, R.A., Coulthard, E.J.: The impact of ageing reveals distinct roles for human dentate gyrus and CA3 in pattern separation and object recognition memory. *Sci. Rep.* **7** (2017)
45. Doxey, C.R., Kirwan, C.B.: Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. *Hippocampus* **25**, 524–533 (2015)
46. Craik, F.I.M., Tulving, E.: Depths of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* **104**, 268–294 (1975)
47. Kaernbach, C.: The memory of noise. *Exp. Psychol.* **51**(4), 240–248 (2004)
48. Cowan, N.: On short and long auditory stores. *Psychol. Bull.* **96**, 341–370 (1984)
49. Winkler, I., Cowan, N.: From sensory to long-term memory: evidence from auditory memory reactivation studies. *Exper. Psychol.* **52**(1), 3–20 (2005)
50. Green, D.M., Swets, J.A.: *Signal Detection Theory and Psychophysics*. John Wiley, New York (1966)
51. Field, D.J.: Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* **4**(12), 2379–2394 (1987)
52. Hebb, D.O.: *The Organization of Behavior: A Neuropsychological Theory*, vol. 11. [print.]. Wiley, New York (1974)
53. Stent, G.S.: A physiological mechanism for Hebb's postulate of learning. *Proc. Natl. Acad. Sci.* **70**(4), 997 (1973). <https://doi.org/10.1073/pnas.70.4.997>

54. Gerstner, W., Kistler, W.M., Naud, R., Paninski, L.: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge (2014)
55. Bienenstock, E.L., Cooper, L.N., Munro, P.W.: Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* **2**(1), 32 (1982). <https://doi.org/10.1523/JNEUROSCI.02-01-00032.1982>
56. Ziegler, M., Wenger, C., Chicca, E., Kohlstedt, H.: Tutorial: concepts for closely mimicking biological learning with memristive devices: principles to emulate cellular forms of learning. *J. Appl. Phys.* **124**(15), 152003 (2018). <https://doi.org/10.1063/1.5042040>
57. Chua, L.: Memristor-the missing circuit element. *IEEE Trans. Circuit Theory* **18**(5), 507–519 (1971). <https://doi.org/10.1109/TCT.1971.1083337>
58. Yang, J.J., Strukov, D.B., Stewart, D.R.: memristive devices for computing. *Nat. Nanotechnol.* **8**(December), 13 (2012)
59. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. *Nature* **453**(7191), 80–83 (2008)
60. Ziegler, M.C., Riggert, M., Hansen, T.B., Kohlstedt, H.: Memristive Hebbian plasticity model: device requirements for the emulation of Hebbian plasticity based on memristive devices. *IEEE Trans. Biomed. Circuits Syst.* **9**(2), 197–206 (2015). <https://doi.org/10.1109/TBCAS.2015.2410811>
61. Bi, G., Poo, M.: Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10464–10472 (1998)
62. Sjöström, P.J., Turrigiano, G.G., Nelson, S.B.: Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* **32**, 1149–1164 (2001). [https://doi.org/10.1016/S0896-6273\(01\)00542-6](https://doi.org/10.1016/S0896-6273(01)00542-6)
63. Fornito, A., Zalesky, A., Bullmore, E.T.: *Fundamentals of Brain Network Analysis*. Academic, Amsterdam (2016)
64. Buzsáki, G.: *Rhythms of the Brain*. Oxford University Press (2006)
65. Ziegler, M., Kohlstedt, H.: Memristive models for the emulation of biological learning. In: *Memristor Computing Systems*, pp. 247–272. Springer International Publishing, Cham (2022)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

