Master Thesis

Investigating Floating Gate Insertion into Ferroelectric Schottky Barrier Field-Effect Transistors for Neuromorphic Computing

submitted in partial fulfillment
of the requirements of the degree of
Master of Science

at Technische Universität Wien
Institute of Solid State Electronics

under supervision of
Univ.Prof. Dipl.Ing. Univ. Dr.-Ing. Walter Michael Weber
and
Univ.Ass. Dr.techn. Daniele Nazzari

by
Viktor Wahler, BSc
Student ID 12019257

Vienna, April 2025

Statutory Declaration

I declare, that I have authored the present work independently according to the code of conduct, that I have not used other than the declared sources and that I have explicitly marked all material quoted either literally or by content from the used sources. This work was not yet submitted to any examination procedure neither in Austria, nor in any other country.

Erklärung zur Verfassung der Arbeit

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct – Regeln zur Sicherung guter wissenschaftlicher Praxis – ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder in ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Vienna, April 2025 .....

Viktor Wahler

Abstract

The recent rise of artificial intelligence (AI) has been facilitated by the ongoing developments in semiconductor technologies. Current implementations rely on graphical processing units (GPUs), while future performance improvements are limited by the slowing down of Moore's law and the "von Neumann bottleneck". In particular, the latter poses a hard barrier to the reduction of power consumption for the existing hardware platforms. A suitable alternative direction appears to be neuromorphic computing, utilizing emerging memory devices, like the ferroelectric field effect transistors (FeFETs), to mimic neurons and synapses that are the basic building blocks of the human brain. Thanks to the recent discovery of HZO, FeFETs can be efficiently implemented in established frameworks. The reconfigurable FET (RFET), which does not use doped regions, but exploits the electrostatic modulation of Schottky Barriers to determine the charge carrier type, represents a feasible alternative that can provide more functionality in a reduced wafer area. Evolving this concept towards a ferroelectric-based neuromorphic architecture is extremely promising for the realization of extremely flexible and power-efficient AI networks.

Founded on an already established ferroelectric SBFET (FeSBFET) process, this thesis studies how the introduction of a TiN floating gate (FG) layer in between the SiO_2 interlayer and the HZO impacts the switching behavior. It is expected that the FG insertion will make the devices less trap-dependent. Within this work, a wet-etching process is developed for the patterning of thin TiN layers. Analysis of the final gate structure employing atomic force microscopy (AFM) and transmission electron microscopy (TEM) indicates an excellent control of the performed processing.

Electrical measurements of test structures validate the ferroelectricity of the deposited HZO and the correct formation of SBs beneath the top gate (TG). Transfer characteristics of the FG FeSBFET devices show a shift of $\pm 1.4\,\mathrm{V}$ depending on the sweeping direction. A simple mathematical model derived from the one describing non-ferroelectric FGFETs used in flash storage devices allows the estimation of the effective charges in the FG layer. These charges, which determine the device characteristics, are the result of the balance between the charges coming from the ferroelectric polarization and those injected into the layer. It is observed that, in the case of slow sweeping, a majority of injected charges are present. Pulsed measurements with varying pulse times indicate that after 6 ms, enough charges are injected to enable the ferroelectric switching mechanism, which, in turn, reduces the injection rate.

The promising characteristics of the fabricated devices for future application in neuromorphic computing platforms are clearly demonstrated by the long-term potentiation (LTP) and long-term depression (LTD) curves extracted from the pulsed measurements. Adaptation of the surface ratio of the involved layers is needed to bring the pulse time in a similar range as what is displayed by other neuromorphic device concepts currently under investigation.

Kurzfassung

Der Aufstieg der künstlichen Intelligenz (KI) wurde speziell durch Fortschritte in der Halbleitertechnologie ermöglicht. Momentane Realisierungen sind auf Graphics-Processing-Units (GPUs) angewiesen, die von der Moore's-Law-Verlangsamung und dem "von Neumann Bottleneck" betroffen sind. Speziell der letzte Punkt stellt ein Hindernis für die Reduktion des Stromverbrauchs der etablierten Herangehensweise dar. Eine valide Alternative scheint Neuromorphic-Computing zu sein, welches Emerging-Memory-Devices wie ferroelektrische Feldeffekttransistoren (FeFETs) verwendet, um Neuronen und Synapsen, die Bausteine des menschlichen Gehirns, zu imitieren. Aufgrund der Entdeckung von HZO können FeFETs effizient in vorhandene Technologien integriert werden. Zusätzlich kann der rekonfigurierbare FET (RFET), der keine Dotierung benötigt, sondern Schottky-Barrieren nutzt, um den Landungsträgertyp einzustellen, mehr Funktonalität auf einer reduzierten Chipfläche ermöglichen. Dieses Konzept in einer ferroelektrisch neuromorphen Architektur zu entwickeln, ist vielversprechend, um extrem flexible und effiziente KI-Netzwerke zu realisieren.

Aufbauend auf einem bereits etablierten ferroelektrischen SBFET (FeSBFET) Fertigungsprozess untersucht diese Arbeit die Einführung einer TiN Floating-Gate (FG) Metalllage zwischen dem SiO_2 Grenzflächendielektrikum und dem HZO, um das Schalten der einzelnen Domänen zu verbessern. Es wird davon ausgegangen, dass das FG den Einfluss von Traps reduziert. Ein nasschemischer Ätzprozess wird entwickelt, um TiN zu strukturieren. Atomic-Force-Microscopy (AFM) und Transmission-Electron-Microscopy (TEM) des finalen Aufbaus zeigen eine gute Kontrolle über die durchgeführte Fertigung.

Elektrische Messungen an Teststrukturen validieren die ferroelektrische Eigenschaft des gewachsenen HZOs und die korrekte Etablierung der SBs unter dem Top-Gate (TG). Transfer Charakteristiken der FG-FeSBFET-Transistoren zeigen eine Verschiebung von $\pm 1.4\,\mathrm{V}$ in Abhängigkeit von der Sweep-Richtung. Ein einfaches mathematisches Model, abgeleitet von der Beschreibung normaler nicht ferroelektrischer FGFETs, die in Flash-Speichern verwendet werden, erlaubt, die effektiven Ladungen in der Metallschicht abzuschätzen. Diese Ladungen, die die Charakteristik beeinflussen, setzen sich aus der ferroelektrischen Polarisation und den injizierten Ladungen zusammen. Es wird beobachtet, dass im Fall einer Transfermessung ein Überhang an injizierten Ladungen vorherrscht. Weitere gepulste Messungen mit variierender Pulslänge zeigen, dass nach 6 ms bereits ausreichend viele Ladungen injiziert werden, um das Ferroelektrikum zu schalten, was die Injektionsrate reduziert.

Die vielversprechenden Charakteristiken der fabrizierten Transistoren für die Anwendung in Neuromorphic Computing Plattformen wird durch die Long-Term-Potentiation (LTP) und Long-Term-Depression (LTD) Kurven, die aus den gepulsten Messungen extrahiert wurden, hervorgehoben. Die Adaptierung der involvierten Flächenverhältnisse ist notwendig, um die Pulsweiten an die von anderen Konzepten anzupassen.

Acknowledgement

I would like to acknowledge and give my warmest thanks to the large number of people who supported me during this project. Especially to the following:

Walter M. Weber
For offering me the opportunity to join his research group as a student after my *Integrierte Bauelemente* exam a year ago. During a small project, I gained experience in semiconductor research alongside other university courses. This allowed me to hit the ground running with my thesis and accelerated my studies. I am also very grateful for the opportunity to attend the 2025 edition of the High-k Workshop in Dresden, which inspired me and further fuelled my interest in the final stages of my work. Of course, his guidance in shaping the project and the numerous insights he provided should also be acknowledged.

Daniele Nazzari
For teaching me the aspects of cleanroom fabrication needed for this undertaking. As the second supervisor of this thesis, I worked very closely with him and came to appreciate his tremendous material science knowledge of HZO and beyond, which was invaluable. Furthermore, his calm and encouraging demeanor kept me going even after a failed sample or a new device model I devised turned out to be incorrect. I also want to thank him for proofreading my work and for his countless valuable inputs.

Lukas Wind
For providing the many tools and processes used in this work. During his PhD, which he completed while I was beginning to write my thesis, the fabrication process on which my samples are based, and many measuring techniques that I was able to adapt, were developed. The findings of his thesis were tremendously useful.

Masiar Sistani
For supporting me during my electrical measurements and providing many remarks that continuously improved both the fabrication and understanding of the investigated devices.

The Emerging Nanoelectronics Group
For challenging my ideas during group meetings, which motivated me to develop them further, and for the companionship in the lab.

FKE and ZMNS
For providing the cleanroom facilities and measuring equipment.

My friends
For the countless moments spent together. During movies, gaming sessions, or ski trips, I was able to recharge and shift my focus. I especially want to thank my dear friend Elias, who has been by my side throughout my entire time at school and university over the past ten years. I truly enjoyed learning, struggling, and discussing the many topics we tackled together.

My family
For their unconditional support throughout this thesis and beyond. In particular, my parents, who gave me the opportunity to focus solely on my studies and provided a loving childhood that shaped who I am today. My sister, with whom I shared many laughs, and my grandparents, as well as close relatives, who supported me in numerous ways.

Thank you very much!

Contents

Chapter 1

Introduction

The British mathematician Alan Turing introduced the idea of machine simulating human intelligence in 1950 [1]. Inspired by this groundwork and the understanding of biological neural networks (BioNNs) in the human brain of the time, the American psychologist and computer scientist Frank Rosenblatt developed the first artificial neural network (ANN), the Mark I Perceptron, based on analog components [2]. The matrix description founded on the work from John von Neumann, still in use [3], and significant improvements in parallel computing with graphical processing units (GPUs) as well as ongoing developments in semiconductor manufacturing facilitated the rise of artificial intelligence (AI) decades after its initial conception, by providing efficient resources for extensive training.

The weakening of Moore's law, i.e. feature size scaling becoming more difficult for classical sub 22 nm CMOS devices, hinders further development of the current hardware, and the "von Neumann bottleneck", i.e. the physical separation between memory and processor, causes most computational power to be spent on transferring data, resulting in high energy needs. The International Energy Agency (IEA) reports that "Electricity consumption from data centers, artificial intelligence, and the cryptocurrency sectors could double by 2026," reaching 1000 TW h roughly the consumption of Japan [4]. This underlines the need for alternative approaches.

One of the most promising directions appears to be neuromorphic computing. The human brain carries out exceptionally complex tasks within a limited energy spectrum. Ranging from vision to speech and problem-solving while only consuming 20 W around 20% of the body's power budget. The brain had to evolve around this boundary because mammals can only consume as much as they can hunt and eat [5]. Modern computing hardware consumes 1000 kW h for similar tasks [6, 7]. Neuromorphic computing tries to implement the main characteristics of the brain in hardware.

Major players like IBM with their TureNorth & NorthPole [8–10] and Intel with their Loihi 1 & 2 [11, 12] chips have already demonstrated such devices, which implement the basic building blocks of the human brain neuron and synapse into CMOS circuits by leveraging decades of research and infrastructure, the semiconductor industry has built for CMOS fabrication. These concepts reduce the "von Neumann bottleneck" and energy consumption, but are still powerless against the weakening of Moore's law. In contrast, other concepts using emerging devices have received a great amount of attention lately, especially non-volatile memories (NVMs) such as phase-change memory (PCM), resistive random access memory (RRAM), conducting bridge random access memory (CBRAM), ferroelectric tunneling junction (FTJ), trapping field effect transistors, and ferroelectric field effect transistors (FeFETs). A neuron or synapse can be represented by only one of these devices compared to the multiple needed in CMOS. This reduces the footprint of the NN and, theoretically, improves energy efficiency due to the reduced number of devices [13].

The industry is now shifting towards the "More-than-Moore" domain, necessary because of development limitations in the currently used CMOS devices due to random dopant fluctuations and the employed lithography fabrication processes [14, 15]. This means incorporating additional functionality into devices that are not necessarily limited by these factors [16]. Following this path, one can find the reconfigurable field effect transistor (RFET), which exploits the electrostatic tunability of the Schottky barriers (SBs) to select the charge carrier type. This enables the runtime reconfigurability of the device behavior between electron and hole conduction modes.

The first concept of this kind was presented by Lin et al. [17] in 2001. Many different devices followed this first example, using different channel materials, such as silicon nanowires (NW), and demonstrating the first real electron/hole-type operation [18–20]. Furthermore, state-of-the-art devices replace silicon NWs with novel silicon on insulator (SOI) based FETs or FinFETs and implement silicon-germanium combinations as channel material, which show higher performance [21–29]. Also, logic gates utilizing RFETs became possible with the mentioned advanced concepts [30, 31]. Studies predict that a 42% area reduction compared to 10 nm FinFET is possible for a 1-bit full adder circuit [32].

CMOS-compatible ferroelectric high-k dielectrics like $Hf_{0.5}Zr_{0.5}O_2$ (HZO) are used in the aforementioned FeFETs to implement nonvolatile behaviors. These approaches can be utilized in RFET technologies to enable a wide range of additional application fields, such as data storage, logic in memory (LiM), nonvolatile RFET gate programming, and neuromorphic computing. Initial work in this direction has proven the viability of such concepts [33, 34].

1.1 Motivation and Scope

The dual top-gated ferroelectric RFET design demonstrated by Nazzari et al. [34] clearly shows the possibility of programming the conduction mode of the device in a nonvolatile manner. It is based on a top-down process in which Si nanosheets are etched out of an SOI device layer. SiO_2 and HZO form the gate dielectric, which is capped by TiN to induce a ferroelectric phase during heating. The two SBs are established by an Al-Si exchange process between the Si nanosheet and the Al source/drain pads. By utilizing pulsing schemes, intermediate states can be set, and the possible application of the device in neuromorphic computing is demonstrated. Endurance and retention measurements indicate a reliable characteristic.

The investigated sample features a reduced electron conduction mode on current in comparison to the hole current. Also, off currents are shifted in a similar manner. It is believed that this behavior is connected to the charging of traps or an insufficient switch of the ferroelectric for the individual modes. When examining the device physics of RFETs it becomes clear that it is mainly governed by the injection of charges through or over the SB [32]. In light of this, precise control of the HZO polarization above the SB is essential. Simulations have shown that only partial switching of the HZO, which consists of individual domains, can cause the formation of current paths [35]. Furthermore, due to the unfavorable electric field situation exhibited on the ferroelectric above the SB caused by the Al to Si interface polarization switching can be harmed. It is believed that only the HZO between the Al source/drain pads and the TiN TG switches. Over the Si channel, the HZO remains unpolarized. These aspects can cause the conduction mechanism to be drastically degraded.

In this regard, the insertion of a TiN metal layer is investigated to enhance the electric field situation on the HZO and prevent the formation of current paths. The previously employed fabrication process is adapted to include this additional layer.

1.2 Outline

Chapter 2 gives the theoretical background needed for this thesis. It starts with motivating neuromorphic computing using the inner workings of the human brain. The interplay between neurons and synapses, as well as how emerging devices can be used to mimic their behavior, is explained. The chapter goes on to discuss the formation of HZO and its implementation as FeFET. Floating gate (FG) FETs are studied from a device physics standpoint since a similar structure is constructed by the TiN metal layer insertion. Finally, conduction modes in SB and their implementation in SBFETs, as well as RFETs, are shown.

Chapter 3 presents the experimental techniques used. First, the fabrication flow of the investigated devices is outlined in detail, explaining each process step and stating the individual process parameters. Secondly, the electrical measurement setups employed to conduct transfer characteristics and pulsed measurements are shown.

Chapter 4 discusses the fabricated devices in detail. Different techniques are used to evaluate them from a physical point of view. Layer quality and heights are investigated. Also, the needed process adaptations compared to [34] are motivated in detail. This chapter further discusses the electrical measurement and introduces a proposed device-switching mechanism model.

Chapter 5 summarizes the major findings and gives an outlook on future research directions.

Chapter 2

Theory

2.1 Neuromorphic Computing

The current understanding of the human brain is still incomplete, and any ANN is based on or inspired by simplified models for BioNN. To further enhance and develop artificial realizations, one should leverage the decades of research conducted by neuroscientists. An in-depth knowledge of the different structures and operations happening in the brain is needed and will be discussed in this section. Further, it goes on to compare the two main approaches to describing BioNNs and how a variety of different emerging devices known from memory applications can be used as building blocks for neuromorphic computing realizations.

2.1.1 Biological Neurons and Synapses

The human brain contains roughly 86 billion neurons that form thousands of connections with each other [37]. They can process and transmit signals from or to other neurons, receptors, or muscles in the body by generating electrical and chemical signals. Figure 2.1(a) illustrates two neurons connected by the axon. The upper one is called the presynaptic cell, and the lower is the postsynaptic cell. When the upper cell generates or fires a signal, also known as an action potential (AP) (Figure 2.1(b)), this AP travels from the axon hillock down the axon itself and splits eventually into different axon terminals. Where it generates neurotransmitters, which are passed on to the next neuron. Each neuron has multiple short branches or dendrites that spread from the cell body and act as inputs. When axon terminals and dendrites come close to each other, synapses form. Because the connection length between neurons can differ from micrometers to meters, the AP's signal integrity must be ensured by isolating myelin and signal-boosting nodes of Ranvier that wrap around the axon. Neurons can be categorized as multipolar, bipolar, unipolar, and pseudounipolar. The presynaptic neuron in Figure 2.1(a) has only one connection to

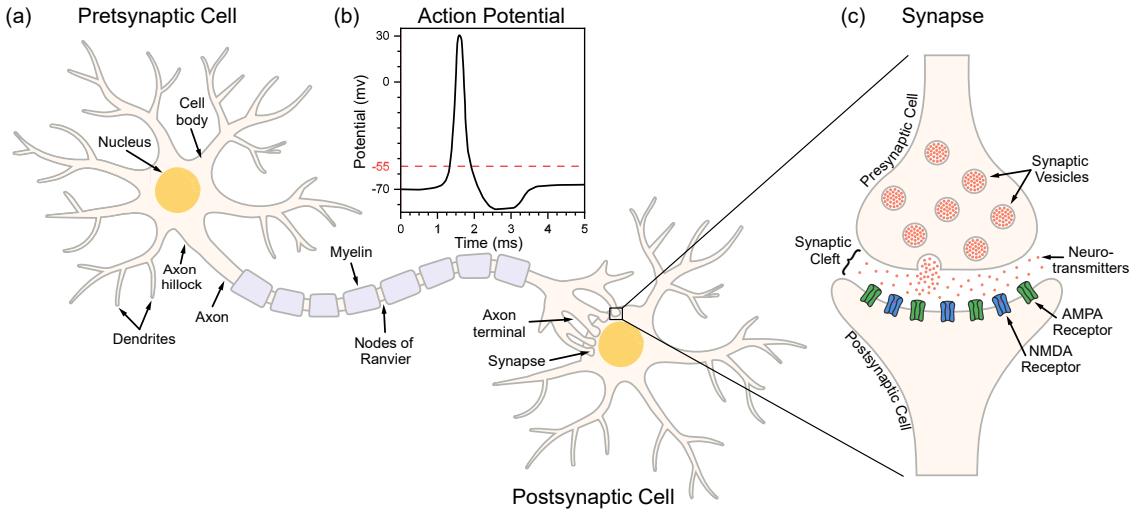

Figure 2.1: Representation of biological neurons. (a) The connection between pre- and postsynaptic neurons via the axon. (b) Diagram of the action potential over time. (c) Inner workings of a synapse after AP stimulation. Movement of synaptic vesicles and diffusion of neurotransmitters toward the postsynaptic cell are depicted. Image adapted from [13, 36].

another neuron, making it bipolar [38, 39].

The average neuron is believed to have connections via 1,000 synapses, bringing the number of synapses in the trillions [40]. In Figure 2.1(c) the inner workings of a synapse are shown. The space between them is called the synaptic cleft, only around 20-40 nm wide. When the AP reaches the axon terminal, synaptic vesicles, which contain the aforementioned neurotransmitters, are generated. During exocytosis, the vesicles fuse with the presynaptic axon terminal's cell membrane, releasing neurotransmitters into the synaptic cleft, which diffuses towards and interacts with protein receptors on the postsynaptic cell. This interaction allows ions to enter the postsynaptic cell and resemble the neuron's input [39, 41].

The main mechanism responsible for generating the AP surrounds a membrane in the axon hillock that separates ions. In its resting state, the membrane potential is $-70\,\mathrm{mV}$ because of an equilibrium between diffusive and electrostatic forces. If the receptors in the different synaptic inputs generate enough ions, the potential can increase to $-55\,\mathrm{mV}$, by which the threshold for the opening of voltage-gated ion channels is reached, which allows further ions to pass through the membrane. This feedback loop increases the potential to $+30\,\mathrm{mV}$, called depolarization. Eventually, the feedback stops because the ion channels close and new ones open, allowing ions to travel back through the membrane, starting repolarisation, which overshots called hyperpolarisation and settles back to the resting membrane potential of $-70\,\mathrm{mV}$. During this time, the neuron is in a refractory period, during which intense stimulation is needed to start a new pulse since the potential falls below the resting state. The axon itself can be understood as a sequence of membranes, and

the APs imbalance of ions propagates towards the axon terminals, triggering neurotransmitter generation. In general, the APs do not vary in size or length and are independent of the amount of stimulation. When the threshold is reached, they will always produce the same pulse [39].

Long-term potentiation (LTP) and long-term depression (LTD) are essential for plasticity and learning in BioNNs. In the postsynaptic synapse membrane, NMDA and AMPA receptors are present (Figure 2.1(c)), and in general, NMDAs are blocked by magnesium ions. When a synapse is activated frequently, LTP can occur. This process involves the removal of the voltage-dependent NMDA receptor blockages, leading to the insertion of additional AMPA receptors into the membrane through cellular mechanisms. By doing so, the amount of ions entering the postsynaptic cell from one AP is improved, increasing the likelihood of a signal over this synapse triggering an AP in the neuron. In contrast, during LTD, the synapse is not activated for a prolonged time. This causes an accumulation of blocked NMDAs and a reduction of AMPA receptors, hindering ions from entering the cell and making it harder to trigger an AP by this synaptic connection. It is believed that LTP and LTD are responsible for long-term memories in the human brain [39, 41].

Different brain sections are responsible for tasks ranging from vision to memory, using slight variations of these fundamental mechanisms. Learning, the updating of the weight of each individual synapse, is done over the span of a lifetime and during every waking hour of a human life [13, 36].

2.1.2 Artificial and Spiking Neural Networks

The discussed concepts of the human brain's inner workings are already a harsh simplification of its true complexity. In order to transfer them to digital software or analog hardware realizations, different models are developed that are able to imitate the essential mechanisms.

Artificial neural networks (ANN) are widely used by AI researchers. In this model, the complex biological AP generated by the neuron after reaching a certain threshold is replaced by a Heaviside step function, which switches the neuron's output from 0 to 1. The binary output of the presynaptic cell is multiplied by a fixed weight value in the synapse, and then the different inputs to the postsynaptic cell are summed up. If a threshold is reached, the postsynaptic cell fires. This mimics the behavior in BioNNs to a limited extent, neglecting the influence of the firing points in time and further effects like the refectory period and hyperpolarisation. Nevertheless, it allows powerful backpropagation algorithms that use gradient descent for learning to be employed [42].

Generally, a neural network contains multiple inputs and one or more outputs. For an exemplary application like image recognition, each pixel of a picture represents one input fed into a hidden layer of neurons. Each neuron receives an input from each pixel. The hidden layer outputs are routed into, for example, two output neurons, which differentiate

between cats and dogs (Figure 2.2(a)). Modern networks use multiple hidden layers, improving accuracy but increasing processing time since the output of each neuron has to be calculated and stored until it is needed as input for the next neuron. In addition, the number of synaptic weights that need to be stored grows. If the network exceeds the capacity of the central processing unit (CPU) or graphics processing unit (GPU) cache, it has to be offloaded to random access memory (RAM) storage, further increasing fetching times and hindering efficiency. These constructs are, in general, called deep neural networks (DNNs). Especially for image recognition, convolutional neural networks (CNNs) are favorable, as they use convolutional layers that can share weights. They are used in modern implementations because of their improved performance in grid-like input data sets. CNNs can learn to use filters across the entire image and use pooling layers that reduce the sensitivity to slight shifts of the input data [42].

Spiking neural networks (SNN) are an alternative approach to realizing BioNNs. Because of their more analog nature and utilization of the AP, SNNs come closer to replicating the human brain than ANNs. Leaky integrate and fire (LIF) is one of the most widely used models for describing SNNs. It is relatively simple to implement and has a low computational cost. Compared to ANN, the Heaviside step function is replaced by a Dirac pulse, which reduces power consumption. The neuron again sums up all inputs, but additionally integrates them. After reaching a certain threshold, the neuron fires an AP, represented now by a Dirac pulse, followed by a refractory period, which mimics the hyperpolarisation in BioNNs. The neuron potential can be modeled by the voltage across a capacitor parallel to a leaky conduction path. In LIF, because of the integration, the moment in time at which the plus reaches the neuron is also taken into account. If the weight of a single input is high enough, it can theoretically trigger the neuron to fire. However, if the weight is too small to reach the threshold, multiple pulses or single pulses from multiple presynaptic cells must reach the postsynaptic cell in quick succession to cause it to fire. Compared to ANNs, SNNs have been less investigated. Powerful training algorithms like backpropagation have not yet been established for SNNs. In addition, suitable data sets are missing, unlike ANNs, which would allow a similar quick development [13, 42].

It has been proven that in digital applications, SNN does not show significant energy efficiency improvement compared to ANN, because in these memory-intensive tasks, reading data is the main power consumption contributor. A hybrid ANN-SNN architecture appears to be desirable [43, 44]. For the realization of neural networks using emerging devices that are able to imitate synapses and neurons using a single device, SNNs are more suitable, and such networks have already been implemented [45, 46].

2.1.3 Realisation using Emerging Devices

Different approaches can be taken to realize neuromorphic systems. Firstly, as discussed in Section 2.1.2, software solutions using standard von Neumann architecture can be used, suffering from memory bottleneck issues. Secondly, ANN can be built to mimic the human

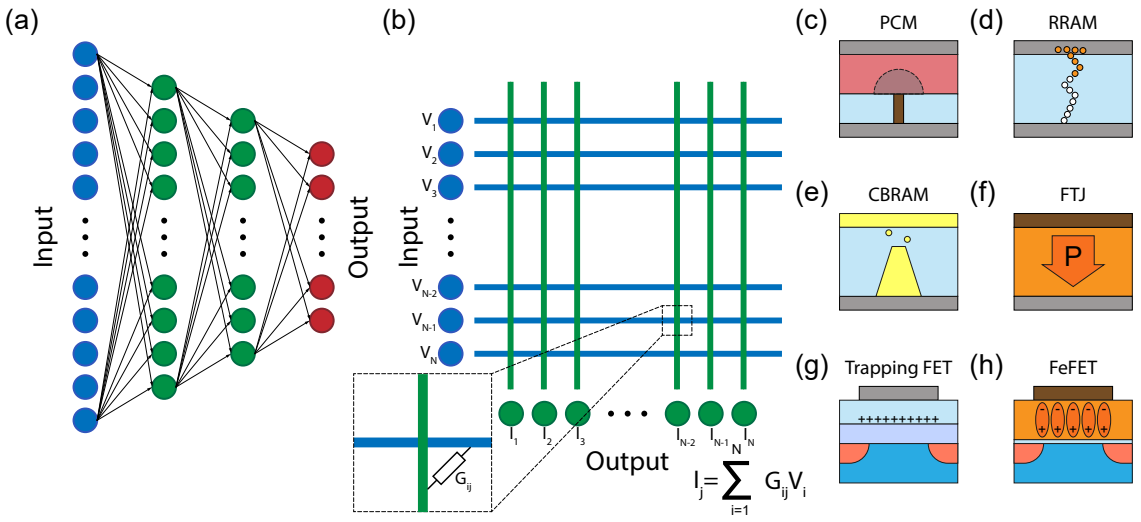

Figure 2.2: (a) ANN with input neurons (blue), output neurons (red), and hidden layer neurons (green). The connection of the other model neurons of each layer shows the number of connections needed. (b) Crossbar arrangement between the input and hidden layer in a neuromorphic system. Inset indicates connections via a memristive device. Image adapted from [7]. Structure of (c) phase change memory (PCM) [47], (d) resistive switching random access memory (RRAM) [48], (e) conducting bridge random access memory (CBRAM) [49], (f) ferroelectric tunnel junction (FTJ) [50], (g) trapping FET and (h) ferroelectric FET (FeFET).

brain synapses and neurons using modern CMOS technology, like IBM's TrueNorth & NorthPole [8–10] and Intel's Loihi 1 & 2 [11, 12] chips. These concepts require multiple CMOS transistors, using a lot of wafer area to represent one synapse. However, they are able to avoid the von Neumann bottleneck by utilizing a logic-in-memory approach, where the synaptic weights are stored near the processing. Finally, emerging devices can be integrated as single synapses or neuron devices. Reducing device numbers but losing space because of the less-scaled state of these structures.

Figure 2.2(b) illustrates a crossbar arrangement commonly used for implementing emerging devices as memristors. In this case, the input layer and first hidden layer are connected, identical to the first connections in Figure 2.2(a). As shown in the inset (Figure 2.2(b)), the horizontal input bar and vertical output bar are connected by a resistive element or memristor, which acts as a current valve. The output current is determined by summing all input voltages with their connecting resistive weights. So-called sneak paths, where current flows from a row to a column through multiple memristors, limit the size of crossbar arrays to one hundred times one hundred devices [51]. CMOS buffers can best be used to reduce the impact of sneak paths [52]. The conductance of the memristors resembles the synaptic weight, it can be set by specific pulsing schemes or a third terminal, making it similar to a transistor [7, 53]. A wide variety of devices is under investigation for application as memristors in neuromorphic computing. Hereafter, a quick overview of the major ones will be presented.

Phase change memories (PCMs) utilise the amorphous to crystalline phase transition in chalcogenides. They are two-terminal devices consisting of a phase change material and insulator layers stacked between two electrodes. The isolator is punctured by a metallic cylinder connected to the bottom electrode, allowing for current crowding. This "heater" has the ability to melt the phase change material. The resistive difference between the crystalline and amorphous phases is used as a low and high conductive state. Because of the high processing temperature, the phase change material is crystalline after fabrication. A short high-voltage pulse is applied to reset the PCM in an amorphic phase, which melts and quenches the material, resulting in a programmed high-resistive amorphous region around the "heater" in series with the low-resistive crystalline base layer. Set programming is done with a longer, lower voltage pulse that crystallizes the layer again completely [47, 54].

Resistive switching random access memory (RRAM) is a two-terminal metal-insulator-metal device featuring a binary transition between high and low resistive states (HRS & LRS). Soft breakdown is believed to be the main contributor during forming after fabrication. If sufficiently high electric fields are applied, oxygen atoms are knocked out of the dielectric and drift to the top electrode, leaving a conductive filament (CF) path of oxygen vacancies or metal precipitates in the case of metal-based oxides behind. Current flows through this CF in the LRS of the RRAM. Oxygen ions are stored at the top electrode, which acts as a "reservoir". During reset with an inverted field, oxygen ions migrate back and recombine with vacancies or oxidize metal precipitates, reducing the CF's length and increasing the oxide's resistance, bringing the RRAM into an HRS. Again, Soft breakdown takes place for set operations, similar to forming, but lower fields are needed. In general, the fields for set and reset are not identical. For a switch from unipolar to bipolar programming operation, the right material combination of the dielectric, top, and bottom electrodes is needed [48, 55].

Similar to RRAM, conducting bridge random access memory (CBRAM) uses a CF path for switching. CBRAM consists of an oxidizable active top electrode, a solid electrolyte isolator, similar to PCMs chalcogenides, and an inert bottom electrode. The set operation works by applying a voltage that causes the oxidation and drift of ions from the top electrode. When they hit the bottom electrode, they are reduced and form a CF, which grows until it reaches the top electrode, bringing the CBRAM into an LRS. Reset is performed by a reversed voltage that causes thinning of the CF until it separates, changing to an HRS [49, 56].

Ferroelectric tunnel junctions (FTJ) use a metal-ferroelectric-metal device structure. Detailed properties of ferroelectric will be discussed in the next section 2.2. FTJ ferroelectrics can be reduced to a layer with two different polarisation directions that allow for an additional potential across the dielectric, which influences the band structure of the device. Screening can be understood as the charges that are forced by the polarization at the metal/ferroelectric interface that attracts or repels charges over a short distance, allowing

an additional electrostatic potential at the interface. The screening length depends on the electrode material and can be between 0.1 nm and 10 nm. The additional potential of these charges is directly proportional to the screening length. For FTJ, the influence of both mechanisms on the band structure of the metal-ferroelectric-metal arrangement needs to be considered. Starting with a rectangular barrier, as known from dielectrics, adding the ferroelectric potential introduces a polarisation-dependent slope. By choosing different materials for the bottom and the top electrodes, the screening length differs on the two sides of the barrier, resulting in asymmetric potentials at the two interfaces. This asymmetry causes the effective barrier height to be not identical for the two polarisation directions, allowing the tunneling probability of the barrier and, accordingly, the device's resistance to be changed by a high voltage set or reset pulses that invert the polarisation [50, 57].

Additionally, trapping mechanisms in classical field effect transistors (FET) can be used as depicted in Figure 2.2. By using high voltages during set and reset defects in the semiconductor/oxide interface can be populated by charges from the channel, shifting the threshold voltage of the device for smaller operating voltages [58, 59]. Also in this direction are ferroelectric field effect transistors (FeFETs) which feature similar to FTJ, a ferroelectric. In this case, the threshold voltage is shifted by the ferroelectric remnant polarization (details will be discussed in Section 2.2). Since those devices are based on transistors, they have three terminals, helping to reduce sneak paths since the input and the output neuron are not directly connected (s. Figure 2.2).

The explained device can be engineered towards application as a neuron or a synapse. To be used as a neuron, the device must incorporate the SNN puls mechanics for time-dependent APs (Subsection 2.1.2), and an abrupt switch from HRS to LRS is needed. In contrast, for synaptic use, the emerging devices need to be able to tune their resistivity on a broad scale with high repeatability. Still, slight variations are negligible since NNs themself work with probabilities. Generally, two-terminal devices are more faithful to the BioNNs since the LTP and LTD are introduced using only one terminal, but three-terminal devices might be more suitable since they reduce sneak points and do not require additional amplification logic between network layers.

2.2 Ferroelectric Field Effect Transistor

The ferroelectric field effect transistor (FeFET) is one of the main building blocks of this thesis. It has seen wide research interest in non-volatile memories (NVMs) and in the field of logic-in-memory (LiM) computing. This section will explain the ferroelectric material used and its implementation in FETs. Further, FeFETs in neuromorphic computing will be discussed.

2.2.1 Ferroelectric Gate Dielectric

Ferroelectricity is understood as a permanent reorientable polarization. The naming is derived from the hysteretic relationship between polarization and an applied electric field, similar to hysteresis curves known from ferromagnets. It can only be found in noncentrosymmetric crystals, reducing the possible number of crystal structures from 32 crystal symmetries to 21 [60]. Of these, 20 are piezoelectric and 10 are pyroelectric, allowing them to support a stable electric dipole in an unstrained condition, understood as spontaneous polarization. From a crystallographic standpoint, pyroelectric and ferroelectric crystals are indistinguishable. Ferroelectrics are able to withstand the electric field needed for switching without the breakdown of the dielectric itself, making additional knowledge of the dielectric properties necessary for differentiating between the two crystal categories [61].

By implementing a phenomenological thermodynamic theory based on Landau's theory [62–64], ferroelectricity can be further examined. The transition from a nonpolar phase to a polar phase under the simplest condition of zero stress is derived by using Gibbs's free energy G:

$$G = (\frac{\alpha}{2})P^2 + (\frac{\beta}{4})P^4 + (\frac{\xi}{6})P^6 - EP \tag{2.1}$$

In which P is the polarization and α is a temperature-dependent coefficient. The minimum of G lies at $P = 0$ for positive α-values. In contrast, for negative values, two minima form symmetrically around the zero polarisation point and are separated by an energy barrier, giving rise to the two stable states and allowing the polar phase. In pyroelectric materials that do not exhibit ferroelectricity, this barrier is too high, preventing switching of polarization before dielectric breakdown.

Ferroelectric materials tend to form domains, which are regions of opposing polarization. Ferroelectrics allow, in the case of uniaxial materials, the dipole moment to be switched by 180° using an applied electric field. In contrast, orthorhombic ferroelectrics allow formation towards the <110> crystal phase edge, resulting in a 60°, 90° or 180°. Walls are spaces in between differently oriented domains. For uniaxial materials, lattice parameters are identical for the two sides of the wall, unlike in orthorhombic, where domains in line with the applied electric field grow in size, causing strain in the differently oriented domains by the ferroelastic effect and giving domains themself interfacial energy. Domains form to overcome depolarization fields (Subsection 2.2.3) and to minimize strain energy, as long as the energy penalty for not forming them exceeds. The spontaneous polarization forces charges on the surface of the material, causing a field, which is prevented by the formation of differently oriented domains. Additionally, domains form to minimize strain introduced by electrode clamping or nonuniform cooling [65].

Figure 2.3(b) shows the polarization P over the electric field E hysteresis curve of a

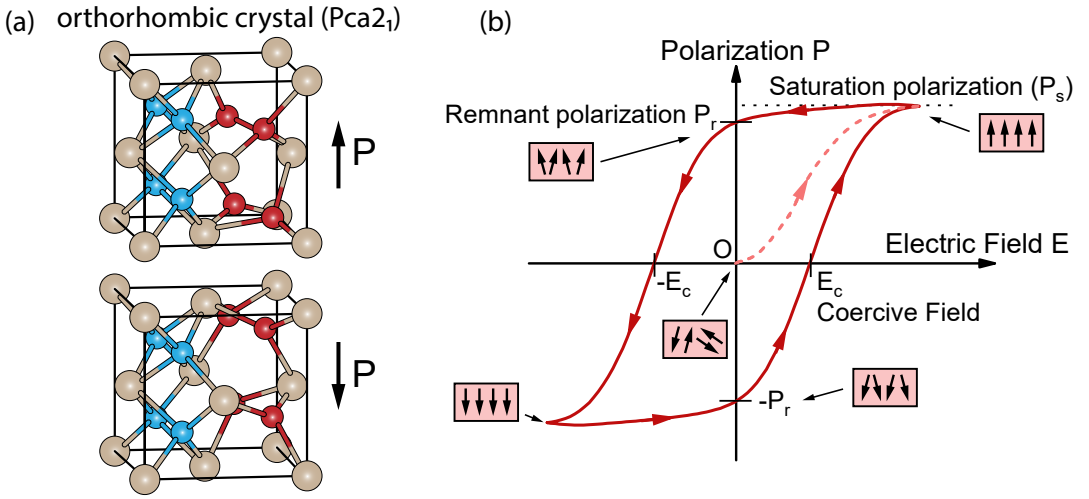

Figure 2.3: (a) Orthorhombic crystal phase ($Pca2_1$) of HZO in up and down ferroelectric polarisation states. O atoms are represented in blue for stationary and red for shifting during the polarisation. Hf/Zr are grey. Image adapted from [66]. (b) Hysteresis curve of standard metal-ferroelectric-metal (MFM) structure. Indicating domain polarisation directions as deposited, saturation, and remnant polarization states. Image from [67].

standard metal-ferroelectric-metal (MFM) device. In the initial state, the domains are polarized, as explained in different directions, ensuring charge neutrality. By applying an electric field the material gets polarized, following the dotted red line in Figure 2.3(b) until the saturation polarization P_s is reached the first time, at this point an increase in E does not result in an increase in P, because all domains are polarized in the same direction, higher fields would result in dielectric breakdown and destruction of the device. By reducing E back to zero, the polarization can be slightly reduced, but a very significant remnant polarization P_r remains since a majority of the domains are still in the minimum of Gibbs's energy landscape. Only after reaching fields that allow polarization to go over the energy barrier and become negative, a reduction of the device's polarization can be induced. The E value at which the P reaches zero is called the coercive field E_c. After passing the $-E_c$ point, the curve again saturates for a negative $-P_s$ value. By going back toward positive E the polarization reaches the $-P_r$ and E_c points ending in the P_s. Without external intervention, like heating, the point of origin can't be reached, and the device can now only be controlled in this non-linear hysteresis. The area enclosed within this curve represents the energy loss from each switching cycle, introduced by dipole switching and domain wall motion.

Schrödinger first proposed the two spontaneous polarisation states of ferroelectrics in 1912. Its name is derived from the German word "ferroelektrisch" [61]. Later, Valasek verified this crystallographic property in 1920, using Rochelle salt [68]. Ferroelectricity has seen a vibrant development over the past 110 years. The first crystal system suitable for practical applications, the perovskite structure, was discovered during World War II [69]. $BaTiO_3$ and $PbZrO_3$ (PZT) have been the two main materials of this kind. Integration of PZT in

FETs as NVM was limited to 130 nm technology due to its incompatibility with CMOS processing [70].

Böscke discovered ferroelectricity in fluorite-structured Si-doped HfO_2 ultra-thin films in 2011 [71]. This sparked a renewed interest since HfO_2 is the most common gate insulator used for modern transistors and can be fabricated using well-established atomic layer deposition (ALD) industrial-scale processes. Multiple different reports using Zr, Y, Gd, Sr, and La followed shortly, of which $Hf_{0.5}Zr_{0.5}O_2$ (HZO) is the most promising. In its orthorhombic crystallographic phase structure, polarization works in contrast to PZT in HZO by the displacement of four of the eight oxygen atoms in the unit cell, as depicted in Figure 2.3(a). Depending on the relative amount of Hf and Zr, HZO can transform from a paraelectric to a ferroelectric to an anti-ferroelectric characteristic [72]. Compared to PZTs, which has an E_c of 0.1 MV/cm, HZO features depending on the Hf to Zr ratio and layer thickness an E_c of 0.8-2.0 MV/cm. This allows for memory windows of around 1 V at thicknesses below 10 nm, underlying the potential "drop in" capability of HZO for almost every modern CMOS architecture [73].

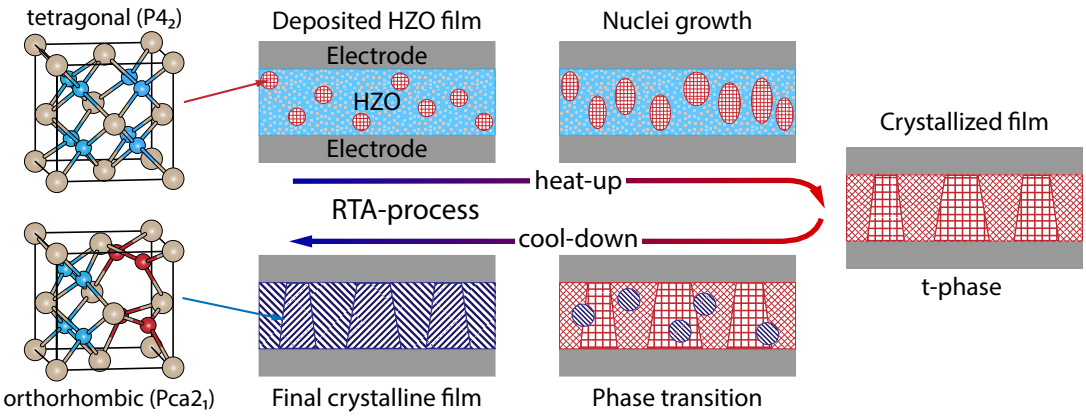

Figure 2.4: Crystallization cycle of HZO during RTA-process. (1) Small t-phase nuclei formation after ALD deposition. (2) An increase in temperature causes the growth of nuclei, which act as seeds for surrounding materials. (3) The entire film crystallizes in a mainly t-phase with a limited amount of m-phase. (4) During cooling, ferroelectric o-phase nuclei form. (5) The entire film remains in the o-phase at room temperature. Image from [67, 74].

In addition to ALD, subsequent rapid thermal annealing (RTA) and cooling process steps are required, in which the material undergoes different phase transitions, to form a ferroelectric phase (Figure 2.4). HZO can exhibit three relevant phases for this process. Firstly, the non-centrosymmetric orthorhombic phase (o-phase, $Pca2_1$) is the origin of ferroelectricity [75]. Secondly, in contrast to the o-phase, which is not thermodynamically stable, the monoclinic phase (m-phase, $P2_1/c$) is always stable [76]. Thirdly, the tetragonal phase (t-phase, $P4_2/nmc$) is a second unstable phase [72]. In the as-deposited state after ALD, it is understood that t-phase nuclei form (2 nm radius) because of very high surface mobility and surface-free energies, which are randomly distributed in the layer. During the RTA-process heating, the nuclei act as seeds for the surrounding material, resulting in the

growth of the individual nuclei until the entire film is completely crystallized in a t-phase. From a thermodynamics standpoint, it is expected that during RTA the t-phase transforms to m-phase, but this transformation is due to a high activation energy kinetically reduced allowing only a limited amount of m-phase. During cooling, the t-phase transitions to the o-phase [74].

To allow ferroelectrics to be used in memory applications, it is necessary to increase the layer height of the ferroelectric layer in order to decrease the leakage current. HZO exhibits a maximum thickness of 10 nm, after which the maximum remanent polarisation P_r is reduced [77]. For layers beyond this limit, the explained mechanism preventing m-phase formation becomes less efficient and allows only a limited amount of t-phase to remain during RTA, decreasing the o-phase generated at cooling and effectively capping the achievable P_r. By inserting Al_2O_3 interlayers splitting the HZO into multiple separate 10 nm layers one can prevent m-phase formation and reduce the leakage [78].

2.2.2 Implementation of HZO in Field Effect Transistor

In recent years, three major devices have been investigated in connection with neuromorphic computing and ferroelectric HZO. Firstly, ferroelectric capacitors (FeCAP) used in ferroelectric random access memory (FeRAM), FeCAPs are either implemented in a simple memory array with perpendicularly oriented bit and word lines [79] or in a 1T1C configuration, which allows a better array distribution. In the 1T1C concept, one terminal of the FeCap is connected to a drain of a standard transistor, while the other is connected to a plate line used for write operations. Word and bit lines are connected to the gate and source contacts. In contrast to current RAM approaches, FeRAM features very high endurance and non-volatile behavior, creating a new emerging memory platform [80, 81]. Secondly, ferroelectric tunnel junction (FTJ), which is very similar to FeCaps. FTJs utilize an HZO layer between two metal plates, but the dielectric is thinned down to a degree that enables quantum tunneling through the isolator. In a classical capacitor, this is unwanted. By switching the polarisation, the height of the barrier can be altered, changing the tunneling probability (Subsection 2.1.3). Thirdly, Ferroelectric field effect transistors (FeFET) will be further discussed in detail in the following subsection.

The main building block of the widely used backbone of the semiconductor industry, the MOSFET transistor, is the metal-oxide-semiconductor (MOS) capacitor. It can be constructed by oxidizing a silicon forming SiO_2 and further deposition of a metallic gating electrode. It is precisely the oxidation process, which allows the formation of a defect-free oxide, that gave silicon the advantage needed over germanium, which has been used for the first transistors and features a higher carrier mobility but does not form a stable oxide. The insulator layer has a larger bandgap compared to the semiconductor channel, which introduces a barrier for charge carriers. When stacked and no voltage is applied, the materials behave according to their different work functions, as will be discussed in more detail in subsection 2.3.1. If a negative voltage is applied to the metal, its energy level rises. This attracts holes at the interface, in the case of a p-type semiconductor, and is

called accumulation. When a small positive voltage is applied, band bending downwards is observed because the region at the interface is depleted. If the voltage is further increased until the intrinsic level of the semiconductor falls below its Fermi level, inversion occurs, and the semiconductor now shows n-type behavior at the interface region to the oxide. In an n-channel MOSFET two n-type regions representing the source and drain contacts are separated by a p-type region called the channel. Above the channel, an oxide and a metal gate contact are placed, forming together with the p-type semiconductor of the channel a MOS structure. In a band diagram along the channel, a barrier is formed, and its height can be controlled with the gate voltage, by bringing the channel into inversion, resulting in an n-type behavior, electrically connecting the source and drain regions on which a bias voltage is applied, allowing charge to flow [82].

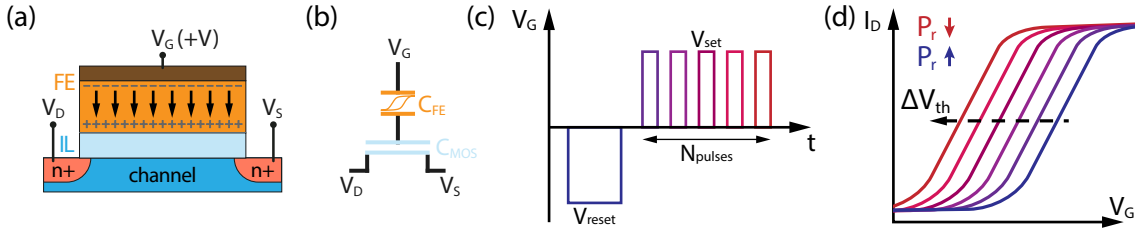

Figure 2.5: (a) Schematic illustration of FeFET featuring an isolating gate dielectric layer (IL) and ferroelectric layer (FE). Polarization charges and direction are indicated by arrows. (b) Simplified model of FeFET consisting of a standard MOSFET and a ferroelectric capacitor. (c) FeFET pulsing scheme using a single reset pulse and multiple set pulses. (d) Impact of the pulsing scheme on transfer characteristic. Image adapted from [67].

Figure 2.5(a) depicts schematically a FeFET. It is identical to the explained MOSFET except for the ferroelectric layer inserted between the oxide and the metal top gate (TG) contact, resulting in a metal-ferroelectric-insulator-semiconductor (MFIS) structure. As shown in Figure 2.5(b), this device can be modeled by adding a ferroelectric capacitor in series to a standard MOSFET. As long as E_c of the ferroelectric layer is not exceeded, the device will behave as a traditional transistor showing fixed threshold voltage V_{th}. By implementing different pulsing schemes in which the individual voltages exceed the E_c, the ferroelectric can be partially switched, meaning a limited amount of domains are inverted by each pulse. Figure 2.5(c) shows a common pulsing technique. First, the device is brought into a well-defined state by using, for example, a long negative pulse of the height V_{reset}, polarising all domains in an up direction. Following are identical positive pulses with the same V_{set} voltage. After N_{pulses}, the transfer characteristic is shifted towards negative values, as a result of a change in V_{th}, as indicated by Figure 2.5(d), storing a state in the device. Logically, the V_{th}-shift is limited since, at one point, all domains are polarized. The voltage separation between maximum up and down polarization states is referred to as the memory window (MW), the range in which the characteristic can be adjusted. The MW is often approximated by [83, 84]:

$$MW = 2 \cdot \alpha \cdot E_c \cdot t_F \tag{2.2}$$

where α is an ideality factor dependent on polarization and dielectric constant ϵ_F of the FE, E_c the coercive field and r_F the thickness of the FE. The standard maximum MW for HZO is 2 V. A nondestructive readout operation can be performed either by sensing the conductance at a constant gate value, commonly 0 V, or by measuring V_{th} using a limited voltage range. Both methods intend to influence the polarisation state as little as possible [73, 85].

2.2.3 Charge Injection through Dielectric Layer

Devices utilizing an MFM structure can be considered a fully charge-compensated configuration since enough charges can be supplied by the electrode contacts to mitigate the depolarization field, which can be as high as the polarization itself. In FeFETs, as discussed in Subsection 2.2.2, a dielectric IL layer (commonly SiO_2) is used, creating an MFIS structure that is to a certain extent charge-uncompensated, because one side of the FE is electrically separated from the corresponding electrode, enabling the build-up of an depolarization field due to unscreened bound charges, reducing the achievable MW with the given ferroelectric [86].

SiO_2 ILs have to be integrated into FeFETs because, like most high-k dielectrics, HZO features a lot of interface traps to silicon, resulting in reduced mobility for charge carriers in the transistor channel [87]. Thermal SiO_2 has superior interface quality and is commonly grown between 0.9 nm and 2 nm, allowing most top gate voltage to be applied to the ferroelectric due to the capacitive voltage divider. Additionally, the influence of oxide vacancy traps that act as fixed charges in the HZO is reduced [88].

Since the IL is relatively thin, quantum-mechanical Fowler-Nordheim tunneling is possible for high enough voltages. The SiO_2 can be modeled by a capacitor parallel to the resistor, which starts conducting when the band bending reaches a triangular shape, allowing tunneling. Before injection through the IL, depolarization charges build up at the SiO_2/HZO interface, reducing band banding of HZO and increasing it for the SiO_2. After a threshold is reached, charges opposite to the depolarization charges are injected through the IL, screening bound charges at the interface. This reduces the depolarization field and the voltage across the IL. Further, it allows the ferroelectric to sustain its polarization state more effectively, improving the endurance of the device. Depending on the height of the applied voltage across the structure, the charges can be under, exact, or over-compensated [89]. Simulations have shown that the explained mechanism and the presence of interface traps at the SiO_2/HZO interface are fundamental for achieving ferroelectric polarization switching [90]. Other concepts like the in Subsection 2.2.5 explained metal-ferroelectric-metal-isolator-semiconductor (MFMIS) structure provide infinite states in between the IL and the FE, making good control over the amount of charge injected in the floating metal layer highly important [86, 91–93].

2.2.4 Usage in Neuromorphic Computing

Section 2.1 motivated the usage of emerging memory devices from a biological standpoint, in which one device can be used as a neuron or synapse in SNNs. Compared to others (Subsection 2.1.3), FeFETs are compatible with current CMOS processing (Subsection 2.2.1), allowing quick scaling of this architecture. Figure 2.5(c) and (d) have already outlined the basic switching mechanism in FeFETs, using a simplified pulsing scheme. In Figure 2.6, three more pulsing approaches are compared, and their influence on the transistor is schematically outlined. As discussed in Subsection 2.2.2, the readout operation can be performed using a fixed gate voltage and measuring the current, which is then transferred into a conductance value. In general, neurons and synapses have different requirements in terms of pulsing. Neurons receive multiple pulses, which are weighted by synapses. When they arrive in quick succession, the transistor should switch from an LRS to an HRS. In contrast, the conductance in synapses is changed during learning, which resembles the change of weight or importance, making precise control necessary (Subsection 2.1.1) [94].

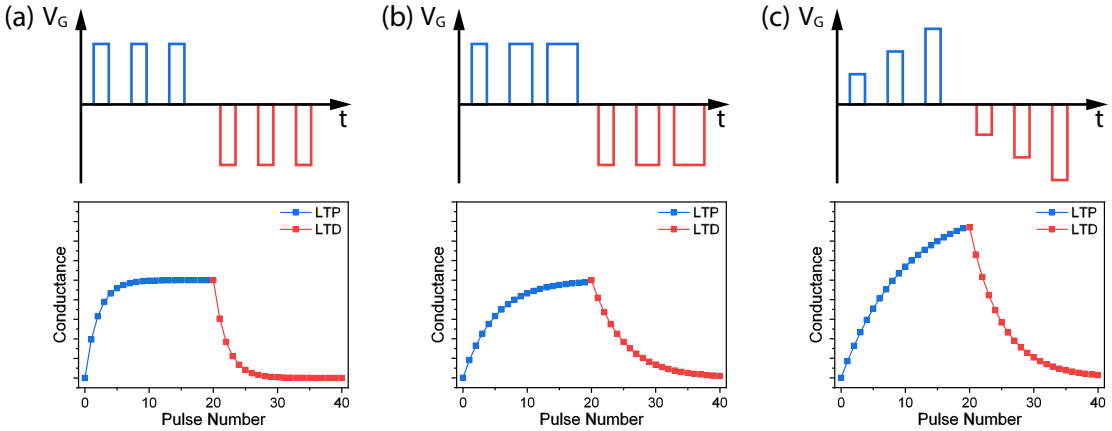

Figure 2.6: FeFET pulsing schemes and conductance values for neuromorphic applications. (a) Schematic illustration of pulsing scheme with identical length and height. Conductance LTP/LTD curves feature only a limited number of intermediate states between the HRS and the LRS. (b) Pulsing scheme with identical pulse height but sweeping pulse length, resulting in more intermediate states in LTP/LTD curves. (c) Pulsing scheme with identical pulse length but sweeping pulse height, increasing conductance maximum, and number of states. Image adapted from [95].

The first pulsing scheme in Figure 2.6(a) consists of equal pulses of identical duration and voltage amplitude. For LTP, positive pulses are used, and for LTD, negative pulses. As evident from the conductance graph, only a limited number of intermediate states between the LRS and the HRS are present, making it ideal for use as a neuron. Pulsing scheme number two (Figure 2.6(b)) uses again pulses of identical height but sweeps the length of the pulses to longer times for ongoing pulse number. Since this scheme can start with lower pulse widths than scheme number one, the initial conductance increase can be reduced, allowing for more intermediate states and gradual control needed for synapses. Pulsing scheme number three (Figure 2.6) has again identical widths but sweeps the pulse height, enabling many intermediate states and unlocking higher conductance states since higher voltages are used. The change in the voltage level makes this scheme difficult to implement

in integrated circuits (IC), requiring more wafer area than changing only the widths, making scheme number three, despite its greater conductance control, unsuitable [96, 97].

The variety of pulsing schemes presented in Figure 2.6 shows that FeFETs can be used as neurons and synapses, allowing for a possible FeFET-only implementation of SNN [95].

2.2.5 Floating Gate FeFET

In contrast to the emerging HZO-based NVM technology, classical approaches implementing charges stored in the gate stack are used in commercial flash memory products at the moment [98]. The two main groups of this kind are charge trapping and floating gate (FG) devices. Both work by tunneling and trapping charges from the silicon at an interface region. Charge trapping devices utilize traps at a nitride/oxide interface, and FG devices insert a metal layer between two oxides. In both cases, the oxide near the silicon is thinned down to enable tunneling. The stored charges in the gate stack give rise to a threshold-voltage shift, allowing for a programmed and erased state of the device. Modern memory designs feature low leakage of charges out of the interface, enabling retention times of over 100 years [99].

Two charge injection mechanisms are used to introduce the threshold-voltage shift in FGFETs, either by hot electron injection (HEI) or Fowler-Nordheim (FN) tunneling. In HEI, electrons traveling from source to drain can be subjected to a lateral electric field. If this field exceeds 100 kV/cm [100], the electrons are no longer in equilibrium with the lattice, causing them to gain energy in relation to the conduction band edge, also known as "heating". A small fraction of electrons acquires enough energy to surmount the barrier between the silicon and FG. Two conditions must hold for an electron to cross the barrier. Firstly, it has a kinetic energy higher than the potential barrier. Secondly, its velocity direction is in a cone directed toward the gate. Factors like the distance of the electron to the barrier and the shape of the potential barrier itself can prevent FG injection, causing only a fraction of electrons to cross [101, 102]. Additionally, the high electric field can induce impact ionization, which can generate secondary hot electrons that can also be injected into the FG [99]. For this reason, degradation of the oxide is a major consideration when HEI is used. The oxide region above the channel outside the drain is affected the most by carrier trapping and the generation of interface states, which can cause an increase in leakage current and decreased endurance [103, 104].

Because of this degradation, modern devices mainly rely on FN tunneling. The nonzero solutions of the Schöndinger equation imply the possibility for electrons to penetrate the forbidden region of a barrier and allow tunneling from one classical-allowed region to another in MOS structures. The tunneling probability depends on the shape, height, and width of the barrier. When a bias is applied to the MOS structures, the oxide is bent into a triangular shape, which reduces the tunneling distance and enables FN tunneling. Using WKB approximation, one obtains the following expression for the current density [105]

$$J = \frac{q^3 E^2}{8\pi h \phi_B} \exp\left(-\frac{4\sqrt{2m^*_{ox}\phi_B^3}}{3\hbar q E}\right) \tag{2.3}$$

where ϕ_B is the barrier height, m^*_{ox} is the effective mass in the dielectric, h is the Planck constant, q is the charge of the electron, and E is the electric field through the oxide. Because the field is directly proportional to the applied bias voltage V_{bias} and indirectly proportional to the oxide thickness t_{ox}, the oxide can be engineered to allow the most current with acceptable voltages. In addition, leakage of charges while stored has to be considered since it determines retention time. An optimum thickness of 10 nm is chosen in modern devices [102].

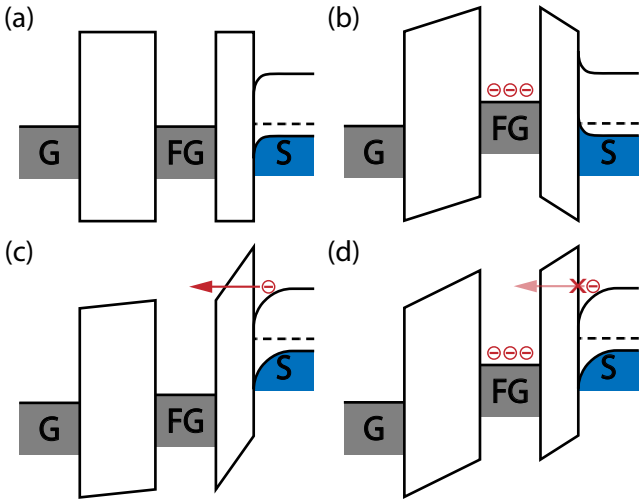

Figure 2.7: Floating gate (FG) transistor band diagrams. (a) Uncharged and (b) charged FG at equilibrium. (c) High electric fields enable Fowler-Nordheim tunneling from the semiconductor into the FG. (d) FG gets charged, preventing further tunneling through the oxide. Image adapted from [99].

The band structure of FG devices changes because of charging over the duration of a pulse, altering injection mechanisms and device behavior. Figure 2.7 depicts schematically the band structure of an FG device at different stages in a simplified 2D region over the silicon channel. The thickness difference of the thicker control oxide between the gate and FG, as well as the thinner tunneling oxide between the FG and semiconductor, is illustrated. When the charge Q in the FG is zero, and no voltage is applied, the band structure in Figure 2.7(a) is established. The Fermi level of the gate, FG, and semiconductor are in line. No electric fields are present at the dielectrics, and the semiconductor bands are pinned accordingly to the tunneling oxide. The two oxides form capacities in different regions of the device. C_C, C_S, C_D, and C_B are the capacitances between the TG and FG as well as the FG and the source, drain, and bulk regions of the transistor. Using Gauss' law

$$Q = C_C(V_{FG} - V_G) + C_S(V_{FG} - V_S) + C_D(V_{FG} - V_D) + C_B(V_{FG} - V_B) \tag{2.4}$$

where V_{FG} is the potential on the FG, V_G, V_S, V_D, V_B the potential of the gate, source, drain, and bulk respectively. The sum of all capacitances can be defined as the total capacitance $C_T = C_C + C_S + C_D + C_B$. For the assumption that $V_S = V_D = V_B = 0\,V$ Equation 2.4 can be rearranged as

$$V_{FG} = \frac{C_C}{C_T}V_G + \frac{Q}{C_T} = \alpha_G V_G + \frac{Q}{C_T} \tag{2.5}$$

in which α_G is the gate coupling coefficient. α_G is normally > 0.6 in standard FG devices. Figure 2.7(c) depicts the beginning of the FG charging using a positive voltage. Because the FG itself is still uncharged, one can see using Equation 2.5 that the potential V_{FG} is closer to V_C than that of the semiconductor, which is, in this case, zero, resulting in significant band bending of the tunnel oxide. FN tunneling of electrons takes place in accordance with Equation 2.3, charging the FG. Since Q gets progressively more negative, V_{FG} is reduced, decreasing the amount of band bending of the tunnel oxide, until FN tunneling is suppressed, as depicted in Figure 2.7(d). If V_G is reduced back to zero, the accumulated charges Q in the FG result still in a V_{FG}. Figure 2.7(b) shows that now electric fields of opposite directions are present at the control and tunneling oxide. The latter one bends the semiconductor bands up, resulting in a threshold voltage shift ΔV_{th} according to the charges Q in the FG

$$\Delta V_{th} = -\frac{Q}{C_T} \tag{2.6}$$

as long as the applied V_G voltage range does not alter Q. This mechanism is used for storing information in NVM memories using FGFETs. As described, positive V_G results in the injection of electrons and a positive ΔV_{th} shift for program operation. Negative V_G, on the other side, results in the injection holes, making ΔV_{th} negative during the erase operation [99, 102].

The discussed approach can be additionally implemented into FeFETs by insertions of metal layers into the already discussed metal-ferroelectric-isolator semiconductor (MFIS) stack (Figure 2.8(a)), converting it to a metal-ferroelectric-metal-isolator-semiconductor (MFMIS) stack (Figure 2.8(b)). Such approaches have been reported in different architectures. Devices with an InGaZnO (IGZO) active channel have been shown by [106, 107] and devices based on a standard FET structure by [108–111].

MFMIS can improve the FeFET in two diffrent directions. Firstly, only the semiconductor region in the FET channel that is located beneath a switched domain of the ferroelectric

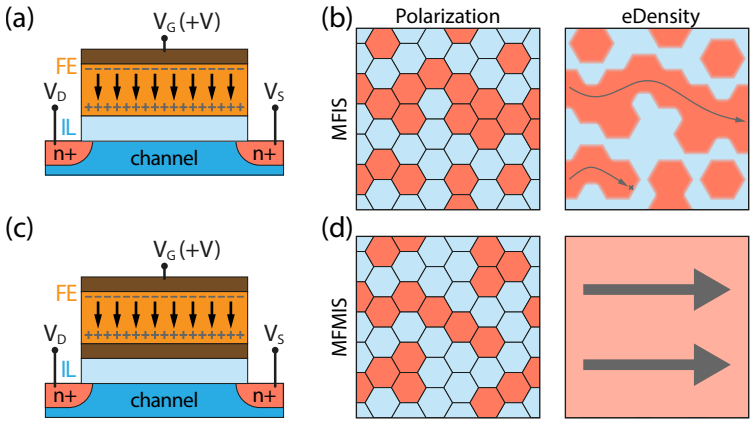

Figure 2.8: FeFET domain polarization and electron density for MFIS and MFMIS stacks. (a) Illustration of FeFET with MFIS stack, indicating polarization direction and resulting charges. (b) Domain polarization and electron density for the MFIS stack, showing the formation of current paths. (c) Illustration of FeFET with MFMIS stack. (d) Different domain polarization and electron density for the MFMIS stack, averaging of individual domains over the entire channel. Image adapted from [112].

goes into inversion for classical MFIS stacks. Figure 2.8(b) shows schematically the Polarization of the individual domains and the resulting electron density in the channel of an MFIS FET. Domain switching is randomly distributed. The number of domains that switch can vary from pulse to pulse. Also, the total number of available domains themself can be different between devices, resulting in variability. The electron density shows the formation of current paths connecting the source and drain regions. Again, this formation underlies a distribution since the switched domains need to be on a single line. Depending on the applied pulses, not enough domains might switch, prohibiting current flow. This behavior is desirable in neurons, where a step-like behavior of the current flow is needed. But for synapses where the current flow needs to be gradually adjusted depending on the pulses, this is a problem. By implementing an MFMIS stack, as depicted in Figure 2.8(c), the number of switched domains is averaged over the channel, reducing the impact of domain count variations, eliminating the formation of current paths, and allowing gradual switching [35, 112]. In addition, another degree of freedom for device design is created. By altering the ratio of the overlap between the top and bottom metals, the memory window can be adjusted at the expense of a reduced on/off ratio [113].

Secondly, as discussed in Section 2.2.3, the injection of charges into the ferroelectric/oxide interface is needed to compensate the depolarization field, enabling sustained switching of domains. Subsequently, a sufficient number of traps needs to be present at the interface to accommodate the needed charges. Engineering their density is one of the main hurdles for ferroelectric devices. The FG insertion can store the needed charges for compensation. One has to distinguish between devices where the oxide to the silicon channel does not allow tunneling and where it does, bringing the resulting devices and their switching mechanism close to the discussed FGFET. For a completely isolating oxide, it is assumed

that during ferroelectric switching, charges of opposite signs to the polarization in the metal are moving right below the domains to compensate for the depolarization field. This causes a charge separation between the top and bottom of the FG. After pulsing, this separation remains, enabling the program and erase operations [114]. For oxide layers that allow FN tunneling, different considerations, like injection amount and time, have to be considered. The charge Q in Equation 2.5 is replaced by a net polarization P_{net} which sums up the polarization P of the ferroelectric and the injected charges Q_{ML} in the metal layer

$$P_{net} = P + Q_{ML} \tag{2.7}$$

In general, P and Q_{ML} are of opposite sign. Charge separation does not appear, the entire layer is of the same charge, which inverts the normally counterclockwise operation of FeFETs into clockwise for $P < Q_{ML}$ [115].

2.3 Schottky Barrier based Devices

The following section discusses the formation and transport mechanisms in Schottky barriers (SBs), their usage in Schottky barrier field effect transistors (SBFETs), and a more novel evolution based on similar methods, the reconfigurable field effect transistor (RFET).

2.3.1 Formation and Transport Mechanism

The work function $q\phi_m$ of a metal is defined as the minimum amount of energy needed to remove an electron from the metal to the vacuum (Figures 2.9(a)). In metals, charge carriers are bound to the Fermi level E_F, allowing the work function to be written as

$$q\phi_m = E_{vac} - E_F \tag{2.8}$$

Due to different electron configurations and lattice structures, this value differs from metal to metal. In a semiconductor, a similar definition is made for $q\phi_s$, even dough in a semiconductor E_F is between the conduction band energy E_C and the valence band energy E_V known as the band gap E_g or forbidden region, in that no charge carriers are allowed. Again, these values change between different semiconductors. Additionally, the electron affinity $q\chi_s$ is the energy difference between the E_{vac} and E_C with

$$q\chi_s = E_{vac} - E_C \tag{2.9}$$

By doping the semiconductor, turning into a n-type (Figure 2.9(a)) or p-type (Figure 2.9(b)) material, the position of E_F in relation to E_C and E_V is changed, but $q\chi_s$

stays unaffected, resulting in a new value for $q\phi_s$. At the equilibrium, the bands of the semiconductor are horizontal [116].

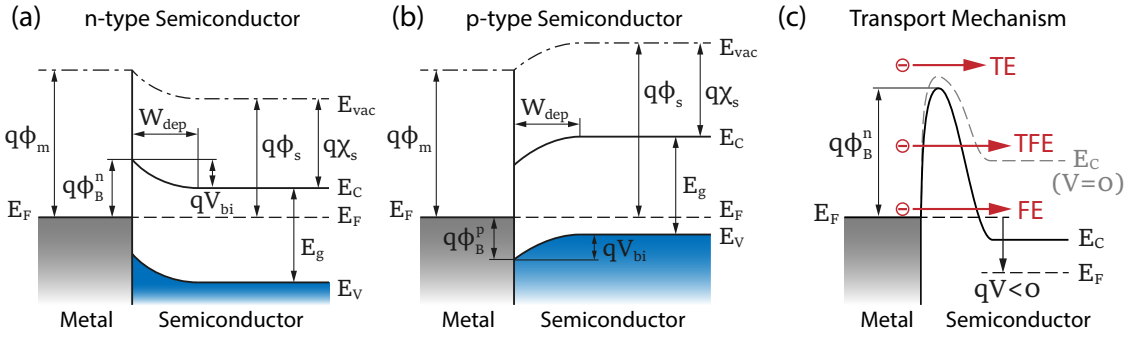

Figure 2.9: Schottky barrier band diagram with indicated energy levels and work functions. (a) Band diagram of a metal brought into contact with an n-type semiconductor. Work function condition $\phi_m > \phi_s$ fulfilled, allowing the formation of the Schottky barrier. (b) Metal in contact with a p-type semiconductor with fulfilled $\phi_m < \phi_s$ condition. (c) In-depth view of Schottky barrier transport mechanism. Thermionic emission (TE), thermionic field emission (TFE), and field emission (FE) are depicted. Image adapted from [67].

When the semiconductor is brought into contact with a metal, a barrier forms because of the difference in work functions. Thiar Fermi levels align and a charge transfer creates a depletion region of the width W_{dep}, resulting in band bending in the semiconductor. If $\phi_m > \phi_s$ for n-type semiconductors (Figure 2.9(a)) and $\phi_m < \phi_s$ for p-type semiconductors (Figure 2.9(b)) the formed barrier is classified as Schottky contact or Schottky barrier (SB) depleting their representative majority carriers. If ϕ_m and ϕ_s switched places in the prior conditions, an Ohmic contact would form, which will be not discussed going forward. The flat band voltage V_{FB} is referred to as the needed potential difference to compensate for the built-in voltage V_{bi} introduced by band banding. This voltage would bring back the bands of the semiconductor to the equilibrium state. The SBs height ϕ_b measured relative to the Fermi level is given for an ideal case without interface charges by

$$q\phi_B^n = q(\phi_m - \chi_s) \tag{2.10}$$

for an n-type semiconductor and

$$q\phi_B^p = E_g - q(\phi_m - \chi_s) \tag{2.11}$$

for p-type semiconductors [82, 116].

The aforementioned interface charges alter the barrier formation drastically by changing the pinning of the valence and conduction bands to the metal. In the ideal case, without interface charges and without interface dielectric, they simply pin to the same levels as

they would in the equilibrium. Metal-induced gap states (MIGS) allow electrons to tunnel into the bandgap of the semiconductor at the interface in association with the tails of their conduction wave function. At the interface, this allows extra states near the E_F, which extend a few atomic layers into the semiconductor. This results in a lower pinning position for E_C and E_V because of screening by the additional charges, reducing the influence the work function has on the $q\phi_B$. For Si, most metals pin near the mid-gap, creating similar barriers for electrons and holes. In Ge, pinning is concentrated near the valence band, hindering the fabrication of ambipolar transistors (Section 2.3.2) [116–118].

Furthermore, it is generally required that electric fields must be perpendicular to a metal surface, causing a force on charges that approach it, known as the image force, since calculations are done using the mirror-image method. This attractive force gives the charge potential energy, which has to be added to the potential barrier of the SB. Figure 2.9(c) depicts the effect an electron has on the SB in detail. For electrons, the image force works subtractively, reducing the height of the SB by lowering the conduction band. Holes bend the valence band upwards, working additively but still decreasing the SB height for them. Additionally, the barrier becomes rounded with its peak value shifted inside the semiconductor. Unlike other effects like MIGSs, which affect the barrier independent from the presence of charges in the valence or conduction band of the semiconductor, the impact of the image force depends on the presence and the number of charges [116].

For current to flow, one has to apply an external voltage on the SB, as shown in Figure 2.9(c) for the conduction band of a metal to n-type semiconductor junction. The E_F of the semiconductor is lowered, allowing charge carriers to flow from the metal to the semiconductor using three transport mechanisms: thermionic-emission (TE), field-emission (FE), and the combination of both thermionic-field-emission (TFE).

In thermionic-emission theory, it is assumed that the current-limiting process across the interface is represented by the barrier itself. In the depletion region, the effects of drift and diffusion are neglected. The quasi-Fermi levels for electrons and holes coincide with the Fermi level of the bulk semiconductor. The total current over the barrier by thermionic-emission J_{therm} can be split into two parts. J_{sm} represents the current density from the semiconductor to the metal. The electron concentration is dependent on the applied bias field since the barrier height is altered, making multiplication with $\exp(qV/kT)$ necessary. The second part is the current density from the semiconductor to the metal J_{ms}. For an SB, J_{ms} remains independent from the bias voltage. Because these two components are in opposite directions, they have to be subtracted

$$J_{TE} = J_{sm} - J_{ms} = A^* T^2 \exp\left(\frac{-q\phi_B}{kT}\right)\left\{\exp\left(\frac{qV}{kT}\right) - 1\right\} \tag{2.12}$$

where

$$A^* = \frac{4\pi m^* q k^2}{h^3} = 1.2 \times 10^6 \left(\frac{m^*}{m_0}\right) \mathrm{A\,m^{-2}\,K^{-2}} \tag{2.13}$$

A^* is the Richardson constant for which the effective electron mass m^* has been substituted for the free-electron mass m_0. Logically barrier lowering mechanisms like MIGS and image-force effect have to be taken into account when implementing the $q\phi_B$ in TE [116].

As depicted in Figure 2.9(c), by applying a bias voltage, not only does $q\phi_B$ get lowered, allowing for more TE, but also the width of the barrier is thinned down, allowing for quantum mechanical tunneling to happen. In a low-temperature regime, only charges near E_F are excited, which have the possibility to tunnel through the barrier known as FE. If the temperature is slightly increased, higher energy charges are excited, allowing tunneling in higher regions of the barrier to happen, named TFE. This mechanism sees a reduced barrier width because the SB gets thinner going upwards, allowing for more probable tunneling of these charges. If the temperature is further increased, the higher current density TE takes mainly over, and tunneling can be neglected. The current/voltage relationship for FE can be written as

$$J_{FE} = J_s \exp\left(\frac{V}{E_0}\right)\left\{1 - \exp\left(\frac{qV}{kT}\right)\right\} \tag{2.14}$$

Which is similar to Equation 2.12. At low temperatures, the factor E_0 remains independent of temperature, representing the FE case. In the higher temperature TFE situation $E_0 = nkT/q$, in which n is an ideality factor. Allowing a smooth transition from FE to TFE. J_s is dependent on temperature, barrier height, and other semiconductor parameters [116, 119, 120].

2.3.2 Schottky Barrier Field Effect Transistor

Schottky barrier field effect transistors (SBFETs) can be motivated using the first FET concept patented by Lilienfeld in 1930 [121], the junctionless accumulation mode (JAM) FET. His research was limited by the smallest possible structures at the time, but today, with channel widths in the nanometre spectrum, JAM devices can be realized. Unlike a normal MOSFET, the channel, drain, and source regions of the device are all of the same homogeneously doped semiconductor. Either a gate-all-around (GAA) approach, known from bottom-up nanowire transistors, or a fully depleted silicon-on-insulator (FDSOI) approach, as shown in Figure 2.10(a), is used, in which a MOS capacitor is formed below the gate. Due to the narrow channel, the depletion or accumulation regions controlled by the gate can occupy the entire width of the channel, effectively allowing modulation of its band structure and giving rise to a barrier separating the source/drain contacts. By gate voltage modulation, the height of this barrier can be controlled. High voltages raise the barrier until pinch-off, the point at which the entire semiconductor below the

gate is inverted, preventing current flow, occurs. Thus, this device shows a normally open characteristic, allowing current to flow when no voltage is applied to the gate, which is not ideal for applications like power electronics [122]. As already mentioned, one advantage of JAM FET is that only one material is needed for the channel, drain, and source, improving device-to-device variability since doping using ion implantation is not entirely reproducible. Short-channel effects such as drain-induced barrier lowering (DIBL) are reduced. Because, unlike an inversion mode device, where the effective channel length L_{eff} is in the off-state shorter than in the on-state, in JAM FETs L_{eff} is larger in the off-state, improving DIBL. Also in comparison, the electric field perpendicular to the current flow in the on-state is lower, resulting in a higher channel mobility and higher current drive [123].

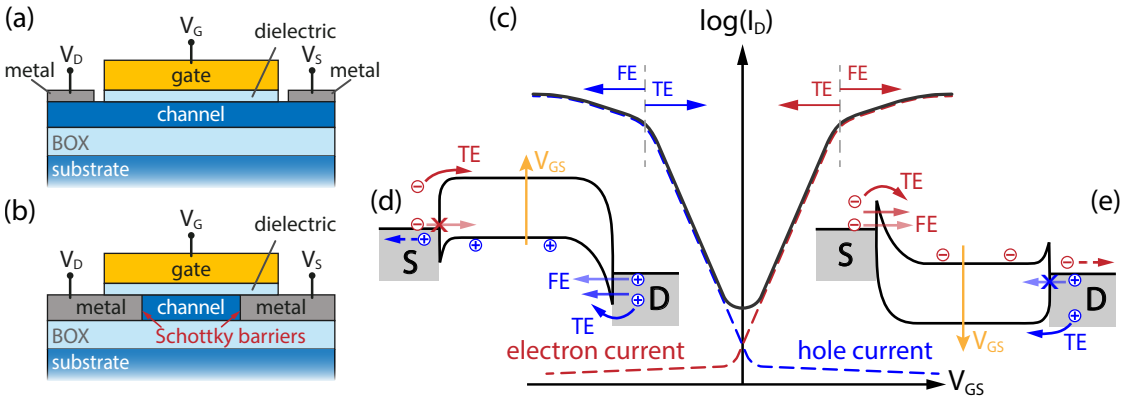

Figure 2.10: (a) JAMFET on SOI substrate with only semiconductor channel below gate contact. (a) SBFET featuring metal source/drain contacts, Schottky barriers at the interface to silicon, and gate contact on the top. (b) Ambipolar transfer characteristic of SBFET. Indicated contributions from electron and hole currents as well as transition from TE to FE transport mechanism. Band diagram for (c) hole and (d) electron conduction. Image adapted from [67].

Figure 2.10(b) depicts schematically the structure of an SBFET. Instead of the metal ohmic contacts being used only in the JAM FET for electrically connecting the source and drain to the silicon, they become in the SBFET essential to the device characteristic, forming SBs and being placed below the gate, resulting in a band structure that can be similarly controlled as the JAM FET. Figures 2.10(d,e) show this for a negative and positive gate voltage V_{GS}. The SBs are oriented opposite to each other, one in the forward direction and one in the backward direction, allowing for a normally-off characteristic. In Figure 2.10(c) a simplified exemplary drain current I_D over V_{GS} transfer characteristic is shown. In SBFETs, the contributions from electrons and holes can be separated due to their different transport mechanisms over the V_{GS} range, but the perceived total current at the drain is the sum of both. In Figure 2.10(d), the band structure is depicted for $V_{GS} \ll 0\,V$, where hole current dominates because the source site SB for electrons is bent upward, increasing $q\phi_B$, reducing TE, and preventing FE. $q\phi_B$ at the drain site SB for holes is constant since the bands are pulled up, keeping TE constant as well. Additionally, FE is enabled, allowing holes to be injected into the channel. Because of the applied source/drain bias voltage V_{SD} holes diffuse and accelerate toward the source, gaining energy, until they reach the SB, which they can easily surpass due to their increase in energy and the reduced

$q\phi_B$ also as a result of V_{SD}. When voltage is increased to $V_{GS} < 0\,V$ FE by the drain SB becomes less probable transitioning into a TE-dominated regime, indicated by a kink in the transfer characteristic. At $V_{GS} = 0\,V$, $q\phi_B$ reaches the same value on the source site for electrons and on the drain site for holes, resulting in a change from hole conduction to mainly electron conduction. Again for $V_{GS} > 0\,V$ TE and for $V_{GS} \ll 0\,V$ FE of electrons takes place.

The explained characteristic resembles the ideal case in which the $q\phi_B$ for holes and electrons is identical, requiring precise midgap pinning of the conduction and valence band, which is not the case for most metal/semiconductor combinations. MIGS, as well as the influence of the image force, have been neglected. Additionally, the V_{SD} results in a shift of the characteristic. Normally, a reduced on-current for electrons is observed in non-ideal devices.

The addition of metal-drain and metal-source regions does not immensely increase complexity in comparison to JAM FET devices, particularly in comparison to the differently doped regions in the classical MOSFET design. Short-channel effects like GIDL are suppressed because no p-n junction is needed in an SBFET. The fabrication and variability of junction sharpness are the biggest challenges in the implementation of SBFETs [122].

2.3.3 Reconfigurable Field Effect Transistor

The reconfigurable field effect transistor (RFET) is an emerging transistor technology that allows switching between a hole conduction p-mode and an electron conduction n-mode during runtime by the implementation of an additional terminal, creating a four-terminal device. Since RFETs can be based on SBFETs (Section 2.3.2) they exhibit similar device physics. Figure 2.11(a) depicts an RFET that is identical to the SBFET shown in Figure 2.10(b) except for the top gate, which is split into three individual structures, representing a triple top gate (TTG) design. Two connected gates are placed over the SBs specified as polarity gates (PGs) and the control gate (CG) is placed over the channel. Figure 2.11(b) represents a simplified dual top gate (DTG) structure, in which the CG is removed from on top of the device, and an intentional ungated region remains. The control gate voltage V_{CG} is now connected to the back gate (BG) of the device, allowing control of the ungated region. Because the PGs remain over the SBs and their influence is stronger than that of the BG, the polarity gate voltage V_{PG} stays in control over the SBs despite the slight influence from V_{CG}.

Figure 2.11(c) shows schematically the individual p and n-mode transfer characteristics of an RFET. If both V_{PG} and V_{CG} are $\ll 0\,V$ the device band structure depicted in Figure 2.11(d) looks identical to the band structure of the SBFET during hole conduction in Figure 2.10(d), making this the p-mode on state. By increasing the V_{CG} to 0 V, a barrier rises in the middle of the band structure (Figure 2.11(e)), due to band bending of the conduction and valence band beneath the CG, which hinders conduction. Under the CG, a similar situation to the JAM FET is created, in which the band diagram is in its

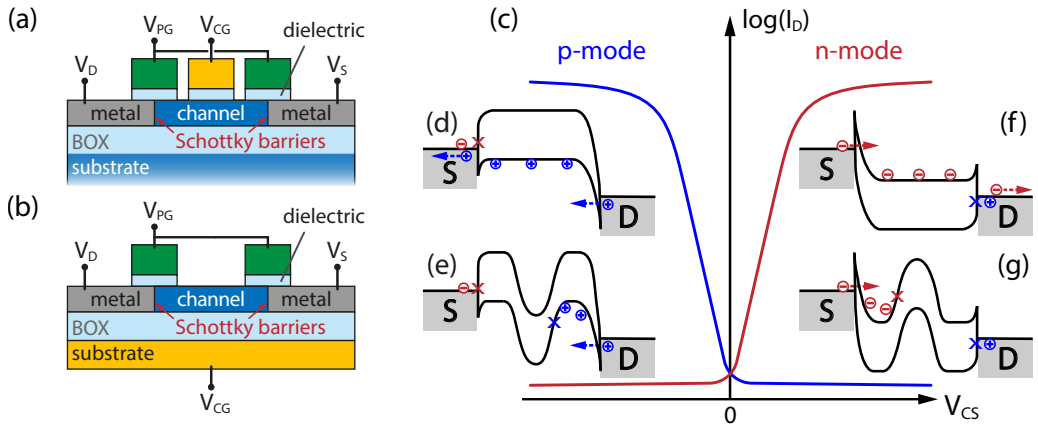

Figure 2.11: (a) Triple dop gate (TTG) SB-based RFET design using two connected polarity gates (PG) over SB to set conduction mode and one control gate (CG) over the channel for current modulation. (b) Dual top gate (DTG) design, CG over back gate contact. (c) RFET transfer characteristics for p and n conduction mode. Band diagram for p-mode (d) on state (e) off state and n-mode (f) on state (g) off state. Image adapted from [67].

equilibrium position, but due to the height of the surrounding semiconductor controlled by the PGs, the barrier for holes is created, making this device normally off. Since the width of this barrier hinders tunneling, the transport mechanism is mainly governed by TE, meaning that after a certain height is reached, the device will turn off. For the n-mode, an identical behavior is present. For V_{PG} and $V_{CG} \gg 0\,V$ the device is in the on state allowing electron conduction (Figure 2.11(f)) similar to an SBFET (Fgiure 2.10(e)). By lowering V_{CG}, the current can again be reduced by the barrier-blocking electrons. PG acts as a filter setting if electron or hole conduction should take place, and CG is used for current modulation [122].

These characteristics are widely utilized in today's complementary metal oxide semiconductor (CMOS) circuits. n-FETs are used for pulling down the network to GND and p-FETs are used for pulling up the network to V_{dd}. In CMOS, it is important that the complementary transistors have a similar driving strength and performance in order to ensure energy-efficient switching, requiring a channel width increase for p-FETs due to their reduced mobility for holes in silicon. In RFETs, the dimensions for the n- and p-program are identical, and only the PG has to be set accordingly. Furthermore, entire logic gates, such as inverters, can be constructed with RFETs that can switch their behavior while in operation. Simulations show that by utilizing RFETs, improvements in path delay, power consumption, and chip area can be made [124–126]. Exclusive-OR and Majority gates are especially difficult to implement in CMOS, requiring a significant number of transistors and interconnects. Using RFETs, these gates can be realized with half of the devices [31, 127].

2.3.4 Ferroelectric SBFET and RFET

When combining the concepts discussed for SBFETs and RFETs with the ferroelectric properties of ferroelectric HZO (Section 2.2) non-volatile variations can be realized. In Figure 2.12(a), the gate stack of a FeFET is used with the Schottky barriers of an SBFET, forming a FeSBFET. In its unpolarized neutral state, indicated by the dashed gray line in Figure 2.12(b), the device would exhibit an ambipolar transfer characteristic assuming identical SB height for holes and electrons. When pulsing this device with sufficiently high voltages, the ferroelectric is polarized, which influences the transfer characteristic in two superimposed ways. Firstly, the threshold voltage V_{th} is altered due to remanent polarization over the transistor channel, shifting the entire characteristic away from the origin position. Secondly, the injection of holes and electrons is altered at the SBs because the ferroelectric increases or decreases the achievable band bending shown in Figure 2.12(c,d). For negative V_G, the up polarization increases the injection, and the down polarization decreases the injection. For positive V_G, this behavior is inverted. Such devices have been fabricated and analyzed in [33]. Here, $NiSi_2$ SB contacts are used with an HZO gate dielectric and TiN TG. Transfer characteristics reaching beyond the E_C of the HZO show a counterclockwise hysteresis. Additionally, a set of measurements is performed, including the mentioned LTP/LTD in Subsection 2.2.4, showcasing the potential of the device for neuromorphic applications.

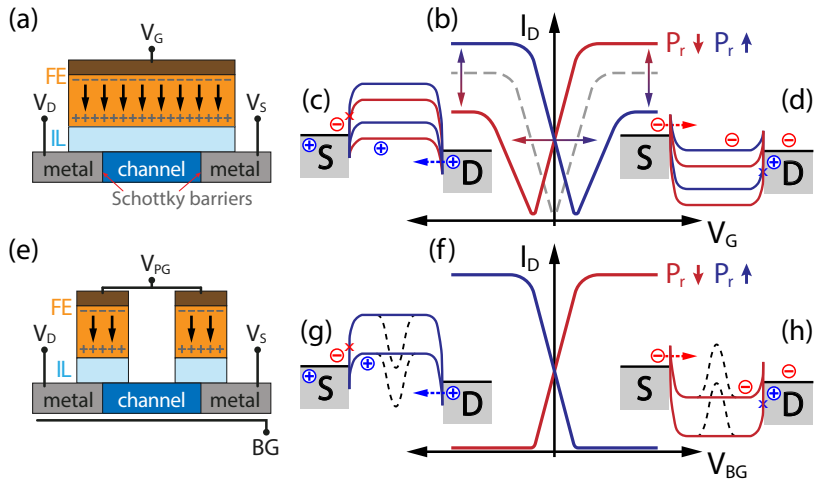

Figure 2.12: (a) Ferroelectric integrated into SBFET forming FeSBFET. Polarization charges and direction are indicated by arrows. (b) Influence of polarization on FeSBFET transfer characteristic, with band diagram for (c) hole and (d) electron conduction. (e) Ferroelectric integrated into DTG RFET forming FeRFET. Ferroelectric barrier modulation over PG and current modulation over BG. (f) FeRFET transfer characteristic for both polarization states. Band diagram for (c) hole and (d) electron conduction with the influence of BG indicated. Image adapted from [67]. Image adapted from [67].

Furthermore, RFETs can be equipped with ferroelectric HZO, allowing similar mechanisms, as shown in Figure 2.12(e). Here, a DTG device is shown, where the SB is modulated with the MFIS polarity gate stack, and the channel is controlled by a BG contact

creating a FeRFET. Identical to standard RFETs, the transfer characteristic of such a device is shown in Figure 2.12(f). FeRFETs present a similar behavior, in which the PG determines if the electron or hole injection is enabled and the BG gate modulates the resulting current, but significantly, the HZO enables the state of the SBs to be programmed non-volatile. Figure 2.12(g,h) depicts the resulting band diagrams. The programmed state by the polarization and the rise of the BG-controlled barrier is indicated. In [34], such a device was demonstrated using Al SB contacts and an MFIS HZO gate stack with TiN TGs. A clear dependency of the injection on the polarization and high retention times for intermediate states was shown.

Both these concepts outline the potential of ferroelectric implementation in SB-based devices. On the one hand, they can be integrated into RFET circuits where the PG contact is used for switching between behaviors. Here, HZO could be used to program the behavior non-volatile, making the PG line unnecessary after the program operation, which allows it to be relocated to other gates using a multiplexer. This reduces the needed integrated circuit (IC) infrastructure for implementing such circuits on chip because only one switching unit and the multiplexer are required. Additionally, applications in neuromorphic computing are possible, in which FeSBFETs and FeRFETs feature some key advantages. They are very reproducible in an industrial setting, outperforming current transistors in device-to-device variability, and feature low leakage currents, possibly further improving the energy efficiency of the proposed SNN.

Chapter 3

Experimental Techniques

This chapter outlines the fabrication of the examined device. As described in the motivation, the main objective of this project is to add a TiN metallic interlayer, also known as a floating gate (FG), into an already established gate stack used in [34]. The physical analysis of the FG implementation is done in the results and discussion chapter 4. The tests needed to allow this adaptation of the process are examined there. This chapter will focus solely on the final process flow employed for the fabrication of the analyzed devices. Used equipment and detailed settings are summarized.

Additionally, the second part of this chapter will discuss the electrical measurements performed on the devices. Used semiconductor analyzers and probing stations, as well as in-depth pulsing schemes and analysis techniques, are mentioned.

Finally, project partners Dr. Lilian Vogl and Dr. Peter Schweizer at the Max Planck Institute for Sustainable Materials in Düsseldorf have performed transmission electron microscopy (TEM), which is used for physical analysis of the fabricated devices.

3.1 Device Fabrication

Figure 3.1 shows the fabrication overview for an STG FG FeSBFET design. The following section will go into detail about each step explaining the used equipment and processes.

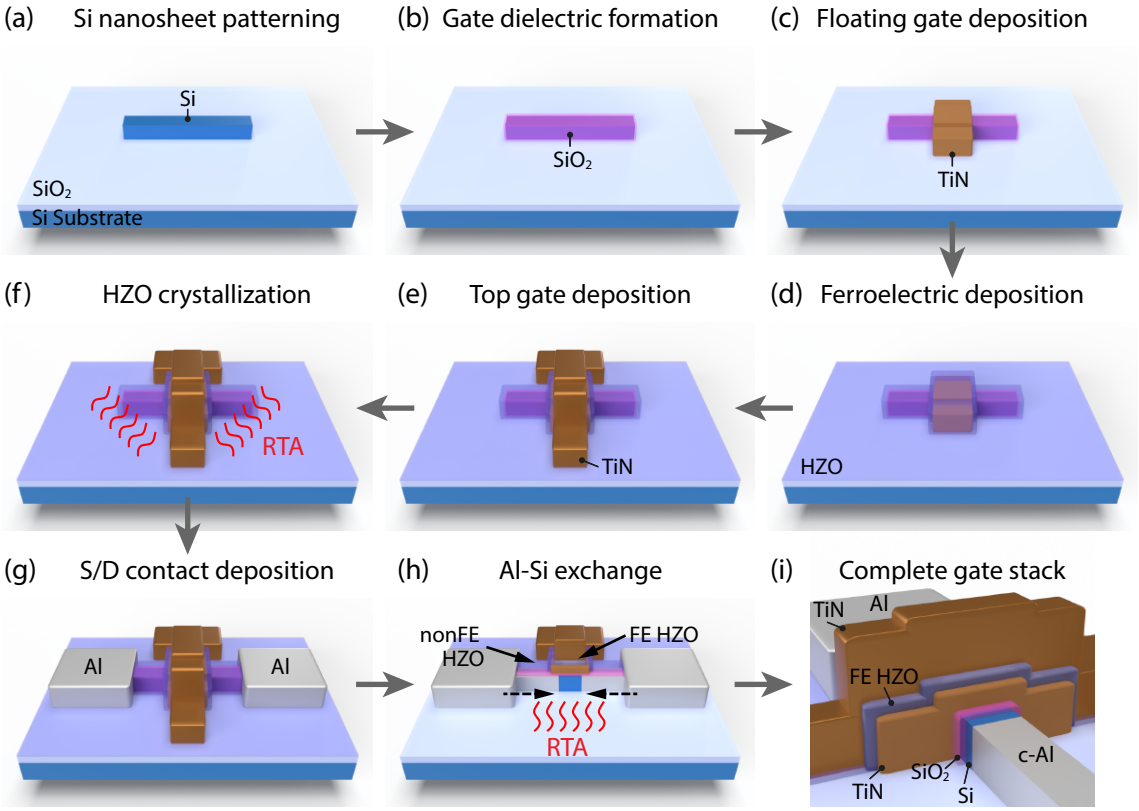

Figure 3.1: Overview of process flow. (a) Si nanosheet patterning using RIE. (b) SiO_2 Gate dielectric formation using oxidation oven. (c) TiN floating gate deposition using reactive sputter and RCA1 etch. (d) Ferroelectric HZO formation using ALD. (e) TiN TG formation. (f) HZO crystallization using RTA oven at 823 K. (g) Al source/drain contact deposition using sputter. (h) Cut through nanosheet showing Al-Si exchange using RTA oven at 773 K and indicating ferroelectric as well as non-ferroelectric HZO. (i) Complete gate stack. Image based on [67].

3.1.1 Silicon Nanosheet Structuring

A silicon-on-insulator (SOI) wafer, featuring 100 nm SiO_2 barrier oxide (BOX) and 20 nm lightly p-doped (001) Si device layer, is utilized as a substrate, which is first cleaved in a 10 mm by 10 mm piece by hand. Sonication and acetone/isopropanol are used to clean the surface of the sample thoroughly. For patterning of the silicon nanosheets, AZ5214 image-reversal photoresist is applied through a spin coating, and a pre-exposure bake is performed at 373 K for 60 s on a hot plate, to initially harden it.

For lithography, a *MLA 150* laser lithography system from *Heidelberg Instruments* is used [128]. In contrast to standard lithography utilizing masks, such systems implement

a laser UV-light source to directly expose the light-sensitive photoresist, employing a digital micromirror arrangement to project a pattern given by a computer-aided design (CAD) file. On a stage, the sample moves continuously from left to right on one axis, and after each swing, the normal axis increases in accordance with the write field width [129]. Because of this movement, the CAD design layer of the nanosheets is rotated so that the swing is in line with their orientation, allowing the *MLA 150* system to pattern evenly with the minimal feature size. Nanosheets with widths of 400 nm are exposed at a dose of 140 J/cm^2 and developed with AZ726MIF for 14 s.

SF_6/O_2-based reactive ion etching (RIE) is used to remove the not-needed Si of the device layer, leaving behind the nanosheets, which are below the photoresist. In such a system, SF_6 and O_2 are ionized in a plasma, which generates fluorine and oxygen radicals. The fluorine radicals are highly reactive with silicon, forming SiF_4, which is volatile and can be evacuated from the chamber. The presence of oxygen facilitates the formation of SiO_xF_y passivation layers, which are anisotropically removed by the physical etching of ions in the plasma, leaving it behind on the side walls generated by the etching process. This combination of chemical and physical processes enables high anisotropy and the creation of features with steep etching profiles [130]. For this sample an *Oxford Instruments PlasmaPro 100 Cobra ICP* system [131] with 50 sccm/4 sccm of SF_6/O_2 gas flow, 15 W of RF power and 308 K of process temperature is used.

Additionally, plasma ashing is employed to strip the damaged resist and passivation layers, implementing a *Pink V10-G* with O_2 plasma at 300 W for 5 min [132]. This process is followed by rinsing in acetone and isopropanol.

3.1.2 Gate Dielectric Formation

As gate dielectric SiO_2 is formed using thermal dry oxidation in a *ATV PEO-601* furnace [133]. The oxide is thereby grown at 1173 K with 50 sccm of O_2 flow. According to the Deal-Grove model, the oxide thickness t_{ox} is controlled by the oxidation time. In such a system, the Si needed for SiO_2 formation is supplied by the substrate, in this case, the nanosheets [134]. This results in a reduction of the sheet height, which can be estimated with $t_{Si} \simeq 0.46 \cdot t_{ox}$, in which t_{Si} is the height of the Si converted to SiO_2 [135]. In order to reduce interfacial defect density, the oxidation is followed by an in-situ post-oxidation anneal in N_2 for the same time duration as the oxidation [136]. For this sample 4 min of O_2 and 4 min of N_2 are performed, which should result in roughly 6 nm of SiO_2 and reduce the silicon nanosheet layer height from 20 nm to 17 nm.

3.1.3 Floating Gate Deposition

A TiN FG layer is deposited using reactive magnetron sputtering, which is categorized as a physical vapor deposition (PVD) process. An Ar plasma is generated by a negative voltage in the vacuum chamber of the system. Positively charged Ar^+ ions are accelerated towards a Ti target because of the applied electric field. When they hit the target, mo-

mentum is transferred, causing a collision cascade that ejects titanium atoms, which then travel through the chamber and condense onto the substrate positioned opposite to the target, forming a layer. Compared to other PVD processes like evaporation, sputtering is understood to have more edge coverage and, due to the higher impact energies, better adhesion, which can, on the other hand, cause damage to the substrate [137]. In magnetron sputtering, a magnetic field is placed behind the target, which brings the Ar^+ on a spiral path, that tilts the impact direction away from solely vertical, increasing yield, because the collision cascade, which is else mainly confined inside the Ti, can more likely reach the surface and eject a Ti atom [138]. To deposit TiN, reactive sputtering is implemented, where N_2 is introduced into the chamber in addition to Ar. Sputtered Ti atoms can react with N^+ ions either in the chamber on their way from the target to the substrate or after already being deposited on the substrate in accordance with

$$2\mathrm{Ti} + \mathrm{N_2} \rightarrow 2\mathrm{TiN} \tag{3.1}$$

In this regard, the N_2/Ar ratio is crucial for achieving quality layers. Too much N_2 can lead to over-saturation and the formation of brittle films, and too little N_2 can result in unwanted stoichiometries of TiN that have a reduced hardness or conductivity [139]. A *Creavac Creamet 750 S10* system is used [140]. First, the Ti target is cleaned, and then sputtering of the sample is performed with for 900 s at 100 W of power, 6 sccm of N_2 flow and $6 \cdot 10^{-3}$ mbar of chamber pressure, resulting in an even 40 nm thick layer over the entire sample.

Subsequently, etching of the TiN layer has to be done to pattern the required FG structure, for which an identical lithography process is used as in Subsection 3.1.1 for the silicon nanosheets, except for a longer developing time of 17 s, which is the value recommended by the manufacturer, and logical a different CAD design. Because the wet etching can lift the photoresist, good adhesion to the layer below is required. Developing this part of the process is one of the key steps introduced in this thesis and will be discussed in detail in Subsection 4.1.2 of the results and discussion chapter. A 30 min hard bake [141] of the resist at 393 K is performed on a hot plate followed by a 2 h rest period.

For etching, a standard clean 1 (SC1), also known as RCA1, which is a reference to the Radio Corporation of America, the company that standardized the semiconductor cleaning process in the 1960s [142], is used. It consists of ammonium hydroxide (NH_4OH), hydrogen peroxide (H_2O_2) and deionized water (H_2O) giving it also the name APM. Normal ratios are 1:1:5, and the solution is heated to 343-353 K. It is employed to remove organic contaminants and to etch some metals like TiN. The chemical process can be split into two partial reactions: oxidation and dissolution. Hydrogen peroxide acts as a strong oxidizer, which converts the TiN into titanium dioxide (TiO_2) and nitrogen-based byproducts

$$TiN + H_2O_2 \rightarrow TiO_2 + N_2 + 2H_2O \tag{3.2}$$

Ammonium hydroxide dissociates partially in water as a weak base or alkaline

$$NH_4OH \leftrightarrow NH_4^+ + OH^- \tag{3.3}$$

Because of the elevated pH created by the NH_4OH, the concentration of hydroxide ions (OH^-) in the solution is increased. Dissolution of the TiO_2 if performed by the OH^-

$$TiO_2 + 2OH^- + H_2O \rightarrow [Ti(OH)_4] \tag{3.4}$$

creating a soluble complex that dissolves in the liquid phase. Elevated temperatures are needed to increase the rate of oxidation and dissolution [135, 143–145]. The implemented process and procedure is based on [146] and has been slightly adapted, which will be discussed in subsectio 4.1.2 in detail. An APM ratio of 1:2:10 is used. First the 10 parts of H_2O and 1 part of NH_4OH are heated to 348 K for 15 min on a hot plate. After the elapsed time, the solution is removed from the hotplate, 1 part of H_2O_2 is added, and stirred lightly for 2 min. The sample is put into the solution for 3 min 45 s and stirred in order to remove the 40 nm of TiN without the build-up of residues. After the completed etching time, the sample is immediately rinsed under a deionized water flow from the tap to remove as efficiently as possible any left behind parts of the solution still on the sample and prevent further etching. This process is followed by an additional rinsing in acetone and isopropanol to remove the remaining resist.

3.1.4 Ferroelectric Formation

The ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ (HZO) is formed using atomic layer deposition (ALD). In this process, the sample is exposed to alternating precursors and surface reactions, that are self-limiting, allowing only one atomic layer of a given speice to be added to the layer stack and enabling atomic scale control over the thickness. In the case of HZO tetra(ethylmethylamido)hafnium (TEMAHf), tetra(ethylmethylamido)zirconium (TEMAZr), and H_2O are used as precursors. The first step consists of a hafnium precursor pulse in which TEMAHf reacts with surface hydroxyl groups (OH) to form the first monolayer

$$Hf(NMe_2)_4 + 2OH \rightarrow HfO_2(NMe_2)_2 + 2H(NMe_2) \tag{3.5}$$

followed by a purge step that removes excess precursor by keeping a continuous flow of

N_2, to remove the reaction byproducts and precursor leftovers. H_2O is then injected to add the OH groups

$$HfO_2(NMe_2)_2 + 2H_2O \rightarrow HfO_2(OH)_2 + 2H(NMe_2) + H_2 \tag{3.6}$$

After a second purge step, the cycle could repeat, which would be the case for pure HfO_2 formation. For HZO, the next cycle is for Zr, which is identical except that the TEMAHf precursor is replaced with the TEMAZr precursor, resulting in very similar stoichiometric equations

$$Zr(NMe_2)_4 + 2OH \rightarrow ZrO_2(NMe_2)_2 + 2H(NMe_2) \tag{3.7}$$

$$ZrO_2(NMe_2)_2 + 2H_2O \rightarrow ZrO_2(OH)_2 + 2H(NMe_2) + H_2 \tag{3.8}$$

The Hf and Zr cycles are alternated to form $Hf_{0.5}Zr_{0.5}O_2$ [135, 147, 148]. Before HZO is grown, the sample is reverse sputtered with a power of 50 W for 240 s using the *Creavac* system to remove oxides that have formed on the FG, further details are outlined in Subsection 4.1.3. Around 9.5 nm are grown in a *Beneq TFS 200* [149] at 523 K and starting off with two cycles ZrO_2 as a seed layer, which promotes lateral growth, reduces island formation and tensile stress [150, 151]. This is then followed by 58 double cycles of HfO_2 and ZrO_2.

3.1.5 Top Gate Deposition and HZO Crystallization

Immediately after the ALD process is finished, the sample is placed into the *Creavac* sputter vacuum chamber, and the TG TiN layer is deposited. This is done to minimize contamination and is enabled by the RCA1 chemical wet etch. Identical procedures to the FG layer are performed for the TiN sputter, the laser lithograph with hard bake, and the RCA1 etch, creating an MFMIS structure in which the ferroelectric HZO is placed between two TiN layers.

In this state, the sample is annealed using a *Unitemp UTP-1100* rapid thermal annealing (RTA) system [152] at 823 K for 120 s, which is necessary to bring the HZO in a ferroelectric phase, as explained in Subsection 2.2.1. The process is performed in a N_2 atmosphere in order to prevent oxidation of metals like the TiN TG electrode. In general, annealing in an atmosphere containing H_2 is preferred, which is not done in samples containing HZO because the resulting H-incorporation degrades significantly the ferroelectric properties [65]. Additionally, it is expected that the TiN FG and TG layers are getting lightly oxidized by the presence of oxygen in the regions in contact with the HZO.

3.1.6 Source/Drain Contact Formation and Al-Si Exchange

Finally, the source/drain contacts are added by sputter deposition. For this process, lift-off is used. The entire device is coated with photoresist, and the area above the two ends of the nanosheets is removed. First, the now crystalline HZO, which exhibits a non-ferroelectric phase since it was not in between the TiN metals during RTA, is reverse sputter etched with 100 W of power for 695 s in the *Creavac* system, which should also slightly etch the SiO_2 below, ensuring full removal of the HZO. Secondly, buffered hydrofluoric acid (BHF, 7:1) is used for 15 s to remove the remaining SiO_2. This chemical etch is self-limiting and should not attack the Si of the nanosheet. Next, the sample is placed back into the sputter vacuum chamber to reduce oxidation of the now-unprotected Si. Finally, Al is deposited by magnetron sputter, which works identically to reactive magnetron sputter but without the N_2 atmosphere. Initially, the sample is again reverse etched at a power of 50 W for 120 s to remove any unwanted SiO_2 which might have formed during handling. Then the Al is deposited with 60 W and 60 s in six cycles. To perform the lift-off, the sample is first submerged into a 333 K acetone bath to break up the photoresist. Sonication is used to remove any remaining metals, which is followed by an additional rinsing in acetone and isopropanol.

Figure 3.1(h) shows a cut along the nanosheet during diffusion of the Al, which happens during RTA at 773 K. This process is performed in a series of 60 s anneals until the desired diffusion length is reached. Due to the semi-transparent characteristic of the TiN FG and TG, this process can be precisely controlled. Al replaces the Si in the nanosheet completely, because their diffusion coefficients are vastly different. Si diffuses very quickly in Al, and on the other hand, Al diffuses very slowly in Si [153, 154]. This causes the gradual propagation of the Al front in the nanosheet. When the front absorbs a Si atom, it gets immediately redistributed in the entire Al pad because of the height diffusion constant. Since the concentration of Si in the pad is very low, the Al can still be considered pure in the nanosheet. The interface is very abrupt from one atomic plane to the next, and the diffused Al becomes crystalline, creating a high-quality metal-semiconductor interface [155–158].

3.1.7 Sample Overview

Figure 3.1 shows the fabricated sample, and different important features are outlined. In every corner, alignment markers are placed, which are necessary for working with multiple layers. Additionally, an autofocus area is added in the center of the sample, which helps the *MLA 150* laser lithography system to focus more easily and can prevent collisions of the write head with the sample. It consists of the metal layers, which have been stacked on top of each other. In order to troubleshoot the fabrication process, reference pads for each layer are added. Firstly, the Al and TiN TG are on top of all other layers, of which the TiN is colored according to the HZO, SiO_2, and BOX below, due to its semi-transparency. Secondly, the TiN floating gate layer is below the HZO but above the SiO_2. Thirdly, the Si of the device layer is below both oxides. On-device CV pads are fabricated to validate and characterize the HZO. They consist of a TiN top electrode on the TG layer and a TiN

bottom electrode on the floating gate layer, which is accessed during the Al source/drain deposition. Because of the high hardness of TiN, damage from the reverse sputter or HF dip should be minimal. Four rows of STG SBFETs and DTG RFETs, as well as two rows of TTG RFETs, are fabricated. There are two columns of STG and DTG devices in which the FG layer is accessible similarly to the CV pad bottom electrode. For the TTG RFET, multiple variations of the control gate layer positioning are implemented. In general, some devices and CV pads are still missing TGs in order to enable tests with a different metal.

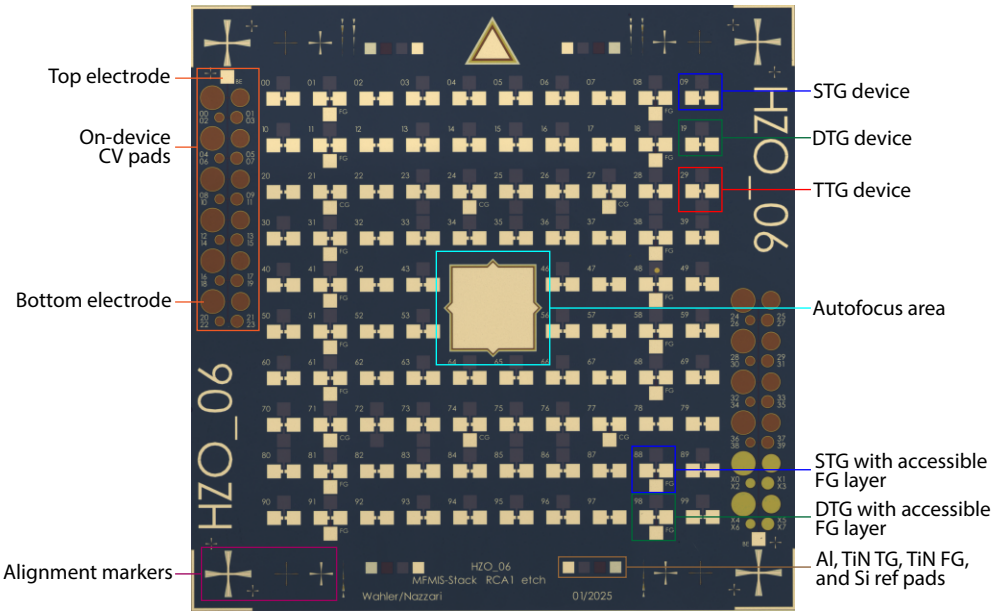

Figure 3.2: Overview image of designed sample outlining important features for fabrication and subsequent electrical device analysis.

3.2 Electrical Characterization Techniques

In order to electrically analyze the fabricated device, three setups are used, which allow different kinds of measurements. The used equipment and performed measurement techniques are explained in the following subsections.

3.2.1 C/V Measurement

As depicted in Figure 3.2, the produced sample is equipped with multiple metal-ferroelectric-metal (MFM) capacitors, which allow, on one hand, to characterize the ferroelectric material and, on the other, to validate that the deposited layer is indeed ferroelectric. Measurements are performed using a *Keithley 4200-SCS* semiconductor analyzer, which is equipped with a CVU (Capacitance Voltage Unit), in a *Cascade microtech summit 11000 AP* probing station. Figure 3.3(a) shows the applied voltage to the MFM capacitor during capacitance-voltage (C/V) measurements. A DC voltage sweep is combined with a small AC signal with a frequency between 1 kHz and 10 MHz as well as a magnitude in the mV range. By integrating the current flow over time, the electric charge Q, bound on the plates of the capacitor, can be measured, and the capacitance C can be calculated with

$$C = \frac{\Delta Q}{\Delta V} = \frac{\varepsilon_0 \varepsilon_r A}{t} \tag{3.9}$$

in which V is the applied voltage, ε_0 is the vacuum permittivity, ε_r is the relative permittivity, A the area of the capacitor, and t the dielectric thickness. In standard non-ferroelectric materials, these values are constant, resulting in a constant C when neglecting charge trapping, interface states, fringing electric fields, stress effects, and electrode polarization [159]. In a ferroelectric capacitor, ε_r depends on the polarization P of the material

$$\varepsilon_r = \frac{\varepsilon_0 E + P}{\varepsilon_0 E} = 1 + \frac{1}{\varepsilon_0} \cdot \frac{P}{E} \tag{3.10}$$

resulting in the C/V measurement shown in Figure 3.3(b). The nonlinearity of the polarization switching causes a counterclockwise butterfly curve with two peaks at the point at which the domains flip [160, 161].

3.2.2 Transfer Characteristics

For transfer characteristics, the *Keysight B1500A* semiconductor analyzer connected to a *Lake Shore PS-100* cryo-probe station is used. It features four SMUs (Source Measure Units), which are connected in the case of an SBFET to all four terminals of the device, including the BG. The absolute value of the drain current $|I_D|$ is recorded for a range of different gate voltages. The top gate voltage V_{TG} is symmetrically varied around 0 V.

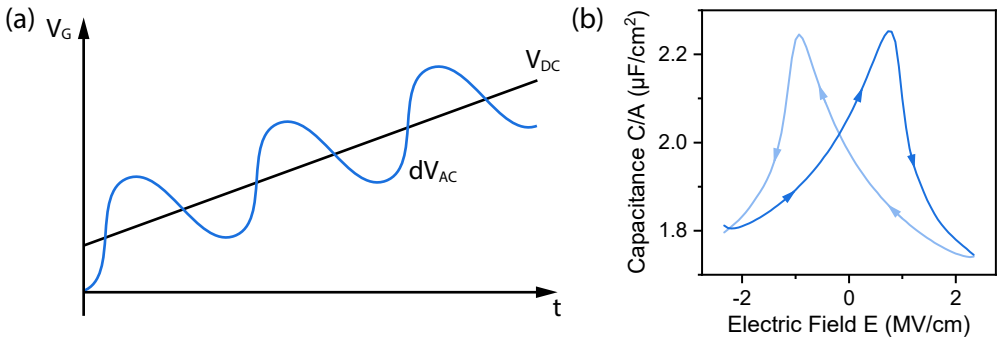

Figure 3.3: (a) Applied voltage on capacitor during C/V measurement. (b) Exemplary data for ferroelectric material featuring a counterclockwise butterfly curve because of nonlinear domain switching. Image adapted form [67, 160, 161].

Additionally, a source voltage V_S and a drain voltage V_D are symmetrically applied to the respective terminals, meaning $V_D = -V_S$. The back gate voltage V_{BG} is set to 0 V. This measurement provides information on the controllability of the device under test. It can be utilized to compare different devices with each other.

3.2.3 Pulse Measurements

The first set of pulse measurements is performed using as well the *Keysight B1500A* semiconductor analyzer in the *Lake Shore PS-100* cryo-probe station. As explained in the HZO theory section 2.2 pulsing schemes can be used to partially flip the domains of the ferroelectric and allow gradual control over the induced ΔV_{th} shift of the device. In this setup, only SMU pulsing is possible, the pulses are generated with the SMU units of the analyzer, which are limited to a minimum pulse width t_{pulse} of 500 µs. The pulse heights for reset V_{reset} and set V_{set} are chosen to induce a field beyond the coercive field E_C of the ferroelectric. First, an exemplary reset pulsing scheme is performed, as depicted in Figure 3.4(a), to bring the device into a well-defined state. During pulsing, the device source, drain, and back gate contacts are connected to ground, as shown in Figure 3.4(b), to ensure similar electric field situations in the oxides. This is followed by a double transfer characteristic, which sweeps in a very limited regime to not disturb the polarization of the HZO. A source-drain bias voltage V_{SD} has to be applied. In general, the transfer after the reset is optional and may not be executed. It gives additional information about the effectiveness of the reset pulsing scheme. Next, a single set pulse is executed to bring the device into the desired state. Different approaches can be taken, ranging from pulse width to pulse height modulation, as discussed in Subsection 2.2.3. The final transfer holds the key threshold voltage shift ΔV_{th} information of this measurement.

The aforementioned SMU pulsing technique is limited in pulse width. To implement shorter pulses, the *Keysight* analyzer is replaced with a *Keithley 4200A-SCS* using the same *Lake Shore* prober. This analyzer is equipped with a two-channel 4225-PMU (Pulse Measure Unit), which is able to generate pulses with widths as low as 10 ns. Additionally, two 4225-RPMs (Remote Preamplifier/Switch Modules) are used. Figure 3.5 shows a block

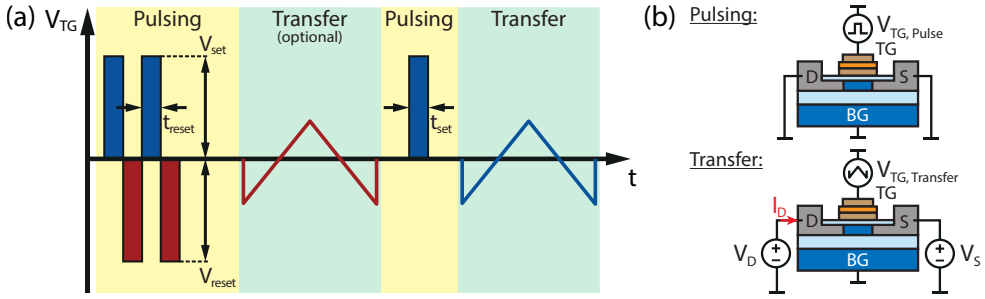

Figure 3.4: (a) Implemented pulsing technique performing reset pulse operation, optional first transfer, set pulse, and final transfer. (b) Device under test connections during different phases. Source and drain are grounded during pulsing, and a bias voltage V_{SD} is applied during transfer. Image adapted from [67].

diagram of the implemented setup. The sense ports of SMU1 and SMU2 are connected to RPM1 and RPM2. Also, channels 1 and 2 of the PMU are connected to RPM1 and PRM2. The RPM1 sense port is connected to the gate contact, and the RMP2 sense port is connected to the drain contact of the device under test. The source and back gate are connected to the GNDU (Ground Unit). This setup allows one to switch between SMU and PMU operations using the RPMs. During pulsing, the PMU is used, enabling shorter pulse widths, and during the transfer measurements, the SMU is used, allowing precise current measurements.

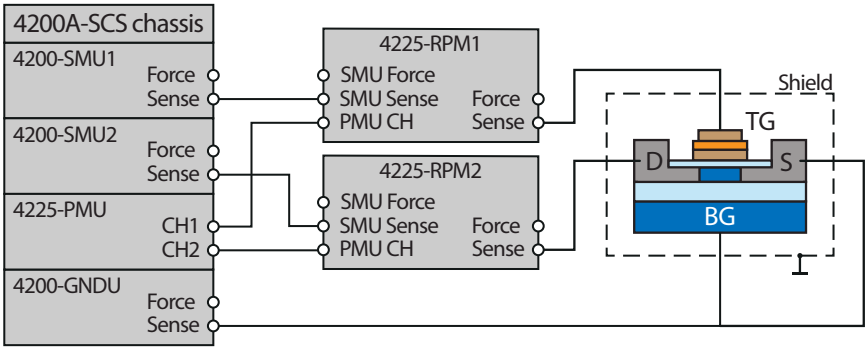

Figure 3.5: High speed pulsing setup using *Keithley 4200A-SCS* semiconductor analyzer. RPMs allow switching between SMU and PMU operations. Image adapted from [162]

Chapter 4

Results and Discussion

In this chapter, the fabricated samples are investigated physically and the obtained electrical measurement results are evaluated.

4.1 Device Physical and Process Analysis

The previously presented fabrication process is discussed in this subsection. Reasons for switching from a lift-off process originally intended for the floating gate implementation to wet chemical etching are given. Atomic force microscopy (AFM), transmission electron microscopy (TEM), and optical microscopy are employed to analyze the produced samples. Aspects ranging from stack composition, layer properties, and fabrication accuracy are investigated.

4.1.1 Floating Gate Lift-Off

The first adaptation of the previously [34] used process to obtain a functional FG is shown in Figure 4.1. The FGs are placed beneath the PG and over the SBs to improve the electric field uniformity around them (Figure 4.1(c)). The employed process is similar to the one outlined in Chapter 3. Instead of using an RCA1 etch for the FG and TG, a TiN lift-off process is used, identical to the one implemented for the Al sour/drain pads (Subsection 3.1.6). In which TiN is deposited not only on the desired area but also on the resist, which is then stripped together with the not-needed material during sonication in heated acetone. Figure 4.1(b) shows an optical image of the fabricated device. Left and right are the Al source/drain pads. In the center, the Si nanosheet is placed. It features additional areas at both ends to improve diffusion that are covered by Al, which are lightly visible in the source/drain regions. The FG and TG form a TTG structure. The PGs are on top of all oxide layers, giving them a lighter color, clearly visible in the center, top, and bottom parts of the image. In contrast, the CG is between the SiO_2 and

the HZO, reducing its intensity and creating a darker color. When the PG covers the FG, the resulting overlap becomes brown, which can be seen in some areas. It is clearly visible that these areas deviate from the intended appearance, which should be a single strip of TiN-HZO-TiN as depicted in Figure 4.1(c). Instead, they form groups of irregular shapes. Furthermore, the edges of the TiN seem to have residual structures.

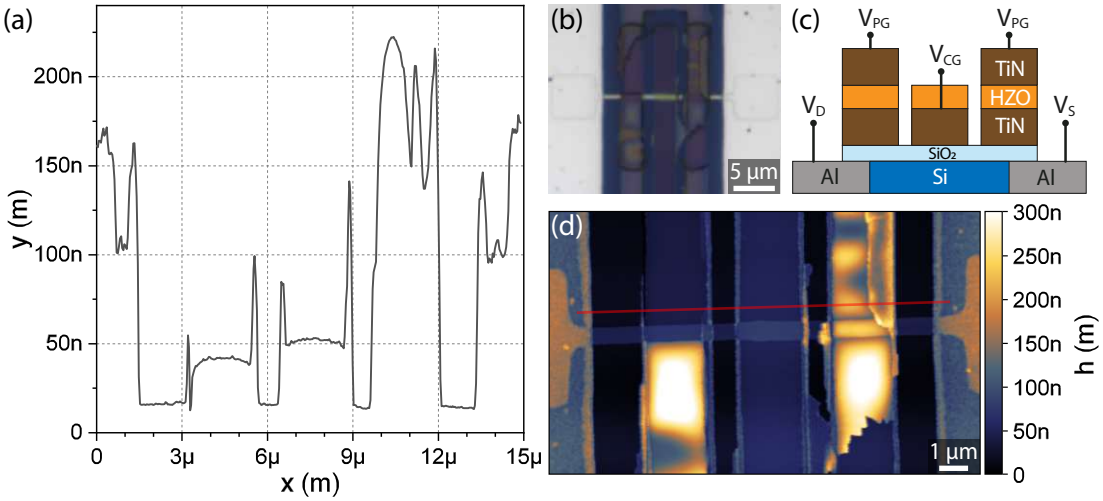

Figure 4.1: FG sample fabricated using a lift-off process. (a) Layer height along the indicated cross-section of the sample. (b) Optical image of the investigated sample. (c) Shemitic illustration of the sample, indicating different layers. (d) AFM image of nanosheet region with marked cross-section.

Figure 4.1(d) illustrates a recorded AFM image of the same device shown in Figure 4.1(b). The areas of interest have different heights and a gradient. The red-marked cross-section is plotted in Figure 4.1(a). Multiple things become directly clear. Firstly, since lift-off is used in combination with sputter, it can not be prevented that TiN is lightly deposited on the edges of the resist. During stripping and sonication, this very thin layer is broken, and a spike is left on the edge of the layer, clearly visible for the left and center gates. This is something inherent to lift-off, as it is also present for the Al source/drain pads and even exaggerated due to the greater layer height. Secondly, the gate layer heights do not match the intended design in Figure 4.1(c). The center gate height corresponds to the intended TiN and HZO layers. In contrast, the right gate is higher than what is possible with the TiN-HZO-TiN stack, and the left gate seems to only have the TiN layer without HZO or TG TiN. When investigating the left gate further in the AFM image in Figure 4.1(d), it is visible that parts of the TG are still remaining on the other side of the nanosheet, and the lower region of the right gate spots a zigzag edge pointing toward a breaking of the layer. When brought together with the potential deposition of the photoresist, it can be hypothesized that because TiN is compared to Al, a relatively hard material, the connection to the resist is so strong that sonication is not capable of breaking it, resulting in the delamination of the layer as it gets removed together with the resist during stripping. Regions like the right gate form, where the TiN is bent upwards or the TG is removed completely.

To counteract this, numerous directions can be taken. For one, the TG can be slimmed down to prevent overlap of the mentioned spikes of the FG layer with the new spikes of the deposited TiN. Additionally, reactive magnetron sputter could be replaced with reactive evaporation [163], which is a strictly directional PVD process, preventing coverage of the photoresist. This process is very uncommon for TiN because it gives less control over film composition, stress, and density. Finally, an etching process can be used, in which the entire sample is first covered with TiN, and then the unwanted material is etched away.

4.1.2 RCA1 Wet Chemical Etch

Since such a process was not established at the institute, it had to be developed as part of this thesis. RCA1 is chosen as a wet chemical etching solution since it is capable of etching TiN, as discussed in Subsection 3.1.3. All subsequent tests were performed using a Si wafer with 100 nm oxide on top to mimic the gate stack present on the devices, where TiN is deposited on the BOX of the SOI wafer. Either a 20 nm or 40 nm thick layer of TiN is sputtered on the sample, using the settings in Subsection 3.1.3 accordingly. The intended sample structure is shown in Figure 4.2(c), mimicking the TG of a TTG device.

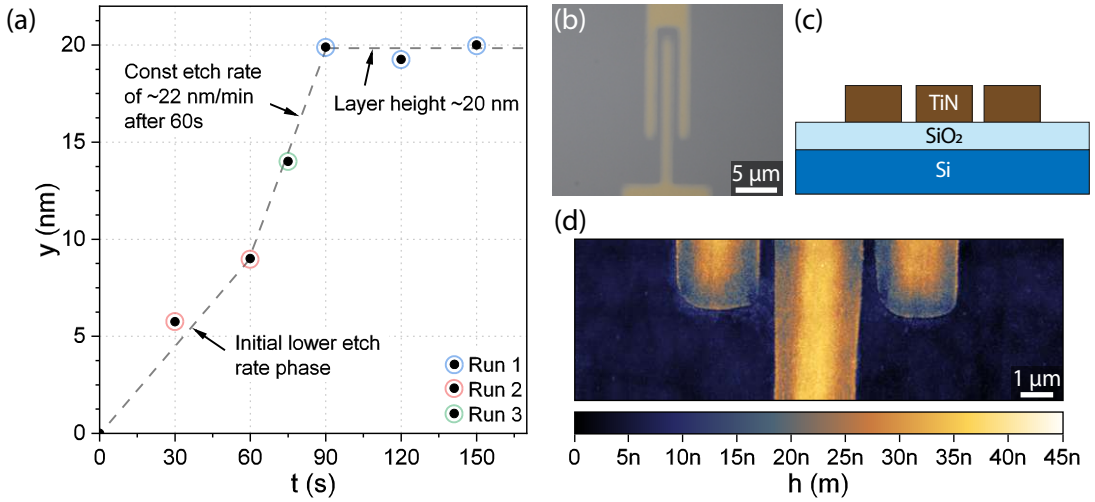

Figure 4.2: RCA1 etching test. (a) Etching with a (1:2:10) ratio during three different runs, recording edge height and etch rate. (b) Optical image of the investigated under-etched sample. (c) Shemitic illustration of the sample, indicating different layers. (d) AFM image of an under-etched sample.

Initially, a standard (1:1:5) APM ratio was used with a resist structured by laser lithography (Subsection 3.1.1) as an etch mask. Because of the weak adhesion of the resist to the TiN the mask was lifted, and the entire sample was etched, which prompted the introduction of a hard bake procedure, in which the sample is placed after development on a hot plate at 393 K for 30 min. This increases the adhesion and resilience of the layer to the etching solution.

RCA1 is normally used to remove biological contaminants from the substrate. The used

AZ5214 image-reversal photoresist consists, among others, of organic components, which can be attacked by ammonia in the solution and result in under-etching of the mask. Figure 4.2(b) shows an optical image of a sample that suffered from this. The edge of the TiN structure has a different, lighter color compared to the center of the stripe. AFM imaging reveals that the layer height is reduced at the edges of the TiN structure. To investigate the influence of the ammonia content on the solution, samples have been etched with (1:1:5), (1:2:10), and (1:4:20) ratios for 2 min. (1:2:10) showed the best compromise between resistance against under-etching and etch rate.

The etchability of the resist additionally highlights the need for good control or knowledge of the etch rate to prevent subsequent under-etching after reaching the end of the layer. Figure 4.2(a) shows the height of a structure similar to Figure 4.2(c) for different etch times during three different runs. A (1:2:10) ratio is used, and all samples were sputtered in the same process. Three etch runs were performed in which the samples were placed in 15 s intervals into the solution and removed according to the stated time. Multiple runs are performed instead of doing a single one to ensure that the quality of the solution is identical for all samples, because RCA1 potency degrades starting with the time of mixture. In total, six samples were investigated. The first run shows that the layer height is 20 nm because all samples featured this height. The second and third runs are below. Initially, the etch rate seems to be lower, which can be explained by some residues of the resist or, in general, contaminants being removed. Then it reaches a constant rate of roughly 22 nm/min.

By considering the discussed aspects of RCA1 etching, the procedure stated in Subsection 3.1.3 is developed. This allows patterning of the FG and TG structure with no spikes on the etches or delamination of the layer, as has been seen in Subsection 4.1.1 for the lift-off process.

4.1.3 Floating Gate Reverse Sputter

Firstly, to investigate the ferroelectric properties of the RCA1 etch, CV pads have been fabricated and electrically investigated. A Si wafer is covered completely by TiN sputtering, HZO is deposited with ALD, crystallization is performed at 823 K with RTA, and RCA1 wet chemical etch or lift-off is used to pattern TiN sputtered top pads, forming an MFM device. It is assumed that the capacitance of the resulting structure is governed by the size of the top pad, neglecting fringing electric fields. The measurement is performed using the Si wafer as a back gate. Both samples showed counterclockwise butterfly CV curves similar to the one in Figure 3.3(b), which is a clear indication of ferroelectric behavior due to the nonlinearity of the polarization switching, causing peaks at the points at which the domains flip. Verifying that the exposure of the HZO to the RCA1 solution does not degrade or destroy ferroelectricity.

Secondly, similar CV pads are fabricated that use, instead of a fully covered TiN layer as the bottom electrode, an RCA1 etch patterned TiN pad. Figure 4.3(b) schematically

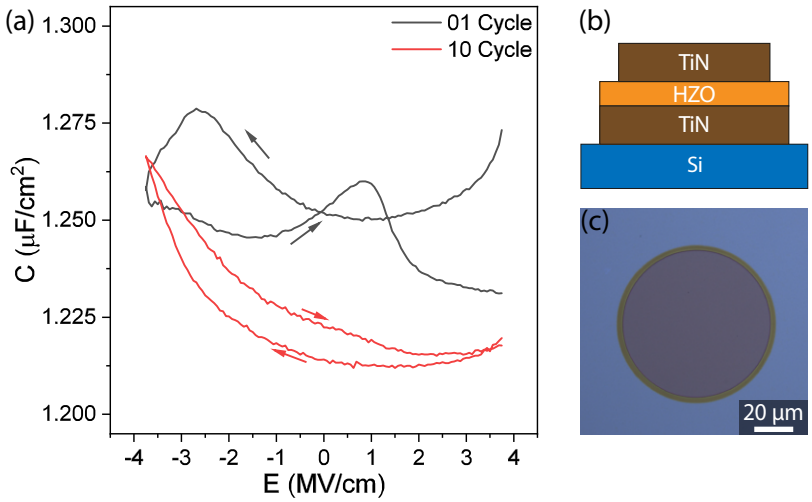

Figure 4.3: (a) CV measurement of MFM capacitor with top and bottom electrode patterned with RCA1 etch, no reverse sputter. (b) Shemitic illustration of the MFM structure, indicating different layers. (c) Optical image of the investigated sample.

illustrates the resulting device structure, in which the bottom electrode is chosen intentionally larger to prevent edge overlap. The optical image in Figure 4.3(c) shows different colors for the different layers, again depending on how they overlap. They differ from the ones in Subsection 4.1.1, because a Si substrate is used instead of the Si wafer with SiO_2 on top. Figure 4.3(a) plots the initial CV curve, which shows ferroelectricity indicated by the appearance of two peaks. They are not of the same height, which is a sign of a degraded switching mechanism. After performing the same measurement ten times, known as cycling, the ferroelectric effect is no longer visible. Measurements in between show that the polarization remains pinned in one direction already for the second cycle. The MFM capacitor gets stuck in one polarization state, which is the case for both samples, regardless of whether an RCA1 etched or lift-off top electrode is used.

Oxidation of the bottom electrode during prolonged exposure to air, introduced by the lengthy RCA1 etch process, is believed to be responsible for this behavior. The oxide can act as a trapping layer that traps only one kind of charge and blocks the accumulation of charges needed to switch the polarization. Further, the compensation of the depolarization field is additionally prevented, which is needed for sustained domain switching, as described in subsection 2.2.3. It can not be excluded that the RCA1 etch process itself is responsible for the oxidation because the previously discussed samples with the bottom electrode on the entire Si wafer have been directly placed into the ALD after the TiN sputter. It is possible that these samples would also exhibit the same behavior if left in ambient air for a sufficient amount of time.

The superficial oxidation of TiN can not be prevented due to the nature of the RCA1 process. Two approaches have been investigated to remove the oxide. First, the sample was dipped in BHF for 15 s, which did not yield in ferroelectric samples. Second, the

sample is reverse sputtered to physically remove the oxide on top before ALD processing, which resulted in ferroelectric MFM capacitors. Due to the FG layer insertion into the gate stack, an MFM substructure is formed in the fabricated device, making implementation of the reverse sputter process mandatory.

4.1.4 Overview Floating Gate Stack Fabrication

Figure 4.4 illustrates a simplified fabrication overview in comparison with the one discussed in the device fabrication Section 3.1. First, Si nanosheets are fabricated on SOI and thermally oxidized (TOX) to 6 nm. The entire sample is covered with 40 nm TiN, and the FG structures are patterned employing an RCA1 wet chemical etch. Using ALD 10 nm of ferroelectric HZO is deposited conformally, which is simplified in Figure 4.4. Again, the entire layer is covert with 40 nm of TiN, and the TG structures are patterned using RCA1. RTA is utilized to crystallize the HZO into a ferroelectric phase at 823 K. Finally, Al is deposited in a lift-off sputter process and the Al-Si exchange is initiated at 773 K.

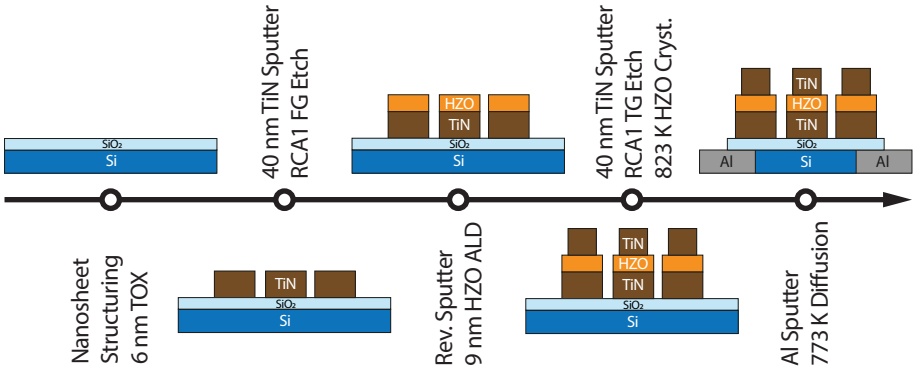

Figure 4.4: Overview floating gate stack fabrication. Oxidized nanosheet base. TiN FG structureing using RCA1 etch. HZO deposition with beforehand reverse sputter. TiN TG structureing using RCA1 etch. Al deposition with lift-off and Al-Si defusion to form SBs.

The subsequent Subsections will analyze the deposition of the floating gate (Subsection 4.1.5), top gate (Subsection 4.1.6), Al-Si exchange (Subsection 4.1.7), and the entire gate stack (Subsection 4.1.8) in detail, implementing mainly AFM and TEM measurements.

4.1.5 Floating Gate Deposition Analysis

As outlined, the floating gate is deposited with reactive magnetron sputter of TiN and subsequent RCA1 wet chemical etch. Figure 4.5(c) shows the schematic illustration of the structured TiN FG layer for a DTG device above the Si nanosheet. It consists of two TiN strips separated by a significant distance to allow control of the channel with the back gate. Figure 4.5(b) is an optical image of the fabricated layer. The color of the TiN changes when it overlaps with the Si nanosheet due to its semitransparency. The in Subsection 4.1.1 mentioned rectangular Si parts of the sheet are clearly visible. They improve the Al-Si exchange of the later added Al to form the SBs and the layer adhesion.

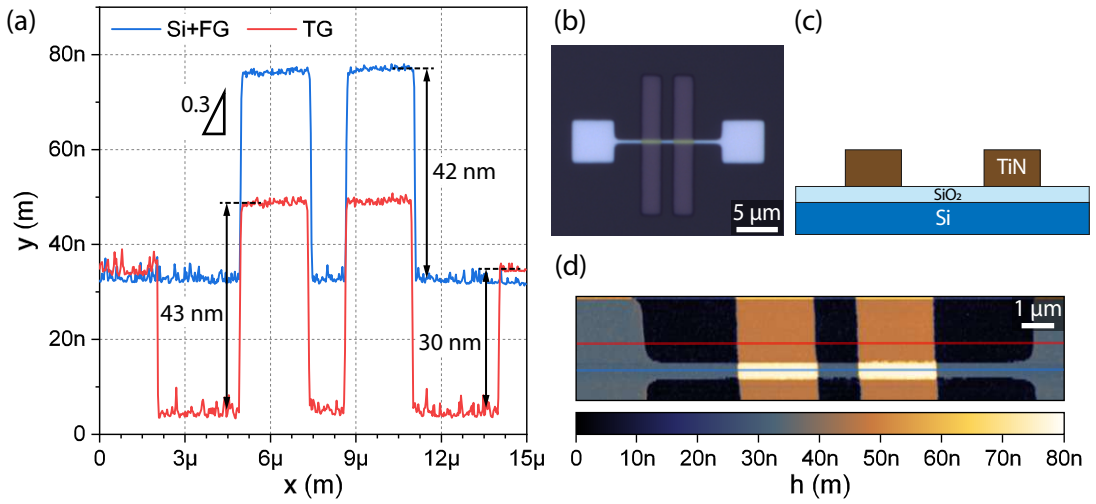

Figure 4.5: Floating Gate layer deposition using RCA1 wet chemical etch process. (a) Layer height along the indicated cross-sections of the sample on top of the Si nanosheet and the BOX. (b) Optical image of the investigated sample. (c) Schematic illustration of the sample, indicating different layers. (d) AFM image of the region near the nanosheet with marked cross-sections.

AFM measurements show the different layer heights in Figure 4.5(d). The two extracted cross-sections, one on top of the Si nanosheet and one over the BOX, are plotted in Figure 4.5(a). A rectangular edge is present, and roughness is minimal. The blue curve on top of the sheet can be used to estimate the TiN layer height because the Si device layer can be considered perfectly flat. In this case, the height is 43 nm. When measuring the height of FG using the read curve 44 nm are extracted. This 1 nm difference can be, on one side, contributed to measurement inaccuracies, and on the other, the RCA1 solution features a relatively low etch rate in SiO_2, which could be responsible. When investigating the height of the Si nanosheet visible on the left and right side of the red curve, a height of 30 nm can be extracted. The 20 nm of the device layer were thermally oxidized to grow 6 nm of SiO_2. According to the Deal-Grove model, $t_{total} = t_{\mathrm{Si}} + t_{\mathrm{SiO_2}}$ should be 23 nm, resulting in a 7 nm difference that is due to the RIE etching, where the used process deliberately etches also parts of the BOX to ensure complete etching of the layer. The slope of the edges is around 0.3 steep. Due to the limited resolution of the used AFM measurement, only one point is in between the min and max value of the step, making slope measurements unreliable. This is further investigated in the following subsections. It should be expected that this process is isotropic.

4.1.6 Top Gate Deposition Analysis

The top gate deposition is analyzed in Figure 4.6 on a TTG device, which features three gates, and to accommodate them, the length of the Si nanosheet is increased. In Figure 4.6(c), the schematic device structure is depicted. Compared to the previously analyzed DTG sample (Figure 4.5(c)) in which only the FG is present, the ferroelectric HZO has been grown using ALD, and TiN TGs have been added utilizing TiN sputter of the

entire sample, followed by RCA1 wet chemical etching. The TiN TGs are slimmed down to ensure that alignment issues do not cause the layer to be over the edge of the floating gates. The optical image (Figure 4.6(b)) shows clearly the overlap between the two TiN layers because of their semitransparency, making it a dark yellow color. When the Si nanosheet is additionally below, it turns into a brown shade.

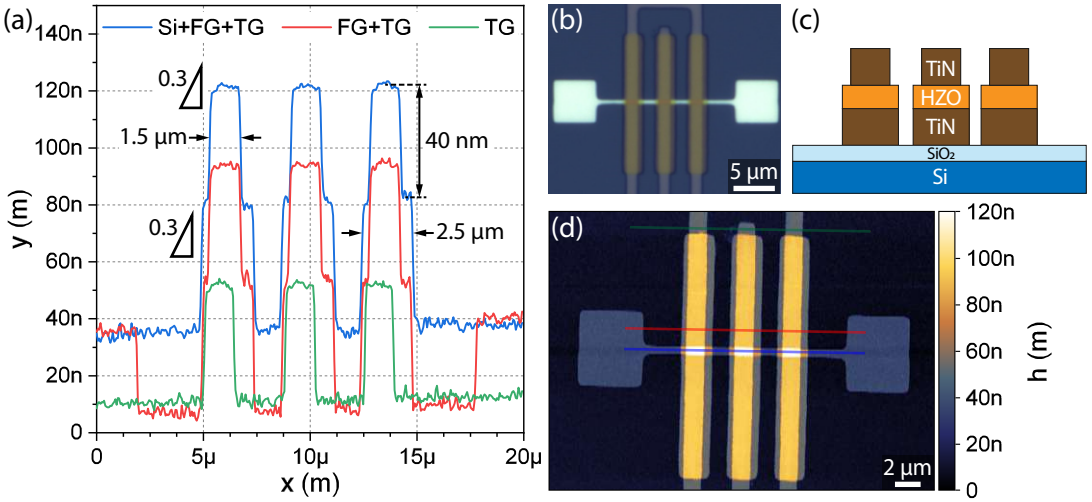

Figure 4.6: Topg Gate layer deposition using RCA1 wet chemical etch process. (a) Layer height along the indicated cross-sections of the sample on top of the Si nanosheet and the BOX. (b) Optical image of the investigated sample. (c) Schematic illustration of the sample, indicating different layers. (d) AFM image of the region near the nanosheet with marked cross-sections.

The different layer heights of the sample are shown in the AFM measurement (Figure 4.6(d)). The clear height difference between the overlapped regions is visible, and a slight misalignment of the TG to the FG layer is present. Cross-sections are indicated and plotted in Figure 4.6(a). In blue, the entire stack is shown. The different features measured before in Figure 4.5(c) are the same because HZO grows conformal, which adds the same amount of HZO both on top of the FG and on the BOX. The added top layer height is 40 nm. The slopes of the TG and the FG are 0.3. The thickness of the TG is 1.5 µm and of the FG 2.5 µm, which is in accordance with the intended CAD file. In red, the FG stack is shown relative to the BOX, including the ends of the Si nanosheet, and in green is the TG depicted without the FG below on the BOX.

4.1.7 Al-Si Exchange Analysis

As described in Subsection 3.1.6, Al is added by first reverse sputtering the HZO, which is in a non-ferroelectric phase above the ends of the nanosheets, and followed by BHF etching to remove the SiO_2 below. Only then is the Si nanosheet exposed, and Al can be sputtered. The Al-Si exchange happens during RTA at 773 K and allows control of the SB position as needed.

Figures 4.7(a-d) are a collection of optical microscopy images of fabricated devices after

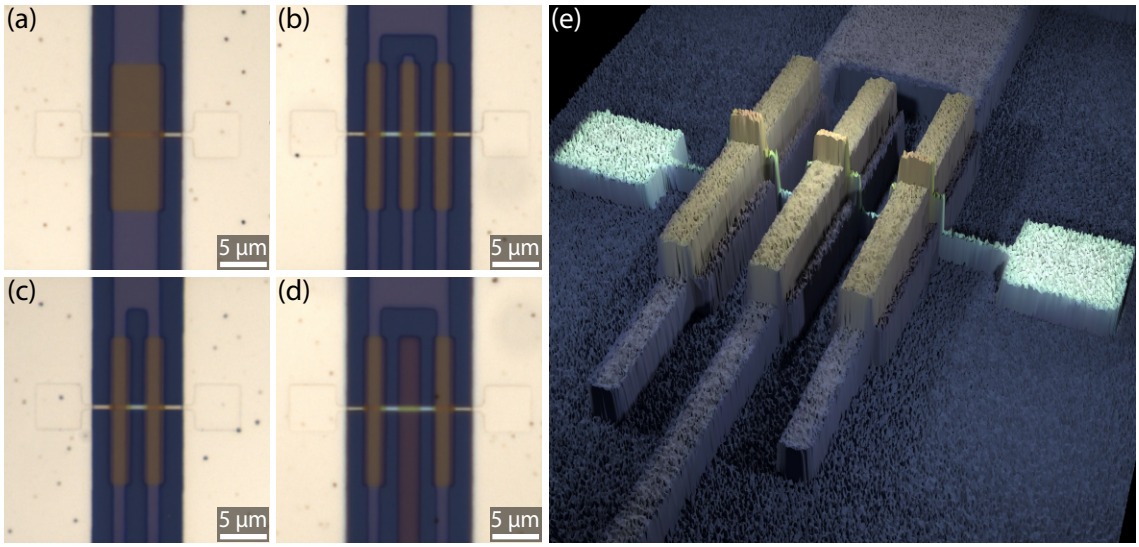

Figure 4.7: Complet device after Al-Si exchange. (a) FG STG FeSBFET (b) FG TTG FeRFET (C) FG DTG FeRFET and (d) FG TTG FeRFET with CG on FG layer. (e) 3D view of TTG device. 2D optical microscope image overlayed with AFM scan data.

the exchange. All feature the Al source/drain pads that cover the rectangular regions of the Si nanosheet. Figure 4.7(a) is a FG STG FeSBFET. The exchanged Si is clearly visible because the nanosheet features the same color as the pad. Beneath the gate stack, a clear distinction can be made between the Al and Si parts of the sheet. In this comparison, the increased nanosheet length of the TTG design, in contrast to the STG and DTG, is clearly shown. The TTG design in Figure 4.7(b) allows identification of the aforementioned color difference between the Al and Si parts of the sheet without other materials on top because the diffusion is controlled in a way that positions the SB below the two PG, left and right. The two gaps between the CG and PG show the Si nanosheet color. This is also the case for the DTG design in Figure 4.7(c). Figure 4.7(d) is a variation of the Figure 4.7(b) design in which the FG is replaced by the CG, visible by the color difference between the PG and the CG. This design is meant to program the n- or p-conduction mode of the RFET and does not interfere with the CG modulation. The 3D illustration in Figure 4.7(e) of a TTG device obtained before Al deposition shows the different layers of the gate stack and their individual heights clearly. Especially the overlap of the FG and TG, as well as the overlap of the entire gate stack with the Si nanosheet, is highlighted.

When comparing the fabrication results in Figure 4.7 with the device that was initially fabricated using a lift-off process instead of the developed RCA1 wet chemical etch in subsection 4.1.1 the advantages of this approach become clear. For one, the spikes on the edges are prevented, and because the TiN can not be deposited on the resist, it can not be delaminated. Furthermore, it is theoretically possible to move directly from SiO_2 or HZO growth to TiN deposition, which results in pristine interfaces.

4.1.8 TEM Analysis

Figure 4.8(a) is a detailed schematic illustration of an FG STG FeSBFET device cross-section, including the BOX and handle Si wafer. The TEM analysis regions are indicated. In Figure 4.8(b), an optical microscope image of the investigated device allows identification of the SB positions.

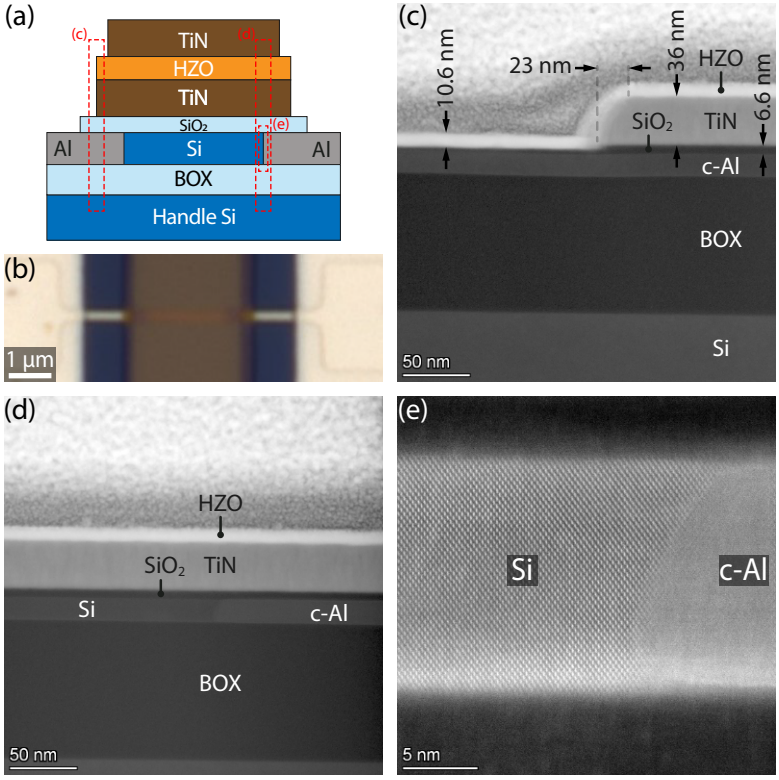

Figure 4.8: (a) Detailed FG STG FeSBFET device cross-section, indicating TEM imaging regions. (b) Zoomed-in optical microscope image of the investigated device. HAADF STEM image of (c) FG edge, (d) gate stack over SB, and (e) Si/Al interface.

A high-angle annular dark field (HAADF) scanning transmission electron microscope (STEM) image of the FG layer edge is shown in Figure 4.8(c). The different layers are indicated. In this image, the edge of the TiN is rounded. From the classical wet chemical etching a sharp edge would be expected. This is most likely connected to the performed reverse sputter, which damages the edge. The slope can be measured, and a gradient of 0.6 is determined, which is double what has been analyzed using AFM measurements. This means that the RCA1 wet chemical does not classify as isotropic and is most likely connected with the TiN phase deposited. The impact of the reverse sputter can be seen not only by the reduced FG layer height but also by the reduced thickness of the SiO_2 not covered by the FG. Originally the SiO_2 was 6.6 nm which is in line with the expected value from Section 3.1.2. The HZO is grown conformal around the FG since it features an identical height through the layer, it is around 10.6 nm thick. Figure 4.8(d) shows the gate stack around the SB. When the Al diffuses into the Si nanosheet, it forms crystalline aluminum

(c-Al). The Si to c-Al interface is clearly visible and enhanced in Figure 4.8(e). Here, the different Si atoms are identifiable, which highlights the abruptness of the interface.

4.2 Electrical Analysis

The first FG approach mentioned in Subsection 4.1.1, utilizing lift-off instead of RCA1 wet chemical etching for the TiN metal layer insertion, shows that the desired results could not be achieved due to the presented insufficient processing. A short between the TG and FG metal layers is observed for all devices. Transfer characteristics are not ferroelectric since the HZO is bridged and presents the same behavior regardless of whether the TG or the FG is swept. CV test pads, which are identical to the ones that will be presented in the following Subsection 4.2.1, were also investigated on this sample and showed clearly ferroelectric behavior.

DTG and TTG devices have been fabricated as outlined in Section 4.1 and the previous Chapter 3 but not tested in detail. TTG samples exhibit because of a large ungated region, almost no controllability of the device, and for DTG samples, it is believed that the introduction of the FG alters the device behavior since this additional layer can be influenced by the BG.

The following subsections will present and discuss the electrical measurements conducted from FG STG FeSBFET devices implementing the RCA1 etch. First, the samples test CV pad and non-ferroelectric test structures are investigated. Then, transfer characteristics are used to derive a possible device behavior mechanism. Finally, pulsed measurements in a slow and high-speed regime are examined, which showcase the possible application of such a device in neuromorphic computing or logic in-memory applications.

4.2.1 CV Pad Teststructures

Figure 4.9(a) illustrates CV measurements on test structures that allow to validate and characterize the ferroelectric HZO. The pad structure is depicted schematically in Figure 4.9(b). An MFM structure is formed by the metal TG and FG layer in which the HZO is placed in between (Section 3.1). By also keeping the Si device layer below the pads, this structure becomes, from a processing point of view, identical to the situation on top of the Si nanosheets. Because the TiN FG layer is normally covered by HZO, it is connected using the Al metal layer introduced during source/drain deposition. In contrast to the CV measurements performed operating the BG of a Si wafer, like in Subsection 4.1.3 for the reverse sputter investigation, this approach creates a true MFM because only metals are between the HZO and the prober needles. Figure 4.9 is an optical microscope image of the measured circular CV pads. Sizes including 37.5 µm, 50 µm, 75 µm, and 100 µm diameters are fabricated. The pad is visible as three circles placed inside of each other. The outermost is the TiN FG layer, which features additionally a contact line to the Al BE pad as illustrated in Figure 4.9(b). This is followed by the Si device layer, which is actually below the FG layer, but its color is altered due to the semitransparency of TiN. The smallest circle is the TG layer.

The measurements are recorded after Al deposition and after Al diffusion, which gives

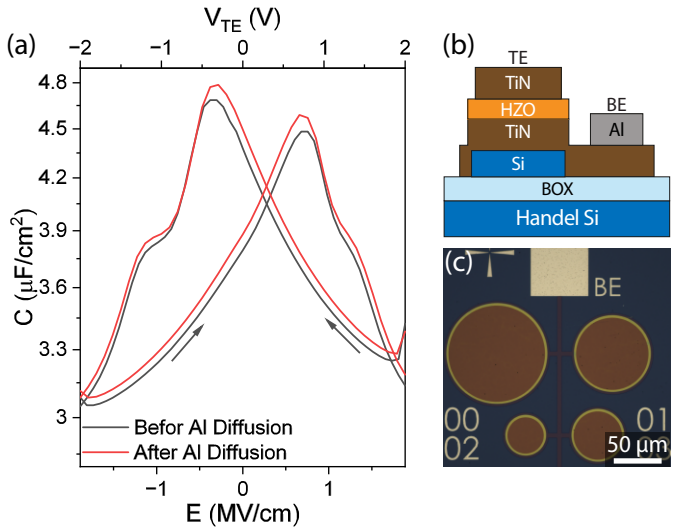

Figure 4.9: (a) CV test structures before and after the Al-Si exchange diffusion. (b) Schematic device structure. (c) Optical microscope image of the tested pads.

information regarding how the ferroelectric response changes when it is subjected to an RTA process of 773 K in addition to the one at 823 K originally done to crystalize the HZO. As depicted in Figure 4.9(a), the CV measurements only slightly differ from each other. This clarifies that the additional RTA does not affect the HZO. A slight shift of the curves is present and the E_C is slightly below 0.5 MV/cm. The curves show, after surpassing E_C, a small plateau that can be connected to the wake-up of the device. These measurements are the first conducted on the pad, and further cycling could prevent this behavior. It can be assumed that the second RTA brings the CV pads back into the as-deposited state, removing any domain polarization present.

4.2.2 Non-Ferroelectric STG SBFET

In Figure 4.10, an FG STG FeSBFET is investigated, which allows its FG to be connected and, by doing so, operate the device without interference from the HZO. The measurement setup is shown in Figure 4.10(c). A V_{DS} of 0.5 V is applied, and the BG as well as the TG are grounded. On the FG, a double sweep from -3 V to +3 V is performed while recording the drain current I_D. The STG device behavior explained in Subsection 2.3.2 is observed in Figure 4.10(a). First, field emission (FE) of hole charge carriers is present for negative voltages, which is replaced by thermionic emission (TE) above -1 V. The device turns off due to the increased tunnel barrier around 0 V. An off-current of below 100 fA can be observed, which is also the noise floor of the used semiconductor analyzer. Next, TE of electrons takes over for positive voltages, and at 1.5 V again, TE is the dominant conduction process over the SB. This device is able to modulate the current over eight orders of magnitude with subthreshold slopes S_{th} of 80 mV/dec for the hole and 280 mV/dec for the electron conduction regime. Which is an increase compared to other devices known from the literature. Especially the 80 mV/dec are very close to the physical

limit given for planar field effect transistors of 60 mV/dec.

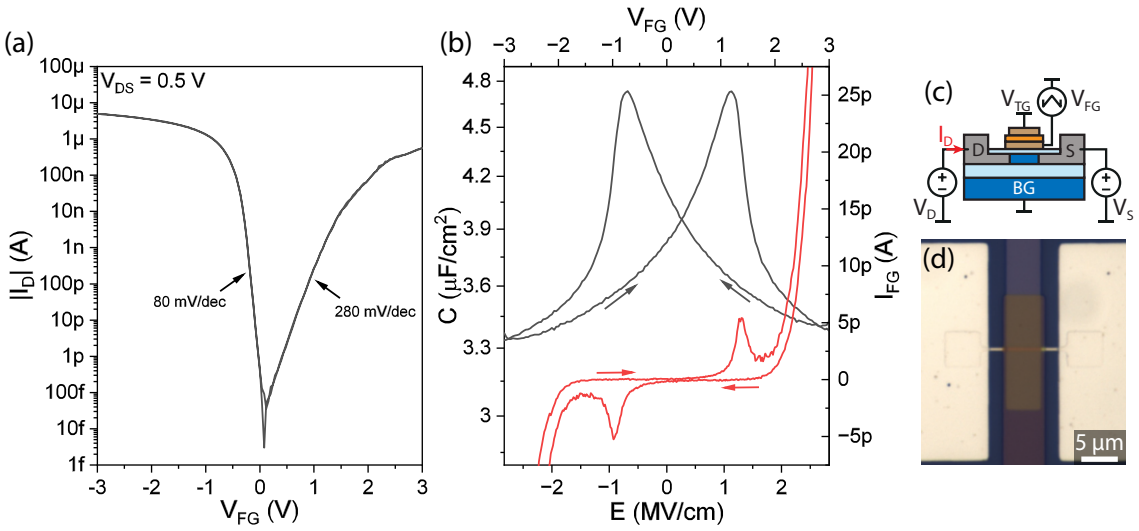

Figure 4.10: (a) Transfer characteristic of FG STG FeSBFET by sweeping the FG, S_{th} values indicated. (b) CV and IV measurements of FG to TG connection creating an MFM capacitor. (c) Schematic illustration showcasing measurement setup. (d) Optical microscope image of the investigated device.

The optical microscope image in Figure 4.10(d) allows one to distinguish between the TG electrode (center top) and the FG electrode (center bottom). The overlapped region presents a dark yellow color and features, according to the CAD file and evaluations using the image, an area of around 80 µm^2. By letting V_D and V_S be floating in Figure 4.10(c) and connecting the FG as well as the TG contact to a semiconductor analyzer the measurement plotted in Figure 4.10(b) can be performed. The CV measurement is comparable with the ones obtained for the test CV pads (Figure 4.9(a)), which is interesting because the size of the used capacitors is almost one order of magnitude different. The capacitance value at the peaks is identical between the two device structures, but the values for higher voltages do not match. This can be connected to the smaller size of the capacitor in Figure 4.10(d), which results in low currents and the noise floor of the analyzer being reached for these voltages. Also, the device butterfly curve features higher E_C values and is more symmetric around zero. The symmetry improvement can be explained by top and bottom electrode work functions, since in Figure 4.9(a) an Al/TiN bottom electrode is used in contrast to the TiN only one in Figure 4.10(b). Additionally, was cycling performed before the shown measurement, which explains the more pronounced device characteristic.

The IV measurement also indicates the ferroelectric switching at -0.9 MV/cm and 1.6 MV/cm by two clearly distinguishable current peaks. Beyond these points, the current increases rapidly, indicating a soft-breakdown of the dielectric. The top x-axis of Figure 4.10(b) converts the electric field E into the applied floating gate voltage V_{FG} by multiplying with the HZO thickness of 10.6 nm extracted from the TEM measurements in Subsection 4.1.8. The dielectric constant of the capacitor in its two fully switched

conditions can be estimated by using the capacitance value recorded at the highest positive or negative applied electric fields which are 2.83 MV/cm or 3 V and features around 3.3 µF/cm^2, this allows to calculated ε_r as 39. Furthermore, by employing Equation 3.9 the polarisation P of the device in the switched state can be estimated as roughly 10 µC/cm^2.

4.2.3 Tranfer Characteristic

A similar transfer characteristic is conducted for the normal FG STG FeSBFET devices. As depicted in Figure 4.11(b), a source/drain voltage of 0.5 V is applied symmetrically, and the BG is connected to the ground. On the TG the sweep signal of the semiconductor analyzer is attached. Figure 4.11(c) shows the optical microscope image of the investigated device, which features an FG layer that is not accessible and limited to around the Si nanosheet, unlike the sample studied in Subsection 4.2.2.

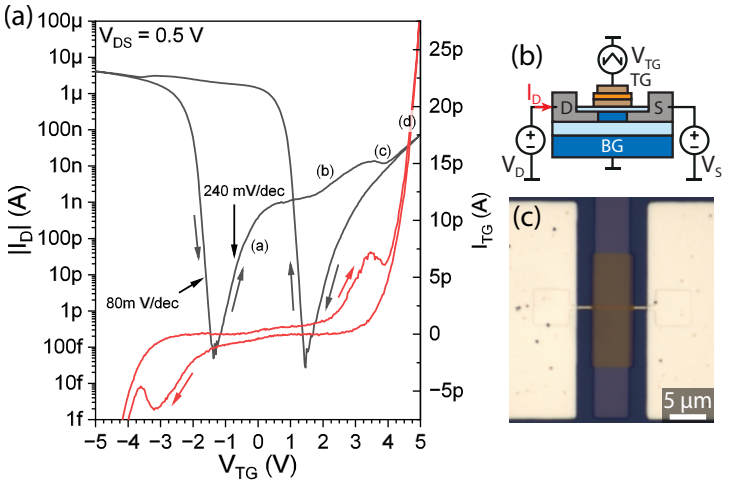

Figure 4.11: (a) Transfer characteristic of FG STG FeSBFET by sweeping the FG, S_{th} values indicated. (b) Schematic illustration showcasing measurement setup. (d) Optical microscope image of the investigated device.

A double sweep from -5 V to 5 V is performed and plotted in Figure 4.11(a). Compared to the transfer shown of the test device, which does not have a ferroelectric HZO in Figure 4.10(a), a clockwise hysteresis is present, which causes the entire characteristic to be shifted either towards the negative or positive, depending on whether the sweeping direction is positive or negative. The resulting measurement is identical to the one obtained before but is shifted by ±1.4 V. It is capable of modulating I_D over eight orders of magnitude. Off currents of the two paths are as the one before, limited by the noise floor of the analyzer, but the on currents become asymmetrical. The hole conduction still reaches the same value, but the electron current is roughly two orders of magnitude smaller, possibly due to some increased electron trapping. Higher voltages are used in contrast to the previously investigated device, which only featured a SiO_2 layer between the modulating electrode and the nanosheet. The device analyzed in Figure 4.11 also incorporates the HZO as well as the FG layer, which increases the needed electrical fields.

Interestingly, similar S_{th} values of 80 mV/dec for the hole and 240 mV/dec for the electron conduction are obtained. A reduction of these values would be expected.

In addition to the transfer characteristic, the current through the TG contact I_{TG} is recorded and plotted. It shows a similar behavior to the MFM structure investigated. For high voltages, the device goes into a soft-breakdown, and the current through the gate stack increases and is added to the I_D. Slightly before, two peaks of opposite values are present on both sides, which indicates ferroelectric domain switching. This current can also be identified in the transfer characteristic. At point (c), an increase in the I_D is notable, which is in line with the I_{TG} peak. Also, a small reduction in I_D can be seen in the negative sweep direction at -3 V.

Figure 4.12(a) shows the transfer characteristic of an FG STG FeSBFET for a V_{DS} of 0.1 V, 0.5 V, 1 V, and 2 V by double sweeping the accessible FG between -3 V and 3 V. A change in the on- and off-state values is visible. The hole current on-state increases by more than one order of magnitude and by three orders of magnitude for the electron current. The off-state increases by six orders. This behavior is connected to the alteration of the band structure for changing V_{DS} voltages. For high V_{DS}, the thickness of the tunneling barrier for FE is, in addition to V_G, thinned down, enhancing on-state currents. Also, the barrier height for TE at V_G values of 0 V is reduced, degrading the off-state currents. Hole current S_{th} remains constant for different V_{DS} voltages except for 2 V. On the other hand, the electron current S_{th} seems to first improve and then back off for higher values.

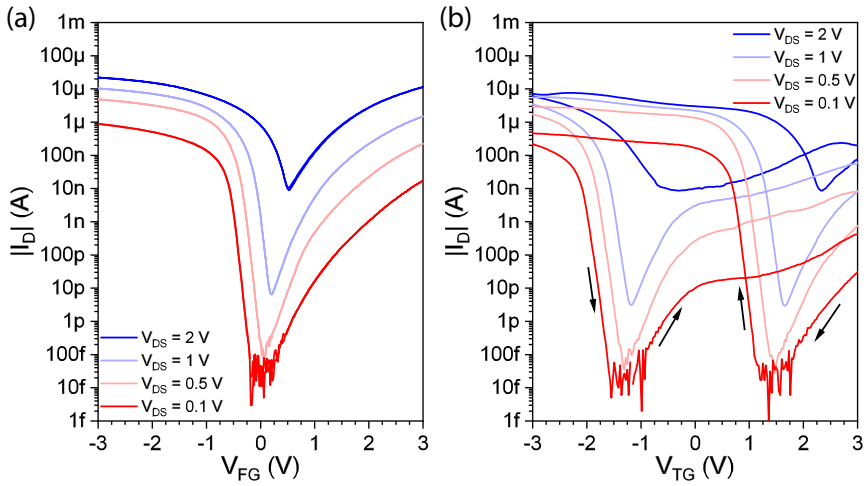

Figure 4.12: Transfer characteristic for different V_{DS} values ranging between 0.1 V and 2 V by sweeping (a) the FG and (b) the TG for separate devices.

The normal FG STG FeSBFET device in Figure 4.12(b) features in some aspects a similar characteristic. Measurements were conducted between -4 V and 4 V. In the plot only the -3 V to 3 V range is depicted to allow clearer comparison to Figure 4.12(a). On- and off-states vary in a similar manner. The previously discussed variation of the electron conduction mode in Subsection 4.2.3 is present for all V_{DS} voltages. For this type of

device, 0.5 V features the best compromise between good off-state and electron current on-state. For 2 V the device degrades a lot due to the reduced electron current. S_{th} values are as before.

4.2.4 Proposed Mechanism

The transfer characteristic data can be combined with the physical mechanism of FG devices in the theory Subsection 2.2.5. By using the optical microscopy image in Figure 4.11(c) as well as the intended CAD file, the surface of the FG and TG can be estimated as 80 µm^2 and 96 µm^2. The overlap of the FG and the Si nanosheet has an area of 2.4 µm^2. From the TEM analysis in Subsection 4.1.8 a SiO_2 thickness of 6.6 nm and HZO thickness of 10.6 nm are obtained, making the gate coupling coefficient α_G in Equation 2.5 almost 1, for a choice of dielectric constants of 3.9 for ε_{SiO_2}, which is book value, and of 39 for HZO, as estimated earlier. This high α_G value is a result of the reduced C_{SiO_2} due to its small surface of 2.4 µm^2 compared to the HZO and means that for an uncharged FG layer and unpolarized ferroelectric, the entire V_{TG} is applied to the SiO_2. The gained measurement does not indicate such a field situation since ferroelectric switching is observed, which requires a certain voltage across the HZO. On the other hand, it explains why S_{th} is unaffected by the addition of the HZO and TiN FG layer, compared to the SiO_2 only device.

By taking into account the observed shift from the origin position of the transfer characteristic in Figure 4.11(a) of ±1.4 V and together with Equations 2.5 and 2.7, the net polarization P_{net} can be calculated as 3.7 pC. It can be assumed that P_{HZO} is around 10 µC/cm^2 (Subsection 4.2.2) resulting in 8 pC, in accordance to the TG and FG overlap. Since P_{HZO} and Q_{ML} in Equation 2.7 are of opposite sign, approximately 11.7 pC or 12.2 µC/cm^2 of Q_{ML} charge is injected into the FG layer. This indicates that the metal layer gets charged beyond the limit set by the HZO.

Figure 4.13 shows the band structure of the gate stack over the Si nanosheet at different positions in the transfer characteristic. The labels of the individual images correspond with the marked points in Figure 4.11(a). Since the sweep begins at -5 V, it can be assumed that the ferroelectric is polarized upwards and the FG layer has been injected fully with holes, resulting in the situation depicted in Figure 4.13(a) at 0 V. The decreased FG potential compared to the TG bends the bands downwards, inducing an electron current. Because the polarization and the charges in the FG layer keep the voltage drop across the HZO low, the SiO_2 band structure gets bent drastically as shown in Figure 4.13(b) when a positive V_{TG} voltage is applied, which enables the injection of electrons replacing the holes in the layer. The FG layer gets progressively charged, but unlike the FG in standard flash memory (Subsection 2.2.5), the injection currents get less impacted by this due to the polarization keeping the band structure constant. After reaching a certain value of injected charges, the ferroelectric switches downwards, as depicted in Figure 4.13(c). This abruptly determines a higher field across the HZO and reduces the band bending of the SiO_2, preventing further injection of electrons. Now, the amount of charges is governed by

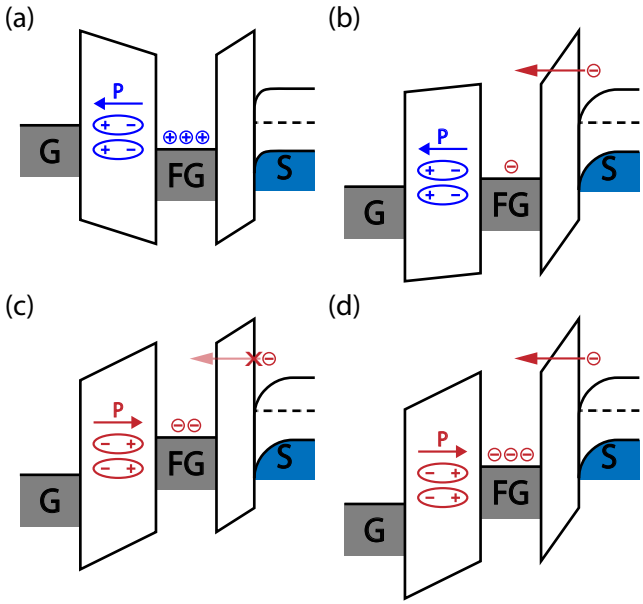

Figure 4.13: Band structure of gate stack at different positions in transfer characteristic (Figure 4.11). (a) Programmed state at 0 V. (b) High charge injection rate. (c) Injection is prevented by ferroelectric domain switching. (d) Extremely high voltages can cause further injection.

the amount needed to switch the HZO. If V_{TG} is increased beyond this point a situation in which tunneling through the SiO_2 is possible can be reached again, which is shown in Figure 4.13(d). This can explain the previously observed inconsistencies of the injected Q_{ML} and the polarization P_{HZO} gained from the shift in Figure 4.11(a).

The proposed mechanism clearly shows multiple advantages compared to a standard FGFET explained in Subsection 2.2.5. It reduces the needed voltage for charge injection because the HZO polarization additionally flattens out the dielectric band structure and, by doing so, increases the voltage across the SiO_2, allowing for smaller write voltages. Additionally, the polarization switching stops the injection abruptly, which prevents the potential overcharging of the FG layer as long as V_{TG} is not increased beyond the switching point.

4.2.5 Pulsed Time Measurements

To investigate the usage of the examined device for storage applications, pulsed measurements are performed. Figure 4.14 shows the effect that the set pulse time t_{set} has on the device characteristic. The measurement setup, explained in Subsection 3.2.3, is used to first send a -5 V amplitude reset pulse of 10 ms followed by a 5 V amplitude set pulse of varying duration. Figure 4.14(a) shows the transfer characteristics recorded after the set pulse by sweeping the gate between -0.5 V and 0.5 V with a V_{SD} of 0.5 V. The t_{set} is rising from left to right, and its corresponding value is indicated by the arrow color. In the black region $t_{pulse} = 0.25 \cdot x$ from 0.5 ms to 1.5 ms, in the red region $t_{pulse} = x$ from 2 ms

to 8 ms, and in the blue region $t_{pulse} = 2^x$ from 16 ms to 256 ms, allowing to investigate a wide range of t_{set} values. The transfer characteristic data is displayed differently in Figure 4.14(b) by taking each V_{TG} value and plotting their corresponding drain current I_D in relation to the t_{set} time. This measurement needs to be carefully interpreted since in Figure 4.14(a) I_D is displayed in a logarithmic scale and in Figure 4.14(b) in a linear scale.

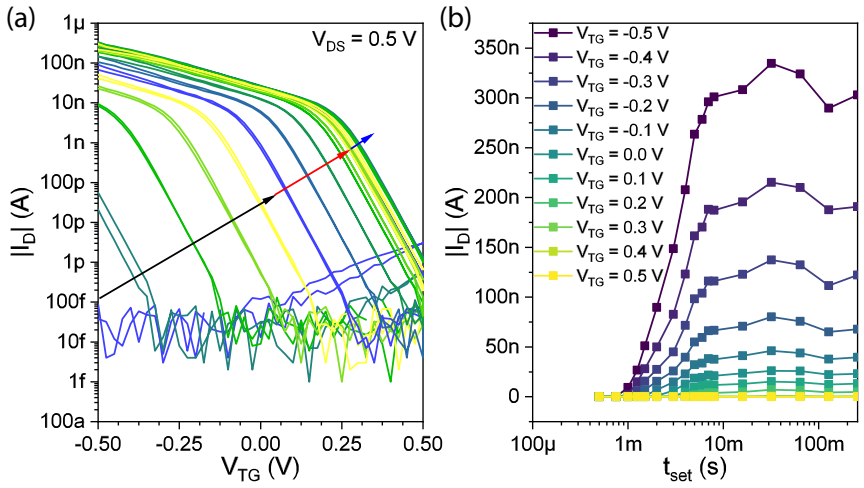

Figure 4.14: Pulsed measurement with varying t_{set} time. (a) Transfer characteristic in a limited V_{TG} range and marked regions indicating t_{set}. (b) LTP graph for different V_{TG} values extracted from transfer.

Both graphs show that the device state can be programmed by the explained pulsing scheme. A shift of the transfer characteristic from left to right depending on the used t_{set} is shown in Figure 4.14(a). This is in line with the expected behavior given by the proposed model in Subsection 4.2.4. The amount of charge injected is directly proportional to the injection time. The already observed flattening of the shift (Figure 4.14(a)) is clearly visible in Figure 4.14(b). Initially, before surpassing 1 ms of t_{set}, the curve remains at very low values because the transfer is mainly in its off-state. After surpassing 2 ms, it increases and flattens again at 6 ms. When examining the 6 ms transfer characteristic in Figure 4.14(a) it can be estimated that the off-state has been shifted to 0.6 V. Similar to what has been done in Subsection 4.2.4, P_{net} can be extracted as 1.6 pC, this results in a relative charge of the metal layer Q_{ML} of almost 10 µC/cm^2, which is identical to P_{HZO}. This shows that after reaching a t_{set} of 6 ms the ferroelectric switches and further injection is reduced as explained in Figure 4.13(c) of Subsection 4.2.4. Further and longer pulsing can, as long as the required V_{TG} voltages are reached, inject additional charges in the FG layer with a lower efficiency due to the reduced SiO_2 band bending, shown in Figure 4.13(d).

This explains the charge deviation between the P_{HZO} and the Q_{ML} in the transfer characteristic in Figure 4.11(a). During this measurement, the V_{TG} voltage is applied to the device for a prolonged time, depending on the present I_D, reaching hundreds of millisec-

onds or seconds. This means that, especially for the extreme values at -5V and 5V, the band structure is bent strongly, and injection is sustained for a long time, which allows, on one side, to switch the ferroelectric but also enables the overcharging mechanisms. In the framework of the proposed switching mechanisms, it should be possible to find a pulse amplitude that is capable of switching the device without subsequent overcharging.

4.2.6 Pulsed Number Measurements

In Figure 4.15, pulsed measurements are presented. Here, instead of varying the t_{set} time, the number of set pulses n_{set} is increased with a fixed t_{set} of 0.5 ms. The pulse amplitude for reset and set is chosen as -5 V and 5 V. Two measurement runs are shown, one for positive set pulses and one for negative set pulses. The pulse number is again indicated by the arrow color. In the black region $n_{pulse} = x$ from 1 to 16 and in the red region $n_{pulse} = 2^x$ from 32 ms to 512 ms, resulting in a cumulative maximum pulse time of 256 ms which is identical to the one used before.

Figure 4.15(a) shows the transfer characteristic recorded after the set pulse by sweeping again V_{TG} between -0.5 V and 0.5 V for a V_{SD} of 0.5 V. Compared to Figure 4.14(a), the transfer characteristic is more gradually shifted, which could indicate that the injection starts delayed to the rising edge of the pulse and reduces the effective injection time when using multiple individual pulses instead of a long one. The previously outlined 0.6 V shift of the transfer characteristic at which P_{HZO} and Q_{ML} are equal in relation to their respective area is present in this measurement run for $n_{set} = 64$. Before, this point was reached for a t_{set} of 6 ms and the cumulative value of n_{set} in this measurement is 32 ms, since one pulse features a t_{set} of 0.5 ms. This difference speaks again towards the pulsed number measurements not injecting the same amount of charges as the time measurement. In Figure 4.15(b), the different individual measurements are plotted as a function of n_{set}, for different values of V_{TG}. Again, three individual regimes are present, and the curves start to flatten at the $n_{set} = 64$ point.

The inverted direction in Figure 4.15(c) shows similar behavior for the individual transfer characteristics. The resolution or distance between measurements is similar, but the max value is lower, which becomes clear in Figure 4.15(c). Compared to before, where the graph saturated at around 325 nA, now only a starting point 225 nA is observed. This is connected to the t_{set} time, which is already long enough to influence the injection tremendously. A lower t_{set} is needed to bring back values above the 225 nA, which will be investigated in Subsection 4.2.7.

Figure 4.15(b) and (d) are very similar to the LTP and LTD graphs discussed in the theory Subsection 2.2.4 on the application of established FeFETs as neuromorphic devices. Compared to those, the LTP Figure 4.15(b) has the capability to have much more intermediate states since n_{set} step size was increased after 16. Furthermore, as explained by using smaller t_{set} values, the number of intermediate states in the LTD Figure 4.15(d) can also be improved.

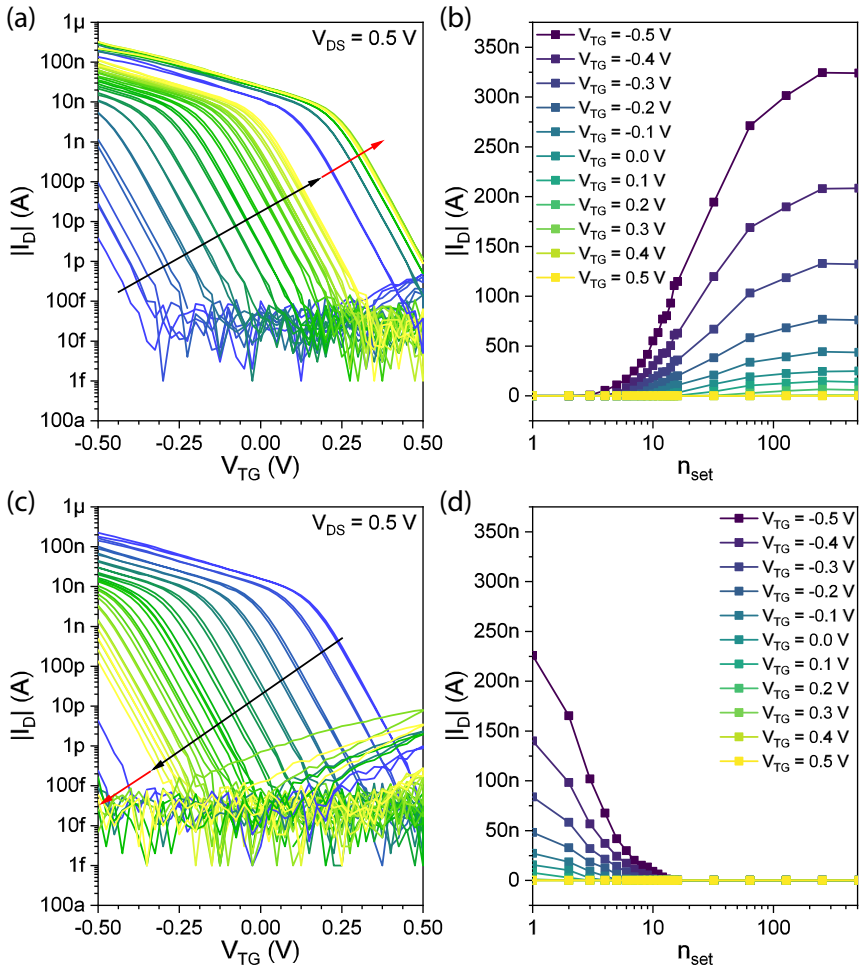

Figure 4.15: Pulsed measurement with varying n_{set} pulse number. (a) Transfer characteristic in a limited V_{TG} range and marked regions indicating n_{set}. (b) LTP graph for different V_{TG} values extracted from transfer. (c) Transfers for inverted pulsing direction and (d) LTD graph.

Throughout the conducted measurements, relatively high pulse times are used in order to switch the ferroelectric compared to other similar works. This can be explained by the size of the FG layer itself, which features an area around forty times larger than the overlap with the Si nanosheet. On the one hand, this results in a high gate coupling coefficient α_G and, on the other, requires a lot of charges to be injected into the FG layer for ferroelectric switching since, in this regard, only the amount relative to the area is important. If the FG area is reduced to a similar size as the overlap, the amount of injected charge needed is decreased, which increases switching times.

4.2.7 Fast Pulsed Time Measurements

To further analyze the devices, faster pulsed measurements are performed by employing a different setup, as described in Subsection 3.2.3. Pulses with t_{set} between 500 ns and 20 µs

are performed with an amplitude of 5 V. A 10 µs reset pulse scheme with -5 V amplitude is used. Transfer characteristics are measured by sweeping V_{TG} between -0.5 V and 0 V. For each pulse length, five identical pulsing runs are combined, and the mean value, as well as the standard deviation for each point, is calculated.

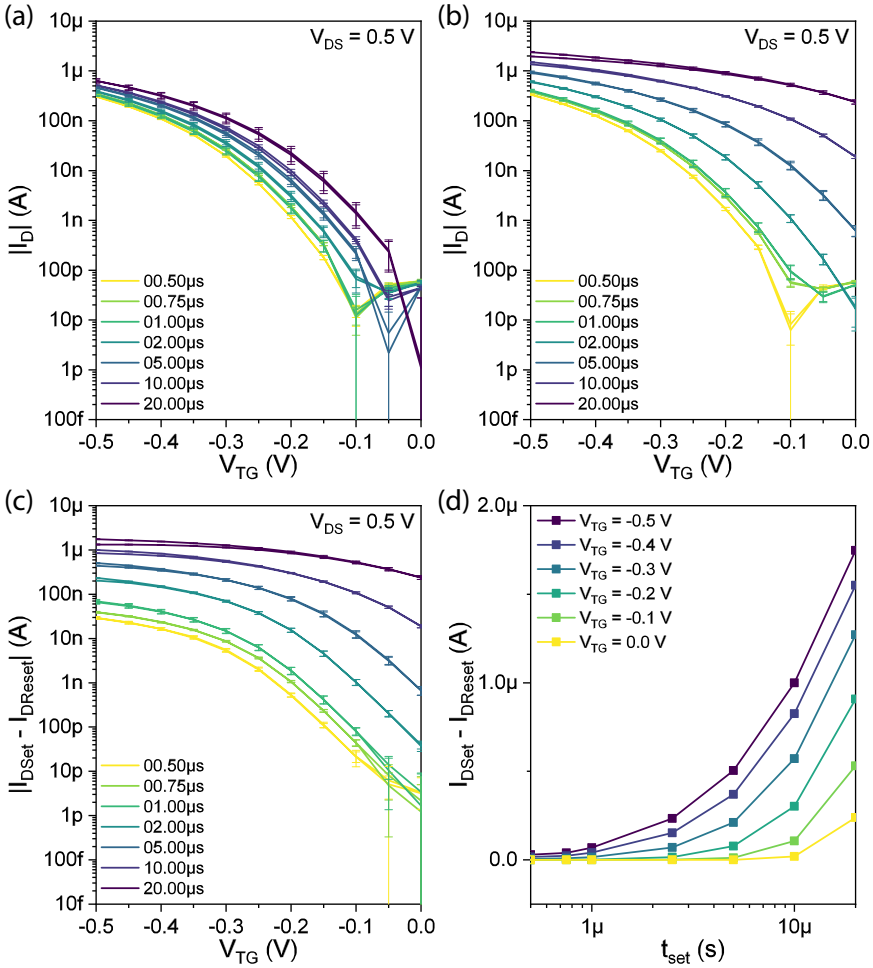

Figure 4.16: Pulsed measurement with varying t_{set} time in microssecond range. Transfer characteristic after (a) reset pulse and (b) set pulse. (a) $I_{Dset} - I_{Dreset}$ difference transfer characteristic and (b) corresponding LTP plot.

Figure 4.16(a) shows the transfer characteristic after the reset, and Figure 4.16(b) before the set for the corresponding t_{set}. The characteristics measured after different reset pulses do not perfectly match, which indicates that the used pulsing scheme does not produce a reliable starting point. In general, this is something that has also been observed during the previous pulsed measurements. Because the proposed mechanism relies on the injection of charges into the metal FG layer, it is crucial to know the amount present before and after the set pulse in order to correctly interpret the set pulse's impact on the device. Certain pulsing schemes are sufficient for low t_{set} values, but for ever-increasing times, the previously sufficient scheme might not be capable of bringing the device to the desired

reset state, noticeable by a boost in voltage shift due to the remaining charges in the layer. This becomes clear in Figure 4.16(a), where the reset state of higher t_{set} set pulses is shifted slightly to the positive due to previously injected charges. This changes the situation after the set in Figure 4.16(b). Excluding this aspect, a good controllability of the device characteristic can be observed. The reset pulse shifts the characteristic to around -0,1 V, which means that only some of the domains are switched, and the set pulse shifts the curves towards the positive. In general, a good reproducibility of the curves was found, indicated by the low error bars.

To circumvent the variation observed after the reset pulse, the difference between I_{Dset} and I_{Dreset} is plotted in Figure 4.16. This relates the set pulse transfer to the reset pulse transfer. Figure 4.16(c) shows the resulting transfer characteristics. A clear shift of the curve is visible, bringing them to the positive. The LTP plot in Figure 4.16(d) indicates a similar picture. When comparing these results to the ones described before, it has to be considered that here, a different device is used, which explains the slightly shifted characteristic that enables higher current values. It is observed that the point at which P_{HZO} and Q_{ML} are equal in relation to their areas is never reached, meaning longer pulses, like in Subsections 4.2.5 and 4.2.6, are required. Pulses in this short t_{set} regime are needed for devices that feature a smaller FG layer area, since here the injection would saturate earlier.

Chapter 5

Summary and Outlook

In this thesis, a TiN floating gate (FG) metal layer is integrated into ferroelectric Schottky barrier field effect transistors (FeSBFETs) with the goal of further enhancing device characteristics and showcasing possible adaptations in logic in memory (LiM) or neuromorphic computing applications. Previous works realized FeSBFETs by using a SiO_2 interlayer and ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ (HZO) in an established framework, consisting of Si nanosheets etched out of a silicon on insulator (SOI) device layer, Al source/drain pads, and an Al-Si exchange process which forms the Schottky barriers (SBs). Additionally, TiN top gates (TGs) have been employed to induce a ferroelectric HZO phase during crystallization. These devices showed inefficient switching of the ferroelectric above the SBs, which governs the conduction mechanism and results in a degraded performance. The aforementioned FG insertion is intended to improve switching and average the ferroelectric evenly over the entire SB.

The first initial sample was fabricated employing a lift-off process for the TiN FG and TG. Optical microscopy images and atomic force microscopy (AFM) data reveal an unsatisfying layer patterning. Due to the high hardness of TiN and the partial deposition on the photoresist edges during reactive magnetron sputtering, the subsequent sonication in heated acetone can delaminate multiple layers. This creates irregular patterns of left-behind layers that are slightly bent upwards, resulting in heights beyond the deposited layer thicknesses visible in the AFM data, or broken-off pieces featuring a zigzag edge. Electrical measurements on less impacted samples show a shortening of the TG and FG.

To circumvent the lift-off, a wet etching process is developed as part of this thesis. Initial tests with an ammonium hydroxide (NH_4OH), hydrogen peroxide (H_2O_2) and deionized water (H_2O) APM solution, also known as standard clean 1 (SC1) or RCA1, indicate that a 30 min hard bake of the resist on a hot plate at 393 K is required to prevent lifting of the resist during etching. Furthermore, it was found that an APM ratio of (1:2:10) is

optimal. An investigation of the TiN etch rate containing six samples reveals roughly 22 nm/min. Additionally, it is believed that because of the long time needed to pattern structures using RCA1, in which the TiN layer is exposed to air during handling, oxides form on its surface, which prevents the ferroelectric behavior of HZO if deposited on top, necessitating the introduction of a reverse sputter process step. AFM and transmission electron microscopy (TEM) show an excellent integration of the FG layer into the gate stack.

Test structures placed on the sample are used to validate and characterize different aspects of the fabricated device. CV pads are employed to investigate the ferroelectric phase of the HZO, and transistors in which the FG is connectable to the semiconductor analyzer show a working SBFET. Electrical measurements on the complete FG STG FeSBFETs reveal a shift of the transfer characteristic by ± 1.4 V depending on the sweeping direction. A simple mathematical model derived from the one describing non-ferroelectric FGFETs used in flash storage devices allows the estimation of the effective charges in the FG layer. These charges, which determine the device characteristics, are the result of the balance between the charges coming from the ferroelectric polarization and those injected into the layer. Subsequently, pulsed measurements with pulse times ranging between microseconds and hundreds of milliseconds demonstrate that after reaching a certain charge in the FG, the injection changes abruptly. This charge is identical to the one associated with the remnant polarization of the HZO, indicating that, when the ferroelectric switches, it alters the band structure, reducing the injection. The pulsed measurements allow the extraction of long-term potentiation (LTP) and long-term depression (LTD) curves, which display the potential of these devices to be used in neuromorphic computing.

A balance between the injected charges and the ferroelectric polarization charges, relative to the electrode area, is reached for a pulse length of 6 ms, compared to other devices currently under investigation. This value is two to three orders of magnitude higher compared to other similar neuromorphic devices investigated in the literature and is caused by the currently selected surface ratio of the two gate capacitors. This imbalance is the reason why a high amount of charges needs to be injected to reach the required surface charge. Future samples could first address this issue by investigating a variety of different FG sizes, which should be in the range of the SiO_2 capacitor. The discussed imbalance additionally affects the gate coupling coefficient α_G, which is around 1 for the produced samples. Varying the ratio between the TG and FG could be investigated as a practical method to change α_G. Furthermore, by choosing a smaller TG, the amount of polarized ferroelectric is reduced, which alters the needed amount of charge to be injected, reducing the pulse time and achievable voltage shift. All these aspects can be utilized to fine-tune this device concept.

The setup used for pulsing in the microsecond regime should be employed to conduct further measurements towards neuromorphic computing applications. LTP and LTD have been recorded by conducting a transfer characteristic after pulsing and extracting the

corresponding top gate voltage V_{TG} for individual pulse times or pulse numbers. In the neuromorphic community, LTP and LTD are recorded using a different method. These more mature measurements could then be utilized to simulate entire neural networks and would allow comparison of this approach to a broader variety of neuromorphic devices.

Finally, the FG metal layer can be introduced to other SB-based devices like the reconfigurable FET (RFET), where it could be used to program the hole or electron conduction mode in a non-volatile manner. This could reduce the required on-chip infrastructure necessary to facilitate RFET circuits that normally require a static voltage to be applied to the polarity gates. Furthermore, the introduction of HZO and a floating gate layer could be exploited to reduce the total number of terminals from four to three in the RFET architecture.

List of Figures

Bibliography

[1] A. M. Turing, "Computing Machinery and Intelligence," *Mind, New Series*, vol. 59, no. 236, pp. 433–460, 1950.

[2] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.

[3] J. von Neumann, "Theory of self-reproducing automata," *University of Illinois Press*, 1966.

[4] "Executive summary - Electricity 2024 - Analysis." URL: https://www.iea.org/reports/electricity-2024/executive-summary

[5] D. Attwell and S. B. Laughlin, "An Energy Budget for Signaling in the Grey Matter of the Brain," *Journal of Cerebral Blood Flow & Metabolism*, vol. 21, no. 10, pp. 1133–1145, Oct. 2001.

[6] E. Strubell, A. Ganesh, and A. McCallum, "Energy and Policy Considerations for Modern Deep Learning Research," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 09, pp. 13 693–13 696, Apr. 2020.

[7] D. Markovic, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing," *Nature Reviews Physics*, vol. 2, no. 9, pp. 499–510, Sep. 2020.

[8] F. Akopyan, J. Sawada, A. S. Cassidy, R. Alvarez-Icaza, J. V. Arthur, P. Merolla, N. Imam, Y. Nakamura, P. Datta, G. J. Nam, B. Taba, M. P. Beakes, B. Brezzo, J. B. Kuang, R. Manohar, W. P. Risk, B. L. Jackson, and D. S. Modha, "TrueNorth: Design and Tool Flow of a 65 mW 1 Million Neuron Programmable Neurosynaptic Chip," *IEEE TCADIS*, Oct. 2015.

[9] J. Sawada, F. Akopyan, A. S. Cassidy, B. Taba, M. V. Debole, P. Datta, R. Alvarez-Icaza, A. Amir, J. V. Arthur, A. Andreopoulos, R. Appuswamy, H. Baier, D. Barch, D. J. Berg, C. Di Nolfo, S. K. Esser, M. Flickner, T. A. Horvath, B. L. Jackson, J. Kusnitz *et al.*, "TrueNorth Ecosystem for Brain-Inspired Computing: Scalable Systems, Software, and Applications," in *SC '16: Proceedings*

of the International Conference for High Performance Computing, Networking, Storage and Analysis, Nov. 2016, pp. 130–141.

[10] D. S. Modha, F. Akopyan, A. Andreopoulos, R. Appuswamy, J. V. Arthur, A. S. Cassidy, P. Datta, M. V. DeBole, S. K. Esser, C. O. Otero, J. Sawada, B. Taba, A. Amir, D. Bablani, P. J. Carlson, M. D. Flickner, R. Gandhasri, G. J. Garreau, M. Ito, J. L. Klamo *et al.*, "Neural inference at the frontier of energy, space, and time," *Science*, vol. 382, no. 6668, pp. 329–335, Oct. 2023.

[11] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C.-K. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y.-H. Weng *et al.*, "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018.

[12] "Neuromorphic Computing and Engineering with AI | Intel®." URL: https://www.intel.com/content/www/us/en/research/neuromorphic-computing.html

[13] J. Tang, F. Yuan, X. Shen, Z. Wang, M. Rao, Y. He, Y. Sun, X. Li, W. Zhang, Y. Li, B. Gao, H. Qian, G. Bi, S. Song, J. J. Yang, and H. Wu, "Bridging Biological and Artificial Neural Networks with Emerging Neuromorphic Devices: Fundamentals, Progress, and Challenges," *Advanced Materials*, vol. 31, no. 49, p. 1902761, 2019.

[14] S. E. Thompson and S. Parthasarathy, "Moore's law: the future of Si microelectronics," *Materials Today*, vol. 9, no. 6, pp. 20–25, Jun. 2006.

[15] T. Mikolajick, V. Häublein, and H. Ryssel, "The effect of random dopant fluctuations on the minimum channel length of short-channel MOS transistors," *Applied Physics A*, vol. 64, no. 6, pp. 555–560, Jun. 1997.

[16] A. Wolfgang, B. Michel, C. Patrick, G. Mart, H. Bert, and M. Reinhard, "More Than Moore White Paper," in *ITRS*, 2010.

[17] H. Lin, K. Yeh, R. Huang, C. Lin, and T. Huang, "Schottky barrier thin-film transistor (SBTFT) with silicided source/drain and field-induced drain extension," *IEEE Electron Device Letters*, vol. 22, no. 4, pp. 179–181, Apr. 2001.

[18] W. M. Weber, L. Geelhaar, A. P. Graham, E. Unger, G. S. Duesberg, M. Liebau, W. Pamler, C. Chèze, H. Riechert, P. Lugli, and F. Kreupl, "Silicon-Nanowire Transistors with Intruded Nickel-Silicide Contacts," *Nano Letters*, vol. 6, no. 12, pp. 2660–2666, Dec. 2006.

[19] S.-M. Koo, Q. Li, M. D. Edelstein, C. A. Richter, and E. M. Vogel, "Enhanced Channel Modulation in Dual-Gated Silicon Nanowire Transistors," *Nano Letters*, vol. 5, no. 12, pp. 2519–2523, Dec. 2005.

[20] A. Heinzig, T. Mikolajick, J. Trommer, D. Grimm, and W. M. Weber, "Dually

Active Silicon Nanowire Transistors and Circuits with Equal Electron and Hole Transport," *Nano Letters*, vol. 13, no. 9, pp. 4176–4181, Sep. 2013.

[21] T. Krauss, F. Wessely, and U. Schwalke, "Electrostatically Doped Planar Field-Effect Transistor for High Temperature Applications," *ECS Journal of Solid State Science and Technology*, vol. 4, no. 5, p. Q46, Apr. 2015.

[22] J. Trommer, A. Heinzig, U. Mühle, M. Löffler, A. Winzer, P. M. Jordan, J. Beister, T. Baldauf, M. Geidel, B. Adolphi, E. Zschech, T. Mikolajick, and W. M. Weber, "Enabling Energy Efficiency and Polarity Control in Germanium Nanowire Transistors by Individually Gated Nanojunctions," *ACS Nano*, vol. 11, no. 2, pp. 1704–1711, Feb. 2017.

[23] W. M. Weber, L. Wind, A. Fuchsberger, R. Behrle, D. Nazari, J. Aberl, E. P. Navarrete, M. Brehm, and M. Sistani, "Ultrathin Germanium and Silicon-Germanium Nanosheet Transistors for Runtime Reconfigurable Electronics," in *2023 IEEE Nanotechnology Materials and Devices Conference (NMDC)*, Oct. 2023, pp. 446–447.

[24] R. Böckle, M. Sistani, M. Bažíková, L. Wind, Z. Sadre-Momtaz, M. I. den Hertog, C. G. E. Murphey, J. F. Cahoon, and W. M. Weber, "Reconfigurable Complementary and Combinational Logic Based on Monolithic and Single-Crystalline Al-Si Heterostructures," *Advanced Electronic Materials*, vol. 9, no. 1, p. 2200567, 2023.

[25] L. Wind, A. Fuchsberger, O. Demirkiran, L. Vogl, P. Schweizer, X. Maeder, M. Sistani, and W. M. Weber, "Reconfigurable Si Field-Effect Transistors With Symmetric On-States Enabling Adaptive Complementary and Combinational Logic," *IEEE Transactions on Electron Devices*, vol. 71, no. 2, pp. 1302–1307, Feb. 2024.

[26] A. Fuchsberger, L. Wind, D. Nazzari, L. Kühberger, D. Popp, J. Aberl, E. P. Navarrete, M. Brehm, L. Vogl, P. Schweizer, S. Lellig, X. Maeder, M. Sistani, and W. M. Weber, "A Run-Time Reconfigurable Ge Field-Effect Transistor With Symmetric On-States," *IEEE Journal of the Electron Devices Society*, vol. 12, pp. 83–87, 2024.

[27] E. Bestelink, G. Galderisi, P. Golec, Y. Han, B. Iniguez, A. Kloes, J. Knoch, H. Matsui, T. Mikolajick, K. M. Niang, B. Richstein, M. Schwarz, M. Sistani, R. A. Sporea, J. Trommer, W. M. Weber, Q.-T. Zhao, and L. E. Calvet, "Roadmap for Schottky barrier transistors," *Nano Futures*, vol. 8, no. 4, p. 042001, Dec. 2024.

[28] L. Wind, S. Preiß, D. Nazzari, J. Aberl, E. P. Navarrete, M. Brehm, L. Vogl, A. M. Minor, M. Sistani, and W. M. Weber, "Si/Ge1-xSnx/Si transistors with highly transparent Al contacts," *Solid-State Electronics*, vol. 225, p. 109069, Apr. 2025.

[29] A. Fuchsberger, A. Verdianu, L. Wind, D. Nazzari, E. P. Navarrete, C. Wilfingseder, J. Aberl, M. Brehm, J.-M. Hartmann, M. Sistani, and W. M. Weber, "Electrostatic

Gating in Ge-Based Reconfigurable Field-Effect Transistors," *IEEE Transactions on Electron Devices*, pp. 1–6, 2025.

[30] J. Trommer, A. Heinzig, S. Slesazeck, T. Mikolajick, and W. M. Weber, "Elementary Aspects for Circuit Implementation of Reconfigurable Nanowire Transistors," *IEEE Electron Device Letters*, vol. 35, no. 1, pp. 141–143, Jan. 2014.

[31] L. Wind, M. Maierhofer, A. Fuchsberger, M. Sistani, and W. M. Weber, "Realization of a Complementary Full Adder Based on Reconfigurable Transistors," *IEEE Electron Device Letters*, vol. 45, no. 4, pp. 724–727, Apr. 2024.

[32] T. Mikolajick, G. Galderisi, M. Simon, S. Rai, A. Kumar, A. Heinzig, W. M. Weber, and J. Trommer, "20 Years of reconfigurable field-effect transistors: From concepts to future applications," *Solid-State Electronics*, vol. 186, p. 108036, Dec. 2021.

[33] F. Xi, A. Grenmy, J. Zhang, Y. Han, J. H. Bae, D. Grützmacher, and Q.-T. Zhao, "Ferroelectric Schottky Barrier MOSFET as Analog Synapses for Neuromorphic Computing," in *ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC)*, Sep. 2022, pp. 121–124.

[34] D. Nazzari, L. Wind, M. Sistani, D. Mayr, K. Kim, V. Wahler, and W. M. Weber, "Nonvolatile Reconfigurable Transistor via Ferroelectrically Induced Current Modulation," *ACS Applied Materials & Interfaces*, vol. 17, no. 7, pp. 10 784–10 791, Feb. 2025.

[35] S. Lee, G. Kim, Y. Lee, H. Shin, Y. Jeong, L. Zhang, S.-O. Jung, and S. Jeon, "Design Guidelines of Hafnia Ferroelectrics and Gate-Stack for Multilevel-Cell FeFET," *IEEE Transactions on Electron Devices*, vol. 71, no. 3, pp. 1865–1871, Mar. 2024.

[36] M. Dingman, "Neuroscientifically Challenged Neuroscience made simpler." URL: https://neuroscientificallychallenged.com/

[37] F. A. C. Azevedo, L. R. B. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. L. Ferretti, R. E. P. Leite, W. Jacob Filho, R. Lent, and S. Herculano-Houzel, "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain," *The Journal of Comparative Neurology*, vol. 513, no. 5, pp. 532–541, Apr. 2009.

[38] S. M. Breedlove and N. V. Watson, *Behavioral Neuroscience*, 8th ed. Sunderland, Massachusetts: Sinauer Associates is an imprint of Oxford University Press, Sep. 2016.

[39] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *Principles of Neural Science*, 5th ed. New York: McGraw-Hill Education Ltd, Oct. 2012.

[40] J. Zhang, "Basic Neural Units of the Brain: Neurons, Synapses and Action Potential," *Cornell University*, May 2019.

[41] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, R. D. Mooney, M. L. Platt, and L. E. White, Eds., *Neuroscience*, 6th ed. New York Oxford: Sinauer Associates is an imprint of Oxford University Press, Oct. 2017.

[42] K. Yamazaki, V.-K. Vo-Ho, D. Bulsara, and N. Le, "Spiking Neural Networks and Their Applications: A Review," *Brain Sciences*, vol. 12, no. 7, p. 863, Jul. 2022.

[43] M. Dampfhoffer, T. Mesquida, A. Valentian, and L. Anghel, "Are SNNs Really More Energy-Efficient Than ANNs? an In-Depth Hardware-Aware Study," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 7, no. 3, pp. 731–741, Jun. 2023.

[44] S. Negi, D. Sharma, A. K. Kosta, and K. Roy, "Best of Both Worlds: Hybrid SNN-ANN Architecture for Event-based Optical Flow Estimation," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2024, pp. 2696–2703.

[45] P. Wijesinghe, A. Ankit, A. Sengupta, and K. Roy, "An All-Memristor Deep Spiking Neural Computing System: A Step Toward Realizing the Low-Power Stochastic Brain," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 5, pp. 345–358, Oct. 2018.

[46] Y. Wang, D. Wu, Y. Wang, X. Hu, Z. Ma, J. Feng, and Y. Xie, "A Low-Cost Hardware-Friendly Spiking Neural Network Based on Binary MRAM Synapses, Accelerated Using In-Memory Computing," *Electronics*, vol. 10, no. 19, p. 2441, Jan. 2021.

[47] H.-S. P. Wong, S. Raoux, S. Kim, J. Liang, J. P. Reifenberg, B. Rajendran, M. Asheghi, and K. E. Goodson, "Phase Change Memory," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2201–2227, Dec. 2010.

[48] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, "Metal-Oxide RRAM," *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, Jun. 2012.

[49] S. Yu and H.-S. P. Wong, "Compact Modeling of Conducting-Bridge Random-Access Memory (CBRAM)," *IEEE Transactions on Electron Devices*, vol. 58, no. 5, pp. 1352–1360, May 2011.

[50] V. Garcia and M. Bibes, "Ferroelectric tunnel junctions for information storage and processing," *Nature Communications*, vol. 5, no. 1, p. 4289, Jul. 2014.

[51] F. M. Bayat, M. Prezioso, B. Chakrabarti, H. Nili, I. Kataeva, and D. Strukov, "Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits," *Nature Communications*, vol. 9, no. 1, p. 2331, Jun. 2018.

[52] P. Narayanan, A. Fumarola, L. L. Sanches, K. Hosokawa, S. C. Lewis, R. M. Shelby, and G. W. Burr, "Toward on-chip acceleration of the backpropagation algorithm using nonvolatile memory," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 11:1–11:11, Jul. 2017.

[53] J. Li, H. Abbas, D. Shenp Ang, A. Ali, and X. Ju, "Emerging memristive artificial neuron and synapse devices for the neuromorphic electronics era," *Nanoscale Horizons*, vol. 8, no. 11, pp. 1456–1484, 2023.

[54] M. Le Gallo and A. Sebastian, "An overview of phase-change memory device physics," *Journal of Physics D: Applied Physics*, vol. 53, no. 21, p. 213002, Mar. 2020.

[55] S. Yu, W. Shim, X. Peng, and Y. Luo, "RRAM for Compute-in-Memory: From Inference to Training," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 7, pp. 2753–2765, Jul. 2021.

[56] M. Kund, G. Beitel, C.-U. Pinnow, T. Rohr, J. Schumann, R. Symanczyk, K. Ufert, and G. Muller, "Conductive bridging RAM (CBRAM): an emerging non-volatile memory technology scalable to sub 20nm," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, Dec. 2005, pp. 754–757.

[57] M. Y. Zhuravlev, R. F. Sabirianov, S. S. Jaswal, and E. Y. Tsymbal, "Giant Electroresistance in Ferroelectric Tunnel Junctions," *Physical Review Letters*, vol. 94, no. 24, p. 246802, Jun. 2005.

[58] L. Wang, W. Liao, S. L. Wong, Z. G. Yu, S. Li, Y.-F. Lim, X. Feng, W. C. Tan, X. Huang, L. Chen, L. Liu, J. Chen, X. Gong, C. Zhu, X. Liu, Y.-W. Zhang, D. Chi, and K.-W. Ang, "Artificial Synapses Based on Multiterminal Memtransistors for Neuromorphic Application," *Advanced Functional Materials*, vol. 29, no. 25, p. 1901106, 2019.

[59] S. Kunwar, Z. Jernigan, Z. Hughes, C. Somodi, M. D. Saccone, F. Caravelli, P. Roy, D. Zhang, H. Wang, Q. Jia, J. L. MacManus-Driscoll, G. Kenyon, A. Sornborger, W. Nie, and A. Chen, "An Interface-Type Memristive Device for Artificial Synapse and Neuromorphic Computing," *Advanced Intelligent Systems*, vol. 5, no. 8, p. 2300035, 2023.

[60] F. Jona and G. Shirane, *Ferroelectric crystals*, ser. International series of monographs on solid state physics. Oxford [u.a.]: Pergamon Pr., 1962.

[61] E. Schrödinger, "Studien über Kinetik der Dielektrika, den Schmelzpunkt, Pyro- und Piezoelektrizität," *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-naturwissenschaftliche Klasse 2a*, vol. 121, pp. 1937–1972, 1912.

[62] A. F. Devonshire, "Theory of ferroelectrics," *Advances in Physics*, Apr. 1954.

[63] L. D. Landau, J. S. Bell, M. J. Kearsley, L. P. Pitaevskii, E. M. Lifshitz, and J. B. Sykes, *Electrodynamics of Continuous Media*. Elsevier, 1960.

[64] J. Grindlay, *An Introduction to the Phenomenological Theory of Ferroelectricity: International Series of Monographs In: Natural Philosophy*. Elsevier, 1970.

[65] U. Schroeder, C. S. Hwang, and H. Funakubo, *Ferroelectricity in Doped Hafnium Oxide: Materials, Properties and Devices*. Woodhead Publishing, Mar. 2019.

[66] R. Materlik, C. Künneth, and A. Kersch, "The origin of ferroelectricity in Hf1-xZrxO2: A computational investigation and a surface energy model," *Journal of Applied Physics*, vol. 117, no. 13, p. 134109, Apr. 2015.

[67] L. Wind, "Si to SiGe Heterostructure Transistors for Reconfigurable Circuits and Functionality Enhancement via Ferroelectric Gating," PhD thesis, Technische Universität Wien, Feb. 2025.

[68] J. Valasek, "Dielectric Anomalies in Rochelle Salt Crystals," *Physical Review*, vol. 24, no. 5, pp. 560–568, Nov. 1924.

[69] A. von Hippel, "Ferroelectricity, Domain Structure, and Phase Transitions of Barium Titanate," *Reviews of Modern Physics*, vol. 22, no. 3, pp. 221–237, Jul. 1950.

[70] C.-U. Pinnow and T. Mikolajick, "Material Aspects in Emerging Nonvolatile Memories," *Journal of The Electrochemical Society*, vol. 151, no. 6, p. K13, May 2004.

[71] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder, and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Applied Physics Letters*, vol. 99, no. 10, p. 102903, Sep. 2011.

[72] J. Müller, T. S. Böscke, U. Schröder, S. Mueller, D. Bräuhaus, U. Böttger, L. Frey, and T. Mikolajick, "Ferroelectricity in Simple Binary ZrO2 and HfO2," *Nano Letters*, vol. 12, no. 8, pp. 4318–4323, Aug. 2012.

[73] T. Mikolajick, M. H. Park, L. Begon-Lours, and S. Slesazeck, "From Ferroelectric Material Optimization to Neuromorphic Devices," *Advanced Materials*, vol. 35, no. 37, p. 2206042, 2023.

[74] M. H. Park, Y. H. Lee, H. J. Kim, Y. J. Kim, T. Moon, K. D. Kim, S. D. Hyun, T. Mikolajick, U. Schroeder, and C. S. Hwang, "Understanding the formation of the metastable ferroelectric phase in hafnia-zirconia solid solution thin films," *Nanoscale*, vol. 10, no. 2, pp. 716–725, Jan. 2018.

[75] X. Sang, E. D. Grimley, T. Schenk, U. Schroeder, and J. M. LeBeau, "On the structural origins of ferroelectricity in HfO2 thin films," *Applied Physics Letters*, vol. 106, no. 16, p. 162905, Apr. 2015.

[76] O. Ohtaka, H. Fukui, T. Kunisada, T. Fujisawa, K. Funakoshi, W. Utsumi, T. Irifune, K. Kuroda, and T. Kikegawa, "Phase relations and equations of state of ZrO2 under high temperature and high pressure," *Physical Review B*, vol. 63, no. 17, p. 174108, Apr. 2001.

[77] H. J. Kim, M. H. Park, Y. J. Kim, Y. H. Lee, W. Jeon, T. Gwon, T. Moon, K. D. Kim, and C. S. Hwang, "Grain size engineering for ferroelectric Hf0.5Zr0.5O2 films by an insertion of Al2O3 interlayer," *Applied Physics Letters*, vol. 105, no. 19, p. 192903, Nov. 2014.

[78] H. Joh, T. Jung, and S. Jeon, "Stress Engineering as a Strategy to Achieve High Ferroelectricity in Thick Hafnia Using Interlayer," *IEEE Transactions on Electron Devices*, vol. 68, no. 5, pp. 2538–2542, May 2021.

[79] D. Buck, "Ferroelectrics for Digital Information Storage and Switching," PhD thesis, MIT, Jun. 1952.

[80] D. Bondurant, "Ferroelectronic ram memory family for critical data storage," *Ferroelectrics*, vol. 112, no. 1, pp. 273–282, Dec. 1990.

[81] N. Ramaswamy, A. Calderoni, J. Zahurak, G. Servalli, A. Chavan, S. Chhajed, M. Balakrishnan, M. Fischer, M. Hollander, D. P. Ettisserry, A. Liao, K. Karda, M. Jerry, M. Mariani, A. Visconti, B. R. Cook, B. D. Cook, D. Mills, A. Torsi, C. Mouli *et al.*, "NVDRAM: A 32Gb Dual Layer 3D Stacked Non-volatile Ferroelectric Memory with Near-DRAM Performance for Demanding AI Workloads," in *2023 International Electron Devices Meeting (IEDM)*, Dec. 2023, pp. 1–4.

[82] B. J. Baliga, *Fundamentals of Power Semiconductor Devices*. Springer, Sep. 2018.

[83] S. L. Miller and P. J. McWhorter, "Physics of the ferroelectric nonvolatile memory field effect transistor," *Journal of Applied Physics*, vol. 72, no. 12, pp. 5999–6010, Dec. 1992.

[84] H.-T. Lue, C.-J. Wu, and T.-Y. Tseng, "Device modeling of ferroelectric memory field-effect transistor for the application of ferroelectric random access memory," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, no. 1, pp. 5–14, Jan. 2003.

[85] H. Mulaosmanovic, E. T. Breyer, S. Dünkel, S. Beyer, T. Mikolajick, and S. Slesazeck, "Ferroelectric field-effect transistors based on HfO2: a review," *Nanotechnology*, vol. 32, no. 50, p. 502002, Sep. 2021.

[86] H. W. Park, S. D. Hyun, I. S. Lee, S. H. Lee, Y. B. Lee, M. Oh, B. Y. Kim, S. G. Ryoo, and C. S. Hwang, "Polarizing and depolarizing charge injection through a thin dielectric layer in a ferroelectric-dielectric bilayer," *Nanoscale*, vol. 13, no. 4, pp. 2556–2572, Feb. 2021.

[87] N. T. Kimbugwe and E. Yilmaz, "Impact of SiO2 interfacial layer on the electrical

characteristics of Al/Al2O3/SiO2/n-Si metal-oxide-semiconductor capacitors," *Journal of Materials Science: Materials in Electronics*, vol. 31, no. 15, pp. 12 372–12 381, Aug. 2020.

[88] J. Lee, K. Yang, J. Y. Kwon, J. E. Kim, D. I. Han, D. H. Lee, J. H. Yoon, and M. H. Park, "Role of oxygen vacancies in ferroelectric or resistive switching hafnium oxide," *Nano Convergence*, vol. 10, no. 1, p. 55, Dec. 2023.

[89] K. Toprasertpong, M. Takenaka, and S. Takagi, "Direct Observation of Interface Charge Behaviors in FeFET by Quasi-Static Split C-V and Hall Techniques: Revealing FeFET Operation," in *2019 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2019, pp. 23.7.1–23.7.4.

[90] D. Lizzit and D. Esseni, "Operation and Design of Ferroelectric FETs for a BEOL Compatible Device Implementation," in *ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC)*, Sep. 2021, pp. 215–218.

[91] E. Yurchuk, J. Müller, S. Müller, J. Paul, M. Pesic, R. van Bentum, U. Schroeder, and T. Mikolajick, "Charge-Trapping Phenomena in HfO2-Based FeFET-Type Nonvolatile Memories," *IEEE Transactions on Electron Devices*, vol. 63, no. 9, pp. 3501–3507, Sep. 2016.

[92] S. Lancaster, P. D. Lomenzo, M. Engl, B. Xu, T. Mikolajick, U. Schroeder, and S. Slesazeck, "Investigating charge trapping in ferroelectric thin films through transient measurements," *Frontiers in Nanotechnology*, vol. 4, Aug. 2022.

[93] K. Toprasertpong, M. Takenaka, and S. Takagi, "On the strong coupling of polarization and charge trapping in HfO2/Si-based ferroelectric field-effect transistors: overview of device operation and reliability," *Applied Physics A*, vol. 128, no. 12, p. 1114, Nov. 2022.

[94] N. K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, "Emerging Memory Devices for Neuromorphic Computing," *Advanced Materials Technologies*, vol. 4, no. 4, p. 1800589, 2019.

[95] M. Jerry, P.-Y. Chen, J. Zhang, P. Sharma, K. Ni, S. Yu, and S. Datta, "Ferroelectric FET analog synapse for acceleration of deep neural network training," in *2017 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, pp. 6.2.1–6.2.4.

[96] H. Mulaosmanovic, E. Chicca, M. Bertele, T. Mikolajick, and S. Slesazeck, "Mimicking biological neurons with a nanoscale ferroelectric transistor," *Nanoscale*, vol. 10, no. 46, pp. 21 755–21 763, 2018.

[97] H. Mulaosmanovic, J. Ocker, S. Müller, M. Noack, J. Müller, P. Polakowski, T. Mikolajick, and S. Slesazeck, "Novel ferroelectric FET based synapse for neuromorphic systems," in *2017 Symposium on VLSI Technology*, Jun. 2017, pp. T176–T177.

[98] D. Kahng and S. M. Sze, "A floating gate and its application to memory devices," *The Bell System Technical Journal*, vol. 46, no. 6, pp. 1288–1295, Jul. 1967.

[99] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. John Wiley & Sons, Dec. 2006.

[100] P. Cottrell, R. Troutman, and T. Ning, "Hot-electron emission in N-channel IGFET's," *IEEE Transactions on Electron Devices*, vol. 26, no. 4, pp. 520–533, Apr. 1979.

[101] B. Eitan and D. Frohman-Bentchkowsky, "Hot-electron injection into the oxide in n-channel MOS devices," *IEEE Transactions on Electron Devices*, vol. 28, no. 3, pp. 328–340, Mar. 1981.

[102] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells-an overview," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1248–1271, Aug. 1997.

[103] E. Takeda and N. Suzuki, "An empirical model for device degradation due to hot-carrier injection," *IEEE Electron Device Letters*, vol. 4, no. 4, pp. 111–113, Apr. 1983.

[104] T. Tsuchiya, T. Kobayashi, and S. Nakajima, "Hot-carrier-injected oxide region and hot-electron trapping as the main cause in Si nMOSFET degradation," *IEEE Transactions on Electron Devices*, vol. 34, no. 2, pp. 386–391, Feb. 1987.

[105] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim Tunneling into Thermally Grown SiO2," *Journal of Applied Physics*, vol. 40, no. 1, pp. 278–283, Jan. 1969.

[106] S. Shin, Y. Jang, J. Park, and S.-Y. Lee, "P-20: Memory Window and Endurance Improvement of In-Ga-Zn-O-Based Ferroelectric Thin Film Transistors by Inserting In-Ga-Zn-O Floating Gate," *SID Symposium Digest of Technical Papers*, vol. 55, no. 1, pp. 1434–1437, 2024.

[107] D. Kwon, E. C. Park, W. Shin, R.-H. Koo, J. Hwang, J.-H. Bae, D. Kwon, and J.-H. Lee, "Analog Synaptic Devices Based on IGZO Thin-Film Transistors with a Metal–Ferroelectric–Metal–Insulator–Semiconductor Structure for High-Performance Neuromorphic Systems," *Advanced Intelligent Systems*, vol. 5, no. 12, p. 2300125, 2023.

[108] T. Ali, K. Seidel, K. Kühnel, M. Rudolph, M. Czernohorsky, K. Mertens, R. Hoffmann, K. Zimmermann, U. Mühle, J. Müller, J. Van Houdt, and L. M. Eng, "A Novel Dual Ferroelectric Layer Based MFMFIS FeFET with Optimal Stack Tuning toward Low Power and High-Speed NVM for Neuromorphic Applications," in *2020 IEEE Symposium on VLSI Technology*, Jun. 2020, pp. 1–2.

[109] T. Ali, K. Mertens, K. Kühnel, M. Rudolph, S. Oehler, D. Lehninger, F. Müller, R. Revello, R. Hoffmann, K. Zimmermann, T. Kämpfe, M. Czernohorsky, K. Seidel, J. V. Houdt, and L. M. Eng, "A FeFET with a novel MFMFIS gate stack: towards

energy-efficient and ultrafast NVMs for neuromorphic computing," *Nanotechnology*, vol. 32, no. 42, p. 425201, Jul. 2021.

[110] Z. Zheng, C. Sun, L. Jiao, D. Zhang, Z. Zhou, X. Wang, G. Liu, Q. Kong, Y. Chen, K. Ni, and X. Gong, "Boosting the Memory Window of the BEOL-Compatible MFMIS Ferroelectric/ Anti-Ferroelectric FETs by Charge Injection," in *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, Jun. 2022, pp. 389–390.

[111] J. Hwang, G. Kim, H. Joh, J. Ahn, and S. Jeon, "The Effect of Ferroelectric/Dielectric Capacitance Ratio on Short-Term Retention Characteristics of MFMIS FeFET," *IEEE Journal of the Electron Devices Society*, vol. 12, pp. 988–992, 2024.

[112] S. Lee, Y. Lee, G. Kim, T. Kim, T. Eom, S.-O. Jung, and S. Jeon, "Effect of Floating Gate Insertion on the Analog States of Ferroelectric Field-Effect Transistors," *IEEE Transactions on Electron Devices*, vol. 70, no. 1, pp. 349–353, Jan. 2023.

[113] T. Ali, K. Mertens, R. Olivo, D. Lehninger, M. Lederer, F. Müller, M. Rudolph, S. Oehler, K. Kühnel, R. Hoffmann, P. Schramm, M. Czernohorsky, T. Kämpfe, and K. Seidel, "Impact of Stack Structure Control and Ferroelectric Material Optimization in Novel Laminate HSO and HZO MFMIS FeFET," in *2022 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)*, Apr. 2022, pp. 1–2.

[114] Q. Han, P. Aleksa, J. Schubert, S. Mantl, and Q.-T. Zhao, "Subthreshold Behavior of MFMIS and MFIS Transistors caused by Ferroelectric Polarization Switching," in *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Oct. 2018, pp. 1–4.

[115] K. Jang, M. Kobayashi, and T. Hiramoto, "Role of gate current and polarization switching in sub-60 mV/decade steep subthreshold slope in metal–ferroelectric HfZrO2–metal–insulator–Si FET," *Japanese Journal of Applied Physics*, vol. 57, no. 11, p. 114202, Oct. 2018.

[116] E. H. Rhoderick and R. H. Williams, *Metal-semiconductor Contacts*. Clarendon Press, 1988.

[117] K.-W. Ang, K. Majumdar, K. Matthews, C. D. Young, C. Kenney, C. Hobbs, P. D. Kirsch, R. Jammy, R. D. Clark, S. Consiglio, K. Tapily, Y. Trickett, G. Nakamura, C. S. Wajda, G. J. Leusink, M. Rodgers, and S. C. Gausepohl, "Effective Schottky Barrier Height modulation using dielectric dipoles for source/drain specific contact resistivity improvement," in *2012 International Electron Devices Meeting*, Dec. 2012, pp. 18.6.1–18.6.4.

[118] S. G. Louie and M. L. Cohen, "Electronic structure of a metal-semiconductor interface," *Physical Review B*, vol. 13, no. 6, pp. 2461–2469, Mar. 1976.

[119] F. A. Padovani and R. Stratton, "Field and thermionic-field emission in Schottky barriers," *Solid-State Electronics*, vol. 9, no. 7, pp. 695–707, Jul. 1966.

[120] C. R. Crowell and V. L. Rideout, "Normalized thermionic-field (T-F) emission in metal-semiconductor (Schottky) barriers," *Solid-State Electronics*, vol. 12, no. 2, pp. 89–105, Feb. 1969.

[121] J. E. Lilienfeld, "Method and apparatus for controlling electric currents," *United States Patent*, Jan. 1930.

[122] W. M. Weber and T. Mikolajick, "Silicon and germanium nanowire electronics: physics of conventional and unconventional transistors," *Reports on Progress in Physics*, vol. 80, no. 6, p. 066502, Apr. 2017.

[123] T. Solankia and N. Parmar, "A Review paper: A Comprehensive study of Junctionless transistor," *National Conference on Recent Trends in Engineering and Technology*, 2011.

[124] P.-E. Gaillardon, L. Amaru, J. Zhang, and G. De Micheli, "Advanced system on a chip design based on controllable-polarity FETs," in *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Mar. 2014, pp. 1–6.

[125] P.-E. Gaillardon, X. Tang, G. Kim, and G. De Micheli, "A Novel FPGA Architecture Based on Ultrafine Grain Reconfigurable Logic Cells," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 10, pp. 2187–2197, Oct. 2015.

[126] L. Amarú, P.-E. Gaillardon, and G. De Micheli, "A Circuit Synthesis Flow for Controllable-Polarity Transistors," *IEEE Transactions on Nanotechnology*, vol. 13, no. 6, pp. 1074–1083, Nov. 2014.

[127] R. Gauchi, A. Snelgrove, and P.-E. Gaillardon, "An Open-source Three-Independent-Gate FET Standard Cell Library for Mixed Logic Synthesis," in *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2022, pp. 273–277.

[128] "MLA 150." URL: https://heidelberg-instruments.com/product/mla150/

[129] L. Lü, J. Y. H. Fuh, and Y. S. Wong, *Laser-Induced Materials and Processes for Rapid Prototyping*. Boston, MA: Springer US, 2001.

[130] K. R. Ryan and I. C. Plumb, "A model for the etching of silicon in SF6/O2 plasmas," *Plasma Chemistry and Plasma Processing*, vol. 10, no. 2, pp. 207–229, Jun. 1990.

[131] "PlasmaPro 100 Cobra ICP Etching System." URL: https://plasma.oxinst.com/products/icp-etching/plasmapro-100-cobra-icp

[132] "V10-G. Low-pressure plasma system. Removal of photoresists." URL: https://www.pink.de/en/products/low-pressure-plasma-systems-en/v10-g.html

[133] "Furnace ATV PEO-601." URL: https://www.epfl.ch/research/facilities/cmi/equipment/packaging-miscellaneous/atv-peo601/

[134] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," *Journal of Applied Physics*, vol. 36, no. 12, pp. 3770–3778, Dec. 1965.

[135] J. Knoch, *Nanoelectronics: From Device Physics and Fabrication Technology to Advanced Transistor Concepts*. De Gruyter, Mar. 2024.

[136] H. Jin, K. Weber, and A. Blakers, "The Effect of a Post Oxidation In-Situ Nitrogen Anneal on si Surface Passivation," in *2006 IEEE 4th World Conference on Photovoltaic Energy Conference*, vol. 1, May 2006, pp. 1071–1073.

[137] M. Ohring, *Materials Science of Thin Films: Depositon and Structure*. Elsevier, Oct. 2001.

[138] P. J. Kelly and R. D. Arnell, "Magnetron sputtering: a review of recent developments and applications," *Vacuum*, vol. 56, no. 3, pp. 159–172, Mar. 2000.

[139] H. O. Pierson, *Handbook of Refractory Carbides and Nitrides: Properties, Characteristics, Processing and Applications*. William Andrew, Dec. 1996.

[140] "Sputteranlage CREAMET S-Line Single | CREAVAC." URL: https://www.creavac.de/de/3d-modelle/3d-model-sputteranlage-creamet-s-line-single/

[141] "Post Exposure Bake (PEB)." URL: https://www.microchemicals.com/dokumente/application_notes/photoresist_post_exposure_bake_peb.pdf

[142] W. Kern, J. L. Vossen, P. Rappaport, A. I. Stoller, S. Opresko, R. F. Speers, and J. A. Amick, "RCA-Review," *Radio Corporation of America*, vol. 31, Jun. 1970.

[143] National Research Council (US) Committee on Prudent Practices in the Laboratory, *Prudent Practices in the Laboratory: Handling and Management of Chemical Hazards: Updated Version*, ser. The National Academies Collection: Reports funded by National Institutes of Health. Washington (DC): National Academies Press (US), 2011.

[144] J. L. Vossen and W. Kern, *Thin Film Processes*. New York: Academic Press, Jan. 1979.

[145] W. Kern, *Handbook of Semiconductor Wafer Cleaning Technology: Science, Technology, and Applications*. Noyes Publications, Jan. 1993.

[146] "RCA-1 Silicon Wafer Cleaning." URL: https://www.inrf.uci.edu/wordpress/wp-content/uploads/sop-wet-silicon-rca-1.pdf

[147] R. G. Gordon, "Overview of ALD Precursors and Reaction Mechanisms," *Harvard University*, 2011.

[148] W. Chen, Q.-Q. Sun, M. Xu, S.-J. Ding, D. W. Zhang, and L.-K. Wang, "Atomic Layer Deposition of Hafnium Oxide from Tetrakis(ethylmethylamino)hafnium and

Water Precursors," *The Journal of Physical Chemistry C*, vol. 111, no. 17, pp. 6495–6499, May 2007.

[149] "TFS 200: Premium ALD R&D Platform." URL: https://beneq.com/tfs200/

[150] J.-N. Song, M.-J. Oh, and C.-B. Yoon, "Effect of a ZrO2 Seed Layer on an Hf0.5Zr0.5O2 Ferroelectric Device Fabricated via Plasma Enhanced Atomic Layer Deposition," *Materials*, vol. 16, no. 5, p. 1959, Jan. 2023.

[151] T. Onaya, T. Nabatame, N. Sawamoto, A. Ohi, N. Ikeda, T. Nagata, and A. Ogura, "Improvement in ferroelectricity of HfxZr1-xO2 thin films using top- and bottom-ZrO2 nucleation layers," *APL Materials*, vol. 7, no. 6, p. 061107, Jun. 2019.

[152] "UniTemp." URL: https://unitemp.de/

[153] T. E. Volin and R. W. Balluffi, "Annealing kinetics of voids and the Self-diffusion coefficient in aluminum," *physica status solidi (b)*, vol. 25, no. 1, pp. 163–173, 1968.

[154] A. Paccagnella, G. Ottaviani, P. Fabbri, G. Ferla, and G. Queirolo, "Silicon diffusion in aluminium," *Thin Solid Films*, vol. 128, no. 3, pp. 217–223, Jun. 1985.

[155] L. Wind, M. Sistani, Z. Song, X. Maeder, D. Pohl, J. Michler, B. Rellinghaus, W. M. Weber, and A. Lugstein, "Monolithic Metal–Semiconductor–Metal Heterostructures Enabling Next-Generation Germanium Nanodevices," *ACS Applied Materials & Interfaces*, vol. 13, no. 10, pp. 12 393–12 399, Mar. 2021.

[156] M. Sistani, R. Böckle, L. Wind, K. Eysin, X. Maeder, P. Schweizer, J. Michler, A. Lugstein, and W. M. Weber, "Polycrystalline Ge Nanosheets Embedded in Metal-Semiconductor Heterostructures Enabling Wafer-Scale 3D Integration of Ge Nanodevices with Self-Aligned Al Contacts," *Advanced Electronic Materials*, vol. 7, no. 5, p. 2100101, 2021.

[157] L. Wind, "Wafer-scale fabrication and characterization of monolithic Al-Ge heterostructures," Thesis, Technische Universität Wien, 2021, accepted: 2021-02-10T09:37:52Z Journal Abbreviation: Herstellung und Charakterisierung von Al-Ge-Al Heterostrukturen auf Waferebene.

[158] L. Wind, R. Böckle, M. Sistani, P. Schweizer, X. Maeder, J. Michler, C. G. Murphey, J. Cahoon, and W. M. Weber, "Monolithic and Single-Crystalline Aluminum–Silicon Heterostructures," *ACS Applied Materials & Interfaces*, vol. 14, no. 22, pp. 26 238–26 244, Jun. 2022.

[159] D. D. L. Chung, "Pitfalls and Methods in the Measurement of the Electrical Resistance and Capacitance of Materials," *Journal of Electronic Materials*, vol. 50, no. 12, pp. 6567–6574, Dec. 2021.

[160] "C-V Testing for Semiconductor Components and Devices - Applications Guide." URL: https://www.tek.com/en/documents/application-note/c-v-testing-semiconductor-components-and-devices-applications-guide

[161] D. Mayr, "Integration of ferroelectric dielectrics in reconfigurable field-effect transistors," Master's thesis, Technische Universität Wien, Oct. 2023.

[162] "Model 4200A-SCS Pulse Card (PGU and PMU) User's Manual." URL: https://rb.gy/zt072j

[163] A. K. Suri, R. Nimmagadda, and R. F. Bunshah, "Synthesis of titanium nitrides by activated reactive evaporation," *Thin Solid Films*, vol. 72, no. 3, pp. 529–533, Oct. 1980.