Memristor-Based Encryption for Free-Floating Neural Implants

Jan Andrès GALVAN HERNÀNDEZ

MEMRISTOR-BASED ENCRYPTION FOR FREE-FLOATING NEURAL IMPLANTS

Thesis

to obtain the degree of Master of Science
at the Delft University of Technology.

by

Jan Andrès GALVAN HERNÀNDEZ

Master Embedded Systems: Computer Architecture
Student Number: 5193273
Duration: September 1, 2021 - November 4, 2022

Graduation Committee:

Prof. dr. ir. S. Hamdioui	Quantum & Computer Engineering, TU Delft
Prof. dr. ir. W.A Serdijn	Bio-Electronics, TU Delft
Dr. ir. M. Taouil	Quantum & Computer Engineering, TU Delft
Dr. ir. C. Strydis	Quantum & Computer Engineering, TU Delft
Dr. ir. M.A. Siddiqi	Quantum & Computer Engineering, TU Delft
Dr. ir. A. Gebregiorgis	Quantum & Computer Engineering, TU Delft

Keywords: Free-Floating Neural Implant, Neural Dust Attacks, IMD, Security, GIFT, Memristor, Encryption, Cipher, Implants, In-situ, In-memory

An electronic version of this dissertation is available at
http://repository.tudelft.nl/

Success is not final;
failure is not fatal:
It is the courage to continue that counts.

— Winston S. Churchill

Memristor-Based Encryption for Free-Floating Neural Implants

Abstract

The recent advances in the semiconductor industry have given rise to the development of highly scalable, wireless and battery-free neural-implant interfaces that enable brain monitoring and brain stimulation with high spatial and temporal resolution. Such implants are referred to as Free-Floating Neural Implants (FFNI), as the small size and untethered communication allow them to be scattered throughout the cortex. Nevertheless, the plethora of proposed interfaces have failed to mention and act against the potential security implications that may arise in highly-constrained FFNIs even though the U.S. Food and Drug Administration (FDA) has recently acknowledged the possibility of short-/long-range attacks on wireless Implantable Medical Devices (IMD). Hence, in this project, the existing threats in FFNIs are revealed, followed by the proposal of a memristor-based lightweight security approach to secure intracranial electromagnetic transmissions whilst considering the anticipated physical limitations of these constrained topologies. More specifically, a consolidated envisioned system is highlighted for which a read-only GIFT cipher is implemented. This lightweight encryption block primarily consists of a *One-Transistor-One-Memristor* (1T1R) crossbar structure for carrying out operations such as *Substitution*, *Permutation*, and *addRoundKey*, without destroying the resistive states and by only performing ‘read’ operations to maintain low power operation. With a footprint of 0.0034 mm^2 the 1T1R-GIFT cipher reaches an average power and energy consumption of only 60.38 μW and 241.52 pJ, respectively. However, the performance does not exceed a CMOS-based implementation yet, whose footprint is similar but has roughly half the average power and energy consumption. This can be attributed mainly to the immaturity of the memristor technology. This work demonstrates that only after further advancements in memristor logic gates, crossbar topologies and fabrication processes, highly-constrained FFNIs can fully benefit from the scalable memristor-based security paradigm.

Contents

List of Figures

List of Tables

1

Introduction

1.1. Motivation and Problem Statement

For many years now Implantable Medical Devices (IMD) have existed to aid those with physical conditions such as heart disease, hearing problems or diabetes. More recently, implantable neural interfaces have been in development that enable (continuous) brain monitoring or brain stimulation which can effectively treat illnesses within the mental spectrum. These implants can perform Deep Brain Stimulation (DBS) that could be deployed effectively for treating Obsessive-Compulsive Disorder (OCD) [1], [2], depression [3], epilepsy [4], Parkinson's Disease (PD), dystonia and Essential Tremor (ET) [5]. Next to that, the development of Brain Computer Interfaces (BCI) may potentially also assist those that suffer from motor impairment [6].

The continuous advances in the semiconductor industry is slowly allowing the transition from bulky designs towards miniaturization reaching sub-mm footprints. This has given rise to new, highly scalable, untethered and battery-free neural interfaces that are referred to in this thesis as *Free-Floating Neural Implants* (FFNI). An FFNI system consists of a group of ultra-small, battery-free implants, also called *dust motes*, that are scattered throughout the brain cortex to perform monitoring, stimulation, or both. The deployed dust motes send (receive) data to (from) the outside world via a wireless communication protocol. The transmissions between the dust motes and an external interrogator (transceiver), can take place directly, or via a transceiver that is embedded within the body or skull. This results in highly scalable implants that offer a more controllable and larger target area, higher temporal/spatial resolution, and less intrusion of the brain, providing long-term clinical treatment by omitting the need for frequent surgical removal. These improvements also open the door to more advanced levels of data gathering, offering clinicians a better view of the functioning of the brain, which in turn permits better treatment for the patient.

Nevertheless, new technology developments also give rise to new security threats. Especially in the healthcare domain it is important that the technology stands by the principles of the CIA (*Confidentiality, Integrity, Availability*) triad. Even so, healthcare and government organizations have been very naive in the past towards the security of wireless IMDs. It was only in 2013 that the FDA presented guidelines for securing wireless IMDs [7], which must be adhered to in order to get into the market. Quite notably, a commercially available IMD was taken off the market in 2017, after its security was deemed insufficient [8]. More recently, several researchers have disclosed the possibility of remote attacks towards commercial wireless IMDs at short-range (10 cm) [9], [10] and long-range (2 to 5 meters) [11]. The assumption was made that adversaries could (unnoticeably) perform public re-play attacks over a relatively short distance of around 10 cm (e.g., in crowded situations), using only amateur equipment. The FDA considered such a short-range scenario to be a valid assumption and took it up in the 2021 FDA advisory [12].

Since the novel FFNIs also belong to the healthcare domain, they must ensure the same security requirements. However, because this is a relatively new field, it must be disclosed to what extent security is missing and how to resolve the lack of it, while considering the imposed heavy footprint constraints of these small implants. Bringing this to light requires extensive background research and a detailed and iterative design exploration, which leads us to the research question: **What security measure is required in the highly constrained state-of-the-art free-floating neural implants?**

1.2. Thesis Scope and Contributions

This section introduces the defined scope of this thesis project by numerating the main objectives, which is followed by a list of the contributions made in this work.

Scope The key objectives of this thesis project are as follows:

i. Provide an overview and taxonomy of the functionalities of FFNIs and an assessment of the state-of-the-art, which is necessary to better approach these systems from the security perspective.

ii. Conduct a security survey to identify the potential threats and countermeasures of FFNIs.

iii. Propose and implement potential countermeasures with the focus on the limitations effectuated by the small footprint of FFNIs.

Contributions The contributions made in this thesis are:

- A state-of-the-art review of FFNIs, including a detailed taxonomy of such implants and an envisioned model.
- An extensive security survey related to conventional wireless IMDs and FFNIs, to identify potential threats and countermeasures.
- A consolidated novel attack tree classification for FFNIs, which is used to design a security application for the envisioned system.
- The proposal of a consolidated memristive lightweight GIFT cipher for encrypting communication with a sub-dura transceiver.

1.3. THESIS ORGANISATION

The organisation of this thesis is as follows: Chapter 2 continues with the principles behind FFNIs, followed by a dissection of the respective workloads that define these implants, an evaluation of the state-of-the-art with the focus towards present security measures, and a discussion on their future prospects. This chapter concludes with the presentation of the envisioned system. Then, in Chapter 3 the existing security threats in conventional wireless IMDs are discussed, together with their corresponding countermeasures. This is succeeded by an identical evaluation for FFNIs, followed by a composed attack tree which is used to establish the required security measures for the envisioned system. In Chapter 4, a memristor-centric security approach is proposed that is specifically adopted to tackle the highly resource-constrained nature of FFNIs. This is followed by the cipher design exploration to establish the proper encryption category and find a potential encryption scheme that is memristor-compatible. After this, the implementation and evaluation process of the memristor-based cipher is detailed in Chapter 5, finishing with future design recommendations. Chapter 6 concludes this work with a short summary and recommendations for future research.

2

Novel Neural Implants

Understanding the scale and orientation of the FFNIs can be a bit confusing. So, to provide a clearer image of such implants, a collection of different FFNIs, and their respective topologies, is shown in Figure 2.1. This chapter will discuss the different functional categories that exist in the present FFNIs. In addition, a taxonomy is proposed that projects the exact classification, offering a clear overview and understanding which necessary to better approach these systems from the security perspective in Chapter 3. After this, the focus will shift towards existing FFNIs and the assessment of their development state and current position towards securing the interface. Based hereon, an envisioned model is composed which will be used as the thesis' template for which a potential security measure will be evaluated and built.

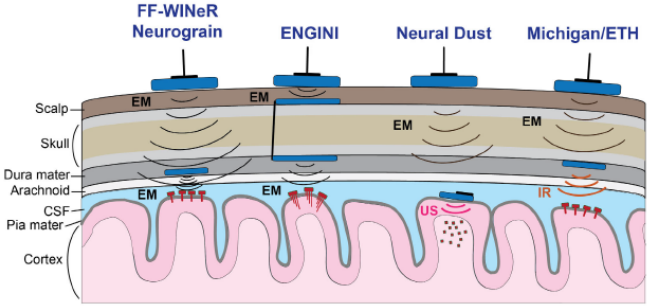


Figure 2.1: FFNIs and their topologies within a human brain. The red dust motes and transceivers communicate by means of *Electromagnetism* (EM), *Ultrasound* (US), or *Infrared* (IR), through the skull, Dura Mater, and Cerebrospinal Fluid (CSF). Image from [13].

2.1. Taxonomy of Free-Floating Neural Implants

As of now, there are two workloads the FFNIs support, i.e. recording of neuronal brain activity and stimulation of the brain (nerves). Of course, within those categories there exist different methods for recording and stimulation, but also power and data transfer. These form a classification under which an FFNI may belong, as discussed in this section. Based on this classification a taxonomy tree is built, which ultimately may be expanded or placed under a subcategory of the more conventional wireless IMDs.

2.1.1. Recording

When recording brain activity it is important to take into account the application the recorded data is meant for. Because depending on the recording signal type one would like to retrieve, the sensing architecture has to be designed accordingly. Besides that, depending on the signal type a higher or lower level of information can be retrieved. Neuronal brain activity can be measured due to current flow within tissues and cells. In the past, intracellular recording was utilized; which retrieves very clear and strong voltage signals by means of electrodes. Unfortunately, it almost instantly destroys the neuron due to the penetration of the membrane. Next to that, it is not scalable in channel count and very difficult to operate [13]. Extracellular signals are formed by synaptic transmembrane currents, Na^+ and Ca^{2+} potentials, intrinsic membrane oscillations and ionic voltage fluxes [14]. Within the domain of FFNIs, there currently exist three types of extracellular electrical recording: electrocorticography (ECoG), Local Field Potential (LFP) recording, and Action Potential (AP) recording. Recording these signals is less invasive and more scalable. In the field of neuroscience, the AP (recording) is often referred to as *Single-Unit* (recording). However, in this work they will be simply referred to as AP (recording).

Electrocorticotgraphy

Unlike electroencephalography (EEG), ECoG measures electric brain activity from the surface of the cerebral cortex, bypassing the signal distortion and attenuation caused by intracranial tissue and bone. It uses closely placed, subdural stainless steel electrodes to perform referential recording to mitigate distortion and improve the spatial resolution. Nonetheless, interference from brain activity and crosstalk is still an issue when monitoring neural signals from afar [13]. ECoG can never achieve the resolution and strength of local neural monitoring.

Local Field Potential Recording

Deeper within the cortex LFP can be measured. LFP is the accumulated average of multiple neural signal artifacts. The amplitude of LFP is typically around 0.5-5 mV_{pp} and it constitutes a low frequency band of 1-250 Hz. Recording LFP minimizes brain tissue damage compared to intracellular microelectrode recording [13], allowing placements of more recording sites and thus increasing spatial resolution [14], as opposed to ECoG. In contrast to APs, LFP measuring has been proven to be more suitable for long and stable recording [15]. In general, long-term placement of recording sites results in the generation of scar tissue caused by a "foreign body response". This, in turn, leads to temporary and spatial averaging of the measured signals and as a result a low-pass filter is formed around the electrode. This causes the high frequency components of APs to become diluted, making them incomprehensible. On the other hand, the low frequency components of LFPs are retained [15], [16].

Action Potentials Recording

APs (also called fast action potentials) are known to produce the strongest currents within the membrane. Due to their short duration they are often referred to as *spike* activity. The Na^+ spikes contribute to the higher-end frequencies of LFPs and usually have a frequency range of 0.8-10 KHz. The voltage amplitude is smaller compared to that of LFPs, i.e. 50-500 μV_{pp} [13]. The AP reveals information about cell communication and neural networks. Moreover, they provide insight into the intracellular dynamics within a brain [14]. As an LFP is a sum of components it is harder to identify the signal of the corresponding neuron. APs, on the contrary, directly record the activity of a specific neuron. Hence, APs exhibit a much larger spatial and temporal resolution.

2.1.2. Stimulating

Majority of the FFNIs that provide stimulation, can do this deep within the brain (DBS). Alternatives offer stimulation at the surface level, peripheral system or vagus nerve. Stimulation of the vagus nerve is also referred to as VNS. As with recording modules there are a couple of main categories within the stimulating implant, i.e.: electrical and optical. Then there are also implants that provide a combination of the two.

Electrical Stimulation

Electrical brain stimulation has been known to effectively provide therapy for a number of disorders and pain treatment [1]–[5]. It does so in a manner that is less intrusive compared to treatments such as lobotomy. The basic idea is to invoke timed current pulses within the region of interest, thereby triggering (suppressing) wanted (unwanted) behaviour of the patient. The stimulation parameters normally consist of repetition rate, pulse width and amplitude of the current, which enables total control over the desired therapy. Ideally the stimulating pulses are excited on a time/event basis, using a Closed-Loop (CL) system. By doing so, the patient receives the proper therapy at the right time with the right intensity. In cases where neural recording is performed, the same configuration can be utilized to perform stimulation; hence, preventing additional procedures. Typically, current stimulation is preferred. This is because voltage stimulation becomes clinically inefficient gradually over time due to the foreign body response which increases the impedance of the electrodes [17]. The injected charge from constant current stimulation is independent from the load impedance, hence it does not lose efficacy and stays constant. When opting for electrical stimulation, there are a couple of challenges that must be faced. An implant that enables concurrent stimulation and monitoring of neuronal activity, must take into account the stimulation artifacts as these can distort recorded signals. During stimulation a voltage transient is generated which could be a lot stronger than neuronal signals [18]. Therefore, selecting the right electrode material and pitch of an electrode pair is essential [13]. Another issue that arises with the downscaling of the size, is the stimulation intensity. The minimum charge for neural activation is around 10 nC (charge per phase) and due to a smaller form factor a higher voltage is required to reach the same stimulation power. This, in return, also increases the power consumption, which is undesired in these ultra-small-scale implants.

Optical Stimulation

A relatively new stimulation method has emerged as an alternative to electrical stimulation: optical stimulation, or optogenetics. Optogenetics is a stimulation method where light, emitted from a Micro-Inorganic Light-Emitting Diode (μ-ILEDs), is used as a means of neuron control. To do this, the targeted neuron needs to be genetically engineered for it to become responsive to the light. This is done by injecting the target neuron with a virus. Once modified, different light sources can control the respective halorhodopsins and opsins, which are light-gated ion pumps or channels that absorb light at a certain wavelengths [19]. As a result, APs are immediately fired when turning on/off the μ-ILEDs. Optogenetic stimulation can be done at a irradiance level of 0.22 mW/mm^2 [20]. A major benefit of optogenetics is its increased spatial/temporal resolution and specificity because only the targeted neurons will respond to light stimulus. Moreover, this will also prevent artifact distortion and any tissue damage to the brain. Lastly, electrodes are not necessary anymore which completely omits the corresponding challenges.

Some works in literature argue about relatively lower stimulation intensity of optogenetics compared to electrical stimulation, therefore obtaining lower efficacy due to loss in signal strength [13]. Yet, it has been proven already that high intensity optogenetic applications can achieve effective and efficient *transcranial* stimulation; allowing the light to emit past the skull [20].

2.1.3. Power and Data Transfer

The Wireless Power Transfer (WPT) and data telemetry within FFNIs is achieved by means of Electromagnetism (EM), Ultrasound (US) and Infrared (IR). Depending on the paradigm used, more depth, spatial resolution and less attenuation may be achieved. The WPT method applied to an FFNI is typically also used for data telemetry.

Electromagnetic Telemetry

EM WPT is the most commonly applied method in wireless IMDs and neural implants because it enables near-field communication and suffers less attenuation across biological tissue. As with other applications, the heat induced from EM waves must stay below the *Specific Absorption Rate* (SAR) limit of 1.6 W/Kg. This translates to a maximum allowed power density level of 10 mW/cm^2 for implants in the human head. In EM two coils are used, i.e. a receiver (RX) coil and transmission (TX) coil. The TX coil uses an AC power source to produce an alternating magnetic field. The receiving coil, which is placed within close proximity, receives the transferred energy from the TX coil when inductively coupled as described by the Maxwell–Faraday equation 2.1. Here, ω denotes the frequency, B the magnetic field and E the electric field. Figure 2.2 shows a basic circuit model for an inductive power transfer link.

$$\nabla \times E = -\frac{\partial B}{\partial t} \propto \omega I \tag{2.1}$$

Near field EM is regarded as a robust WPT method due to the relatively large amount of energy it can transfer at a considerable depth of 10 to 30 mm, suffering little attenuation even across tissue and skull. Nonetheless, when the distance increases, the high absorption rate of EM rapidly decreases its Power Transfer Efficiency (PTE). Moreover, the scaling towards sub-mm implants forces the operating frequency and self resonance frequency of the inductive link to shift towards the high end of the frequency spectrum. At such high frequencies, the absorption rate in biological tissue is very large, which results in the substantial reduction of the EM PTE [21]. Furthermore, WPT using EM waves results in a relatively low PTE (compared to US) due to the lower maximum allowed output power density.

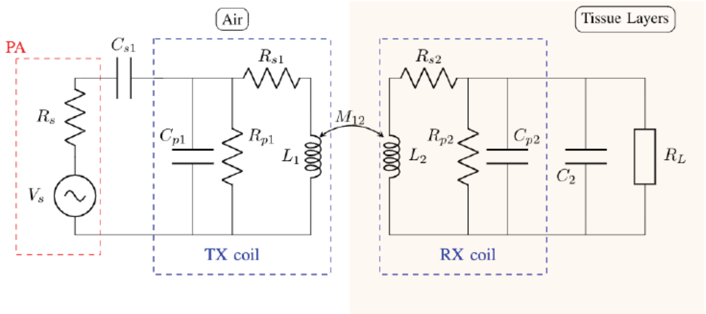


Figure 2.2: Basic schematic of an IPT. Image from [21].

Ultrasound Telemetry

US-based power/communication links for medical implants have been receiving an increasing amount of attention over the past years due to safety and efficiency reasons. US communication is established by utilizing ultrasonic waves in combination with transducers composed of piezoelectric material to send and receive kinetic energy. Piezoelectric material converts this kinetic energy into electrical energy which in turn can be used to power an implant and vice versa. US-base WPT is considered a favourable alternative to EM WPT because of its higher PTE and achievable transmission depth: at the same frequency US has a much shorter wavelength than EM waves; thus, it allows significant downscaling of transducers, i.e. up to several orders of magnitude. Furthermore, US powering allows a much lower frequency which results in significantly less power attenuation across tissue. Consequently, this enables much deeper in-body penetration and higher PTE. The FDA has set a higher limit for the power density of US (720 mW/cm^2) due to the significantly lower attenuation in tissue, allowing more intense signals [22]. All in all, US systems are the most efficient when it comes to power transfer. A downside of US is that it is sensitive to misalignment of the transducers, causing it to decrease PTE. On the other hand it has been proven that it is possible to work around this, enabling sufficient signal strength for deep brain applications [22], [23]. Next to that, US suffers from strong power attenuation through the skull; a larger signal frequency results in higher absorption by the skull, resulting in 65-90% energy loss at the receiving end [24], [25]. Another consideration of US is that it requires physical contact with the target body/patient at high transducer frequencies [22], albeit improving security as discussed in subsequent chapters. The figure below shows the basic structure of a US-based power/telemetry link.

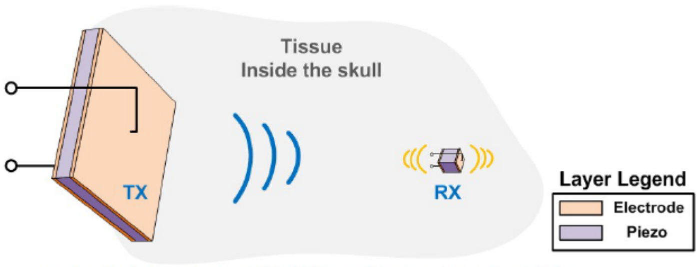


Figure 2.3: Schematic of ultrasound link for Neural Dust. Image from [26].

Infrared Telemetry

Near Infrared (NIR) powering utilizes power coming from externally emitted light, by means of a laser, which is then collected by the RX inside the head to power the implant. With a photovoltaic cell (also used in solar cells) the absorbed light causes electron excitation to a higher energy state, resulting in electric currents [27]. It is not used very often in the field of IMDs; yet, it shows potential for recharging implant batteries not only by laser light, but also sunlight. On top of that, it makes scaling down to the micron level possible; nonetheless, it is still limited to single channel use [28]. The major downside, however, is that it attenuates heavily through tissue [13]. Additionally, Moon et al. have shown that efficacy decreases rapidly after distances of more than 10 mm [29].

N-Tier Hybrid Transmission

As an alternative to all the above, N-tier power harvesting may be deployed. An N-tier structure uses N stacked power/data harvesting schemes to bridge the PTE gap created by tissue and bone attenuation. This allows a neural interface to benefit from the advantages of different harvesting techniques, eliminating some of the respective limitations. For example, it is known that EM waves have a much higher attenuation rate compared to US. Consequently they may not be a good fit for DBS etc. US on the other hand, enables much deeper propagation depths, albeit suffering heavy attenuation through skull. By combining these schemes one could use EM for transcranial transmissions to/from the outside world, and US for intracranial transmissions to/from the dust motes located deep within the cortex. Because two transceivers are used in this case, it would be a 2-tier architecture. Examples of 2-tier and even 3-tier FFNI architectures are illustrated in Figure 2.1.

2.1.4. Uplink Modulation

In the works in literature the uplink and downlink communication is done using Amplitude Modulation (AM) methods such as Amplitude Pulse-Width Modulation (PWM), Amplitude/Load-Shift Keying (ASK/LSK), On-Off-Keying (OOK), Binary Phase-Shift Keying (BPSK) etc. Because the dust mote is extremely small, modulation is usually done using *passive* methods. More specifically, there is no room for *active* back telemetry using a dedicated pulse generator, so FFNIs apply a method referred to as *backscattering*. Backscatter is created by changing the reflectivity of the incoming EM or US pulse. For example, by changing the impedance in a piezoelectric crystal, it would resonate the reflection according to one of the AM methods. Because no pulse generator is involved, it is considered passive. The method for backscattering in EM-based links is the same for US-based links.

2.1.5. Taxonomy Tree of Free-Floating Neural Implants

Considering the functionality, WPT, and all the underlying methods used for realization, a high-level taxonomy tree is composed. Given the fact that FFNIs are still in the midst of development, there is room for expansion of the taxonomy. For example, *Closed-Loop* may become an additional class of the taxonomy, as illustrated in Figure 2.4. Additionally, the taxonomy could be placed under a subcategory of the more conventional wireless IMD.

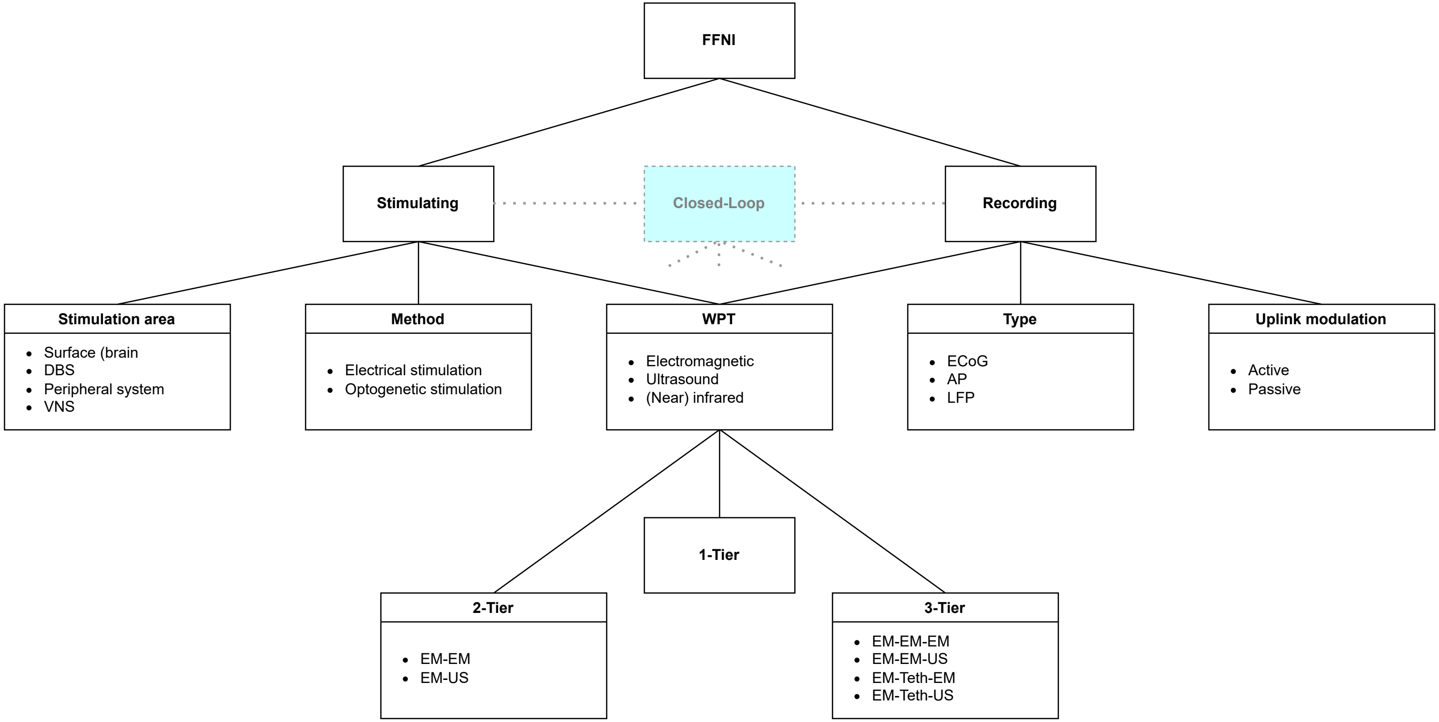


Figure 2.4: taxonomy tree of the free-floating neural implant. The marked class indicates a future potential expansion of the tree.

2.2. STATE-OF-THE-ART FREE-FLOATING NEURAL IMPLANTS

With the taxonomy of Section 2.1 as template, this section will brief the State-Of-The-Art (SOTA) FFNIs. Majority of the research focuses on Deep Brain Monitoring (DBM) and DBS, but those that target the peripheral neural system will also be considered in this section. At the end of this section one can find Table 2.1, which summarizes all the specifics of the SOTA as discussed below. This section is concluded with some thought on the current SOTA and the consolidated future prospects of the FFNIs.

NEURAL DUST

One of the first works that proposed the idea of a chronically implantable, highly scalable and free-floating neural interface, was only published in 2013 [30] and has been in development ever since [31]. The minds behind this introduce the concept of *Neural Dust*, which gets its name from the notion that thousands of μm-scale, independent sensing nodes (or dust motes), are scattered throughout the human brain; recording deep brain activity and ultimately performing CL stimulation. As of now, Neural Dust provides single channel AP recording at couple of mm depth using US WPT and data telemetry. As the description suggests, it provides both downlink and uplink communication through US backscatter. More specifically, Neural Dust only requires to change the reflective parameters of the piëzoelectric material, modulating the uplink signal such that it contains the respective data. This enables low-power and scalable design of the dust mote. Both uplink and downlink happens at 1.85 MHz with a throughput of 0.5 Mbps. Neural dust assumes a 2-tier communication architecture, consisting of a sub-dura transceiver with US communication to the dust motes and EM transcranial communication to the external transceiver. Although the goal is to achieve implants of μm-scale, the authors do point out that current physical limitations in packaging, fabrication and overall mote size, prevents them from further scaling down of the Neural Dust at this moment [31].

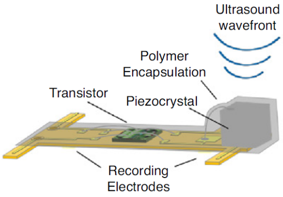


Figure 2.5: Neural Dust. Image from [31].

NIR-BASED RECORDING IMPLANT

In [28] one of the lowest power consuming recording motes is proposed. With a footprint of 0.19×0.17 mm^2 it produces no more than 0.07 μW on average. This is due to the fact that this system computes the spiking band power, which is the absolute average amplitude of a signal within a 300-1000 Hz band. In this case, these signals are APs. By omitting measurement of all AP trajectories, the bandwidth can be significantly reduced to 100 Hz which, in return, lowers uplink telemetry power. Furthermore, the extreme scalability is achieved by the use of NIR for power and data transfer, where an embedded

LED is used for the uplink communication. Nevertheless, it only allows single-channel recording and surface recording. NIR uplink telemetry is realized by firing the LED at a rate proportional to the spiking band power. Because the chip performs feature extraction, some signal filtering is performed during analog signal processing. However, Digital Signal Processing (DSP) is performed externally. Unlike all other FFNIs, this implementation provides an 8-bit hardwired password to prevent unauthorized programming.

FF-WIOS

In contrast to the previous two, Jia et al. propose a 2-tier mm-sized free-floating DBS implant, based on optogenetics (FF-WIOS) [32]. The complete architecture (including transceiver) primarily consists of the FF-WIOS implant, a sub-dermal resonator and a head stage containing power amplification, an off-the-shelf microcontroller (MCU) and a battery. The implant has 16 stimulation channels consisting of a 4×4 μLED array which can be activated individually from a remote distance (Bluetooth) via a graphic user interface. The data and power telemetry across the tiers is inductive only. The relative large resonator enables the placement of multiple motes within the same plane. Figure 2.6 depicts the implant and its placement. The achieved stimulation depth is 100 μm; thus, only superficial stimulation can be provided. Unlike other FFNIs, FF-WIOS contains CL power control. More specifically, the digitized, rectified voltage of the dust mote is sent back to the head stage, using LSK, such that the received power at the implant can be verified and active stimulation can be confirmed.

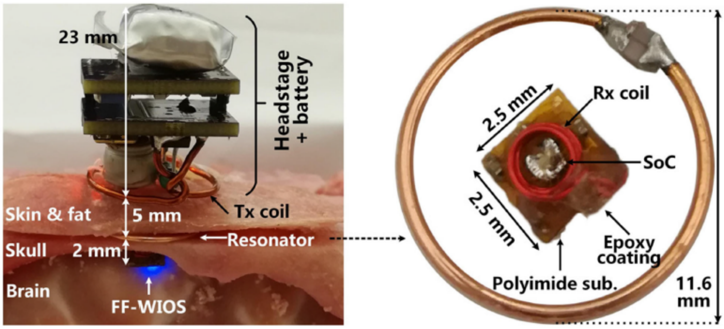


Figure 2.6: FF-WIOS in-situ (left). FF-WIOS implant and resonator (right). Images from retrieved [32].

Opto-electrical stimulating mote

The work in [33] showcases an implant offering both electrical and optogenetic stimulation targeting the peripheral nervous system. The implant is similar to Neural Dust as it uses piezoelectric material and US for receiving data and energy harvesting. More specifically, at 1.314 MHz and 0.011 Mbps, stimulation parameters such as current amplitude, pulse-width and inter-phasic delay are sent from the external transceiver to offer exact stimulation. To guarantee sufficient charge for the worst-case stimulation parameters, a small external capacitor is incorporated. With both the electrical and optical sources the implant exhibits four stimulation channels. The custom-made IC mainly consists of power/data recovery blocks and a Finite State Machine (FSM) for stimulation control.

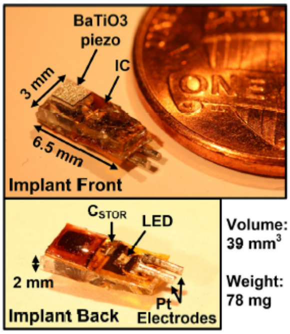


Figure 2.7: The opto-electrical stimulating mote to scale. Image from [33].

Sub-mm^3 multimote recording neural interface

Ghanbari et al. propose a passive FFNI which is closely related to Neural Dust, both in functionality and topology. In addition, their sub-mm^3 FFNI is among the first that offers multimote parallel DBM and recording of the vagus nerve with a US-based wireless link at a depth of 50 mm [23]. This is achieved using a single unfocused US transmitter and a Code-Division-Multiplexing (CDM) code generated by each implant. This code is generated by frequency division, using a ripple counter. The modulated uplink signal is then processed by the external transceiver to extract the corresponding multimote recordings. The unfocused US transmitter enables higher operating frequencies, resulting in a high temporal resolution. Figure 2.8 shows the implants and its IC. Here, one can also observe the method of multimote parallel uplink telemetry.

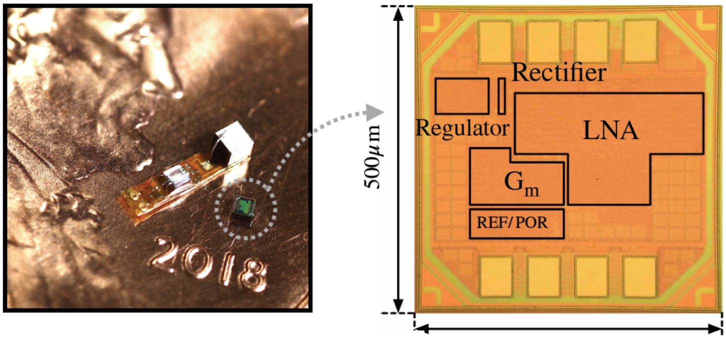


Figure 2.8: Scale image of the sub-mm^3 FFNI (left) and the corresponding IC (right). Image sourced from [23].

Rod-based recording dust mote

A fully passive EM powered mote that can capture extremely small APs (20 μV) is presented in [34]. It does so by penetrating electrode rods into the brain cortex, which may lead to questioning the intrusiveness of this neural interface. As with similar implants, the recording signal is sent by means of backscatter; in this case at 2.4 GHz. In contrast to other solutions, this work utilizes a more traditional antenna setup for the transmission of power and data. Yet, because of this the resulting architecture also generates a significantly larger footprint, i.e. 87 mm^2.

ENIAC

In the same year as [34] the Encapsulated Neural Interfacing Acquisition Chip (ENIAC) FFNI was proposed [35]. ENIAC offers 16-channel electric stimulation or ECoG. Since the implant only utilizes single tier inductive WPT and on-chip electrodes, it can only monitor and stimulate at the surface of the cortex. The ENIAC chip primarily consists of 16 Analog-Front-End (AFE) channels, some filtering, a rectifier and an ASK demodulator for the incoming telemetry. The uplink telemetry is set up using the LSK protocol. For the downlink, ASK communication configures the desired mode of operation and/or stimulation parameters. Data reception synchronization is achieved with a 16-bit ID code which is sent ahead of the Serial Peripheral Interface (SPI) communication between ENIAC and the external transceiver.

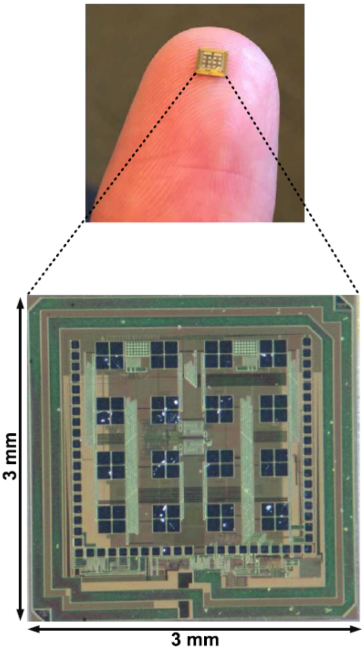


Figure 2.9: The ENIAC FFNI, to scale. Image sourced from [35].

StimDust

StimDust is a small US-based FFNI for precise stimulation of the peripheral nervous system. StimDust is similar in architecture to the other US-based FFNIs. However, the stimulating electrode is integrated in a clamp cuff, which assures that the dust mote stays attached to, for example, the sciatic nerve. This is also illustrated in Figure 2.10. As with Neural Dust, [22] does not present a design for the linked transceiver. More specifically, the in-vivo experimental setup uses a PC-powered microcontroller and off-the-shelf US transmitter with an ASIC for sending stimulation parameters and processing backscatter. The StimDust IC consists of power management, a watchdog-driven FSM, and a downlink demodulator. Similar to [32], it uses the uplink to indicate whether the implant is stimulating or not. Furthermore, the authors claim an ex-vivo implementation depth of 70 mm [22].

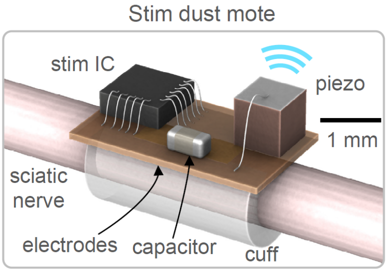


Figure 2.10: StimDust and its placement. Image retrieved from [22].

Microbead

The Microbead [36] is an ultra-small-footprint wireless surface stimulator that uses a 1.18 GHz EM for downlink communication. The authors claim to have optimized the wireless inductive link to such an extent that the Figure of Merit (FOM) remains relatively high compared to the SOTA [36]. However, the exact achievable depth and power consumption information is not disclosed. The Microbead lends its small size from the simple and compact IC design, consisting only of a voltage regulator and reference, a charge pump rectifier and an optimized on-chip coil. Furthermore, by means of individual resonating frequencies of each Microbead, the implants can be addressed separately using a frequency division protocol (FDMA). Here, the spacing between each Microbead is 250 MHz which allows the transmitter to address ten different dust motes within the spectrum of 0.5-3 GHz. Khalifa et al. illustrate how the small mote size makes it possible to implant it using a small syringe [36]. Figure 2.11 shows the actual size of the dust mote.

FF-WINeR

All the previous discussed FFNIs use a passive method for uplink telemetry. Different from this is the Free-Floating Wireless Implantable Neural Recording SoC (FF-WINeR) that uses active back telemetry implemented with 3-layer inductive coupling circuit [37]. The recorded AP signals are digitized and transmitted using TDM data which drives

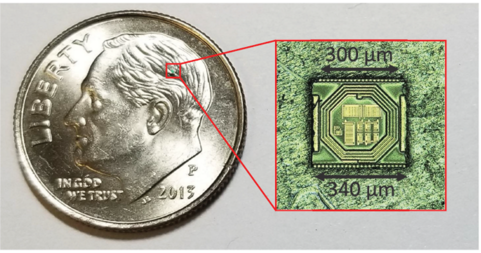


Figure 2.11: Scale photograph of the Microbead dust mote. Image from [36].

the impulse generator. Short bursts of EM pulses are then emitted through the 2-tier communication architecture. The single channel ASIC consists of a power management block, an AFE, active back telemetry, automatic resonance tuning, a clock generator, a FSM, a capacitor bank, and a sampler. With a footprint of 1.1 mm^2 and low power consumption, it is relatively efficient compared to the other designs. Similar to [34], FF-WINeR uses a electrode needle, creating an additional depth of approximately 1.2 mm. Yet, as in [34], the tissue intrusion implications must be considered. Figure 2.12 shows an image of the actual implant.

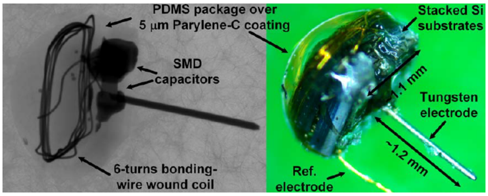


Figure 2.12: Photograph of the FF-WINeR implant. Image from [37].

ENGINI

The *Empowering Next Generation Implantable Neural Interfaces* (ENGINI) dust mote [16] shows similarities with [34] and [37] in terms of topology. To be more concrete, the diskform implant is planted on the brain surface with an array of microwire probes penetrating the tissue at different lengths. In total there are eight probes, or in other words, eight recording channels. Furthermore, uplink telemetry and WPT are realized using a 3-tier architecture: the external transceiver is inductively coupled to a sub-dermal resonator, which is transcranially hardwired to the sub-dura resonator, which is then inductively coupled to the ENGINI motes. This is also shown in Figure 2.13. The implant is aimed at recording LFPs at different depths, with the longest probe reaching the white matter at 6 mm. Nonetheless, compared to some of the depths claimed by other proposals, this is still shallow [22], [23]. The simplicity of the circuit design allows for a medium-small footprint of just above 2 mm^2. The ASIC consists of capacitors, some filtering, power management and AFE.

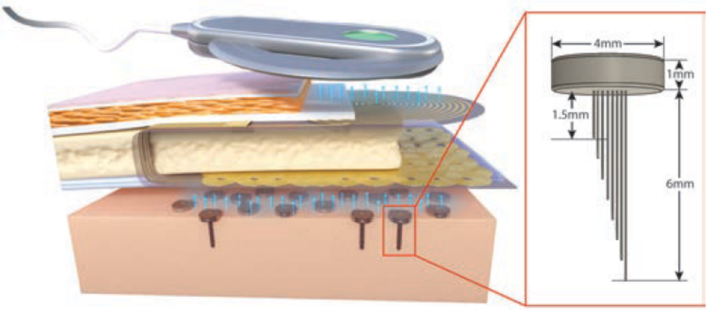


Figure 2.13: ENGINI implanted in the brain using its 3-tier communication architecture. Image from [16].

Neurograin

Neurograin is a 2-tier, inductively coupled, neural interface that is able to autonomously perform ECoG and surface stimulation with up to 1000 independently addressable dust motes [38]. However, during in-vivo experiments the total amount of implanted dust motes was 48. Utilizing a time division multiplexing method (TDMA) a highly distributive sensing/stimulating system may be realized, ultimately opening the doors for real-time CL neural interfaces. The authors explain that the TDMA protocol enable scheduled and sequential communication to/from the dust motes with just a single carrier frequency (1 GHz), ultimately preventing data collision which further enables high channel count. The addressing of each separate dust mote relies on the chip's ID composed by a PUF; it determines the motes' transmission time slots within the queue [38]. With the implant's spatial resolution, ultra-small size, low power consumption and high channel count, it may perhaps be considered the most advanced FFNI interface to date. Be that as it may, the need for deep brain applications may not be answered with this current solution. Figure 2.14 shows the Neurograin topology, retrieved from [38].

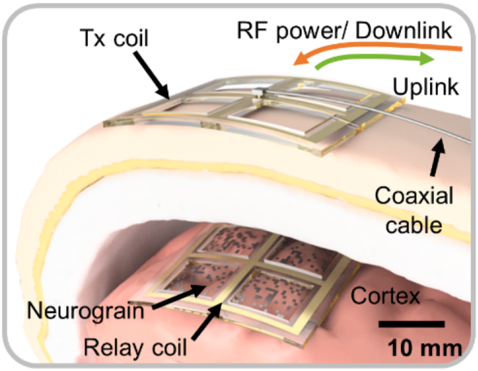


Figure 2.14: The Neurograin scattered on the brain surface and its 2-tier communication architecture. Image sourced from [38].

WiOptND

Another optogenetic solution is the Wireless Optogenetic Nanonetworking Device (WiOptND) [39], which may be considered as the stimulating counterpart of Neural Dust. As with other optogenetic implants, this system relies on the genetic modification of neurons to trigger AP generation. The WiOptND IC is completely analog and consist of a rectifier, a storage capacitor and a voice operated switch which responds to different resonating frequencies, ultimately enabling the use of multiple stimulating motes. The piezo element chosen by the developers is piezoelectric nanowires, unlike the generally used piezo crystal. The main limitation of this design is the packaging and resonator, causing a footprint of 10 mm^2. Some recommendations are made for the transceiver; however, an actual implementation is not provided yet. An illustration of WiOptND is shown in Figure 3.3.

Transcranial stimulating sub-dermal implant

Another very advanced optogenetic neural interface is the wireless, sub-dermal, transcranial implant, which has been demonstrated in an in-vivo experiment with freely moving mice [20]. This interface only uses a downlink for WPT and stimulation programming. Different colours of light can be emitted to exclude neurons or induce different behaviour. For controlling the μLED on-site an off-the-shelf μMCU is used. Next, the implant contains an on-chip capacitor bank, a voltage regulator, an on-chip resonator and the μLED. By placing the mice in a 70 × 70 cm experimental arena surrounded by an antenna, the developers were able to wirelessly control the freely moving mice, individually. Upon inducing stimulation, behavioural changes were observed in the mice [20]. Compared to the SOTA this implant has a footprint on the larger side. Nevertheless, it may also not be suitable for human DBS as the focus in [20] is more on animal behavioural studies. Figure 2.15 shows an in-situ illustration of the optogenetic implant.

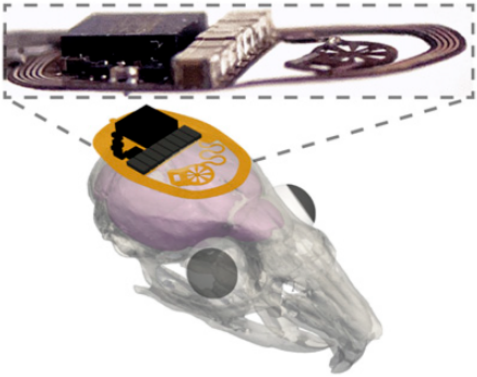


Figure 2.15: The transcranial optogenetic stimulator and its corresponding placement on the mouse skull. Image sourced from [20].

Stimulating/Recording multimote interface

The recently published work in [40] presents a neural implant that enables both recording and stimulation. By utilizing compressive sensing, the required data rate is reduced which, in return, significantly reduces the average duty-cycled power (1.15 μW). This eventually leads to a total power consumption of 1.2 mW. Additionally, the system features a ring-oscillator PUF to generate device IDs and prevent network collision. As a result of this, multiple implants may be used concurrently. The current design is missing onboard processing; this may be implemented in the near future.

Table 2.1: SOTA free-floating neural implants and specifications.

Work	Year	Stimulation	Stimulation method	Cortex depth (mm)	Recording signal	Power /Data telemetry	Communication stream uplink/downlink	No. of channels rec./stim.	Power (mW)	Total area (mm^2) IC/PCB
[28]	2020	-	-	<1	AP	NIR	Symbol interval modulation / PWM	1/-	0.00074	0.0323/0.0323
[32]	2019	DBS	Optical	<1	-	EM	Backscatter (LSK)/OOK	-/16	1	-
[33]	2018	Peripheral	Electrical & Optical	105	-	US	-/ASK-PWM	-/4	3	3.58/-
[23]	2019	-	-	50	AP	US	Backscatter (AM)/ pulse-echo US beam	1/-	0.0377	0.25/3.19
[34]	2017	-	-	<1	AP	EM	Backscatter/-	1/-	-	-/87
[41]	2016	-	-	2	AP	US	Backscatter/ pulse-echo US beam	1/-	92	-
[35]	2017	Surface	Electrical	<1	ECoG	EM	backscatter (LSK)/ ASK-PWM	16/16	0.145	9/9
[22]	2018	Peripheral	Electrical	70	-	US	Backscatter (AM)/ TDC*	-/1	-	-/5.89
[36]	2019	Surface	Electrical	-	-	EM	-/FDMA	-/1	-	0.02/0.02
[37]	2018	-	-	1.2	AP	EM	IR-UWB/-	1/-	0.29	1.1/1.1
[16]	2019	-	-	6	LFP	EM	LSK/-	8/-	0.09	2.1/2.1
[38]	2020	DBS	Electical	<1	ECoG	EM	BPSK/ASK-PWM	1/1	0.04	0.25/-
[39]	2020	DBS	Optical	60	-	US	-	-/1	-	10/10
[20]	2021	DBS	Optical	4.8	-	EM	-/OOK	-/2	92	9/-
[40]	2021	Surface	Electrical	-	LFP	EM	OOK/ OOK	1/1	1.2	1/-

* Time-to-Digital-Converter.

2.3. Evaluating Free-Floating Neural Implants

Taking the above literature research and Table 2.1 into consideration, it may be inferred that the current development of novel FFNIs is still unsettled. Besides the ability to record and stimulate, there seems to be missing a concrete consensus on the performance requirements, safety standards and the design approach for FFNI interfaces. This holds for both the dust mote implants and the transceivers. Next to that, majority of the works do not actually disclose the clinical target application other than claiming that it could offer treatment for certain neurological disorders. It is understandable that, given the novelty of this research field, scientists do not keep themselves occupied with those challenges at this point in time. Yet, they all have one goal in common: Developing a high spatial/temporal resolution FFNI interface with an ultra-small footprint and advanced processing workloads. Nonetheless, the numbers summarized in Table 2.1 do not entail such systems, but merely the individual dust motes. This would imply that further expansion of the workload with a transceiver containing blocks such as on-site DSP, security, a CL function, data logging, and perhaps even learning algorithms [31], would only create more obstacles on the way towards reaching highly scalable implants. On top of that, numerous groups have already pointed the physical limitations they have come across during the development of their implementations [28], [31], [38], [39]. Hence, these observations should be taken into consideration in the further development of the FFNIs.

Another interesting fact is that, despite most of the published works hypothesizing about the necessary workloads for the dust motes and their respective transceivers, none of the papers implement, nor mention, any form of security or protective measures across the neural interface. In addition to that, no work considered the potential security implications that may arise in these highly-constrained wireless neural interfaces. While the 'novelty of field' argument could be applied here as well, it does not hold up quite well: it would be quite insensible to finally achieve the desired ultra-lightweight implant, only to return back to the sketching table to incorporate one of the most important blocks, i.e. security. In the previous chapter, the seriousness of this matter was already emphasized with a real-life example [8]. This can be further supported by the security guidelines issued by the FDA [7], [42], which should be adhered to in order to avoid exclusion of one of the succeeding steps: human trial testing and, ultimately, commercialisation. As this is the ultimate goal of the novel FFNI interfaces, it is crucial to obey the CIA triad and incorporate security .

The next section will focus on the envisionment of the FFNI interface, which is derived from the discussed literature and educated opinion. The envisioned system will be used as the thesis' template for which the security measure will be evaluated and built.

2.4. Prospects for Free-Floating Neural Implants

Section 2.2 showcased many different FFNI systems. Despite the fact that they are still missing unity in design choices, some strong arguments have been made that support the potential functionality and look of the next generation FFNIs. For example, right now there exists a division between FNNIs that utilize EM-based WPT and those that use US, with the majority (by a small margin) inclining towards EM due to the long existing use of the method. Nonetheless, the literature is strongly hinting towards the use ultrasound telemetry [13], [22], [23], [30], [33], [39] not only because of its PTE, but also because of safety and scalability reasons, which was extensively discussed in Section 2.1.3. Nonetheless, as was pointed out earlier, US strongly attenuates through the skull. Thus, unless the skull would be thinned or carved out for transceiver placement, it would probably be fitting to consider a 2-tier architecture in the future. In that case, a transcranial EM link would be formed between the sub-dura and external transceivers. The dust motes would only be linked to the sub-dura transceiver via US. Another point of interest is the applied stimulation method. For a long time electrical stimulation was the assumed method, but the rise of optogenetics may change this. Section 2.1.2 discussed in detail the benefits of optogenetics, i.e. a high spatial resolution, specificity and a decreased intrusion. If advancements in power efficiency allow for the further scaling down of optogenetics, it will most likely cause a shift towards this stimulation method. On the other hand, scientists and clinicians will face a major obstacle if they choose to go down this road. This refers to the ethical discussion regarding genetic modification of the respective neurons, which is required to perform optical stimulation. Overruling the opposed may proof to be too difficult.

Lastly, additional workloads for FFNIs have been briefly discussed. For example, most dust motes apply either stimulation, monitoring or both (but not concurrently). Many have argued the future dust mote to be able to do both, concurrently, in an autonomous CL manner. This would require expansion of most of the existing FNNIs with a recording (stimulating) AFE, increasing both their size and power. At the same time, it would require a CL control at a higher level. The additional CL-processing could be done off-site (on a server), but depending on the application it may require fast response times. Hence, opting for an on-site (in-body) CL system, situated on the transceiver, may be the preferred approach. Consequently, this would increase the FFNI footprint. Alternatively, one could opt for using separate sensing and stimulating motes spread throughout the brain. This idea is highlighted in [31], [38]. Next, there is the common goal of high spatial resolution, implying the need for a large number of dust motes. At the moment, communication of data between the motes and surface/sub-dura transceivers is done either sequentially or in parallel, as in [36], [38], respectively. Yet, with an increasing channel count the bandwidth limitations have to be carefully considered. To do so, the use of sparse data by means of feature extraction and spike-sorting would be the necessary approach. This would also mean expansion of the AFE on the dust mote and DSP on the transceiver above. Together with the above workloads, blocks for data logging, machine learning and security, contribute to the composition of the future autonomous FFNIs.

ENVISIONED SYSTEM

Based on the literature research conducted in Section 2.2 and the arguments above and in Section 2.3, a high-level envisioned FFNI deep brain interface is composed and depicted in Figure 2.16. The envisioned system consists of a 2-tier architecture: a lower-tier US link between the dust motes and sub-dura transceiver, and a higher-tier EM link between the sub-dura and external transceivers. The interface contains both optical stimulating and recording motes, and optional stimulation coming from the surface transceiver. The illustration indicates the existing blocks in the external and internal transceiver, as discussed in the previous section. The envisioned system will be used as the model for which the security will be evaluated and designed in the subsequent chapters. Details on the scope of the countermeasures will be provided in Chapter 3.

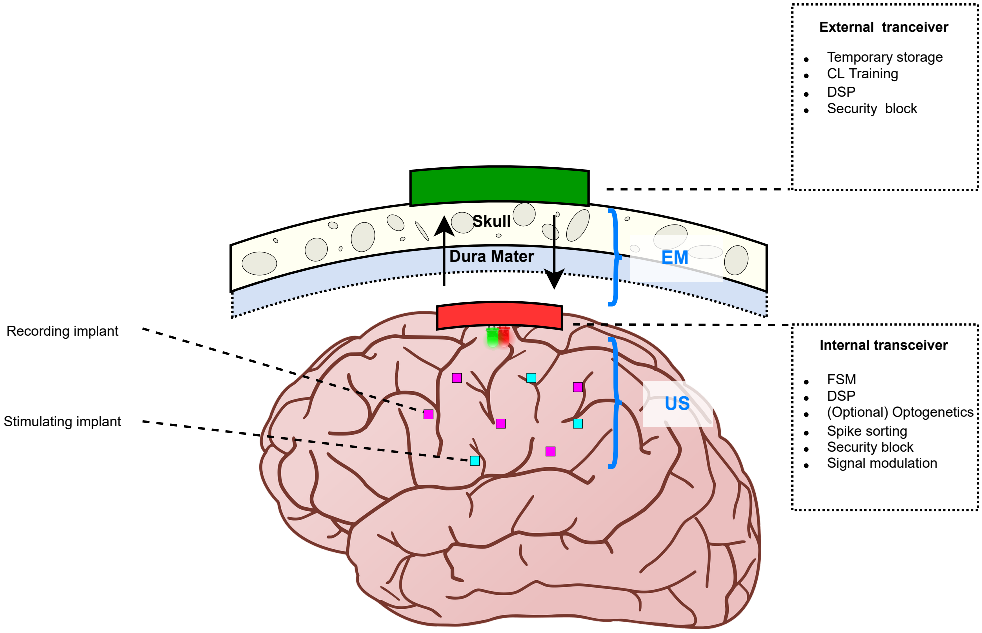


Figure 2.16: The envisioned free-floating neural interface.

3

Securing Highly-Constrained Free-Floating Neural Implants

During the assessment of the SOTA it was revealed that no attention is given to the security implications of the new implantable neural interfaces. More so, the literature highlighting the newest FFNIs does not consider potential vulnerabilities at all, fully omitting the CIA principles that ought to be implemented in wireless implants. As emphasized already, this may be blamed on the novelty of the field, which is focused more towards downscaling. On the other hand, leaving such critical functionality unanswered is rather naive given the fact that it entails personal security and privacy. Moreover, for this technology to mature and ultimately become commercially applicable to the human being, initiating steps towards the development of a truly secure system is all the more important. One of the goals of this chapter is to establish the security implications for FFNIs and, thus, also the envisioned system. What security measures should be taken? Are there any threats? If so, what vulnerabilities cause these security threats. To answer these questions, a detailed literature review has been conducted. This chapter starts with an overview of vulnerabilities and threats present in IMDs. To make sure no important details are omitted, the initial focus will be on the well-established traditional IMDs and BCIs, and their corresponding security threats. Besides that, known IMD attacks that have been conducted in the past, will also be considered. Next, existing countermeasures will be discussed and their effectiveness will be questioned. After establishing the security paradigm for traditional IMDs, the focus will shift towards the SOTA FFNIs, such as the envisioned system. Even though there is little reference material yet, the discussion for the FFNIs will follow a similar structure to that of the traditional IMDs. This is done by pointing out the vulnerabilities, threats and countermeasures. The conclusions drawn from the survey are then used to compose a novel attack tree classification for the FFNIs. Finally, a section is dedicated to the security of the envisioned system, with the objective to narrow down the design considerations for a potential security application of the envisioned system.

3.1. Security in Conventional IMDs

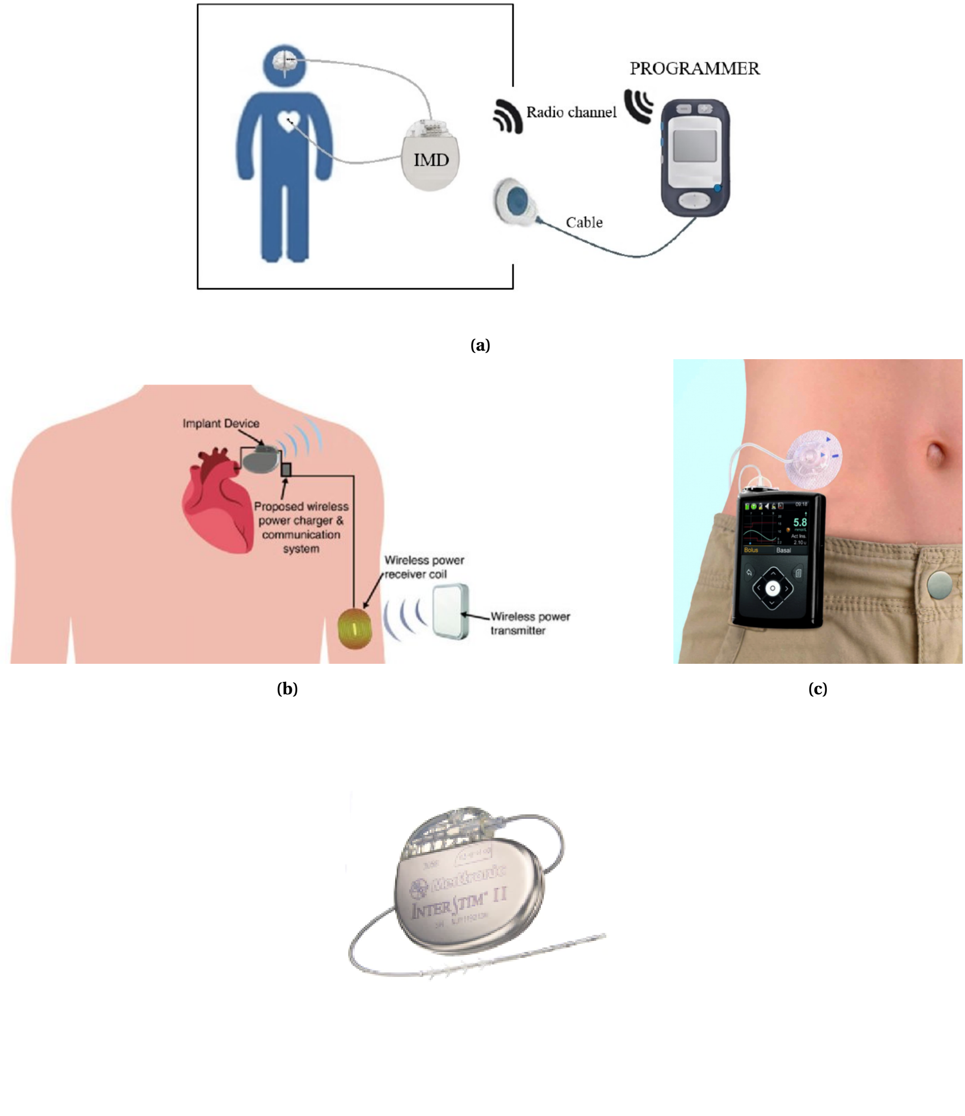
In this section, the security vulnerabilities of the conventional IMDs are discussed. This is done by highlighting real-life examples of attacks and known security breaches. In conjunction to this the corresponding countermeasures are presented. The conventional IMDs are discussed first, so that their vulnerabilities and countermeasures can be extrapolated for FFNIs. Devices that are considered 'conventional' are mainly those that have been commercially available for some while now and better known to the public. More specifically, these implants often require tethered within-body signaling, (chargeable) batteries, and occupy more volume as opposed to the proposed sub-mm works mentioned in Chapter 2. Figure 3.1 shows a collection of conventional IMDs.

3.1.1. Vulnerabilities and Threats in Conventional IMDs

IMDs have been available for commercial use for quite a while now. As IMDs grew more popular, so did the idea of an IMD getting compromised by a remote attacker. There have been multiple cases where standard IMDs, such as cardiac defibrillators and insulin pumps, have been exploited publicly [10], [47]–[49]. These attacks are relevant for all kinds of IMDs, including brain implants. In literature, different names have been given to the notion of targeting BCIs. For example, Pycroft et al. describe the concept of *brainjacking* and the different attacks that lead to the unauthorized control over BCIs [50]. Bernal et al. describe the same phenomenon as *Neuronal Cyberattack* (NCA) [51]. Another term is *brain-hacking*, which is used to describe the 'possibility of co-opting BCIs and other neural engineering devices with the purpose of accessing or manipulating neural information from the brain of users' [52]. Even though these definitions basically describe the same thing, each work may categorize each vulnerability, attack or threat a bit differently. The work in [50] mainly focuses on the effects of attacks on Implantable Pulse Generators (IPG) and DBS. Two major attack methods are highlighted: *blind attacks* and *targeted attacks*. The first comprises of attacks that do not require any prior knowledge of the patient/implant. This includes halting stimulation, draining implant batteries, tissue damaging, and monitoring. The latter comprises of attacks that require physiological knowledge of the patient or knowledge about the IMD. This includes disabling motor functions, impulse control alteration, modification of emotions, pain induction and modulating reward systems. Table 3.1 summarizes the different attacks and consequences.

In [53], a different categorization is used to classify (security) vulnerabilities in IMDs, which, in return, may help with the construction of the appropriate countermeasures. In contrast to [50] the categorization is based on cause. The authors explain that the cause of a compromised vulnerability may be divided into: proximity, IMD activity, and patient state. IMD activity refers to the state the device is in at that moment, i.e.: sensing, actuating (producing some therapeutic effect), information processing, and communicating. The patient's state refers to heart rate, stress etc.

Just like in [53], do Ienca and Haselager divide different kinds of brain-hacking in relation to the different cycles of a BCI: input generation, measurement, decoding and feedback. Furthermore, Ienca and Haselager place brain-hacking as a subcategory of *neurocrime*, which is described as 'offenses against individuals or groups of individuals with a criminal motive to intentionally cause direct or indirect physical and mental harm


PROGRAMMER
Radio channel
IMD
Cable
Implant Device
Proposed wireless power charger & communication system
Wireless power receiver coil
Wireless power transmitter
5.8
Bolus
Basal
Medtronic
INTERSTIM II

(a)

(b)

(c)

(d)

Figure 3.1: Examples of traditional IMDs. a) Brain and heart IMD, from [43]. b) Cardiac IMD. Image from [44]. c) Medtronic insulin pump [45]. d) InterStim II sacral neuromodulation system from Medtronic [46].

to the victim as well as harm to the victim's reputation and property by accessing or manipulating neural information through the use of neural devices' [52]. When taking a closer look at the published literature one may notice that the same type of attacks are mentioned repeatedly, i.e.: 1) data modification, 2) impersonation, 3) eavesdropping,

Table 3.1: Attack categorization and the corresponding consequences. Table adapted from [50]

Attack Category	Attack Type	Condition	Potential Harms
Blind	Switching off IPG	Any	Denial of stimulation, rebound effects
	Draining battery		Denial of stimulation, rebound effects, IPG damage
	Overcharge stimulation		Tissue damage
			Violation of patient privacy, facilitation of further attacks
Targeted	∼10 Hz Subthalamic nucleus stimulation	PD	Hypokinesia/akinesia
	GPi electrode contact change		
	STN electrode contact change		Impulse control disorders, alteration of affect
	Increased frequency PAG/PVG* stimulation	Pain	Increased pain
	Increased frequency VPL/VPM** stimulation		
	Increase voltage/decrease frequency ViM*** stimulation	ET	Exacerbated tremor
	Nucleus accumbens electrode contact change	OCD	Alteration of affect
	Nucleus accumbens stimulation control	OCD, depression	Alteration of reward processing, operant conditioning

[*] Periaqueductal/Periventricular Gray Matter.
[**] Ventroposterior Lateral/Medial Thalamic Nucleus.
[***] Ventral Intermediate Thalamic Nucleus.

4) replaying, and 5) Denial-of-Service (DoS) [54]. Although these are different types of attacks, one may imply the other. The following paragraphs will briefly discuss each of these attacks with related examples.

Data Modification - data modification could be in the form of an attacker remotely altering stimulation parameters (frequency, pulse width etc.) which can change the effect of stimulation [55]. Ultimately, this may lead to the halting of the stimulation to prevent impulse control (or in other words, DoS). Halting stimulation can cause severe *rebound* symptoms in diseases such as PD, ET and OCD [50]. Moreover, an adversary can inflict damage on the brain cells or cause different neuronal pathways [56]. Other forms of data modification include adding noise to manipulate measurements or generate different outputs and feedback [52], [57]. In [10] it is demonstrated how radio-frequency based attacks can send/retrieve commands/information to/from one of the newest neurostimulators available in the market. The authors, however, do not disclose the type and brand of the neurostimulator. They highlight several attacks such as the 'sleep deprivation torture' attack. Here, stimulation pulses are constantly sent, stripping the patient from a good night's rest. Yet another form of modification is the Out-of-Band (OOB) side-channel attack performed in [58], where analog components such as sensors are compromised [55], [59], [60]. Here, multiple attacks on different Cardiac Implantable Electrical Devices (CIED) have been performed, utilizing arbitrary Pulse Generators (PG). Kune et al. have performed low-power baseband attacks as most IMDs use low pass signal filtering at the analog sensing end. The idea behind the attack is to manipulate the sensor readings such that shocks from a defibrillator could be delivered or paces could be stopped. This is also called "back-door" coupling [58].

Impersonation (Spoofing) - Impersonation attacks may also have a similar result. During such an attack, the attacker tries to become authorized by 'impersonating' a permitted user. Unauthorized access to implants can be lethal as it may enable unwanted stimulation or halting [10], [50], [52], [56]. One such attack was demonstrated by Marin et al [10] after performing reverse engineering .

Eavesdropping - Eavesdropping is a significant vulnerability as it endangers privacy and information security [52], [54], [56], [57]. Information extraction could target raw sensor data [52], [57] or personal information. According to [55], privacy issues such as monitoring stimulation settings could be used to get insight into a patient's condition because attackers could establish details about a patient's pathology or mind-state. Martinovic et al. have focused their research on showing the feasibility of performing side-channel attacks using a consumer-grade BCI gaming device [61]. One of the goals of these attacks is to reveal the user's private data containing details about the pin code, area of living, bank and payment cards. To do this, the authors utilized an Event-Related Potential (ERP) signal. These signals are detected after a visual or auditory stimulus is presented to a subject [61]. Another example is the work presented in [10], which reverse engineered a commercially available neurostimulator using just a *black-box* approach. After reverse engineering it became possible to intercept messages (which were transmitted using OOK) and retrieve details such as the stimulator serial number and model, the state of the system, therapy information etc. In a more recent work the same authors have proven, in a similar way, the possibility of compromising the (at the time) latest commercial ICD (Implantable Cardioverter Defibrillator) by performing long-range passive and active attacks from a 2 to 5 meter distance [11]. The same principle, but for a short-range and different IMD, was demonstrated in 2008 by Halperin et al. [9].

Replay - Besides impersonation attacks and eavesdropping, [10] performed replay attacks, utilizing the intercepted communication. The intercepted messages are then re-sent to gain access or initiate unauthorized commands. Pycroft and Aziz [55] stressed that communication between IMDs and the corresponding base-stations or transceivers, could be intercepted remotely when not protected by encryption/authentication protocols. Next to that, they point out that the transceiver could be intercepted as well if not secured properly. For example, an adversary could sabotage or delay functionality, overwrite a transmitted signal by transmitting a previous one, or simulate sensed data and therefore causing unwanted stimulation behavior. One may see this as a sub-division of a spoofing attack, but here it is considered as a separate category since it does not necessarily require any knowledge. In [58] a similar scenario is orchestrated in which the EM baseband is exposed, which is used to replay sensor readings such that it induces cardiac stimulation. A similar approach was taken in the real-life cases mentioned in [47]–[49].

Denial of Service - One of the most frequently mentioned attacks is battery drainage as a form of DoS or denial of treatment [50], [53]–[55]. The idea is that an adversary occupies a communication channel such that proper functionality is blocked [50] or to eventually deplete the powering battery, achieving similar results [60]. As the availability of the IMD is put at risk, it can have detrimental consequences for the patient's health and eventually putting his/her life at risk [56].

All the above attacks and vulnerabilities have shown the potential of enforcing the same result, albeit not with the same attack methods necessarily. When considering BCIs specifically (since this thesis concerns neural implants), these attacks can cause severe pain, fear or psychological distress [50], [52].

3.1.2. Countermeasures Against IMD Attacks

In the previous section it became apparent that different attacks could lead to the same outcome. A countermeasure, on the other hand, may clear multiple vulnerabilities and therefore prevent multiple attacks. The attack categories presented in [54] also come with respective standard countermeasures. Table 3.2 presents an overview of the attack types and corresponding countermeasures as proposed in that work. [50] mentions potential countermeasures for the blind/targeted-attacks. These include rolling code cryptography, server-based cryptographic key management, cloakers and proximity-based authentication. Yet, they do not mention how these could possibly be implemented. Kune et al. describe three simple countermeasures against their attack scheme: shielding, differential comparators and filtering [58].

Table 3.2: Attack risks for IMDs and their preventive measures.

Attack types	Security requirements
Data modification	Data integrity
Impersonation	Authentication
Eavesdropping	Encryption
Replaying	Freshness protection
DoS	Pairing protocol

In contrast to the two aforementioned papers, the work in [53] divides the countermeasures into *protective*, *corrective* and *detective* countermeasures. The authors do mention creating an authorization list to counteract against impersonation attacks. Still, using such a list may be difficult depending on the patient's state, i.e. during emergency. Authentication in this case is guaranteed by providing something known, something possessed and/or something unique about the party like a fingerprint for example. Bonaci et al. propose a whole different strategy: standardization of devices , algorithms and regulations to mitigate problems w.r.t. security [57]. With this standardization, the authors lay emphasis on neurosecurity, which is the protection of the CIA triad of neural devices, from malicious parties. The goal is to preserve the safety of a person's neural computation, free will and neural mechanisms [57]. Anyway, standardization might yet be beyond the grasp of FFNIs as their development is still at the early stages. Another tool mentioned in this work is the *BCI anonymizer*, which basically implies never transmitting and storing *raw* neural signals. However, this may already be the direction we are heading towards with spike sorting and other dimensionality-reduction techniques.

In [56], a deep-learning methodology is presented to predict different attack stimulations in DBSs, to aid in ensuring safety, security, availability and privacy of IMDs. More specifically, *Long Short-Term Memory* (LSTM) is used for predicting Rest Tremor Velocity (RTV) and classifying the type of attack performed on a patient. RTV is a characteristic used to evaluate the intensity of PD. For a patient under stimulation treatment, this value should be around 0. Inducing an attack can increase the RTV up to 200 depending on the type of stimulation attack performed. The experiments were carried out using a PD data set from 16 real patients. Whenever the predicted RTV differed from the true RTV (after inducing an attack, for example) a flag was raised and the attack was classified. Unfortunately, the presented results included significant validation losses for some of the attack types.

Ienca and Haselager list some interesting countermeasures as well, such as the detection of unfamiliar noise during processing and measurement, and an ML self-control mechanism to detect inconsistencies at the data classification stage [52]. Pycroft and Aziz focus more on maintainability and propose four recommendations when designing IMDs [55]:

1. **Auditing** - As most IMDs are not able to log activity, it is difficult to identify the occurrence of an attack.
2. **Bug reporting** - Identify security flaws and make it possible to quickly patch them.
3. **Multi-factor authentication** - Close proximity authentication, biometric information etc. These make it harder for attackers to access IMDs.
4. **Education** - Most clinicians are not aware of cyber security risks. Also, manufacturers have not been putting enough effort in designing secure systems in the past.

Marin et al. [10] propose a clever solution against their series of attacks performed on a commercial grade neurostimulator. They present a security architecture for secure data transmission between a device programmer and a neurostimulator, utilizing a secret OOB channel. It is composed of: 1) Session key initialization, 2) Key transport, and 3) Secure data communication. Key generation is done by using the LFP signal as random source since it cannot be measured remotely by adversaries. The authors state several advantages w.r.t. *Pseudo/True Random Number Generators* (PRNG/TRNG), such as low-cost randomness. Also, the LFP can be collected using the already existing lead configurations. The key is generated by XORing three bits of the signal using two-stage parity filters to increase entropy density. Key transportation is done using a *touch-to-access* protocol [62] by connecting short wires from the microcontroller to the case as shown in Figure 3.2. Touch-to-access implies that access to the stimulator is only granted to the device programmer that can touch the subject's skin for some seconds. Furthermore, the secure data exchange is achieved by using the key to perform any standard authentication protocol. More specifically, they create a MAC and use a counter to prevent replay attacks increasing the communication overhead by <10% compared to the original message format. While this solution seems very promising, it is not clear yet to what extent it may be applicable to more sophisticated free floating, deep-brain neural motes as data transmission takes place on a smaller scale and may not be able to benefit from only the touch-to-access protocol. Nevertheless, in Section 3.3 it will be shown that this same principle is also adapted in [60], using a more deliberate scheme.

3.2. Security in Free-Floating Neural Implants

Despite the expected differences in architecture and communication protocol, many of the vulnerabilities, threats and countermeasures in Section 3.1.1 may be extrapolated to FFNIs as well.

3.2.1. Vulnerabilities and Threats in FFNIs

Bernal et al. [51] are one of the few that have ignited the discussion about the risk of new types of possible cyberattacks in novel wireless neuronal applications. They mention both Neural Dust [30] and WiOptND [39] as architectures consisting of multiple vulnerabilities. In these systems no security functionalities are integrated in the implant,

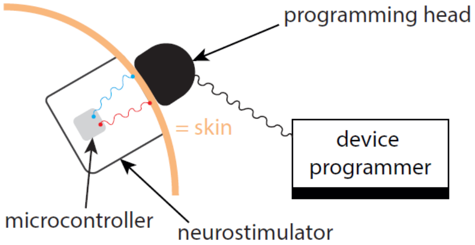


Figure 3.2: Presented touch-to-access protocol for sending a session key. The protocol is based on [62]. Image from [10].

sub-dura transceiver and external transceiver, allowing an attacker to not only collect sensitive data, but to also threaten the patient's health. For example, WiOptND allows an adversary to send a malicious firing pattern which causes unwanted stimulus. Figure 3.3 illustrates the vulnerabilities in both the applications.

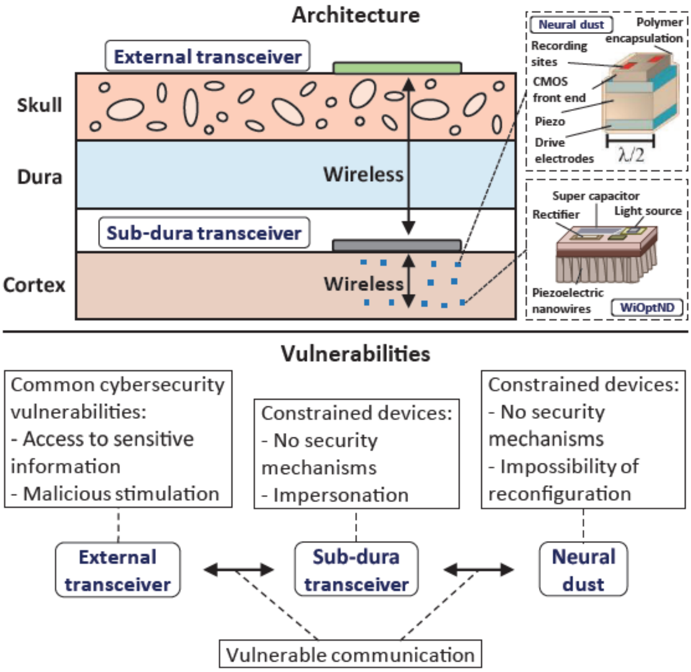


Figure 3.3: Architecture, placement and vulnerabilities of neural dust and WiOptND. Image retrieved from [51].

Moreover, this work focuses more on the types of physical impact of different attacks, rather than the attacks themselves. More precisely, they talk about the effects on neuronal activity and neuronal stress due to certain attacks, hence, their use of the term NCA. They define two types of NCAs: Neuronal Flooding (FLO) and Neuronal Scanning (SCA), both of which can influence the neuronal activity during neurostimulation. Similar to flooding attacks, FLO causes overstimulation and behavioural change by stimulating multiple neurons at a certain time instance. An SCA attack only targets one neuron at the time. To evaluate the effect of both attacks, an abstract of the primary visual cortex of mice was replicated using a Convolutional Neural Network (CNN). With this model they built a simulation of a mouse moving through a maze, to see the impact during this activity. Both attacks are executed by inducing a voltage rise in the corresponding neurons. The main difference between the two attacks is that FLO has an instant effect, as multiple neurons are triggered at the same time, while SCA has a relatively delayed impact. Having said that, both alter the spontaneous behaviour of neuronal signaling. However, they do not talk about the methods of orchestrating such attacks nor the prevention strategies.

With the above in mind, it is apparent that the vulnerabilities in the SOTA are just as real as the ones discussed in Section 3.1.1. Nonetheless, many of the threats and attacks mentioned in literature are either hypothetical or far fetched in the sense that the attacks require too much complexity, extremely optimal attack conditions, physical contact and highly-skilled attackers. For these reasons, an attack tree classification is created that shows threats that are considered most common and plausible for the new implants. The attack tree in Figure 3.4 is based on the discussion in Section 3.1.1 and composed for the FNNI. Instead of using the term ‘targeted attacks’, the term *informed attacks* is used, as it better describes the nature of the attacks: attacks that require knowledge of the device and/or patient. Also, as these attacks are specific to FFNIs, they fall under a new categorization of attacks: *Neural Dust Attacks* (NDA).

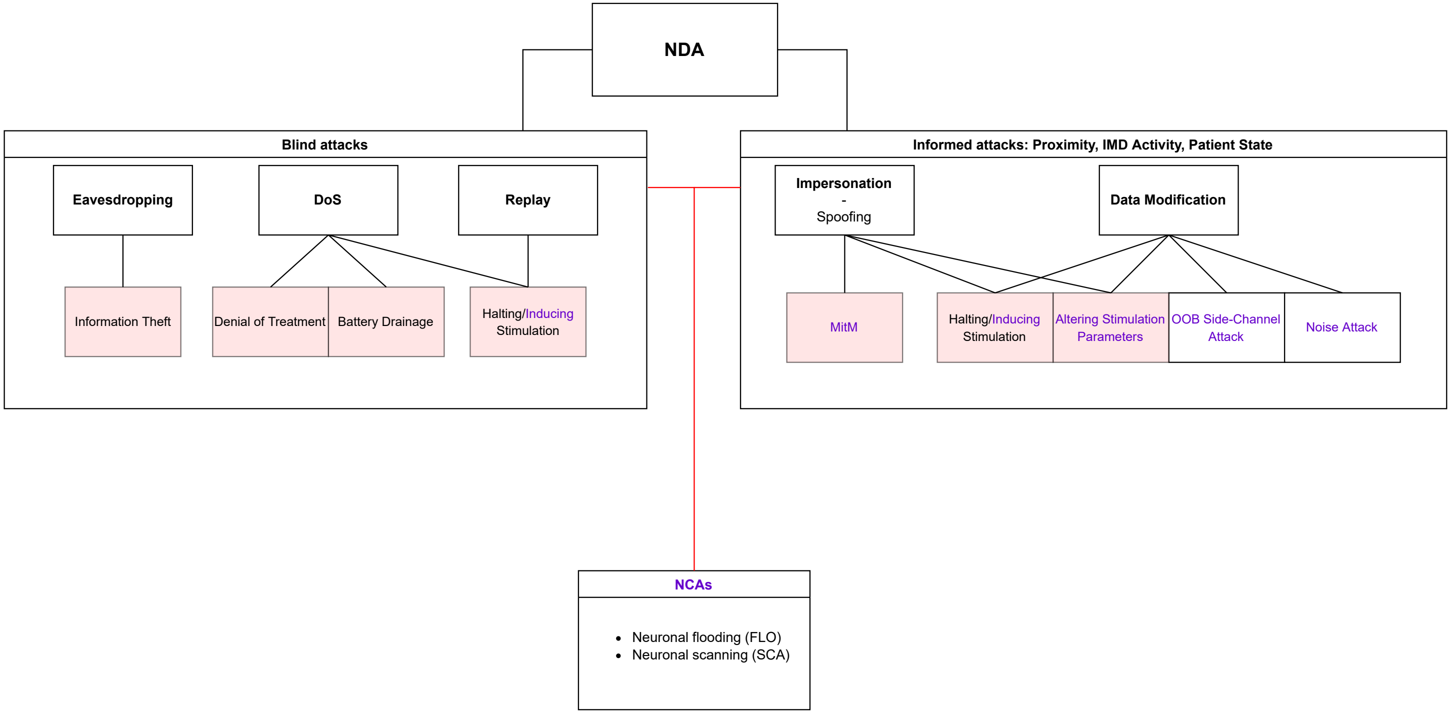


Figure 3.4: Attack tree for the NDA. The marked boxes are deemed as the most critical attacks against patients. The purple attacks may potentially lead to NCAs.

3.2.2. COUNTERMEASURES AGAINST NDAS

In theory, the countermeasures discussed in Section 3.1.1 could be applied against NDAs. However, not all the countermeasures are suitable. For example, the proposal in [56] may be too resource demanding for the FFNIs that are expected to decrease in size and energy consumption. Besides this, a lot of the proposals are not efficient or still show weaknesses. Potential threats like OOB side-channel attacks could be prevented with shielding. However, it depends on the type of sensing circuit (acoustic or not) whether this is necessary. Also, protection by means of obscurity is not considered a desired approach. The most promising and encapsulating countermeasures so far, are proposed in [10] and [63]. One of the reasons is because they use US for key (or data/power) transmissions; which is the expected trend as mentioned in Chapter 2.

Siddiqi et al. propose a robust and lightweight US-based device pairing protocol called SecureEcho [60]. The protocol does not only assure safe key transmission, but also prevents battery drain attacks; such a protection mechanism falls under Zero-Power Defense (ZPD). The touch-to-access-based protocol employs an ultrasonic Body-Coupled-Communication (BCC) channel for authentication and key sharing that is inherently secure. The implementation is completely passive as it does not require energy harvesting; this also enables the communication interface to stay in "sleep mode" before access is initiated using the BCC channel. By doing so, the protocol offers ZPD. Figure 3.5 depicts the SecureEcho architecture. According to the authors the attacker will not be able to use the US communication channel without being noticed. Furthermore, EM/RF communication can never happen before it is enabled by first using the ultrasonic BCC channel. Hence, a battery drain attack is prevented.

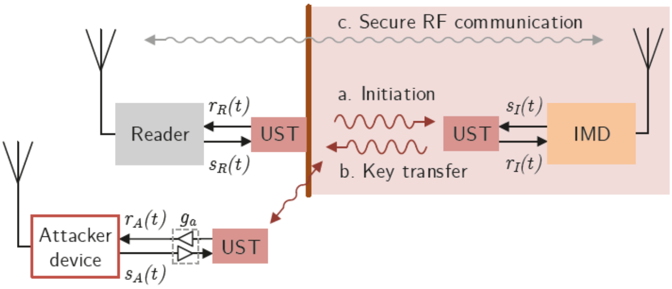


Figure 3.5: Architecture of BCC pairing protocol SecureEcho. Image from [60].

Utilizing US as means of secure transmission, has been promoted in multiple works lately. For example, Siddiqi et al. [60] argue that US is preferred because US transducers allow directional and very-short-range communication, which is perfect for private key transmissions. To further support this claim the authors have conducted a thorough security analysis of US communication. They have proven that it is impossible to eavesdrop on a 2 MHz US IMD transmission over air, at a distance larger than 5 cm. For lower frequencies the eavesdropping distance increases; albeit for an extremely low noise floor of -130 dBm. Next to this, the authors have also proven that even at very close proximity, the attacker has to be strictly aligned with the IMD transducer, making a close proximity attack impossible. Moreover, in [64] it is stated that US-based FFNIs, such as WiOptND, prevent security threats from malicious US signals since the motes are placed underneath the skull. The authors emphasize that a security breach on that level can only be performed by embedding the malware among the existing implanted units. Also, in [24], [25] it is mentioned that US attenuates quickly when passing through the skull.

3.3. Security of the Envisioned System

At the end of Chapter 2 the architecture of the envisioned system was presented with the goal to construct a security architecture for the future FFNIs. This was done by analysing potential threats for FFNIs and, thus, the envisioned system. Earlier, it was disclosed that no security mechanisms currently exist within the dust motes and the transceiver(s). Furthermore, it has been shown that future FFNIs would have probably two types of transmission: EM-based and US-based for the external and subdural transceivers, respectively. In Section 3.2.2 the implicit safety of US was discussed. Other works have also pointed out the benefits of utilizing US transmissions [22], [24], [58], [60], [65], [66]. With this in mind, the US link as described in Section 2.1.3 and 2.4 is assumed to suffice as an indirect countermeasure against remote NDAs. Thus, the lowest communication layer in the envisioned system is deemed secure in this work. This leaves the external and sub-dura transceivers shown in Figure 2.16. This work will develop a security block for **securing the uplink communication** of the **sub-dura transceiver**. It is important to note that once such a block is devised, it may also be applied to the less constrained external transceiver.

One may argue that the design rules for securing the sub-dura transceiver are a lot more relaxed compared to that of the dust motes; this may be true. Be that as it may, it should be taken into consideration that the current research, as presented in Chapter 2, aims for extreme small sizes, also for these transceivers. Hence, taking into account the most crucial threats in Figure 3.4 (marked), the minimal required security that should be applied, is basic encryption. Also, since the target applications are highly-constrained, **lightweight encryption** fits the profile.

The next chapter will discuss the potential implementation technology for realising a lightweight encryption block.

4

Towards Memristor-Based Encryption

Chapter 3 established the basic means for securing the envisioned system. In the chapter's final section, it was also established that basic encryption would cover a multitude of the most critical threats in future FFNIs. However, the development of FFNIs may also face physical limitation due to their target size, putting constraints on the design of the required security block. Because of this, it is important to explore alternative implementation technologies compared to mainstream ones (e.g. CMOS), for realizing such an encryption module. This chapter will briefly explain why memristors could potentially be deployed in a lightweight encryption block, followed by an overview of the principles behind memristor technology and existing memristor-based security models. After this, the discussion will focus on the design exploration for establishing the most suitable encryption method using memristors, followed by a quick background of the respective encryption method. This chapter is concluded with an evaluation of domain-specific Memristor-based Logic Gates (MLG), which will be considered as a tool for realizing the memristor-based encryption block.

4.1. A Memristive Countermeasure Approach

In Section 2.3 it came to light that downscaling may become challenging, considering the technology's physical limitations. Moreover, the footprint consumption highlighted in the SOTA mainly refers to that of single channel dust motes, as was shown in Table 2.1. Besides that, the table shows that the most advanced FFNIs (i.e. the ones with the most comprehensive workload) already have a significantly higher energy and area consumption. Hence, opting for an ultra-small footprint is expected to become even more difficult when expanding the FFNI workloads in the near future. The development of FFNIs is likely approaching extremely constrained designs that require robust security, low complexity, ultra-low power, energy efficiency, and high throughput for multichannel recording/stimulation. For now, the security problem may be approachable with encryption modules based on conventional MOSFET technology. Yet, considering the target miniaturisation of the implant architectures, combined with the nearing limit of achievable density [67], it is wise to look for alternatives. One such potential alternative is the memristor technology. Besides offering new levels of scalability (thereby passing the density limit), it is also extremely energy efficient, inhibits high switching speeds, has low complexity, and is compatible with the CMOS technology. For these reasons, this work considers memristors as a means of devising an encryption module.

4.2. Memristor background

To better understand the principles of operation of the memristor, this section will provide some background on where the memristor comes from and how it works. Furthermore, the memristor crossbar is highlighted and some applications are illustrated, with the focus on the SOTA memristive hardware security.

4.2.1. Basic Theory and Principles of the Memristor Operation

The memristor is the fourth fundamental circuit element and was for the first time proposed by Chua in 1971 as a result of combining the fundamental circuit variables: current I, voltage V, magnetic flux ϕ, charge q and time t [68]. The *memristor* is short for memory resistor and sets a direct relation between magnetic flux (ϕ) and charge (q) in its expression: $M = \frac{d\phi}{dq'}$; this is illustrated in Figure 4.1 and was identified as the fourth element. From Figure 4.1 and Equation 4.1, it can be observed that the memristance depends on the history of its current.

$$v(t) = M(q(t))i(t) \tag{4.1}$$

However, a memristor behaves as any ordinary resistor element at a given instant of time. Furthermore, Chua explains that a memristor acts like a *linear time-varying resistor* as soon as the memristive device current (voltage) $i(t)$ ($v(t)$) is established [68]. The non-volatile memristor has two different states: Low-Resistance State (LRS) and High-Resistance State (HRS). These states depend on the voltage applied to one of the terminals and the time span over which this is done. When the memristor is used as bitcell, the HRS usually denotes the logic state '0' and the LRS the logic state '1'. The process of changing the resistance from LRS to HRS is called 'reset', while changing the memristor from HRS to LRS is called 'set'. The actual resistive values that define these states depend not only on the manufacturer, but more importantly on the non-volatile and resistive switching materials of the memristor. A memristor can be composed in different ways. The most widely used materials are TiO_2, HfO_2 or MoS_2/MoO_2 [69]. In these oxygen-based models the resistance is defined by the oxygen vacancies in the disc region near the active electrode.

Despite Moore's law, the seemingly endless scaling of transistor-based devices is being driven to its physical limits. Due to the decreasing size, problems such as thermal effects, leakage and low reliability are appearing more frequently. For years now the researchers have been looking for alternatives to the conventional silicon transistor [67]. The memristor has a lot to offer as it enables simplicity, fast switching speeds, ultra-low power consumption, high integration density, and high ON/OFF ratio [70]. On top of that the memristor is CMOS-compatible. The non-volatile character of the memristor enables it to retain its state after left without voltage supply. Despite these promising traits, it was only in 2008 that the first physical model of a memristor was realized in the Hewlett-Packard (HP) laboratories [71]. Nonetheless, the memristor has been considered as a potential replacement for CMOS-based circuit design [69], [72], [73] as it shows the potential to overcome the von-Neumann bottleneck and sizing problem of transistors [74]. Because the memristor is only a two-terminal element, it requires CMOS-based control to operate. Due to the its simple structure it is most common to configure memristors in a crossbar structure, optionally combined with transistors. The next section will briefly discuss different memristor crossbars in more detail.

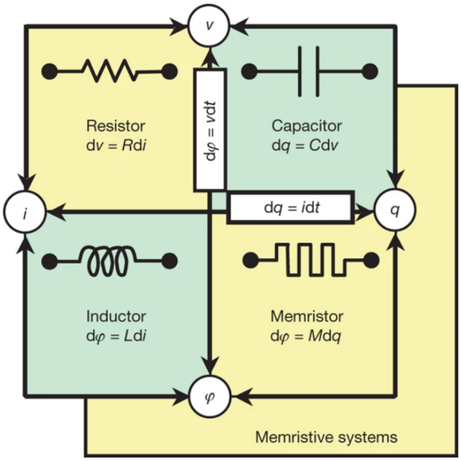


Figure 4.1: The missing fourth element: memristor [71].

4.2.2. MEMRISTOR CROSSBAR ARRAYS

The memristor crossbar is traditionally used as a memory structure to, for example, replace the traditional SRAM. This is also referred to as the Resistive-RAM (RRAM). Additionally, the structure may be used as an accelerator to drive neuromorphic applications by means of Vector-Matrix Multiplications (VMM) [75], [76]. Figure 4.2 shows the general structure of a crossbar. The vertically stacked 3D structure enables high density; the cell size per bit is reduced to 4 F^2 (F is short for minimum feature size) [77].

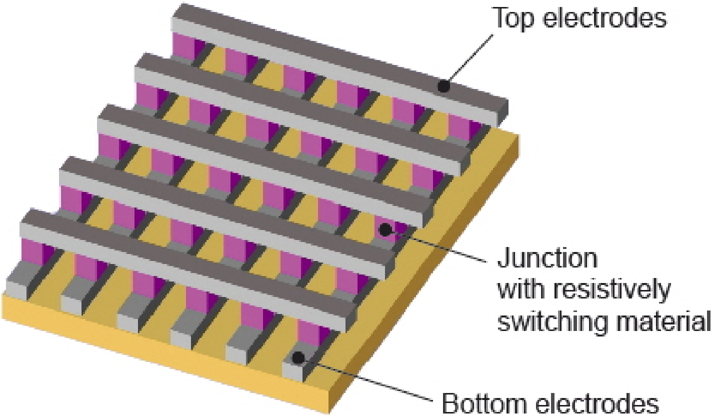


Figure 4.2: The general structure of a memristor crossbar array. The metallic (grey) electrodes sandwich the resistive switching layer (pink) [77].

There are many different crossbar configurations. In general, they are referred to as *n-element-m-resistor* (n_mR) arrays; one crossbar bitcell consist of **n** element(s) (e.g. transistor or diode) and **m** memristor(s). First of all, there is the crossbar only consisting of memristors (1R). Although 1R (Figure 4.2) contains the lowest area overhead, its difficulty lies in selecting the individual bitcells. This is because without a specified selector, current sneak-paths are induced. Hence, sneak-paths are avoided by implementing crossbars with an additional element. The 1-selector-1-resistor (1S1R) crossbar in Figure 4.3 utilizes an integrated graphene selector to drive the memristive elements. Like 1S1R, the simple 1-diode-1-resistor (1D1R) crossbar achieved a footprint of 4 F^2 and potentially even smaller. On the other hand, device performance is not as good compared to other structures since the 1D1R requires uni-polar memristors instead of bipolar ones [74].

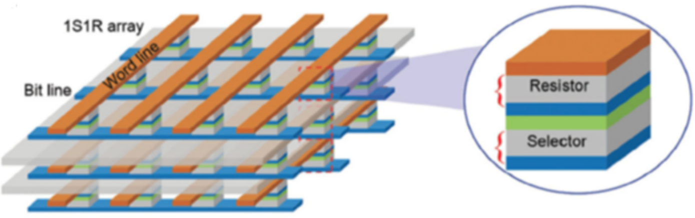


Figure 4.3: Schematic of a 3D 1S1R crossbar. Image from [74].

The most popular and widely used structure is a 1-transistor-1-memristor (1T1R) crossbar as shown in Figure 4.4. This is mainly because it enables easy selection and programming of the bitcells for in-memory computations (in-situ) using Word Lines (WL) and Selection Lines (SL) [74]. Next to that, the transistor and memristor elements can be stacked, allowing greater density. As with the other types, it is also possible to construct crossbars using more (memristive) elements. An example of this is the 2T2R crossbar proposed in [78]. Furthermore, there is the 1BJT1R, CRS and SRC crossbar array; for details please refer to [74].

Even though 1T1R is the preferred structure for now, its scalability is limited by the transistor, unlike some of the other solutions. Be that as it may, every crossbar structure has its individual ceiling. Thus, it is up to the development of new algorithms and resistive switching materials to address these limitations.

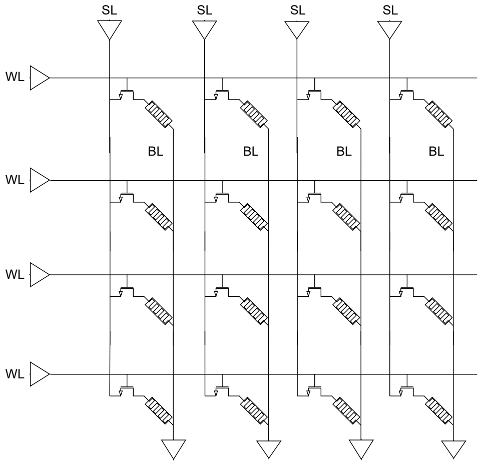


Figure 4.4: Basic structure of a 1T1R crossbar array

4.2.3. MEMRISTIVE HARDWARE SECURITY

The idea of using memristors as building blocks for hardware security is not new. In recent years there has been an increase in papers proposing to upgrade existing countermeasures with memristor components [70], [79]–[90]. Besides the earlier statements, reasons for employing memristors in hardware security, include:

1. Current methods of utilizing non-volatile memory for storing encrypted keys (flash memory, EEPROM, etc.) are susceptible to side-channel attacks and signal analysis in addition to being costly and complex.

2. Conventional signatures cannot meet the lightweight security requirements.

3. It has become more urgent to have high quality TRNGs. Memristors makes this possible.

What supports these claims, is the fact that memristors intrinsically show stochastic behavior [84], [91] like resistance variability [86], [92], probabilistic switching [82], [83] and Random Telegraph Noise (RTN) [93]. On top of that, the Device-2-Device/Cycle-2-Cycle (D2D/C2C) variation is also present in memristive devices [81], [83]. With these stochastic properties, Physical Unclonable Functions (PUF), TRNGs, and chaotic circuits may be built depending on the resistive material [70], [79], [80], [82], [84], [88], [89]. These security schemes are used for authentication, key-generation and encryption, respectively. The conventional PUF is a hardware security solution that relies on the intrinsic process variations and impurities that arise during the semiconductor fabrication. With these variables individual Challenge-Response Pairs (CRP) are created that serve as a hardware fingerprint.

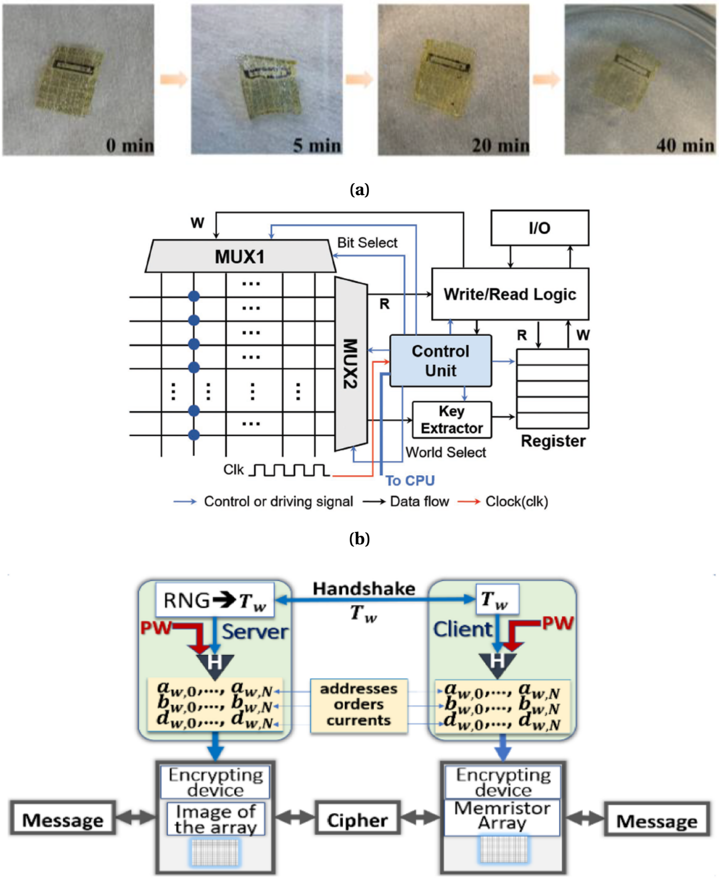
There are other creative solutions that utilize memristors. For example, Sun et al. propose a next generation memristor-based PUF that also prevents compromisation, using triggered solubility when necessary [82] (see Figure 4.5a). The authors achieved this by using water-assisted transfer printing. Another work uses a randomly initialized memristor crossbar to perform VMM for creating hypervectors as means of encryption [83]; relying on crossbar non-idealities and C2C variation. However, for decryption the authors propose a neural network. Similarly, in [81] a 1T1R crossbar is used to store plaintext and perform in-situ XOR operations with key bits for encryption (see Figure 4.5b). The key bits are generated using the subthreshold-slope of each transistor. These vary intrinsically and hence function as PUF [81]. Two other papers discuss the concept of keyless encryption, using memristors as a source of entropy [80], [90]; however, they still require a handshake/secret transaction (see Figure 4.5c), which is susceptible to attacks.

Despite these promising schemes, many of these are not necessarily lightweight due to the required large driver circuits. Besides that, the majority of these solutions require frequent operational switching of the memristors; this should be avoided due increase in energy consumption and decrease in life-time of the memristive device [94].

4.3. LIGHTWEIGHT CRYPTOGRAPHY DESIGN EXPLORATION

In Chapter 3, it was established that lightweight encryption is required to secure the communication channel between the external and sub-dura transceiver. This section will detail the design exploration of the right encryption (cryptography) scheme for this purpose. The interested reader can refer to Appendix A for a brief overview of cryptography.

Multiple extensive literature reviews on SOTA Lightweight Cryptography (LWC) have been done already [95]–[97]. The data compiled in these works assists in narrowing down the search space for finding a suitable block cipher for FFNIs. The ciphers in this search space are evaluated by looking at **throughput**, **power/energy consumption**, **area (GE)** and most importantly **security**. When it comes to block ciphers, there are two main types: Feistel network and Substitution Permutation Network (SPN). Some other works also consider Add-Rotate-XOR (ARX), Non-Linear-Feedback Shift Register (NLFSR), Generalized Feistel Network (GFN) and Hybrid [96]. In general, the SPN cipher is preferred for LWC, as Feistel ciphers are known to suffer from security problems and require more rounds [96], [97]. As the name already suggests, SPN replaces plaintext blocks with different values and subsequently rearranges the sequence. Feistel, on



(a)

(b)

(c)

Figure 4.5: Memristor-based hardware security: a) Crossbar PUF showing triggered dissolvement over time [82]. b) In-situ encryption scheme using 1T1R crossbar and subthreshold-slope for key generation [81]. c) Keyless memristor-based encryption scheme [90].

the other hand, splits the plaintext into two blocks that will interact with each other by means of an XOR operation. Table 4.1 shows a collection of selected Lightweight Block Ciphers (LWBC), based on the parameters highlighted above.

It should be pointed out that the means of data gathering, as done by the corresponding papers, is not entirely clear and sometimes inconsistent. GE, power and energy consumption seem to differ in every literature and also the scope within which these numbers are considered. This is partially due to FPGA type, operation speed or technology that is used for evaluation. Next to that, different evaluation approaches have been used. For example, some works multiply GE with the number of clock cycles to evaluate energy/bit [98], while others use energy equations based on power, frequency and

Table 4.1: LWBCs and specifications. Partially adapted from [95]–[97].

Cipher	Structure	Key size (bits)	Block size (bits)	Rounds	Area (GE)	Technology μm	Power μW	Energy (μJ/bit)
PRESENT	SPN	80	64	31	1570	0.13	2.20	10.63
RECTANGLE	SPN	80	64	25	2063.6	0.13	1.78	7.25
SIMON	Feistel	96	64	52	1216	0.13	1.21	8.55
SPECK	Feistel	96	64	28	1522	0.13	1.52	6.89
GIFT	SPN	128	64/128	28/40	1345/1997	0.09	1.97	8.03
SLIM	Feistel	80	32	32	553**	n.a.	n.a.	n.a.
μ^2	GFN	80	64	15	n.a.	n.a.	n.a.	n.a.
ANU-II	Feistel	80/128	64	25	1322	n.a.	2400*	0.055*
NLBIST	Feistel & SPN	64	64	5 (at least)	n.a.	n.a.	n.a.	n.a.
Piccolo	GFN	80/128	64	25/31	1136/1197	0.13	n.a.	n.a.
BORON	SPN	80/128	64	25	1939	0.18	n.a.	n.a.

* At a frequency of 10MHz. Other frequencies are unknown.
** GE only covers encryption core.

cycle information [99]; resulting in different accuracies of the final numbers. Hence, for the above reasons, Table 4.1 is merely constructed as an indicator for finding suitable ciphers for FFNIs.

The authors in [95] point out that in terms of throughput and energy, the ciphers Speck, Simon and GIFT show the best result. Regarding simplicity, GIFT outperforms the former two. From the duo, Simon is optimized for performance in hardware and, thus, more energy efficient compared to Speck [100]. However, as pointed out in [95], it requires some improvements to be used as LWBCs. The SLIM architecture shows a minimal footprint and simple design and is secured against linear and differential cyberattacks [97]. The fairly new μ^2 scheme is claimed to be more efficient than PRESENT and also robust against all well known attacks [101]. However, there is only speculation about the actual performance. The same holds for NLBIST; it is claimed to hold up against all well known attacks and creates good potential by combining both SPN and Feistel [97], [102]. Yet, the authors only infer low energy consumption based on the small memory size that is used [102]. ANU-II shows very good results in terms of overhead. It also holds up to all well known attacks and, as stated by [97], it is the smallest LWBC as of now w.r.t. execution time, memory requirement and power consumption. The BORON scheme is SPN based and very compact; nevertheless, it only shows reasonable resistance against attacks [97].

Overall, it can be noted that majority of the mentioned LWBCs are very similar in terms of overhead and performance. In terms of area, SLIM seems to be number one, albeit energy and power are unknown. Other than that, GIFT and the Simon and Speck family seem to show the best numbers overall. ANU-II also shows promising energy consumption. Yet, keep in mind that the known frequency of operation is most likely higher compared to the others. Also, it is not clarified what technology is used during testing.

It may be observed that the suitability of a LWBC highly depends on the application at hand. A reference design for a memristor-based cipher must therefore show potential compatibility with memristors. The cipher that does that the best, is GIFT. Not only does it outperform the others in terms of simplicity in structure and operation [95], its simple structure (mainly consisting of substitution boxes) enables a straightforward crossbar implementation. More specifically, a memristor crossbar can be composed in a way such that a GIFT encryption round can be performed by only a single 'read' action (more about this in the subsequent chapter). For these reasons, the GIFT cipher will be used as a reference and inspiration towards implementing a lightweight memristor-based encryption block for the envisioned system. The next section will explain more about the working principle of GIFT.

4.4. THE GIFT CIPHER

Similar to known standardized algorithms such as AES, SKINNY and PRESENT, GIFT is based on SPN (Substitution-Permutation Networks) in which the plaintext nibbles are replaced with other values, followed by rearrangement. The cipher is inspired on its predecessor PRESENT, with the main differences being improved security and efficiency. Most of PRESENT's security against differential attacks relies on its S-boxes (short for substitution box), which comes at high implementation cost as it requires a differential branching number of 3 (a measure of the diffusion power of a permutation) [103]. GIFT removes this obstacle by carefully composing the bit permutation and substitution in conjunction. More specifically, by removing this constraint a much cheaper S-box may be selected. Moreover, compared to PRESENT, GIFT requires half the number of XOR operations as it only performs key addition on 2 bits per nibble in addition to constant adding for some of the nibbles; PRESENT applies key addition to all the bits and also uses a larger S-box . The GIFT family consists of two members: GIFT-64 and GIFT-128. The former takes in 64 bits and uses 28 rounds while the latter encrypts 128 bits, using 40 rounds. Both versions use a 128-bit encryption key. Figure 4.6 shows the architecture of the cipher. GIFT's round function consists of three basic operations:

1. **SubCells**: During substitution, each nibble is fed into a 4-bit S-box, replacing its value with another 4-bit value

2. **PermBits**: The output of the S-box is permuted. Each of the 4 bits is re-routed by means of hardwiring

3. **AddRoundKey**: For both versions a 128-bit key is used. For the 64(128)-bit version a 32(64)-bit Round Key (RK) is extracted from the 'key state', which is partitioned into 2(4) 16-bit words. Using the RK two bits are XOR'ed with the corresponding nibbles. In case of the 64-bit version, this will be the two LSBs of the nibbles and the middle two bits for the 128-bit version. Figure 4.6 illustrates this for a single round of GIFT-64.

Since only two key bits per nibble are used when performing the XOR operations, each RK is updated and shifted after every round, to ensure that the other part of the key state is used. This is done by performing a 32-bit right rotation. Following this is a 2-bit and 12-bit right rotation, performed on the two MSB bytes and the two bytes thereafter, respectively. Figure 4.6 depicts the complete key state update. In addition to the RK, there also exists a 7-bit Round Constant (RC) which is applied to bit positions n-1, 3, 7, 11, 15, 19 and 23. After every round, the round constant is updated by means of a rotational left shift followed up by two XOR operations between the new LSB, MSB and '1'. The RK and RC schedules are the same for both versions of the cipher.

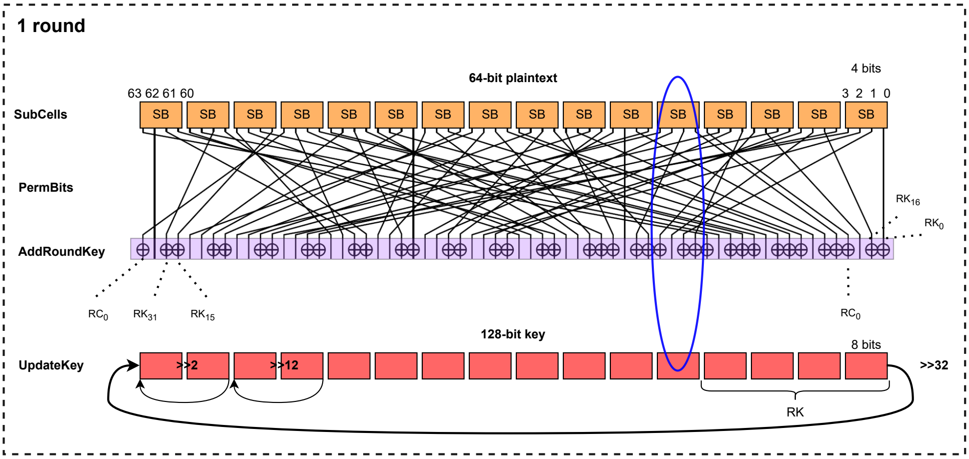


Figure 4.6: Basic architecture of GIFT-64. The encircled section (blue) denotes the single-round encryption of a bit-slice. Illustration adapted from [103].

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$P_{64}(i)$	0	17	34	51	48	1	18	35	32	49	2	19	16	33	50	3
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$P_{64}(i)$	4	21	38	55	52	5	22	39	36	53	6	23	20	37	54	7
i	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
$P_{64}(i)$	8	25	42	59	56	9	26	43	40	57	10	27	24	41	58	11
i	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$P_{64}(i)$	12	29	46	63	60	13	30	47	44	61	14	31	28	45	62	15

Figure 4.7: GIFT-64 permutation mapping. Adapted from [103].

4.5. Overview and Evaluation of Memristor-based Logic Architectures

Any encryption module design will involve a logic computation architecture together with a memory unit. In this section an overview and evaluation of existing memristor-based computation blocks is given. The potential of memristors for lightweight security and IoT applications is being recognized more in recent years [70], [79]–[87], [104]. The development of MLGs and computation blocks is therefore inevitable. Related to this is the development of primitive operations which, ultimately, are the main building blocks for performing computations in digital designs. Examples of such operations are: AND, OR, NAND, NOR, XOR and NOT. By looking at the architecture of these MLGs, one could define four different types:

1. Crossbar: The functionality of a logic gate is realized using a memristor crossbar.
2. Logic: A logic gate is realized using (just) memristors.
3. Hybrid: Combining memristors and CMOS to generate gate functionality.
4. Crossbar compatible: A logic gate is realized using (just) memristors, but could also be integrated in a crossbar structure.

It should be made clear that some architectures can fall under a multiple of these classes. Another distinction can be made between the input/output characterization of an MLG. More specifically, Ielmini and Wong [105] originally define this distinction and it is later also applied in [106]. There are four different characterizations: V-R, R-V, V-V and R-R, corresponding to the input and output, respectively. 'V' implies that an input/output is characterized by a voltage level. Similarly, 'R' indicates an input/output is characterized by the resistive state of the memristors [105]. Besides looking at the architecture of the MLGs, specifications such as the number of memristors, number of steps and delay will be taken into consideration in the subsequent discussion. Notable MLG architectures from literature are discussed below. Note that only two of them, i.e., [107] and [108], discuss power, energy and area overheads.

4.5.1. MAGIC: Memristor-Aided Logic

MAGIC was proposed in [109] and only consists of memristors. It is a sequential logic family consisting of the universal gate set AND, NAND, OR, NOR and NOT. As Figure4.8 shows, the implementation of MAGIC is rather simple. It primarily consists memristors that are connected in series, where one memristor serves as input and the others as output. The number of input resistors corresponds to the number of gate inputs [109], [110]. The method for defining input and output is based on the resistive states, which makes this scheme fall under the R-R category. Ielmini and Wong [105] also define this as 'stateful' because *1)* the logic operation is done in-memory, *2)* it is a true cascadable in-memory operation, and *3)* it relies on the non-volatile states of the memristors. Stateful logic is often preferred because V-V logic is volatile and there is no conversion overhead as opposed to R-V and V-R logic [106]. A MAGIC logic operation requires three steps to be fulfilled. First, the output memristor is initialized to a ohmic value depending on the type of gate. Then, the two input memristors are also initialized to the desired logic input. After initialization, a voltage pulse is applied via a switch which leads to the operation of the voltage divider. Depending on the voltage drop at the output memristor, its resistive state changes, making it a destructive operation; hence, for a follow-up operation, all memristors require to be initialized again by means of a write operation. It should be pointed out that initialisation of the input memristors is done sequentially. So in fact, the entire operation takes (at least) four steps depending on the number of logic inputs. Also, because only the NOR gate can be implemented within memory (i.e. crossbar) this logic family is not considered crossbar compatible.

4.5.2. SIXOR: Single-Cycle In-Memristor XOR

SIXOR is the first stateful XOR operation [107]. Its logic construction is illustrated in Figure4.9. TaheriNejad [107] claims that this operation only requires one cycle, as opposed

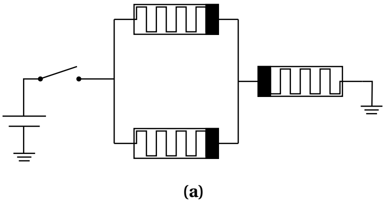


Figure 4.8: MAGIC [109]. a) NOR gate. b) OR gate.

to other stateful architectures; nevertheless, one should also consider the fact that both input and output memristors A/B and F, and auxiliary memristors need to be initialized in some cases. On the other hand, when placing them efficiently within a crossbar the operation could be done in a single step. The power and delay presented in [107] is 20.25 μW and 2 μs, respectively. Even though the author mentions the technology dependency, these numbers are still on the high side for a single operation. A nice addition to this XOR scheme is the 1R crossbar compatibility. However, mapping the MLG into a crossbar may be too difficult. On top of that, 1R crossbar arrays larger than 16 × 16 will introduce sneak paths and leakage in contrast to 1T1R arrays [107].

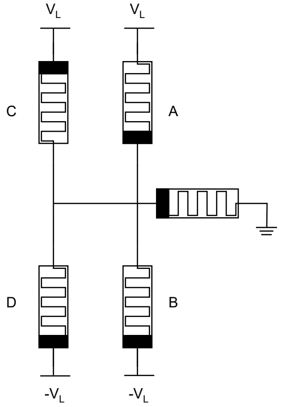
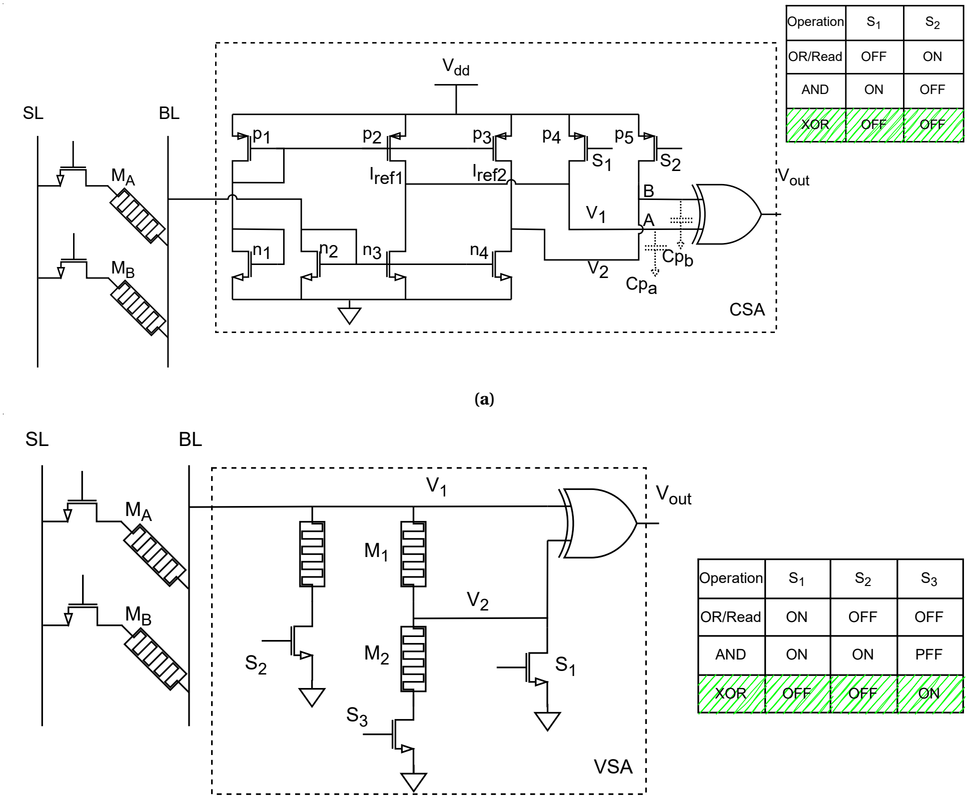


Figure 4.9: SIXOR gate. Adapted from [107].

4.5.3. Scouting Logic

The Scouting Logic family presented in [108] is an R-V based hybrid logic design that offers OR, AND and XOR operations. Unlike the previous two implementations, the logic operations performed with Scouting Logic are non-destructive as the resistive states are preserved across multiple operations. Hence, no switching between resistive states is necessary, which is power efficient. Moreover, a single logic operation only requires one step in which no initialisation and restoring is needed as it primarily consists of a reading operation. The core of the Scouting Logic design varies depending on the sensing scheme one would like to use. [108] propose a Current Sense Amplifier (CSA) and a Voltage Sense Amplifier (VSA) and their respective architectures are illustrated in Figure4.10a and 4.10b.



Operation	S_1	S_2
OR/Read	OFF	ON
AND	ON	OFF
XOR	OFF	OFF

(a)

Operation	S_1	S_2	S_3
OR/Read	ON	OFF	OFF
AND	ON	ON	PFF
XOR	OFF	OFF	ON

(b)

Figure 4.10: Scouting Logic sense amplifiers. (a) Scouting Logic CSA with transistor switch truth table for enabling different gates. (b) Scouting Logic VSA with transistor switch truth table for enabling different gates. Both figures are adapted from [108].

For both architectures the desired operation is enabled depending on the switch that is turned on. CSA generates its reference current using a PMOS and NMOS transistor, while VSA generates a reference using additional memristors [108]. Both SAs use a CMOS XOR gate as a threshold function to pass the proper output pulse depending on the generated reference. So, when applying a read pulse concurrently to selecting the desired operation switch (of CSA or VSA) and M_A and M_B, a current, which is determined by the equivalent input resistance (denoted as $M_A//M_B$), will flow from the Bit Line (BL) to the SA. Depending on whether the current (or resulting voltage) is higher or lower than the gate threshold, the output will become ‘0’ or ‘1’.

One may notice from Figure4.10a that CSA has larger area overhead compared to

VSA. On top of that, CSA has a power consumption of 20 μW while VSA consumes half of that. On the other hand, VSA has a much larger delay of 12 ns while CSA's delay is under the 2 ns [108]. The Scouting Logic design is compatible for 1T1R crossbars which make the overall design less susceptible to leakage and sneak paths. Yet, due to small sensing margins, variability of resistance may cause operation failure. To tackle this, Xie et al. [108] propose a design methodology for creating a more robust version of Scouting Logic CSA/VSA.

4.5.4. 1T1R RRAM In-Situ Boolean Logic

Wang et al. [111] propose a simple 1T1R RRAM Boolean logic scheme which is later also adapted in [81] as an in-situ encryption module for hardware security. The design is proposed as a hybrid standalone MLG but is also crossbar compatible [81], [111]. The proposed design utilizes the V-R method. Although the authors claim it only requires a single step to do the computation, this logic scheme is destructive and therefore needs initialisation and resets. Because of this, three steps are actually required which also influences the time necessary to finish a single operation. Despite the simple structure of this MLG, it supports up to 16 boolean operations [111]. While many of the proposed MLGs lack the possibility to cascade the gates, this design only partially circumvents this problem: for only five logic functions it is possible to directly cascade the output without an intermediate read-out [111]. Figure 4.11 shows an example of how an XOR gate would work. Here, both the gate, drain, and source function as encoded input and the resistive state corresponds to the output [81].

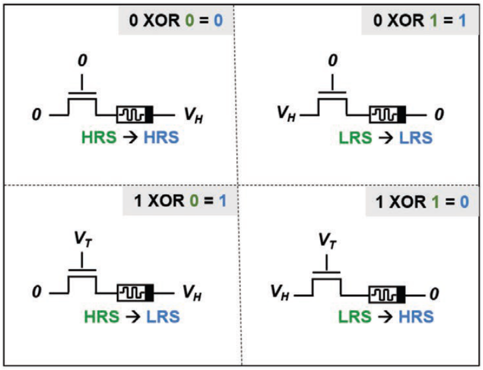
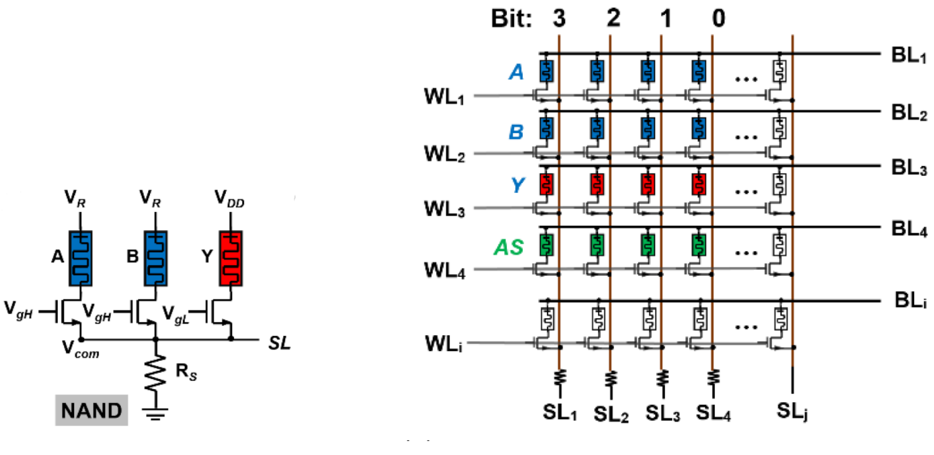


Figure 4.11: 1T1R XOR operation for four different inputs. Adapted from [81]

4.5.5. Stateful 1T1R RRAM NANDs

The design proposed in [106] is constructed for 1T1R crossbar arrays. As such, it addresses the sneak path problem common in 1R arrays [106]. As the name already suggests, the operations belong to the R-R category. The design, i.e. a NAND gate, is a universal MLG because any binary logic function can be realized using a number of NAND gates, even for multiple inputs [106]. With this in mind, the authors demonstrate how to construct an XOR operation using only the proposed architecture shown in Figure4.12.

Also for this design the operations are destructive. Prior to a logic operation, output memristor Y must be initialized to HRS. Depending on whether Y switches, it must be RESET for the next operation. Figure 4.12c shows that in addition to initialisation, performing an XOR operation takes four steps, hence increasing delay. Despite the benefits of this implementation, concocting a simple logic gate may require more elements and bring additional delay. More specifically, the presented 4-bit XOR gate requires a 5 X 4 1T1R array (see Figure4.12b). Also, Shen et al. [106] claim that parasitic capacitance further limits the speed of the logic.



(a)

(b)

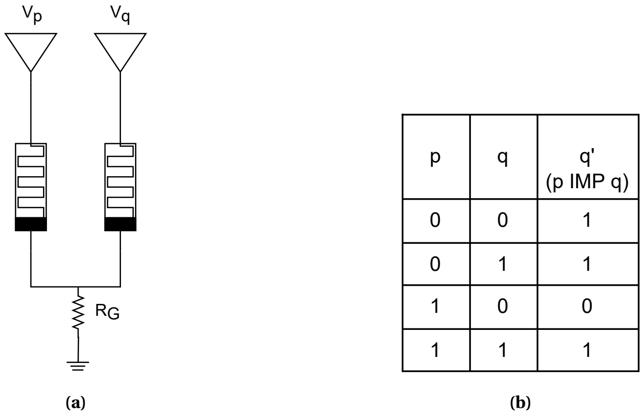
Step	Logic operation
①	*Y = NAND(A,B)*
②	*AS = NAND(A,1)*
③	*AS = NAND(B,1)*
④	*Y = NAND(AS,1)*

(c)

Figure 4.12: 1T1R RRAM NAND. a) A basic NAND circuit. b) 5 X 4 1T1R crossbar configuration for a 4-bit XOR gate. Additional assistant cells (green) are placed for fulfilling the XOR operation. c) Sequence of steps required for performing an XOR operation. Image retrieved from [106].

4.5.6. IMPLY: Material Implication Memristor Logic

IMPLY is another stateful MLG design that can create any Boolean logic. It is considered an addition to the well known operations AND, OR and NOT [110] and its equation can be expressed as "p IMP q" (p implies q) or "if p, then q". Its simplicity enables crossbar integration. Figure 4.13a shows the architecture of IMPLY. The key of performing the operation is by simultaneously applying an auxiliary voltage (V_{cond}) pulse to P and a set voltage (V_{set}) pulse to Q for a conditional switching operation [110]. Depending on the resistive states of both inputs q may switch. Like [106], the downside of this scheme is the sequence length of IMPLY operations to construct a single Boolean function [110]. The authors also discuss the possibility for a multi-input IMPLY operation, but this exponentially increases the number of steps required to perform any Boolean logic operation which, in return, increases power consumption.



(a)

p	q	q' (p IMP q)
0	0	1
0	1	1
1	0	0
1	1	1

(b)

Figure 4.13: IMPLY [110]. a) IMPLY gate architecture. b) IMPLY truth table.

4.5.7. MRL: Memristor-Ratioed Logic

Kvatinsky et al. [112] share a CMOS-memristive logic family named Memristor-Ratioed Logic (MRL). In MRL, the OR and AND gates are made using memristors and additionally a CMOS NOT gate is used for composing a complete logic set while also restoring degraded signals [110]. In contrast to the previous examples, this work utilizes voltages as the logic state variable. The memristors in Figure4.14 are connected in series with opposite polarity and the output node is the common node. Depending on the applied input state, the memristance switches. Even though signal degradation is very minimal when switching from HRS to LRS, it is necessary to perform signal restoration when cascading multiple gates as this degradation may become significant. The authors explain that input size may be extended by means of connecting more elements to the common node. One major benefit of this design is that the result of a single operation is independent of the memristors' initial state. Besides that, computation only takes one step. Nevertheless, MRL uses a linear type of memristor which, according to the authors, is slower than the common threshold type. For this reason they are slower and the switching time depends on applied voltage.

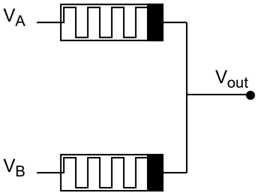


Figure 4.14: MRL. Illustration adapted from [112].

4.5.8. CMOS-Like Logic

The work done in [113] proposes an architecture that is supposed to be a "1-to-1" mapping of the corresponding CMOS transistor circuit. The architecture is "complementary" because it utilizes symmetrical pairs of opposite polarity memristors [113]. The Boolean logic function depends on the topology of the involved memristors and it is possible to make any digital logic circuit. The methodology used is V-V and the output voltage is always a fraction of the supply (read) voltage because the circuit acts as a voltage divider. Because the memristor state may change after operation, this scheme is considered destructive. This design offers the possibility to expand the number of inputs arbitrarily. This does increase the number of required memristors, ultimately increasing the number of steps for a single computation to 2n. This is still better than IMPLY, but cascading is not possible for this scheme. Then there is also the fact that the design of a gate does not necessarily generate less overhead than its CMOS counterpart. This makes the use of such a design perhaps obsolete.

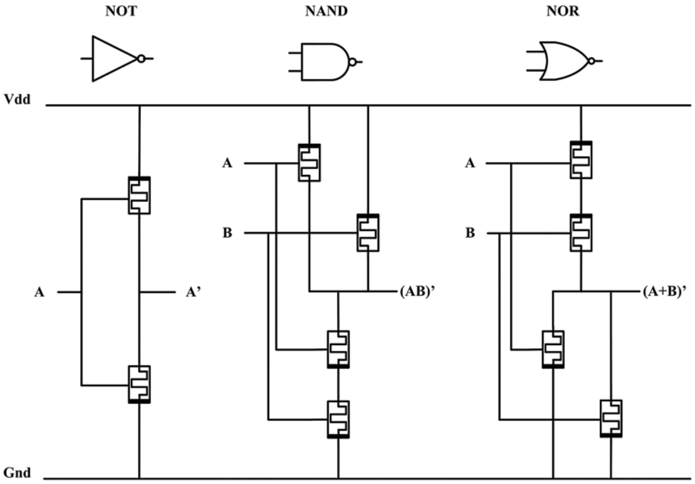


Figure 4.15: CMOS-like logic NOT, NAND and OR gate. Image retrieved from [113].

4.5.9. Parallel Input-Processing Memristor Logic

Up until now, the majority of the discussed designs process input sequentially. In [114] a parallel approach is highlighted. More specifically, an MLG family is presented that enables parallel execution of logic computation while also offering a large set of possible logic gates as opposed to MRL. The number of memristors in a gate depend on the type of gate and the number of inputs. For any of the gates, all the inputs are first summed up. More than two inputs is possible, but the number of inputs is limited by the amount of voltage a memristor can take before being damaged [114]. Like some of the previous designs, the memristors may have to be reset after computation. This design, however, offers a simultaneous reset (using a single pulse) for all the memristors without having to access them separately. Hence, a single computation requires two steps at most. Cascading is also possible, but requires some restoration between stages. Figure 4.16 shows the architecture of different logic gates following this scheme. Figure 4.16a shows the general scheme that is applied to every gate. Figure 4.16c shown an example of the cascading of multiple gates. This scheme is not crossbar compatible.

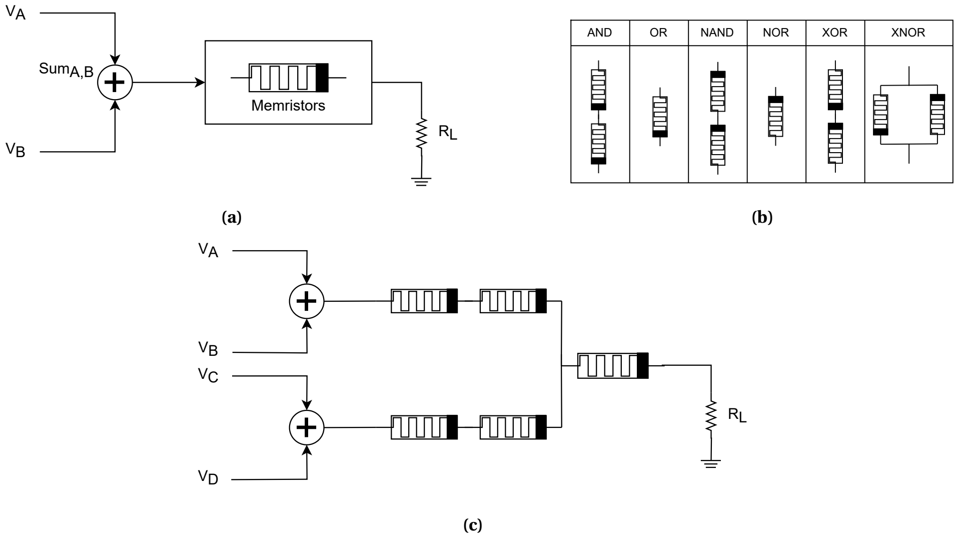


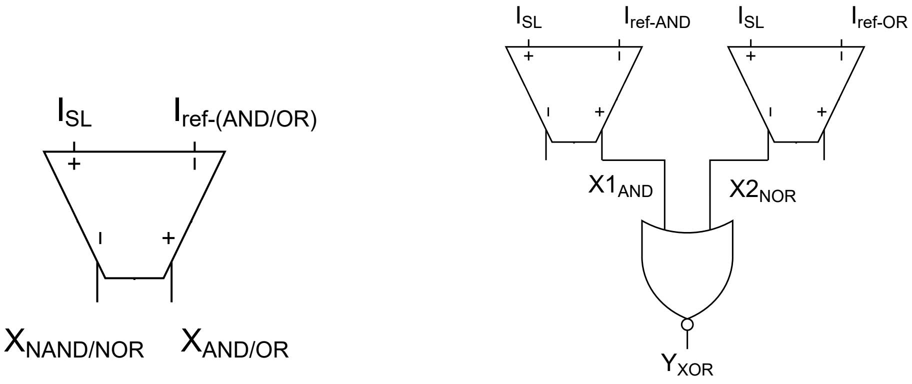
Figure 4.16: Parallel input-processing memristor logic [114]. a) MLG scheme: Input A and B are accumulated and passed through the memristor topology. b) List of synthesizable MLGs. c) Two parallel AND gates connected in series with an OR gate.

4.5.10. Dual Sense Amplifier Crossbar Logic

An alternative approach to all of the above is the utilisation of a (dual) SA setup for the composition of more complex gates or operations such as bitwise XOR and ADD. By combining the reference currents of the NOR and NAND sensing schemes (see Table 4.2 and Figure4.17a), an XOR operation is realized, following the equation $Y_{XOR} = (X1_{AND})$**NOR**$(X2_{NOR})$, as shown in Figure 4.17b; where X1/X2 correspond to the prior output of the respective sensing schemes. The dual SA scheme (DSA) is meant for crossbar operations and follows the R-V method. Due to its simplicity it is able to perform some operations using a single cycle only. However, this comes at the cost of using the two SAs and a single 2-bit NOR gate. To support multiple operations, it requires additional MUXs for selection and a more elaborate global reference generation. Despite the authors specifying the schematic for the SAs, any SA can be used for this scheme and the multi-gate scheme. The work reports high energy and power efficiency improvements compared to a baseline design implemented with a conventional Spin-Transfer Torque (STT)-MRAM. Yet, no exact numbers are discussed for this design.

Table 4.2: Logic '1' truth table

0	0	NOR
0	1	XOR
1	0	XOR
1	1	AND



(a)

(b)

Figure 4.17: DSA crossbar logic. a) Generic (N)AND/(N)OR SA. b) Sensing scheme for single cycle 2-bit XOR operation, adapted from [115].

4.5.11. Assessing The Memristor Logic Gates

Many of the MLGs discussed above are limited due to cascading problems, destructiveness, and long sequences of operations. MLGs implemented in a constrained environment should have very low power consumption and an MLG that requires many sequential resets may therefore be less desirable. As of now, the usage of standalone MLGs such as SIXOR [107], MAGIC [109] and MRL [112] may be too inefficient and excessive to make its use for FFNI encryption justifiable. Crossbar solutions, on the other hand, show good potential for parallel in-memory computing applications. Hence, solutions such as Scouting Logic [108] and DSA [115] are strongly considered in this work.

Preferably one would want to compare numbers such as area, power, energy and delay. Unfortunately, almost none of the papers mention any of those numbers. However, educated estimations can be made, based on the topology and working principle of every scheme. From the evaluation it has become apparent that both Scouting Logic [108] and DSA [115] are front-runners. Alas, due to lack of figures, no direct comparison can be made. Scouting Logic has the advantage of not requiring global reference generation, and its VSA scheme has very low complexity. On the other hand, Scouting Logic VSA shows larger latency and Scouting Logic CSA may have high energy consumption due to static power. Moreover, both still require an XOR gate. DSA is less complex than Scouting Logic CSA and is considered very power efficient and robust. However, it does require additional SA periphery, which may have implication for power and area consumption. Table 4.3 provides a clear overview of all the architectures, types, number of elements, and other attributes.

Table 4.3: Evaluation of MLGs

Information		Architecture								Specs					Attributes								
		Type				Method																	
Work	Year	Crossbar	MLG	Hybrid	CB Compatible	V-R	R-V	R-R	V-V	# of mems	Average Power μW	Energy pJ	# of steps	Op. delay ns	Multiple Bool Op.	Simplicity	Non destructive	No active switching	Single step	No sneak paths	Universal	Multiple (>2) operands	No signal restoration****
[109]	2014	no	yes	no	no	no	no	yes	no	3	-	-	4	-	+	+	-	-	-	-	-	-	-
[107]	2021	no	yes	no	yes	no	no	yes	no	5	20.25	44.55	1	2000	-	-	-	-	+	-	-	-	
[108]	2017	yes	no	yes	yes	no	yes	no	no	2	20 (CSA) / 10 (VSA)	-	1	2 (CSA) / 10 (VSA)	+	+	+	+	+	+	-	-	
[111]	2017	no	yes	yes	yes	yes	no	no	no	1	-	-	3	*	+	+	-	-	-	+	+	-	-
[106]	2019	yes	no	yes	yes	yes	no	no	no	3 (at least)	-	-	4	*	+	-	-	-	-	+	+	+	+
[110]	2016	no	yes	no	yes	no	no	yes	no	2 (at least)	-	-	3	**	+	+	-	-	-	-	+	+	-
[112]	2012	no	yes	yes	no	no	no	no	yes	2 (at least)	-	-	1	***	+	+	-	+	+	-	+	+	-
[113]	2014	no	yes	yes	no	no	no	no	yes	2 (at least)	-	-	2n (operands)	-	+	-	-	+	-	-	+	+	
[114]	2014	no	yes	yes	no	no	no	no	yes	1 (at least)	-	-	2	-	+	+	-	-	-	-	+	+	-
[115]	2017	yes	no	no	yes	no	yes	no	no	2	-	-	1	*	+	+	+	+	+	+	-	-	

[*] Depends on initialization, writing and reading sequence.

[**] Depends on logic operations.

[***] Depends on applied voltage.

[****] If applicable.

5

BUILDING AN IN-SITU MEMRISTOR-BASED LIGHTWEIGHT BLOCK-CIPHER

This chapter will focus on the proposed design, its sub-modules, and the implementation. A comparative analysis (against a CMOS-based implementation) will be presented, followed by a detailed discussion.

5.1. Building a memristor-based SPN Cipher

This section will discuss the road towards implementing the GIFT SPN cipher using memristors. By means of unrolling the devised construction of each of the operations discussed in Section 4.4, a clear view of the entire composition of the proposed design is achieved.

5.1.1. Design Approach and Assumptions

The general idea is to take advantage of the bit slicing topology of GIFT [103] and compress all the operations shown in Figure 4.6 into one lightweight module. Frequent switching of memristors could introduce reliability issues and also shorten the lifetime of the memristor block significantly [70], [79], [82], [87], [94]. So the incentive is to eliminate writing/reset operations where possible. In the previous chapter it was pointed out that memristor technology has not matured sufficiently for the implementation of cascaded MLGs. Thus, creating a memristor-based AND-OR tree architecture, inspired of traditional CMOS-based AND-OR trees, is is not the focus in this work. More so, utilizing a memristive crossbar structure is preferred given the possibility of high-density, in-memory (in-situ) computations. As stated earlier, due to its simple structure, GIFT is a proper candidate for a crossbar based implementation. Moreover, the first operation *SubCells* is realized using 4-bit S-boxes, which are normally implemented using a Look-Up Table (LUT). Hence, this gives reason to follow the same LUT approach using a memristor crossbar. So, to summarise:

- Mapping the three GIFT operations and key scheduling (see Section 4.4) to a memristor crossbar, makes it possible to execute a GIFT encryption round, for one slice (encircled in Figure 4.6), using a single 'read' operation.
- Only at the start of an encryption session would a 'writing' operation be required to program the constant/key-bit values.
- Each crossbar implementation covers the 40 rounds of encryption.
- By minimizing switching activity and by mapping all the operations to the crossbar, smaller energy consumption and footprint may be achieved.

Before diving deeper into the design, a couple of assumptions must be established. First of all, for this design project the 128-bit GIFT version is considered. Nonetheless, it can be easily configured to the alternative version when desired. Second of all, private session keys need to be created and transmitted to perform the encryption. To do so in a secure and efficient manner, a BCC protocol such as SecureEcho will be assumed. It assures both secure key transmission and ZPD as explained in Chapter 3. This enables to focus all the effort towards the design of the encryption module only. Furthermore, this project assumes the robustness of the GIFT cipher as detailed in [95], [103], [116]. Lastly, to demonstrate the applicability of memristors in lightweight encryption, the design and evaluation of the slice marked in Figure 4.6 will suffice. Figure 5.1 shows the top view of the proposed design. To the best of knowledge, no other work has proposed a memristor-based LWBC at the time of writing.

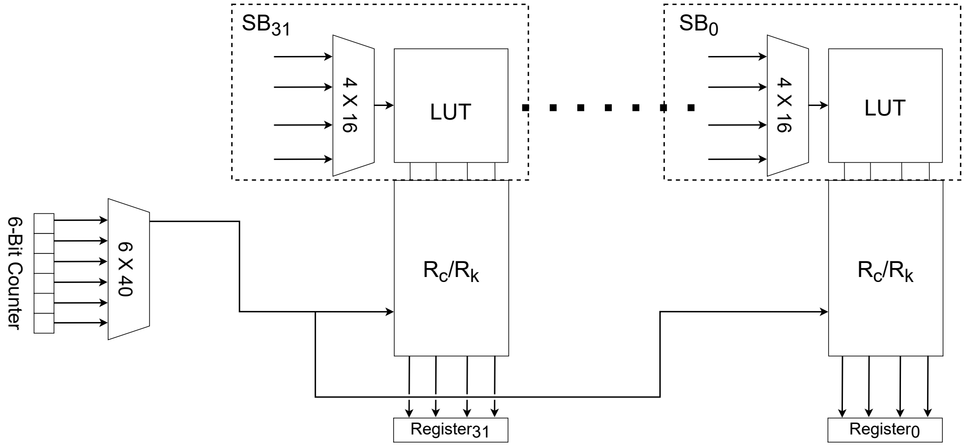


Figure 5.1: Top view of the memristor-based GIFT cipher.

5.1.2. S-BOX

The first operation in the encryption process is substitution of the plaintext nibble with a 4-bit sequence generated from the inverted S-box. For this implementation and its subsequent schematics, the S-box will be referred to as *SB* (as done in Figure 4.6). For the differential properties, linear properties and other S-box heuristics please refer to [103]. Table 5.1 shows the substitution specifications that are used for the 4-bit S-box. The implementation of a memristor based LUT has been done before, as presented in [117]–[119]. The question that remains, is what the suitable type of crossbar structure would be?

Table 5.1: Substitution map of S-box SB as specified in [103].

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
SB (x)	1	a	4	c	6	f	3	9	2	d	b	7	5	0	8	e

Passive crossbar arrays (1R) are an attractive solution for high-density and low power integration. However, their weaknesses include current sneak paths and floating state issues during reading and writing operations. As highlighted already, this leakage becomes more severe when increasing the number of cells [81], [107], [111], [117]. An established answer to this problem is the nTnR structure. Most common is the 1T1R structure, but 2T2R has also shown major benefits [78]. However, as the name already suggests, 2T2R is twice the size of 1T1R. Mapping SB into a 1T1R crossbar results in the structure shown in Figure 5.2. The SB unit has 16 4-bit values, corresponding to Table 5.1, and this is translated in 16 rows of four memristor and transistor elements, where each memristive cell represents a '0' or '1' bit. So, given a 4-bit input, an address decoder will select one of the 16 WLs (in correspondence to Table 5.1) by applying a voltage to the transistor gates of that row. The details of the crossbar operation will be provided in Section 5.1.5.

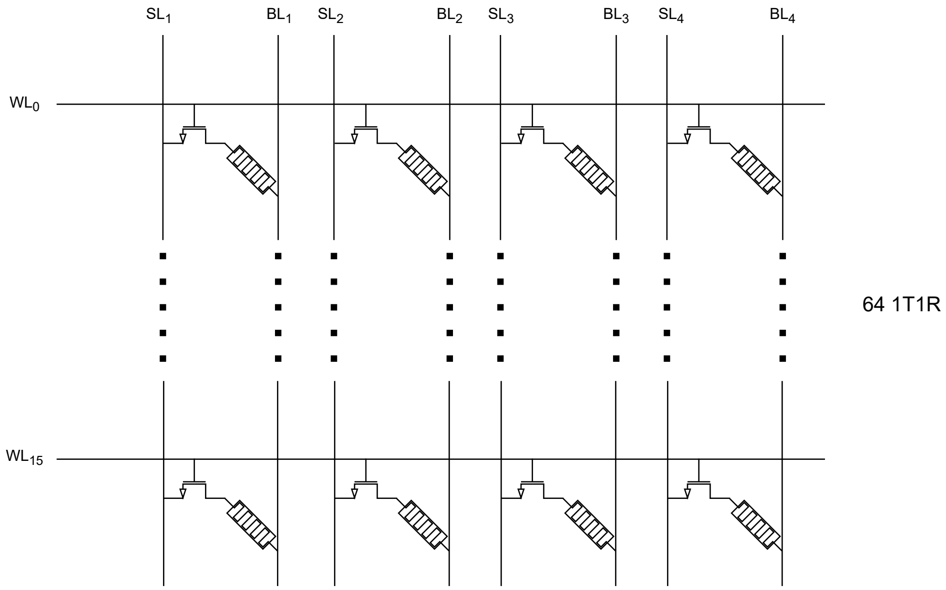


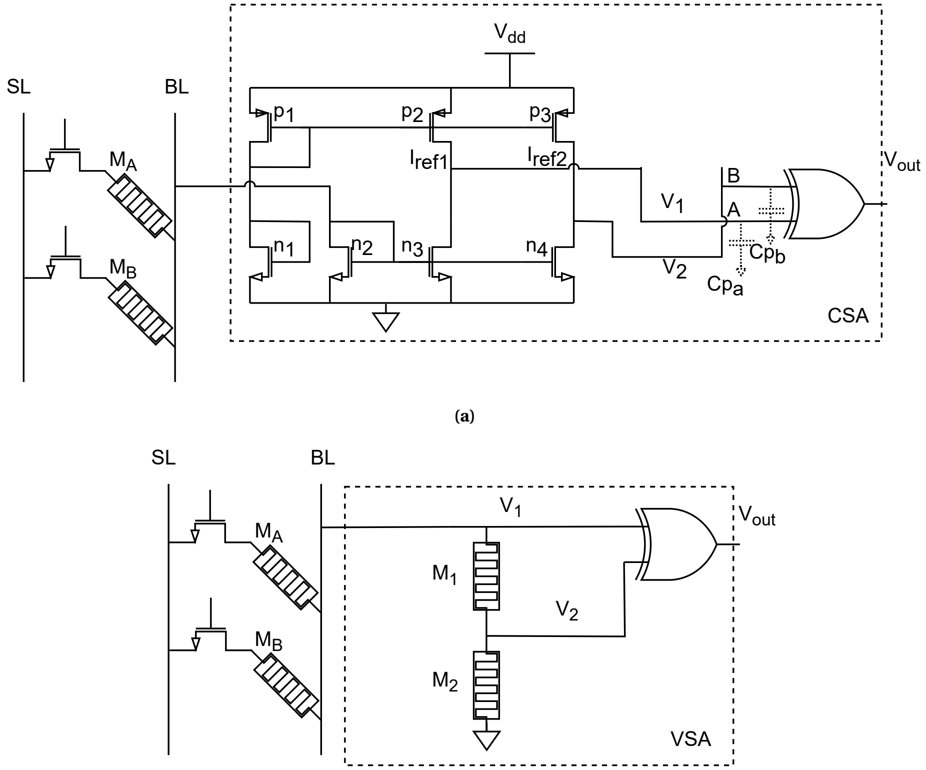
Figure 5.2: The 1T1R S-Box LUT: WLs are used to select the row, SLs for selecting the columns and BL for evaluating ‘read’ operations.

5.1.3. XOR operation

The second operation that is part of the GIFT encryption, is *AddRoundKey*. To do this in an efficient way, advantage should be taken from the crossbar structure proposed for the SB unit. In Section 4.5 it was disclosed that finding the proper MLG is crucial for achieving low energy and area consumption. Not many of the presented MLG schemes offer the possibility to perform an XOR operation, which is required in AddRoundKey. While there are schemes that offer this possibility by means of performing a sequence of MLG operations, these are avoided for the fact that the total delay and power accumulates very quickly. More so, it is highly preferred to apply XOR operations that are stateful. For this reason Scouting Logic [108] and DSA [115] can potentially be a good fit. Hence, both these sensing schemes are considered for the XOR implementation.

Scouting Logic XOR

The shaded rows in the tables in Figures 4.10a and 4.10b shows that all switches are cut off (except for S_3) to realize the XOR operation. Since the remaining switches are not used, the design may be simplified to fit the requirements of the *AddRoundKey* operation. For Scouting Logic CSA this means that the unnecessary reference switches are removed, as illustrated in Figure 5.3a. For Scouting Logic VSA the unused memristors are removed as well, leaving just the voltage divider configuration formed by M_1 and M_2 (see Figure 5.3b. Even though both sense amplifiers fulfill the same functionality, their respective performance shows significant differences. In Section 4.5.3 it was highlighted that the CSA-based Scouting Logic shows the lowest delay to execute an XOR computation. On the other hand, VSA-based Scouting Logic consumes a lot less power. Especially when adapting it to the needs of the GIFT cipher, the power consumption could become even less. Logically, it would then make sense to go for Scouting Logic VSA. Nonetheless, both versions are considered to verify this. Both Scouting Logic schemes still follow the same operation principles that were detailed in Section 4.5.3.



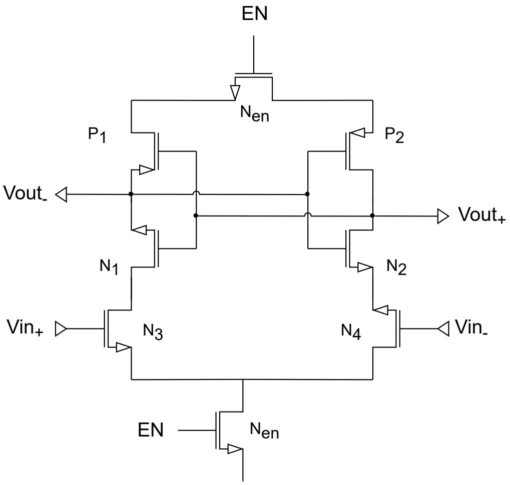
(a)

(b)

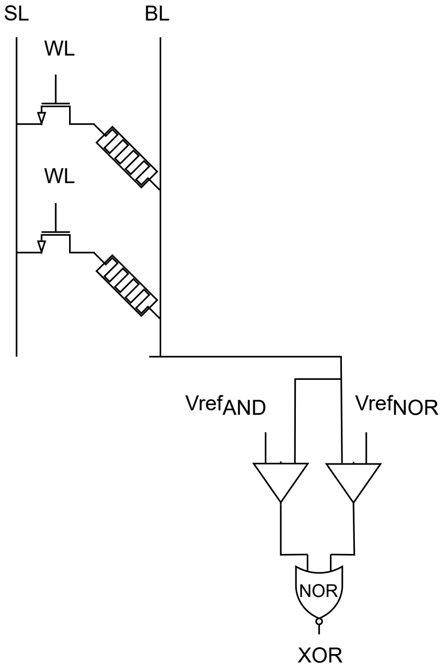
Figure 5.3: Modified Scouting Logic sense amplifiers for the 1T1R GIFT-128 cipher XOR operation. (a) Scouting Logic CSA. (b) Scouting Logic VSA.

DSA XOR

Section 4.5.10 discussed the principles of DSA and its XOR configuration (see Figure 4.17b). In [115], current mirrors are used in the crossbar, to drive the SAs. For the design of the 1T1R-GIFT, a voltage-based SA will be used so that the current mirrors are not necessary anymore, which results in higher energy efficiency. The schematic for the voltage-based SA is adapted from the SA proposed in [115], with minor alterations to target only XOR functionality (see Figure 5.4a).



(a)



(b)

Figure 5.4: DSA XOR scheme: a) Voltage-based sense amplifier. b) DSA 1T1R sensing scheme.

5.1.4. PERMUTATION

Permutation can be done following the same approach as proposed in [103]. Having a more passive design with minimal switching activity, can be enforced by taking a slightly different approach, which consists of two parts:

1. The web of connections in Figure 4.6 show the physical permutation. Figure 5.5 corresponds to the mapping of these wires for each round. Each nth bit of every nibble is connected to the nth bit of another nibble that drives the SB of the next stage. For example, considering the permutation scheme in Figure 4.7, bit 1 of the rightmost SB (in the first round) is connected to the input of bit 1 of the SB encircled in Figure 4.6. This applies to every output and input of the SBs that are connected during permutation. The same holds for GIFT-128.
2. Figure 4.6 shows that the mapping of the XOR operations never changes; it is always the same bits that are subjected to the XOR manipulation. Knowing this and the permutation map in Figure 4.7, it is possible to anticipate the RK and RC bit additions for every round if the encryption key is known. More specifically, instead of performing RK/RC updates every round, these values are pre-computed (offline) and arranged using the permutation table. During initialisation of the encryption module, these values will be uploaded only once to the memristor-based GIFT cipher; saving overhead in the process due to removing the active key schedule operation in its entirety.
3. Each slice now contains all the permuted key-bit values for every round of encryption. At the start of a new encryption round, the output of the slice will be fed back to its input, instead of routing it to a different slice.

5.1.5. ROUND ENCRYPTION

The means of incorporating the permutation and keyscheduling approach is illustrated in Figure 5.5. In addition to the crossbar proposed in Section 5.1.2, a 40 × 2 (40 × 3 for the few nibbles with additional RC) 1T1R crossbar will be connected in series. With each of the 40 rows (WL_{16} to WL_{55}) encoding the RK (and RC) bits, the permutation and key/constant scheduling will be mapped for the complete encryption of a 128-bit plaintext. The crossbar in Figure 5.5 clarifies how the architecture is divided into a *substitution* section (top) and an *RC/RK encryption* part (bottom). For the bottom part, the bitcells are labelled with $\mathtt{R}_n\mathtt{K}_{i(+32)}$, where `n` stands for the round number and `i` for the RK bit as explained in [116]. Here, every third bit of the nibble is encrypted with the `i+32`th roundkey bit at each round, while every second bit is encrypted with the `i`th roundkey bit. It can also be seen how the XOR SAs are connected to the bottom of the crossbar. This connection is the same for Scouting Logic as for DSA. Figure 5.5 illustrates the entire slice operation, including XOR addition, for a single round. A round of encryption goes as follows:

- Following the mapping presented in Table 5.1 the corresponding WL will be driven by a voltage pulse, upon which the NMOS switches in that row will close and the respective memristors are selected. The RC/RK memristors, corresponding to the first round, are selected. The activated rows and columns are marked with blue.

- A read pulse is generated for each SL, allowing current to flow through the memristors and the BL, into the XOR and RO SAs. This read pulse is marked in red.
- With the desired rows selected, an XOR operation is performed between bitcell 2 and 3 of the top part and the bottom part. Depending on resistive states of these bits, a '0' or '1' pulse will be generated at the output. Bitcell 1 and 4 are being read out without performing an XOR operation.
- The output is temporarily stored in the output registers.
- In the next round output is fed back again to the input of the same SB.
- All of the above steps are performed at the same time.

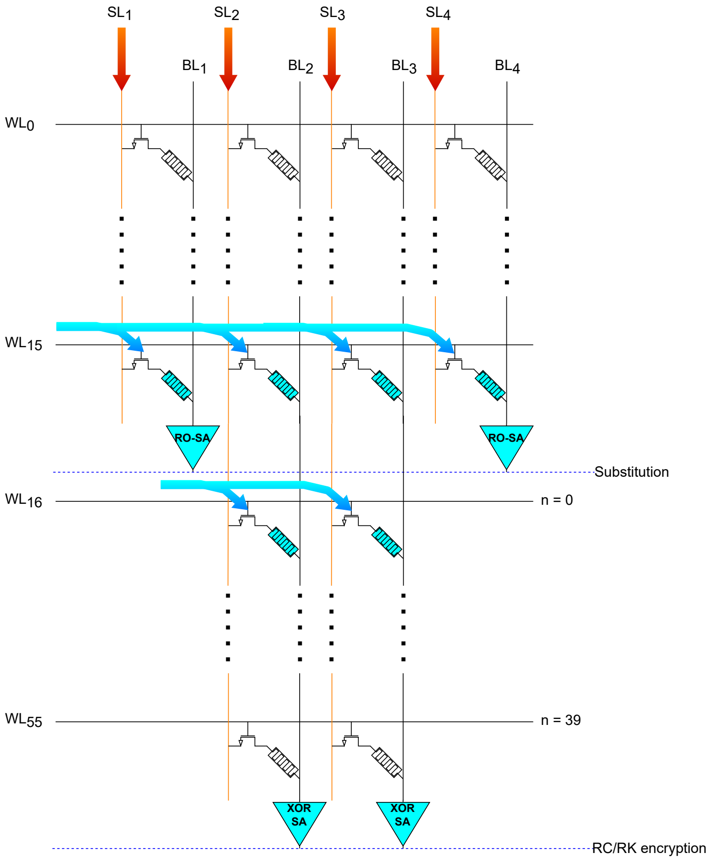


Figure 5.5: The slice architecture of the 1T1R GIFT cipher and operation sequence corresponding to a round of encryption.

5.1.6. CROSSBAR ADDRESS DECODERS

The purpose of the address decoders in Figure 5.1 is to drive the desired WLs in the crossbar, resulting in the selection of the memristors embedded in the respective row. Looking at Figure 5.1 it becomes apparent that the address decoders can cause quite a bit of overhead. This is because this GIFT-128 design requires 32 4-to-16 address decoders and a single 6-to-40 address decoder. Hence, it is important to design address decoders with minimal overhead. Many different designs were considered. For example, one idea was to decode the bits by means of a sequence of PMOS and NMOS transistors connected in series. Each one of them would correspond to the expected input bit. An illustration of this idea is shown in Figure 5.6. The disadvantage of such a circuit is that it is completely analog, which means the transistors have to be sized precisely. Moreover, the caused voltage drop across the transistor would become an issue taking into consideration the transitor overdrive. Consequently, a repeater would have to be placed to enforce enough input drive for the WLs. Nevertheless, this is not an optimal solution. A better solution is to go for a complete logic scheme as with such schemes the smallest possible transistor sizes can be used.

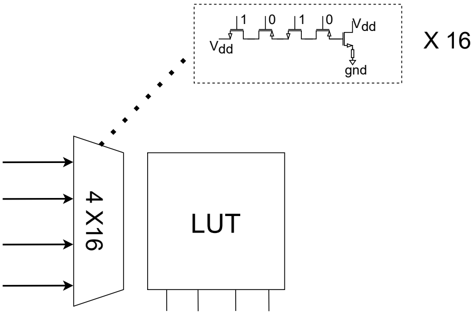


Figure 5.6: Analog 4-to-16 decoder.

There is another well-established approach for decoding addresses in SRAMs [120], [121]. This approach is relatively lightweight, reliable, and can also be applied to memristor crossbar arrays. The main idea is to have a decoder logic scheme, consisting of AND, NAND, some inverters and/or NOR gates. In addition to that there are complement inputs that are necessary for realising the decoder scheme. This would mean, for example, that instead of having four inputs, the decoder will have eight inputs which are segmented in two parts of 2 × 2. The next two parts will clarify the implementation details concerning the two proposed decoders for the memristor crossbars.

4-TO-16 ADDRESS DECODER

Keeping in mind the schemes applied in [120], [121] various design iterations took place. One of the ideas was to reduce the transistor count by implementing PMOS/NMOS transmission gates in combination with the NAND-NOR alternate stages. This is shown in Figure 5.7. Unfortunately, upon simulation this design did not possess the required voltage strength for driving the WLs. Next to that it did show unwanted behaviour and non-idealities such as leakage and untimed switching. For these reasons, the transmission gates have been completely omitted. The final design of the 4-to-16 address decoder can be seen in Figure 5.8. The decoder has an MSB part and an LSB part corresponding to the two MSBs and two LSBs of the inputs nibble. The complementary inputs are indicated with an apostrophe. It should be noted that the gates in the final stage must be sufficiently large in order to properly drive the WLs. All the other gates may be of minimal size. All the details will be explained in Section 5.2.2.

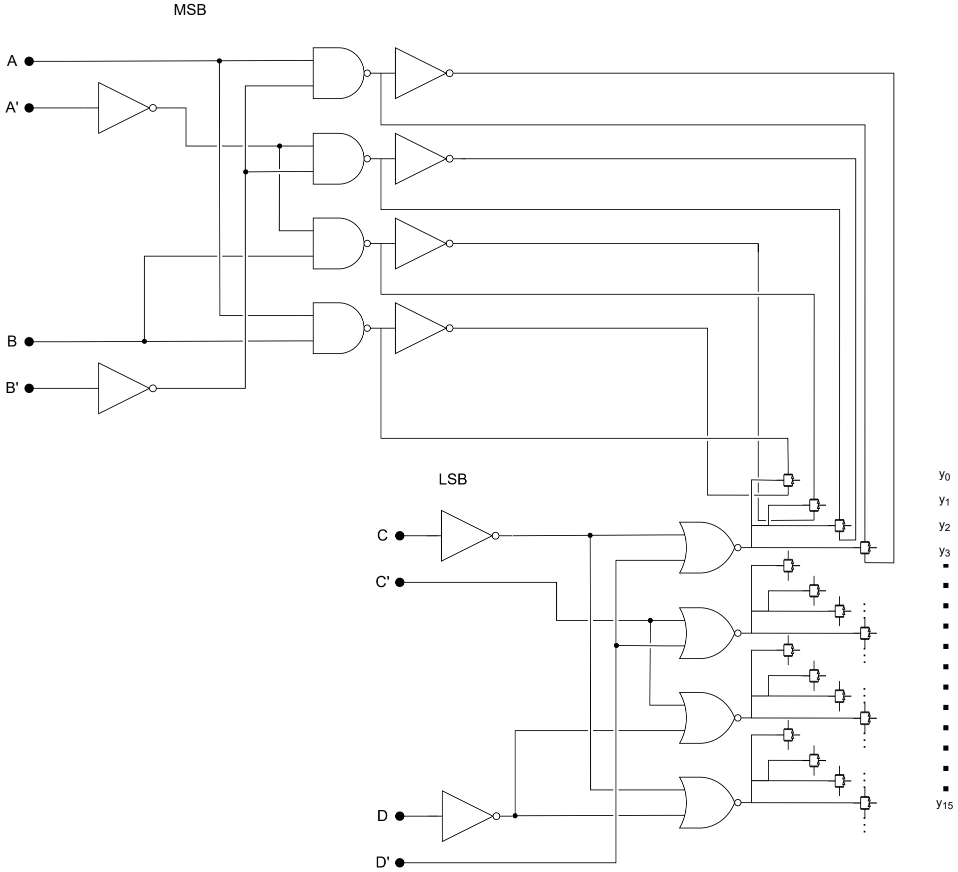


Figure 5.7: NAND/NOR stages combined with transmission logic.

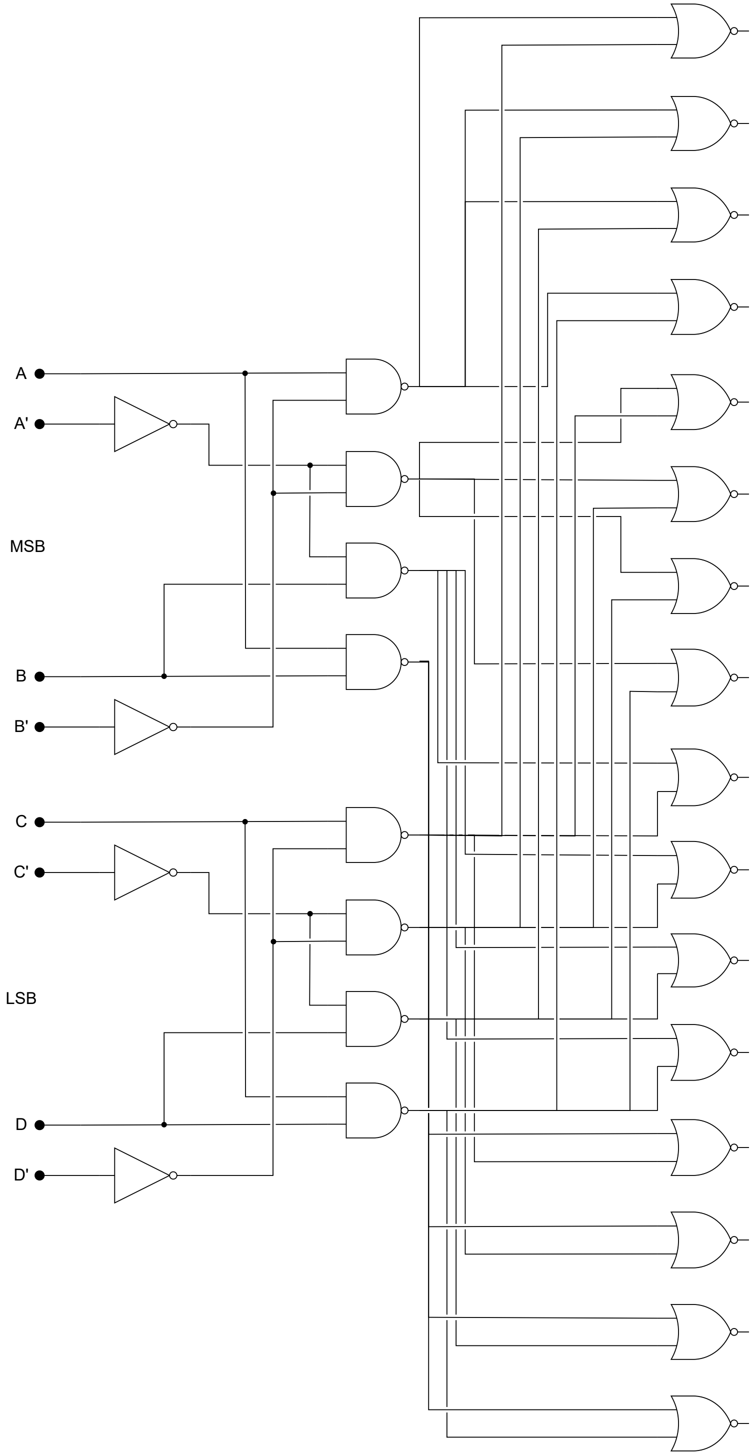


Figure 5.8: 4-to-16 address decoder.

RC/RK Selector

The entire encryption architecture requires a single RC/RK selector for selecting the WLs in all the 32 RC/RK units shown in Figure 5.1. As with the previous address decoder, multiple iterations took place. The goal of this selector is to select a WL every round, starting from round 1 and incrementing up to round 40. The first idea was to use shift registers for selecting each WL by an increment. However, after careful inspection it was found that having 40 shift registers is more costly than using a 6-bit counter that drives a 6-to-40 address decoder. This is mainly because these shift registers are constructed using D-type Flip Flops (DFF) [122]. A common DFF consists of 38 transistors [123], so having 40 of these units accumulates to *1520* transistors. For construction of a counter, a T-type Flip Flop (TFF) is sufficient since it performs a "toggle" operation. Constructing a simple 6-bit counter like the one shown in [124] would accumulate to 264 transistors. The schematic of the 6-bit counter is summarized in Figure 5.9. The address decoder that is driven by the 6-bit counter is built using the same design principle as the 4-to-16 address decoder. The only difference is an additional stage of NOR-gates and the inclusion of another 2 × 2 block for inputs E, F, and their complements (see Figure 5.10).

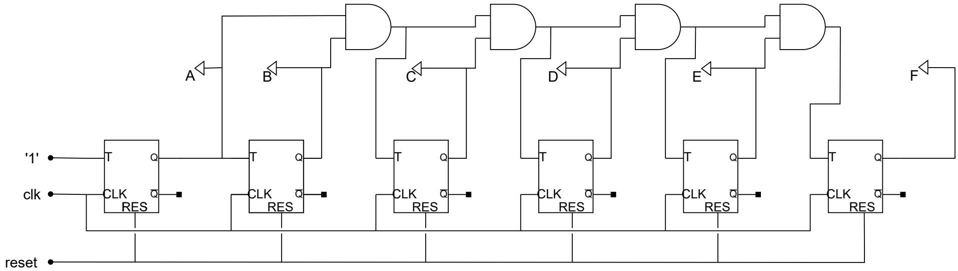


Figure 5.9: Circuit of the 6-bit counter.

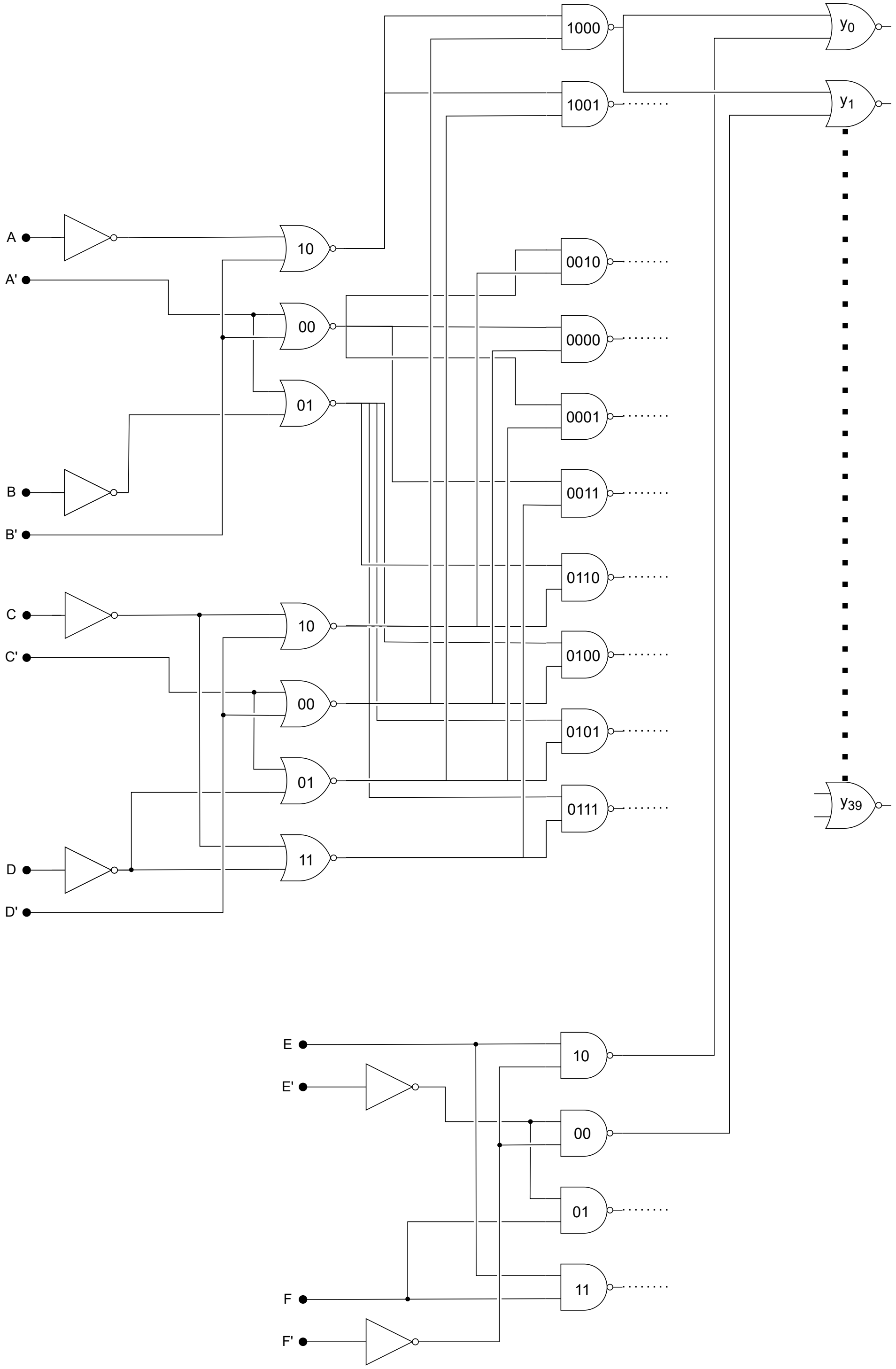


Figure 5.10: 6-to-40 address decoder.

With the above two designs the number of transistors amounts to *520*, which is a lot less compared to using shift registers. As with the 4-to-16 version, the final stage of the decoder must be of sufficient size to supply enough drive to the respective WLs. All other gates in this scheme may be of minimal transistor size. Details about sizing will be highlighted in Section 5.2.2.

5.2. Implementation, Results and Comparison

This section will describe the verification of the proposed design. Then the results, extracted from transient analysis, will be presented and used for the the quantitative comparison between the proposed architectures and a CMOS-based implementation.

5.2.1. Setup and Method

The memristor-based GIFT cipher is constructed using Cadence Virtuoso V.6.1.8 and simulated with Spectre. The design is realized with the TSMC 40nm library. Performance analysis is done using the Spectre calculator tools. The area calculation is based on the bitcell layout, depicted in Figure 5.11, and the gate properties provided by the technology library. Of course, to show how the memristor-based cipher holds-up against a CMOS-based implementation, a comparative analysis has to be made between the two. To do so, an open source GIFT-128 implementation [125] has been adapted to fit the requirements for the comparison. This implementation is written in VHDL code and verified using the Xilinix Vivado design suite simulator. After verifying the implementation, a *Value Change Dump* (VCD) file was generated during simulation, containing the switching activity of the implementation. Together with a written *Unified Power Format* (UPF) file, the RTL was fed into Synopsis SpyGlass [126] to generate accurate area and power estimations using the TSMC 40nm library. Three versions of the 1T1R-GIFT are implemented, using different XOR schemes: Scouting Logic CSA XOR, Scouting Logic VSA XOR, and DSA XOR. For simplicity, the implementations will be referred to as SL(C/V)-GIFT and DSA-GIFT, respectively. The design parameters for both the memristor-based (1T1R-GIFT) and the CMOS-based implementation (CMOS-GIFT) are summarized in Table 5.2. For SL(C/V)-GIFT and DSA-GIFT these parameters are the same.

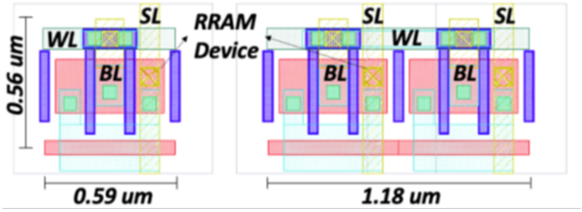


Figure 5.11: Layout of the used 1T1R bitcell. The image on the left is for a single bitcell. The image on the right depicts two bitcells. Image adapted from [78].

Table 5.2: Design parameters.

Implementation	CMOS-GIFT [125]	1T1R-GIFT (this work)
Technology	40nm	40nm
Memristor model	n.a.	HfO_2 *
Operating voltage	0.9	0.9
Operating freq.	10 MHz	10 MHz
t_{pH}/t_{pL} **	10 ps	10 ps

* JART VCM v1b memristor model from [92].
** Rise/Fall time

5.2.2. Implementation

The simulation models are described by SPICE netlist and entail a single slice as highlighted in Section 5.1.1. The 1T1R model is adapted accordingly from [78], which uses the HfO_2-based memristor from [92]. The structure of a single bitcell is described in Figure 5.12. In this model, non-idealities such as wire resistance/capacitance are considered. These are illustrated as capacitive and resistive elements. Implementing a single row for the SB (RC/RK) unit requires four (two) bitcells to be connected to the same WL. By means of a simple script, a multitude of these rows can be connected to each other; SL_B and BL_B of one row are connected to SL_T and BL_T of the next row, respectively. In the same manner, the SB unit and the RC/RK unit are connected. The final row of the RC/RK unit is connected to the XOR SAs via SL_B and BL_B. The interconnection of all the rows contains a small wire resistance as well. In the original work [103], the implementation runs at a frequency of 10 MHz. The same will also be done for these implementations. All the CMOS-based logic that is used in the designs, is implemented using the cells provided by the TSMC standard library. In general, minimal size is used for the gates. However, the gates in the final stages of the decoders are four times larger to ensure sufficient drive strength for the WLs.

The following sections will provide an overview of the implementation details of SL(C/V)-GIFT and DSA-GIFT.

SL(C/V)-GIFT

The first implementation, i.e. SLC-GIFT, is following the architecture detailed in Figure 5.3a. However, during simulation it was found that this structure significantly increases the power consumption of the design. More specifically, each Scouting Logic CSA consumes around 12 μW. With this in mind, Scouting Logic CSA was not considered an option anymore and Scouting Logic VSA was used instead.

In [108], the authors explain that, whilst operating Scouting Logic VSA, the resistive states of the memristors will be retained. Hence, as with all other memristors in this design, they only need to be programmed once. For the voltage divider to work properly, the acting memristors need to be scaled accordingly to guarantee an output of (at least) $0.6V_{DD}$ and $0.4V_{DD}$ for logic ‘1’ and ‘0’, respectively. Since this work utilizes a different structure and technology, the scaling rules, as stated in [108], cannot be applied. For example, the original Scouting Logic VSA requires M_1 and M_2 to be $2 \times LRS$ and $2.5 \times LRS$,

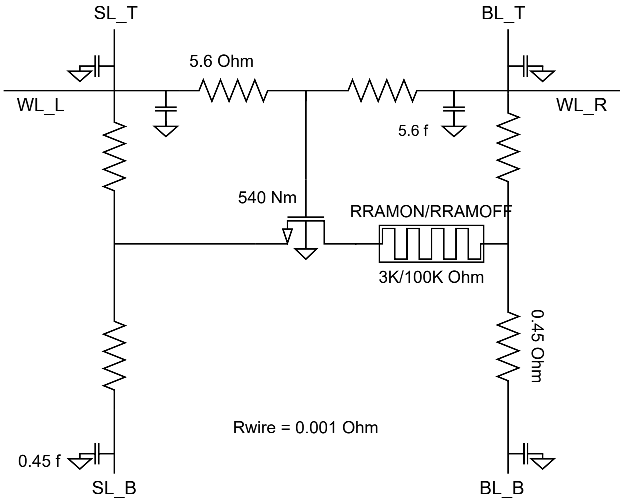


Figure 5.12: Used 1T1R structure based on [78].

respectively. In doing so for the proposed cipher, the small resistive values will result in V_1 and V_2 not passing 0.20-0.30 V, which is way below the required threshold value of 0.45 V. Moreover, keeping the 1:1.25 ratio between the two memristors in Scouting Logic VSA results in too big of a margin between V_1 and V_2. Because of this, V_2 will always stay below the threshold, resulting in operational failure of Scouting Logic. Finding the correct resistive values may quickly lead to an exhaustive search. So, using a simplified model of the (to be used) memristor crossbar, the proper resistor values could be tested. For the final implementation it was found that SLV-GIFT performs well when using 2 kΩ and 250 kΩ for M1 and M2, respectively. Earlier it was clarified how XOR operations are only performed on the middle two bits of each nibble. While this requires the architecture as shown in Figure 5.3b, the remaining LSB and MSB bits just need to be read out using a read-out SA. For this purpose, Scouting Logic VSA is downsized to just a single memristor. More specifically, the truth table in Figure 4.10b shows that a read operation is performed by only selecting S_1. As a result, only M_1 is used and the XOR gate in the SA can be replaced by an OR gate. For the read-out SA, M_1 is programmed to 550 kΩ.

After putting together all the components, a transient analysis is conducted to verify correct behaviour of the encryption slice. The goal of simulation is not to verify the security of this cipher, as this has been done extensively in other works [103], [116], [125]. Figure 5.13 shows the waveforms that summarize the functionality of the design. What is plotted here is a round of encryption and the respective signals: 4-bit input (`ABCD`); 4-bit output, from LSB to MSB, coming from the read-out (RO) SAs and Scouting Logic XOR unit; the selected WL after substitution (`WL11`). After a short initialisation sequence, the first round of encryption starts. An input bit sequence '1010' is fed to the the address decoders. Following the mapping in Table 5.1 this input results in the substitution value '11'. The plot shows that this is indeed the case and WL_{11} is enabled. With the RK bits set to '0', the captured output will be '11' (1011), as shown in the plot. In the second round of encryption the value of ‘1’ (0001) gets encrypted. After substituting with ‘10’ (1010) the middle two bits are XORed with the RK bits set to ‘1’, which results in the output value ‘12’ (1100). It can be observed that a 1⊕1 operation requires some additional time as the BL takes longer to discharge in this case (11.05 ns).

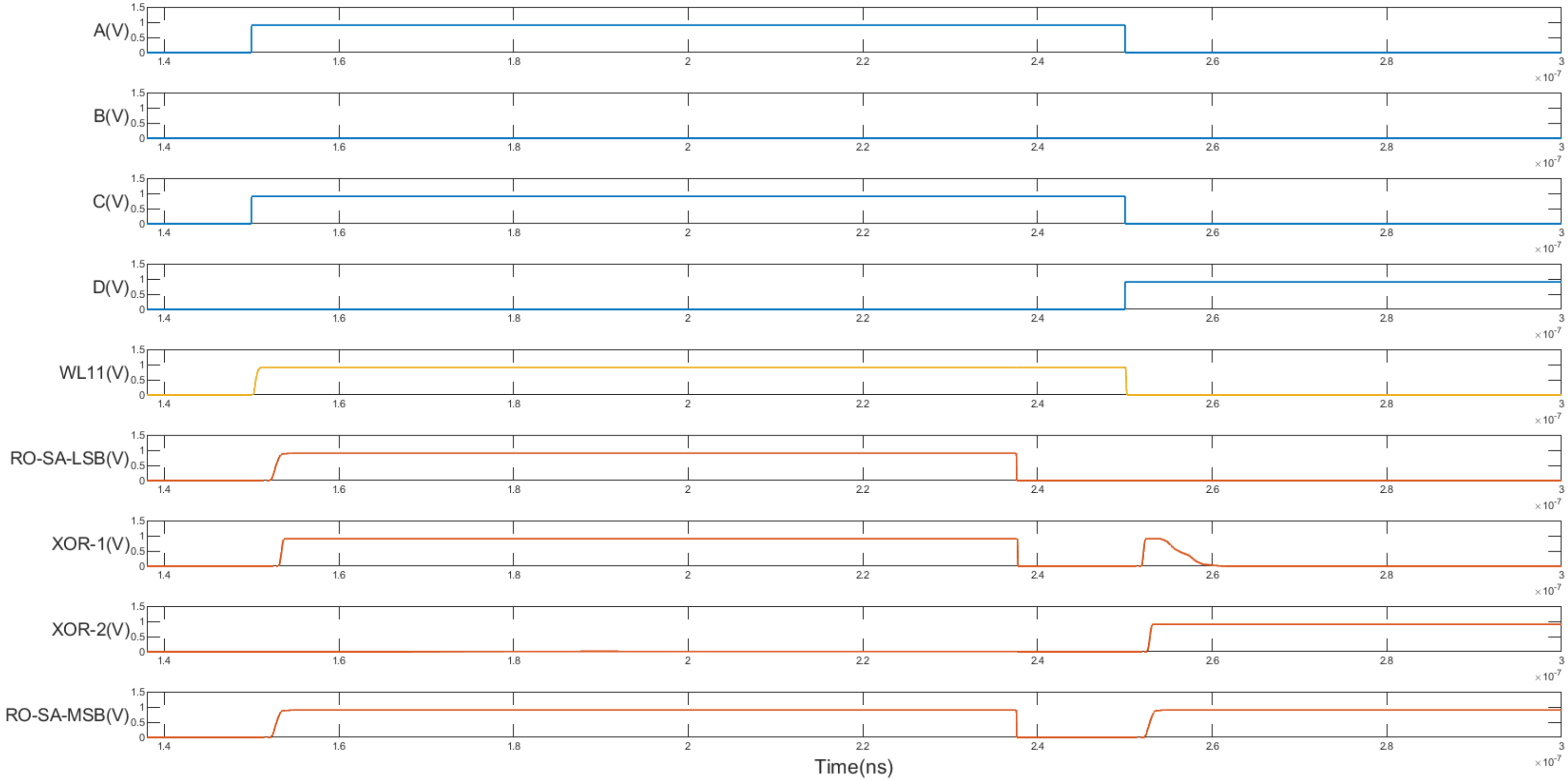


Figure 5.13: Transient response of the implementation.

DSA-GIFT

For DSA-GIFT, the crossbar BLs include a transistor for power gating, which reduces leakage current, resulting in lower power consumption of the bitcells. Other than that, DSA-GIFT has the same 1T1R-GIFT architecture and functionality as SLV-GIFT. Similar to SLV-GIFT, only two XOR sensing schemes are required. This means that for the LSB and MSB, only a single SA suffices. For the AND and NOR SAs, a constant voltage reference of 0.45 V and 0.43V is used, respectively. For the RO SAs a reference of 0.43V is used as well. As mentioned earlier, the implementation for the SA (see Figure 5.4a) is adapted from [115]. However, as only XOR operations are required, the SA is simplified a bit.

5.2.3. Simulation Results and Comparison

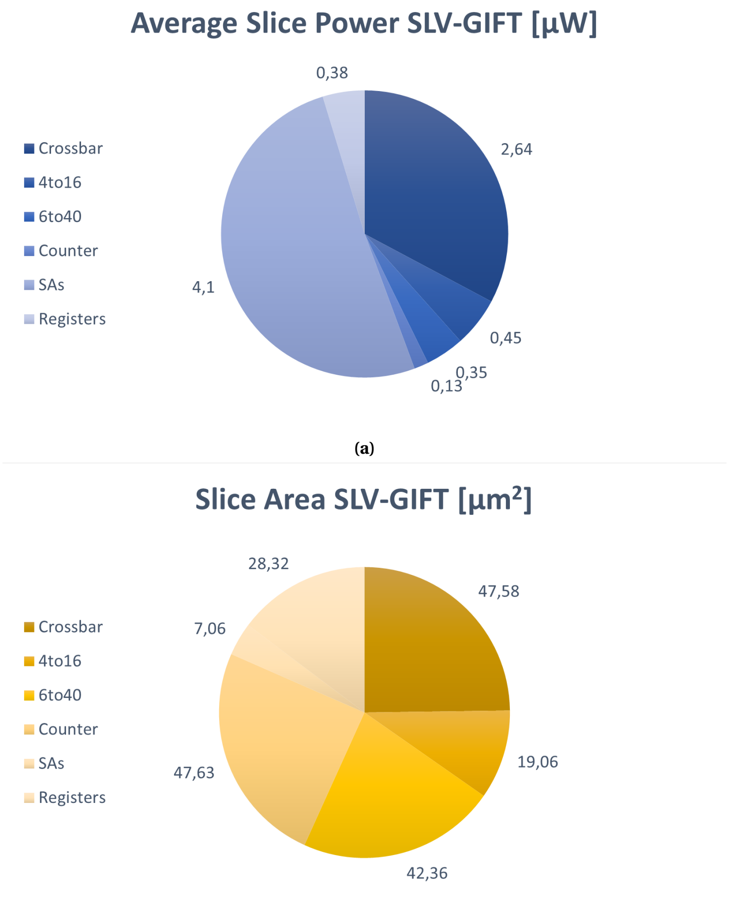
The implementation that is considered for comparison, utilizes the same operational frequency of 10 MHz. It is important to note that this implementation also includes the I/O data processing of both the plaintext and key data. For example, besides the 40 encryption cycles, input and output processing takes 16 cycles each. For fair energy-consumption comparison, the CMOS-GIFT is only evaluated for the 40 encryption rounds, as in this work. The results for all the implementations are summarized in Table 5.3.

Table 5.3: Results.

	CMOS-GIFT [125]	SLV-GIFT	DSA-GIFT
Average Power (μW)	31.77	257.6 *	60.38 *
Energy (pJ)	127.09	1030.40 *	241.52 *
Area (mm^2)	0.0030	0.0034	0.0034
Latency ** (us)		4	

* Extrapolated consumption.
** Latency of 40-round encryption

The average power consumption of a single Scouting Logic VSA is 1.46 μW and for the time simulated the total energy amounts to 5.84 pJ. For the largest decoder, i.e. 6-to-40, the total energy is only 1.42 pJ. As shown in Figure 5.14a, the power consumption of a single SLV-GIFT slice is 8.05 μW, and its energy consumption is 32.20 pJ. Nevertheless, these numbers pertain to a single slice. It was, however, explained earlier that the actual implementation would entail 32 of these slices; extrapolation gives an average consumption of 257.6 μW (1.03 nJ). In Section 4.5.3, it was mentioned that Scouting Logic VSA is slower compared to Scouting Logic CSA. During simulation it was found that the Scouting Logic VSA delay is 11.05, which is more than the delay of Scouting Logic CSA (1.68 ns). Upon further inspection, it was found that it only takes 438.4 ps for a signal to reach the Scouting Logic VSA, with the starting point being the input of the address decoders. Furthermore, the SAs in this implementation are responsible for over 50% of the total power consumption, which is quite significant. With such relatively large delay and energy consumption it is apparent that the system's bottleneck is the Scouting Logic VSA.



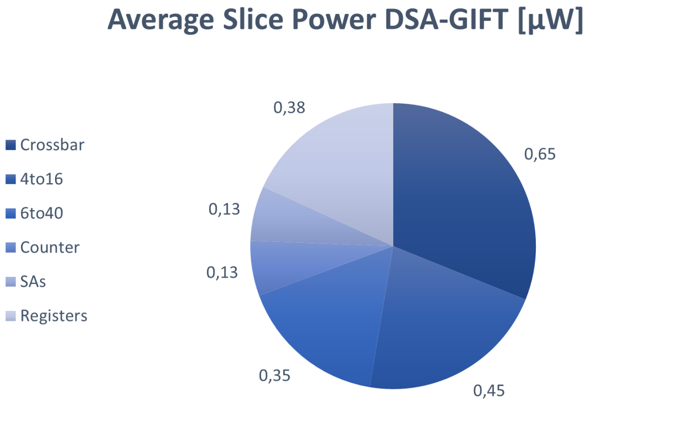
(a)

(b)

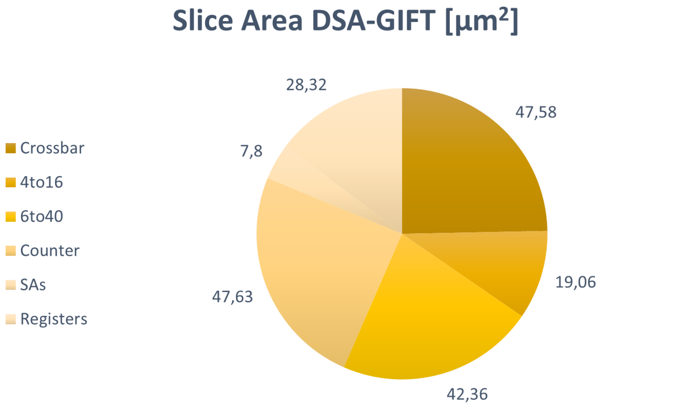
Figure 5.14: Slice breakdown of the SLV-GIFT. a) Power breakdown for a single slice of the SLV-GIFT. b) Area breakdown for a single slice of the SLV-GIFT.

The performance results of DSA-GIFT show drastic improvements. With a 41.40 nW power consumption of the DSA XOR, almost 5× more energy efficiency is achieved. Compared to SLV-GIFT, the increase in area overhead is negligible. The charts in Figure 5.14a and 5.15a show a shift of the bottleneck. More specifically, the DSA XOR only accounts for roughly 7% of the total power consumption and the crossbar now consumes significantly less power due to power gating. Still, compared to CMOS-GIFT, DSA-GIFT produces a bit more overhead in terms if energy.

Now, it can be observed that the energy consumption of the SLV-GIFT is 8× larger than that of CMOS-GIFT. The same holds for the average power consumption. As for area, the proposed architecture also consumes a bit more when extrapolated. Figure 5.16 shows the estimated power breakdowns of all three implementations. This breakdown



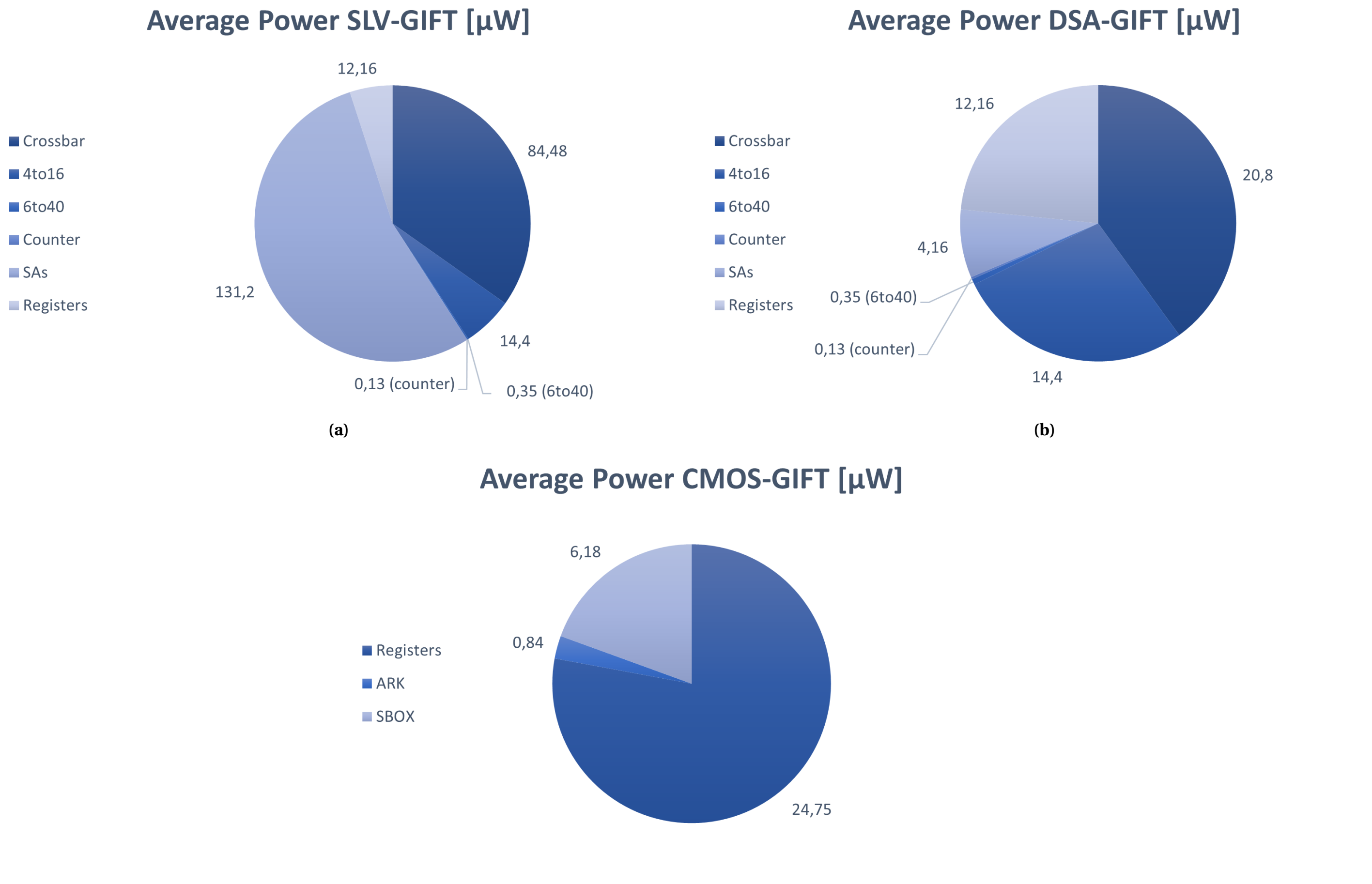
(a)



(b)

Figure 5.15: Slice breakdown of the DSA-GIFT. a) Power breakdown for a single slice of the DSA-GIFT. b) Area breakdown for a single slice of the DSA-GIFT.

confirms the SAs being the bottleneck of the SLV-GIFT design. Figure 5.14b and 5.15b shows that the 6-to-40 address decoder is the largest contributor. However, it should be noted that only one is required as it is reused for all the rounds. Hence, when applying extrapolation, it has the smallest area occupation. For CMOS-GIFT, the registers dominate the accumulated area consumption.

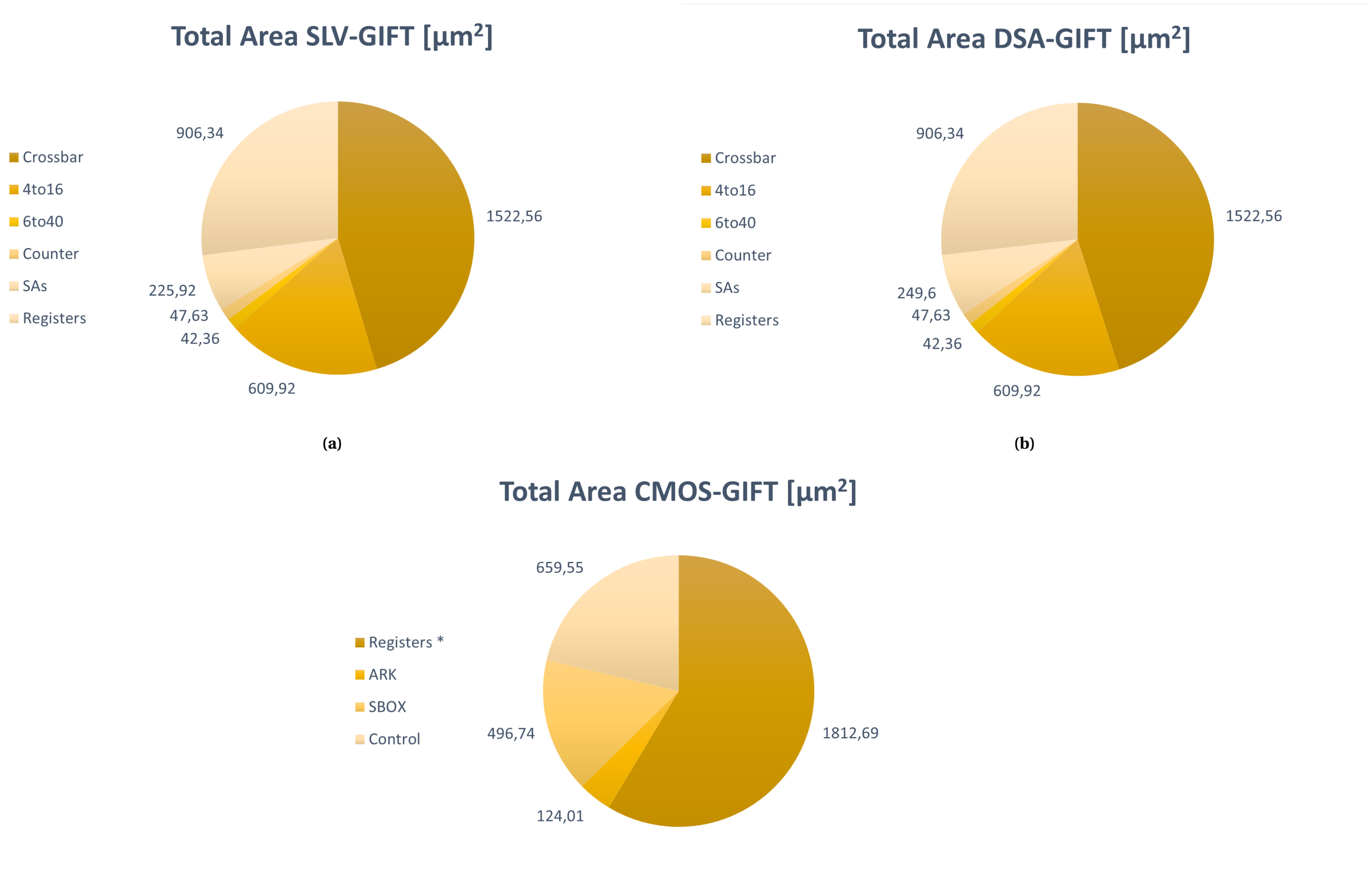


(a)

(b)

(c)

Figure 5.16: Power breakdown for the GIFT implementations: a) Total power consumption of the SLV-GIFT cipher. b) Total power consumption of the DSA-GIFT. c) Total power consumption of the CMOS-GIFT cipher.



(a)

(b)

(c)

Figure 5.17: Area breakdown for the GIFT implementations: a) Total area of SLV-GIFT. b) Total area of DSA-GIFT. c) Area of the CMOS-GIFT (* state/key registers.).

5.3. Discussion

Compared to SLV-GIFT, DSA-GIFT shows significant improvements in efficiency. Scouting Logic CSA and VSA consumes 12 and 1.45 μW for performing XOR operations, respectively, whereas the DSA XOR operation only consumes 41.40 nW. Nonetheless, the analysis shows that the proposed designs do not confirm the benefit of using memristors as building blocks for LWBCs, yet. Still there may be room for improvement in the DSA-GIFT implementation. For example, one could explore the design of the address decoders as they contribute significantly to the total footprint. Right now, static CMOS-gates are used, but dynamic gates may also be considered as they are known for their low power dissipation. However, because these gates are clocked, the current implementation may lose its ability to perform one encryption round per cycle. Besides that, incorporating such gates may also increase complexity.

The goal of this project is to find and answer the neglected security concerns in the novel FFNI interfaces, by means of a solution that accommodates the anticipated downscaling of the highly-constrained implants. The results do not disclose the memristor's imperfections, but rather the technology's need to mature a bit more. For instance, Section 4.5 talked about the different attempts of creating robust and lightweight MLGs for fast computation architectures. Researchers have pointed out the difficulty of creating long MLG chains due to the loss in drive strength over several gate stages and the destructive nature of the operations. Despite numerous solutions, it is still sub-optimal due to limitations relating to the number of operands, the required number of repeaters, and the number of operations. If advancements allow the resolvement of these constraints, it may be extremely beneficial for the implementation in this work: it would allow the construction of the S-box using a simple memristor-based AND-OR tree, thereby significantly reducing footprint consumption.

Another alternative would be the shift towards a 2T2R approach, as in [78]. Yes, it would result in a twofold increase of area, but in-situ operations would be done using differential sensing, doubling the sensing margin compared to [108] and, hence, increasing reliability. Also, the energy consumption with 2T2R operands is significantly lower: 72 fJ for four operands. Unfortunately, this structure does not support XOR operations yet.

An additional design approach would be to look at the crossbar topology. Currently, researchers are exploring the possibility to stack memristor crossbars, thereby significantly reducing the architecture footprint. These are also referred to as *3D vertical RRAMs*. Some successful proposals are presented, exploiting different memristor and transistor compositions to produce 1T1R pillars [127]–[131]. On top of that, one of the works presents a 3D architecture [130] to function in conjunction with the Scouting Logic architecture initially used in this thesis project. Similarly, a 3D vertical RRAM structure could be considered for the 1T1R-GIFT cipher, by stacking the 32 slices for parallel (or sequential) encryption cycles. This approach would primarily influence the area consumption of the chip.

Lastly, the reasons behind opting for 1T1R structures have been discussed extensively in this work. Still, the transistor in this structure is also the limiting factor for scaling down the crossbar. Thus, advancements in resistive switching materials, selection methods, and crossbar structures can enforce enhancements and solve the limitations.

5

So, do the FFNI interfaces benefit from a memristor-based security solution? At the moment, it may not seem so. However, over time further advancements in this field will most probably unlock the full potential of memristor architectures and, consequently, make them a fitting building block for lightweight security applications in FFNIs.

6

CONCLUSION

This final chapter gives an overview of the contributions made during this thesis project and lists the direction for future research. First, a chapter-wise summary is given, restating the main conclusions for each part. Subsequently, a brief outline of the future research topics related to this work is presented.

6.1. THESIS SUMMARY

Chapter 1 introduced the topic of wireless implantable medical care and the novel research field of the FFNIs. I stressed the importance of security in wireless FFNI interfaces and emphasized the main motivation of this work, which is to identify and countermeasure potential security threats in the highly-constrained state-of-the-art FFNIs. Furthermore, research topics related to this motivation were described. Next, a brief overview of the key objectives in this work was presented, followed by an outline of the thesis contributions.

Chapter 2 briefed about FFNI principles and provided a dissection of the entailed workloads. Subsequently, a detailed taxonomy of the state-of-the-art was laid out, which gave a strong insight into the design directions of future FFNIs. Based hereon, an envisioned system was proposed which also served as the project's design template. Notwithstanding the promising state of FFNIs, the complete absence of security considerations and measures in such systems was revealed.

In **Chapter 3** a thorough security survey, covering both the conventional wireless IMDs and FFNIs, was conducted, from which a novel attack tree classification was consolidated. It was then established that a lightweight encryption block, targeting the uplink EM transmissions of the sub-dura transceiver, is required to address the most critical threats of the classification.

Chapter 4 started with the proposal and argumentation for a memristive-based lightweight encryption approach, to accommodate the design constraints (on the future FFNIs) imposed by the anticipated physical limitations of such implants. This was followed by a short overview of memristor technology principles and existing memristor hardware security applications. Next, the chapter dived into the domain specific design exploration of the aforementioned lightweight cipher, with emphasis on a fitting memristor-based LWBC implementation. Here, a variety of highly efficient LWBCs were showcased, of which the the GIFT cipher stood out as the most suitable candidate for building a memristive security application. After a short background of GIFT, an analysis of existing resistive computation architectures was given in order to determine the key memristive components.

Finally, **Chapter 5** detailed the design and implementation of a novel, read-only 1T1R-GIFT cipher, followed by a comparative analysis against a CMOS-based GIFT cipher. Three different design approaches were tried for the XOR addition: Scouting Logic CSA, Scouting Logic VSA, and DSA. Of the three, the latter performed the best, and resulted into an estimated consumption close to that of the CMOS-based implementation. From this analysis it was then inferred that memristor technology requires some more time to mature before highly-constrained FFNIs can fully benefit from the potential of a memristor-based security paradigm.

6.2. Future Research Directions

This section puts forward future research directions that may be pursued. The research directions can be classified under *FFNI Design, Memristive Hardware Security Design* and *In-Memory Computation Design*.

FFNI Design

Memristors may offer a good alternative design approach for lightweight hardware security blocks. The anticipated downscaling of FFNIs, in conjunction with its workload expansion, opens the door to an interesting research field focusing on the realization of the additional workloads by means of memristive computing architectures, circumventing the walls of CMOS-based architectures.

Secondly, only part of the security implications are addressed in this work with the hope to ignite an important topic of discussion. With the continuing development of FFNIs it is all the more recommended to further explore their security aspects.

Memristive Hardware Security Design

The present literature mostly focuses on memristors as a source of entropy for hardware security applications. For future highly distributed application, it would be beneficial to further explore and expand the field of memristor-centric lightweight security blocks.

In-Memory Computation Design

An extensive research and development of MLG optimizations is required to solve problems related to gate cascading, memristor durability, state destruction, and multi-operand and complete MLG sets. Such advancements will enable tremendous improvements for architectures such as the one presented in this thesis and related works.

Secondly, an optimized and non-destructive in-situ crossbar computation architecture will have major benefits for energy consumption, failure rate, and operational speed. Most architectures only support basic operators such as AND, OR and NOT, but a more complete set of single-step gates is required. Hence, further research in this topic is highly recommended.

Lastly, the exploration of 3D vertical RRAMs or alternative crossbar structures will lead to significant improvements in density and efficiency of in-memory computation architectures. Nonetheless, it is important to preserve the controllability and reliability of nTnR crossbar. So, further research is needed to realize this.

Bibliography

[1] Medtronic, *About dbs for ocd.* [Online]. Available: https://www.medtronic.com/be-nl/patienten/behandelingen-en-therapieen/neurostimulator-dwangstoornis-ocd/wat-is-dbs-therapie.html.

[2] A. Vora, H. Ward, K. Foote, W. Goodman, and M. Okun, "Rebound symptoms following battery depletion in the nih ocd dbs cohort: Clinical and reimbursement issues," *Brain Stimulation*, vol. 5, no. 4, pp. 599–604, 2012, ISSN: 1935-861X. DOI: https://doi.org/10.1016/j.brs.2011.10.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1935861X11001525.

[3] S. Delaloye and P. E. Holtzheimer, "Deep brain stimulation in the treatment of depression," *Dialogues in Clinical Neuroscience*, vol. 16, no. 1, pp. 83–91, 2014. DOI: 10.31887/DCNS.2014.16.1/sdelaloye. eprint: https://doi.org/10.31887/DCNS.2014.16.1/sdelaloye. [Online]. Available: https://doi.org/10.31887/DCNS.2014.16.1/sdelaloye.

[4] M. Sprengers, K. Vonck, E. Carrette, A. Marson, and P. Boon, "Deep brain and cortical stimulation for epilepsy," *Cochrane Database of Systematic Reviews*, no. 7, 2017, ISSN: 1465-1858. DOI: 10.1002/14651858.CD008497.pub3. [Online]. Available: https://doi.org//10.1002/14651858.CD008497.pub3.

[5] *Deep brain stimulation for movement disorders.* [Online]. Available: https://www.ninds.nih.gov/health-information/disorders/deep-brain-stimulation-movement-disorders.

[6] M. B. Lee, D. R. Kramer, T. Peng, M. F. Barbaro, C. Y. Liu, S. Kellis, and B. Lee, "Clinical neuroprosthetics: Today and tomorrow," *Journal of Clinical Neuroscience*, vol. 68, pp. 13–19, 2019, ISSN: 0967-5868. DOI: https://doi.org/10.1016/j.jocn.2019.07.056. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967586819302735.

[7] C. for Devices and R. Health, *Radio frequency wireless technology in medical devices - guidance*, Aug. 2013. [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/radio-frequency-wireless-technology-medical-devices-guidance-industry-and-fda-staff.

[8] *Firmware update to address cybersecurity vulnerabilities identified in ...* Sep. 2017. [Online]. Available: https://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/ucm573669.htm?data2=dwnjltx.

[9] D. Halperin, T. S. Heydt-Benjamin, B. Ransford, S. S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno, and W. H. Maisel, "Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses," in *2008 IEEE Symposium on Security and Privacy (sp 2008)*, 2008, pp. 129–142. DOI: 10.1109/SP.2008.31.

[10] E. Marin, D. Singelée, B. Yang, V. Volski, G. A. E. Vandenbosch, B. Nuttin, and B. Preneel, "Securing wireless neurostimulators," in *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, ser. CODASPY '18, Tempe, AZ, USA: Association for Computing Machinery, 2018, pp. 287–298, ISBN: 9781450356329. DOI: 10.1145/3176258.3176310. [Online]. Available: https://doi.org/10.1145/3176258.3176310.

[11] E. Marin, D. Singelée, F. D. Garcia, T. Chothia, R. Willems, and B. Preneel, "On the (in)security of the latest generation implantable cardiac defibrillators and how to secure them," in *Proceedings of the 32nd Annual Conference on Computer Security Applications*, ser. ACSAC '16, Los Angeles, California, USA: Association for Computing Machinery, 2016, pp. 226–236, ISBN: 9781450347716. DOI: 10.1145/2991079.2991094. [Online]. Available: https://doi.org/10.1145/2991079.2991094.

[12] *Ics medical advisory (icsma-19-080-01)*, Apr. 2021. [Online]. Available: https://www.cisa.gov/uscert/ics/advisories/ICSMA-19-080-01.

[13] K.-W. Yang, K. Oh, and S. Ha, "Challenges in scaling down of free-floating implantable neural interfaces to millimeter scale," *IEEE Access*, vol. 8, pp. 133 295–133 320, 2020. DOI: 10.1109/ACCESS.2020.3007517.

[14] G. Buzsáki, C. Anastassiou, and C. Koch, "The origin of extracellular fields and currents—eeg, ecog, lfp and spikes," *Nature reviews. Neuroscience*, vol. 13, pp. 407–20, May 2012. DOI: 10.1038/nrn3241.

[15] R. D. Flint, E. W. Lindberg, L. R. Jordan, L. E. Miller, and M. W. Slutzky, "Accurate decoding of reaching movements from field potentials in the absence of spikes," *Journal of Neural Engineering*, vol. 9, no. 4, p. 046 006, Jun. 2012. DOI: 10.1088/1741-2560/9/4/046006. [Online]. Available: https://doi.org/10.1088/1741-2560/9/4/046006.

[16] N. Ahmadi, M. L. Cavuto, P. Feng, L. B. Leene, M. Maslik, F. Mazza, O. Savolainen, K. M. Szostak, C.-S. Bouganis, J. Ekanayake, A. Jackson, and T. G. Constandinou, "Towards a distributed, chronically-implantable neural interface," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2019, pp. 719–724. DOI: 10.1109/NER.2019.8716998.

[17] H.-G. Rhew, J. Jeong, J. A. Fredenburg, S. Dodani, P. G. Patil, and M. P. Flynn, "A fully self-contained logarithmic closed-loop deep brain stimulation soc with wireless telemetry and wireless power management," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2213–2227, 2014. DOI: 10.1109/JSSC.2014.2346779.

[18] A. Zhou, S. Santacruz, B. Johnson, G. Alexandrov, A. Moin, F. Burghardt, J. Rabaey, J. Carmena, and R. Muller, "A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates," *Nature Biomedical Engineering*, vol. 3, Jan. 2019. DOI: 10.1038/s41551-018-0323-x.

[19] *Optogenetics guide*. [Online]. Available: https://www.addgene.org/guides/optogenetics/.

[20] J. Ausra, M. Wu, X. Zhang, A. Vázquez-Guardado, P. Skelton, R. Peralta, R. Avila, T. Murickan, C. R. Haney, Y. Huang, J. A. Rogers, Y. Kozorovitskiy, and P. Gutruf, “Wireless, battery-free, subdermally implantable platforms for transcranial and long-range optogenetics in freely moving animals,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 30, e2025775118, 2021. DOI: 10.1073/pnas.2025775118. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2025775118. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2025775118.

[21] G. L. Barbruni, P. M. Ros, D. Demarchi, S. Carrara, and D. Ghezzi, “Miniaturised wireless power transfer systems for neurostimulation: A review,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 6, pp. 1160–1178, 2020. DOI: 10.1109/TBCAS.2020.3038599.

[22] D. Piech, B. Johnson, K. Shen, M. Ghanbari, K. Li, R. Neely, J. Kay, J. Carmena, M. Maharbiz, and R. Muller, “Stimdust: A 2.2 mm3, precision wireless neural stimulator with ultrasonic power and communication,” Jul. 2018.

[23] M. M. Ghanbari, D. K. Piech, K. Shen, S. Faraji Alamouti, C. Yalcin, B. C. Johnson, J. M. Carmena, M. M. Maharbiz, and R. Muller, “A sub-mm3 ultrasonic free-floating implant for multi-mote neural recording,” *IEEE Journal of Solid-State Circuits*, vol. 54, no. 11, pp. 3017–3030, 2019. DOI: 10.1109/JSSC.2019.2936303.

[24] C. Baron, J.-F. Aubry, M. Tanter, S. Meairs, and M. Fink, “Simulation of intracranial acoustic fields in clinical trials of sonothrombolysis,” *Ultrasound in Medicine & Biology*, vol. 35, no. 7, pp. 1148–1158, 2009, ISSN: 0301-5629. DOI: https://doi.org/10.1016/j.ultrasmedbio.2008.11.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301562908005504.

[25] A. V. Alexandrov, C. A. Molina, J. C. Grotta, Z. Garami, S. R. Ford, J. Alvarez-Sabin, J. Montaner, M. Saqqur, A. M. Demchuk, L. A. Moyé, and et al., “Ultrasound-enhanced systemic thrombolysis for acute ischemic stroke,” *New England Journal of Medicine*, vol. 351, no. 21, pp. 2170–2178, 2004. DOI: 10.1056/nejmoa041175.

[26] D. Seo, “Design of ultrasonic power link for neural dust,” Ph.D. dissertation, 2016.

[27] *Solar cells*. [Online]. Available: http://www.chemistryexplained.com/Ru-Sp/Solar-Cells.html.

[28] J. Lim, E. Moon, M. Barrow, S. R. Nason, P. R. Patel, P. G. Patil, S. Oh, I. Lee, H.-S. Kim, D. Sylvester, D. Blaauw, C. A. Chestek, J. Phillips, and T. Jang, “26.9 a 0.19×0.17mm2 wireless neural recording ic for motor prediction with near-infrared-based power and data telemetry,” in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2020, pp. 416–418. DOI: 10.1109/ISSCC19947.2020.9063005.

[29] E. Moon, D. Blaauw, and J. D. Phillips, “Subcutaneous photovoltaic infrared energy harvesting for bio-implantable devices,” *IEEE Transactions on Electron Devices*, vol. 64, no. 5, pp. 2432–2437, 2017. DOI: 10.1109/TED.2017.2681694.

[30] D. Seo, J. M. Carmena, J. M. Rabaey, E. Alon, and M. M. Maharbiz, *Neural dust: An ultrasonic, low power solution for chronic brain-machine interfaces*, 2013. DOI: 10.48550/ARXIV.1307.2196. [Online]. Available: https://arxiv.org/abs/1307.2196.

[31] R. M. Neely, D. K. Piech, S. R. Santacruz, M. M. Maharbiz, and J. M. Carmena, “Recent advances in neural dust: Towards a neural interface platform,” *Current Opinion in Neurobiology*, vol. 50, pp. 64–71, 2018, Neurotechnologies, ISSN: 0959-4388. DOI: https://doi.org/10.1016/j.conb.2017.12.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959438817302386.

[32] Y. Jia, S. A. Mirbozorgi, B. Lee, W. Khan, F. Madi, O. T. Inan, A. Weber, W. Li, and M. Ghovanloo, “A mm-sized free-floating wirelessly powered implantable optical stimulation device,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 4, pp. 608–618, 2019. DOI: 10.1109/TBCAS.2019.2918761.

[33] J. Charthad, T. C. Chang, Z. Liu, A. Sawaby, M. J. Weber, S. Baker, F. Gore, S. A. Felt, and A. Arbabian, “A mm-sized wireless implantable device for electrical stimulation of peripheral nerves,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 2, pp. 257–270, 2018. DOI: 10.1109/TBCAS.2018.2799623.

[34] C. W. L. Lee, A. Kiourti, and J. L. Volakis, “Miniaturized fully passive brain implant for wireless neuropotential acquisition,” *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 645–648, 2017. DOI: 10.1109/LAWP.2016.2594590.

[35] S. Ha, A. Akinin, J. Park, C. Kim, H. Wang, C. Maier, P. P. Mercier, and G. Cauwenberghs, “Silicon-integrated high-density electrocortical interfaces,” *Proceedings of the IEEE*, vol. 105, no. 1, pp. 11–33, 2017. DOI: 10.1109/JPROC.2016.2587690.

[36] A. Khalifa, Y. Liu, Y. Karimi, Q. Wang, A. Eisape, M. Stanaćević, N. Thakor, Z. Bao, and R. Etienne-Cummings, “The microbead: A 0.009 mm3 implantable wireless neural stimulator,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 5, pp. 971–985, 2019. DOI: 10.1109/TBCAS.2019.2939014.

[37] P. Yeon, M. S. Bakir, and M. Ghovanloo, “Towards a 1.1 mm2 free-floating wireless implantable neural recording soc,” in *2018 IEEE Custom Integrated Circuits Conference (CICC)*, 2018, pp. 1–4. DOI: 10.1109/CICC.2018.8357048.

[38] J. Lee, V. Leung, A.-H. Lee, J. Huang, P. Asbeck, P. P. Mercier, S. Shellhammer, L. Larson, F. Laiwalla, and A. Nurmikko, “Wireless ensembles of sub-mm microimplants communicating as a network near 1 ghz in a neural application,” *bioRxiv*, 2020. DOI: 10.1101/2020.09.11.293829. eprint: https://www.biorxiv.org/content/early/2020/09/13/2020.09.11.293829.full.pdf. [Online]. Available: https://www.biorxiv.org/content/early/2020/09/13/2020.09.11.293829.

[39] S. A. Wirdatmadja, “Wireless optogenetics nanonetworking device (wioptnd) opto-acoustic brain machine interface,” Ph.D. dissertation, 2020.

[40] B. Chatterjee, K. G. Kumar, M. Nath, S. Xiao, N. Modak, D. Das, J. Krishna, and S. Sen, “A 1.15uw 5.54mm3 implant with a bidirectional neural sensor and stimulator soc utilizing bi-phasic quasi-static brain communication achieving 6kbps-10mbps uplink with compressive sensing and ro-puf based collision avoidance,” *2021 Symposium on VLSI Circuits*, pp. 1–2, 2021.

[41] D. Seo, R. M. Neely, K. Shen, U. Singhal, E. Alon, J. M. Rabaey, J. M. Carmena, and M. M. Maharbiz, "Wireless recording in the peripheral nervous system with ultrasonic neural dust," *Neuron*, vol. 91, no. 3, pp. 529–539, 2016, ISSN: 0896-6273. DOI: https://doi.org/10.1016/j.neuron.2016.06.034. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627316303440.

[42] C. for Devices and R. Health, *Premarket submissions - cybersecurity in medical devices - guidance*, Oct. 2014. [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/content-premarket-submissions-management-cybersecurity-medical-devices.

[43] C. Camara, P. Peris-Lopez, J. M. De Fuentes, and S. Marchal, "Access control for implantable medical devices," *IEEE Transactions on Emerging Topics in Computing*, vol. PP, pp. 1–1, Mar. 2020. DOI: 10.1109/TETC.2020.2982461.

[44] N. B. Asan and R. Augustine, "Fat-ibc a new paradigm for intra-body communication: A pioneering work on fat channel communication in human body from my phd student and dr. noor badariah asan!" Ph.D. dissertation, Nov. 2019.

[45] *Insulinepompen voor kinderen*, Jun. 2018. [Online]. Available: https://www.medtronic-diabetes.nl/kind-met-diabetes/insulinepomp-voor-kinderen.

[46] Medtronic, *Sacral neuromodulation systems - interstim ii*. [Online]. Available: https://www.medtronic.com/nl-nl/zorgprofessionals/producten/urology/sacral-neuromodulation-systems/interstim-ii.html.

[47] J. Radcliffe, *Black hat usa 2011 //briefings*. [Online]. Available: https://www.blackhat.com/html/bh-us-11/bh-us-11-archives.html#Radcliffe.

[48] D. Pauli, *Hacked terminals capable of causing pacemaker deaths*, Oct. 2012. [Online]. Available: https://www.itnews.com.au/news/hacked-terminals-capable-of-causing-pacemaker-mass-murder-319508.

[49] R. Jordan, *Robertson j. mcafee hacker says medtronic insulin pumps vulnerable to attack*, Oct. 2012. [Online]. Available: https://www.bloomberg.com/news/articles/2012-02-29/mcafee-hacker-says-medtronic-insulin-pumps-vulnerable-to-attack#xj4y7vzkg.

[50] L. Pycroft, S. G. Boccard, S. L. Owen, J. F. Stein, J. J. Fitzgerald, A. L. Green, and T. Z. Aziz, "Brainjacking: Implant security issues in invasive neuromodulation," *World Neurosurgery*, vol. 92, pp. 454–462, 2016, ISSN: 1878-8750. DOI: https://doi.org/10.1016/j.wneu.2016.05.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1878875016302728.

[51] S. L. Bernal, A. H. Celdrán, L. F. Maimó, M. T. Barros, S. Balasubramaniam, and G. M. Pérez, "Cyberattacks on miniature brain implants to disrupt spontaneous neural signaling," *IEEE Access*, vol. 8, pp. 152 204–152 222, 2020. DOI: 10.1109/ACCESS.2020.3017394.

[52] M. Ienca and P. Haselager, "Hacking the brain: Brain–computer interfacing technology and the ethics of neurosecurity," *Ethics and Information Technology*, vol. 18, no. 2, pp. 117–129, Jun. 2016, ISSN: 1572-8439. DOI: 10.1007/s10676-016-9398-9. [Online]. Available: https://doi.org/10.1007/s10676-016-9398-9.

[53] J. A. Hansen and N. M. Hansen, "A taxonomy of vulnerabilities in implantable medical devices," in *Proceedings of the Second Annual Workshop on Security and Privacy in Medical and Home-Care Systems*, ser. SPIMACS '10, Chicago, Illinois, USA: Association for Computing Machinery, 2010, pp. 13–20, ISBN: 9781450300940. DOI: 10.1145/1866914.1866917. [Online]. Available: https://doi.org/10.1145/1866914.1866917.

[54] M. Al Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of medical systems*, vol. 36, pp. 93–101, Feb. 2012. DOI: 10.1007/s10916-010-9449-4.

[55] L. Pycroft and T. Z. Aziz, "Security of implantable medical devices with wireless connections: The dangers of cyber-attacks," *Expert Review of Medical Devices*, vol. 15, no. 6, pp. 403–406, 2018, PMID: 29860880. DOI: 10.1080/17434440.2018.1483235. eprint: https://doi.org/10.1080/17434440.2018.1483235. [Online]. Available: https://doi.org/10.1080/17434440.2018.1483235.

[56] H. Rathore, A. K. Al-Ali, A. Mohamed, X. Du, and M. Guizani, "A novel deep learning strategy for classifying different attack patterns for deep brain implants," *IEEE Access*, vol. 7, pp. 24 154–24 164, 2019. DOI: 10.1109/ACCESS.2019.2899558.

[57] T. Bonaci, R. Calo, and H. J. Chizeck, "App stores for the brain: Privacy amp; security in brain-computer interfaces," in *2014 IEEE International Symposium on Ethics in Science, Technology and Engineering*, 2014, pp. 1–7. DOI: 10.1109/ETHICS.2014.6893415.

[58] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, "Ghost talk: Mitigating emi signal injection attacks against analog sensors," in *2013 IEEE Symposium on Security and Privacy*, 2013, pp. 145–159. DOI: 10.1109/SP.2013.20.

[59] I. Giechaskiel and K. Rasmussen, "Taxonomy and challenges of out-of-band signal injection attacks and defenses," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 645–670, 2020. DOI: 10.1109/comst.2019.2952858. [Online]. Available: https://doi.org/10.1109%2Fcomst.2019.2952858.

[60] A. Siddiqi, "On the security and privacy of implantable medical devices," English, Ph.D. dissertation, Erasmus University Rotterdam, Sep. 2021, ISBN: 978-94-6423-357-5.

[61] I. Martinovic, D. Davies, M. Frank, D. Perito, T. Ros, and D. Song, "On the feasibility of side-channel attacks with brain-computer interfaces," in *21st USENIX Security Symposium (USENIX Security 12)*, Bellevue, WA: USENIX Association, Aug. 2012, pp. 143–158, ISBN: 978-931971-95-9. [Online]. Available: https://www.usenix.org/conference/usenixsecurity12/technical-sessions/presentation/martinovic.

[62] M. Rostami, A. Juels, and F. Koushanfar, "Heart-to-heart (h2h): Authentication for implanted medical devices," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, ser. CCS '13, Berlin, Germany: Association for Computing Machinery, 2013, pp. 1099–1112, ISBN: 9781450324779. DOI: 10.1145/2508859.2516658. [Online]. Available: https://doi.org/10.1145/2508859.2516658.

[63] M. A. Siddiqi, R. H. S. H. Beurskens, P. Kruizinga, C. I. De Zeeuw, and C. Strydis, "Securing implantable medical devices using ultrasound waves," *IEEE Access*, vol. 9, pp. 80 170–80 182, 2021. DOI: 10.1109/ACCESS.2021.3083576.

[64] S. Balasubramaniam, S. A. Wirdatmadja, M. T. Barros, Y. Koucheryavy, M. Stachowiak, and J. M. Jornet, "Wireless communications for optogenetics-based brain stimulation: Present technology and future challenges," *IEEE Communications Magazine*, vol. 56, no. 7, pp. 218–224, 2018. DOI: 10.1109/MCOM.2018.1700917.

[65] C. Camara, P. Peris-Lopez, J. M. de Fuentes, and S. Marchal, "Access control for implantable medical devices," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1126–1138, 2021. DOI: 10.1109/TETC.2020.2982461.

[66] D. Seo, "Design of ultrasonic power link for neural dust," M.S. thesis, EECS Department, University of California, Berkeley, May 2016. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-21.html.

[67] S. P. Mohanty and A. Srivastava, *Nano-CMOS and post-CMOS electronics*. The Institution of Engineering and Technology, 2016.

[68] L. Chua, "Memristor-the missing circuit element," *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971. DOI: 10.1109/TCT.1971.1083337.

[69] A. Singh, "Circuit design for memristor based in-memory computing," Ph.D. dissertation, 2019.

[70] R. Wang, J.-Q. Yang, J.-Y. Mao, Z.-P. Wang, S. Wu, M. Zhou, T. Chen, Y. Zhou, and S.-T. Han, "Recent advances of volatile memristors: Devices, mechanisms, and applications," *Advanced Intelligent Systems*, vol. 2, no. 9, p. 2 000 055, 2020. DOI: https://doi.org/10.1002/aisy.202000055. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aisy.202000055. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aisy.202000055.

[71] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, pp. 80–83, 2008. DOI: 10.1038/nature06932.

[72] T. Perez and C. De Rose, "Non-volatile memory: Emerging technologies and their impacts on memory systems," Apr. 2015. DOI: 10.13140/RG.2.1.3037.6486.

[73] J. Rofeh, A. Sodhi, M. Payvand, M. A. Lastras-Montaño, A. Ghofrani, A. Madhavan, S. Yemenicioglu, K.-T. Cheng, and L. Theogarajan, "Vertical integration of memristors onto foundry cmos dies using wafer-scale integration," in *2015 IEEE 65th Electronic Components and Technology Conference (ECTC)*, 2015, pp. 957–962. DOI: 10.1109/ECTC.2015.7159710.

[74] H. Li, S. Wang, X. Zhang, W. Wang, R. Yang, Z. Sun, W. Feng, P. Lin, Z. Wang, L. Sun, and Y. Yao, "Memristive crossbar arrays for storage and computing applications," *Advanced Intelligent Systems*, vol. 3, no. 9, p. 2 100 017, 2021. DOI: https://doi.org/10.1002/aisy.202100017. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aisy.202100017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aisy.202100017.

[75] S. Zhang, G. L. Zhang, B. Li, H. H. Li, and U. Schlichtmann, "Aging-aware lifetime enhancement for memristor-based neuromorphic computing," in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2019, pp. 1751–1756. DOI: 10.23919/DATE.2019.8714954.

[76] B. R. Fernando, Y. Qi, C. Yakopcic, and T. M. Taha, "3d memristor crossbar architecture for a multicore neuromorphic system," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8. DOI: 10.1109/IJCNN48605.2020.9206929.

[77] Emrl, *Redox-based tera-bit memories.* [Online]. Available: http://www.emrl.de/r_a_1.html#Artikel_1.

[78] A. Singh, R. Bishnoi, R. V. Joshi, and S. Hamdioui, "Referencing-in-array scheme for rram-based cim architecture," in *2022 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2022, pp. 1413–1418. DOI: 10.23919/DATE54114.2022.9774571.

[79] S. Lv, J. Liu, and Z. Geng, "Application of memristors in hardware security: A current state-of-the-art technology," *Advanced Intelligent Systems*, vol. 3, no. 1, p. 2 000 127, 2021. DOI: https://doi.org/10.1002/aisy.202000127. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aisy.202000127. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aisy.202000127.

[80] L. Wang, T. Dong, and M.-F. Ge, "Finite-time synchronization of memristor chaotic systems and its application in image encryption," *Applied Mathematics and Computation*, vol. 347, pp. 293–305, 2019, ISSN: 0096-3003. DOI: https://doi.org/10.1016/j.amc.2018.11.017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0096300318309901.

[81] L. Yang, L. Cheng, Y. Li, H. Li, J. Li, T.-C. Chang, and X. Miao, "Cryptographic key generation and in situ encryption in one-transistor-one-resistor memristors for hardware security," *Advanced Electronic Materials*, vol. 7, no. 5, p. 2 001 182, 2021. DOI: https://doi.org/10.1002/aelm.202001182. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aelm.202001182. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aelm.202001182.

[82] J. Sun, Z. Wang, S. Wang, M. Yang, H. Gao, H. Wang, X. Ma, and Y. Hao, "Physical unclonable functions based on transient form of memristors for emergency defenses," *IEEE Electron Device Letters*, pp. 1–1, 2022. DOI: 10.1109/LED.2022.3145487.

[83] J. Cai, A. Amirsoleimani, and R. Genov, *Hyperlock: In-memory hyperdimensional encryption in memristor crossbar array*, 2022. arXiv: 2201.11362 [cs.CR].

[84] J. Singh, "Implementation of memristor towards better hardware/software security design," *Transactions on Electrical and Electronic Materials*, vol. 22, Jan. 2021. DOI: 10.1007/s42341-020-00269-x.

[85] C. Yang, B. Liu, H. Li, Y. Chen, M. Barnell, Q. Wu, W. Wen, and J. Rajendran, "Security of neuromorphic computing: Thwarting learning attacks using memristor's obsolescence effect," in *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2016, pp. 1–6. DOI: 10.1145/2966986.2967074.

[86] W. Dai, X. Xu, X. Song, and G. Li, "Audio encryption algorithm based on chen memristor chaotic system," *Symmetry*, vol. 14, no. 1, 2022, ISSN: 2073-8994. DOI: 10.3390/sym14010017. [Online]. Available: https://www.mdpi.com/2073-8994/14/1/17.

[87] N. Du, H. Schmidt, and I. Polian, "Low-power emerging memristive designs towards secure hardware systems for applications in internet of things," *Nano Materials Science*, vol. 3, no. 2, pp. 186–204, 2021, Nano Energy Materials and Devices for Miniaturized Electronics and Smart Systems, ISSN: 2589-9651. DOI: https://doi.org/10.1016/j.nanoms.2021.01.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589965121000015.

[88] Y. Pang, B. Gao, B. Lin, H. Qian, and H. Wu, "Memristors for hardware security applications," *Advanced Electronic Materials*, vol. 5, no. 9, p. 1 800 872, 2019. DOI: https://doi.org/10.1002/aelm.201800872. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aelm.201800872. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aelm.201800872.

[89] N. Du, H. Schmidt, and I. Polian, "Low-power emerging memristive designs towards secure hardware systems for applications in internet of things," *Nano Materials Science*, vol. 3, no. 2, pp. 186–204, 2021, Nano Energy Materials and Devices for Miniaturized Electronics and Smart Systems, ISSN: 2589-9651. DOI: https://doi.org/10.1016/j.nanoms.2021.01.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589965121000015.

[90] B. Cambou, D. Hély, and S. Assiri, "Cryptography with analog scheme using memristors," *J. Emerg. Technol. Comput. Syst.*, vol. 16, no. 4, Sep. 2020, ISSN: 1550-4832. DOI: 10.1145/3412439. [Online]. Available: https://doi.org/10.1145/3412439.

[91] Y. Bai, H. Wu, R. Wu, Y. Zhang, N. Deng, Z. Yu, and H. Qian, "Study of multilevel characteristics for 3d vertical resistive switching memory," *Scientific reports*, vol. 4, p. 5780, Jul. 2014, ISSN: 2045-2322. DOI: 10.1038/srep05780. [Online]. Available: https://europepmc.org/articles/PMC4105739.

[92] Emrl, *Jart – jülich aachen resistive switching tools.* [Online]. Available: http://www.emrl.de/JART#Artikel_1.

[93] C.-Y. Huang, W. C. Shen, Y.-H. Tseng, Y.-C. King, and C.-J. Lin, "A contact-resistive random-access-memory-based true random number generator," *IEEE Electron Device Letters*, vol. 33, no. 8, pp. 1108–1110, 2012. DOI: 10.1109/LED.2012.2199734.

[94] S. Zhang, G. L. Zhang, B. Li, H. H. Li, and U. Schlichtmann, "Aging-aware lifetime enhancement for memristor-based neuromorphic computing," in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2019, pp. 1751–1756. DOI: 10.23919/DATE.2019.8714954.

[95] A. Joseph and P. V, "An analysis on low cost and performance of hardware and software oriented lightweight block ciphers for iot applications," *SSRN Electronic Journal*, 2021. DOI: 10.2139/ssrn.3883327.

[96] V. A. Thakor, M. A. Razzaque, and M. R. A. Khandaker, "Lightweight cryptography algorithms for resource-constrained iot devices: A review, comparison and research opportunities," *IEEE Access*, vol. 9, pp. 28 177–28 193, 2021. DOI: 10.1109/ACCESS.2021.3052867.

[97] N. Naser and J. Naif, "A systematic review of ultra-lightweight encryption algorithms," *International Journal of Nonlinear Analysis and Applications*, vol. 13, no. 1, pp. 3825–3851, 2022, ISSN: 2008-6822. DOI: 10.22075/ijnaa.2022.6167. eprint: https://ijnaa.semnan.ac.ir/article_6167_d534c6deba02d4a447202377eb079 pdf. [Online]. Available: https://ijnaa.semnan.ac.ir/article_6167.html.

[98] K. Shibutani, T. Isobe, H. Hiwatari, A. Mitsuda, T. Akishita, and T. Shirai, "Piccolo: An ultra-lightweight blockcipher," in *Cryptographic Hardware and Embedded Systems – CHES 2011*, B. Preneel and T. Takagi, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 342–357, ISBN: 978-3-642-23951-9.

[99] L. Dalmasso, F. Bruguier, P. Benoit, and L. Torres, "Evaluation of spn-based lightweight crypto-ciphers," *IEEE Access*, vol. 7, pp. 10 559–10 567, 2019. DOI: 10.1109/ACCESS.2018.2889790.

[100] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, and L. Wingers, *The simon and speck families of lightweight block ciphers*, Cryptology ePrint Archive, Paper 2013/404, https://eprint.iacr.org/2013/404, 2013. [Online]. Available: https://eprint.iacr.org/2013/404.

[101] W.-Z. Yeoh, J. S. Teh, and M. I. S. B. M. Sazali, "μ2 : A lightweight block cipher," in *Computational Science and Technology*, R. Alfred, Y. Lim, H. Haviluddin, and C. K. On, Eds., Singapore: Springer Singapore, 2020, pp. 281–290, ISBN: 978-981-15-0058-9.

[102] A. Alahdal, G. A. AL-Rummana, G. N. Shinde, and N. K. Deshmuk, "Nlbsit: A new lightweight block cipher design for securing data in iot devices," *International Journal of Computer Sciences and Engineering*, vol. 8, no. 10, Oct. 2020. DOI: DOI:https://doi.org/10.26438/ijcse/v8i10.164173.

[103] S. Banik, S. K. Pandey, T. Peyrin, Y. Sasaki, S. M. Sim, and Y. Todo, "Gift: A small present," in *Cryptographic Hardware and Embedded Systems – CHES 2017*, W. Fischer and N. Homma, Eds., Cham: Springer International Publishing, 2017, pp. 321–345, ISBN: 978-3-319-66787-4.

[104] K. Humood, H. Abunahla, and B. Mohammad, "Memchar: Portable low-power and low-cost characterization tool for memristor devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022. DOI: 10.1109/TIM.2022.3144202.

[105] D. Ielmini and H.-S. P. Wong, "In-memory computing with resistive switching devices," *Nature Electronics*, vol. 1, no. 6, pp. 333–343, 2018. DOI: 10.1038/s41928-018-0092-2.

[106] W. Shen, P. Huang, M. Fan, R. Han, Z. Zhou, B. Gao, H. Wu, H. Qian, L. Liu, X. Liu, X. Zhang, and J. Kang, "Stateful logic operations in one-transistor-one- resistor resistive random access memory array," *IEEE Electron Device Letters*, vol. 40, no. 9, pp. 1538–1541, 2019. DOI: 10.1109/LED.2019.2931947.

[107] N. TaheriNejad, "Sixor: Single-cycle in-memristor xor," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 29, no. 5, pp. 925–935, 2021. DOI: 10.1109/TVLSI.2021.3062293.

[108] L. Xie, H. Du Nguyen, J. Yu, A. Kaichouhi, M. Taouil, M. AlFailakawi, and S. Hamdioui, "Scouting logic: A novel memristor-based logic design for resistive computing," in *2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2017, pp. 176–181. DOI: 10.1109/ISVLSI.2017.39.

[109] S. Kvatinsky, D. Belousov, S. Liman, G. Satat, N. Wald, E. Friedman, A. Kolodny, and U. Weiser, "Magic - memristor-aided logic," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, pp. 895–899, Nov. 2014. DOI: 10.1109/TCSII.2014.2357292.

[110] I. Vourkas and G. C. Sirakoulis, "Emerging memristor-based logic circuit design approaches: A review," *IEEE Circuits and Systems Magazine*, vol. 16, no. 3, pp. 15–30, 2016. DOI: 10.1109/MCAS.2016.2583673.

[111] Z.-R. Wang, Y.-T. Su, Y. Li, Y.-X. Zhou, T.-J. Chu, K.-C. Chang, T.-C. Chang, T.-M. Tsai, S. M. Sze, and X.-S. Miao, "Functionally complete boolean logic in 1t1r resistive random access memory," *IEEE Electron Device Letters*, vol. 38, no. 2, pp. 179–182, 2017. DOI: 10.1109/LED.2016.2645946.

[112] S. Kvatinsky, N. Wald, G. Satat, A. Kolodny, U. C. Weiser, and E. G. Friedman, "Mrl — memristor ratioed logic," *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, 2012. DOI: 10.1109/cnna.2012.6331426.

[113] I. Vourkas and G. C. Sirakoulis, "A novel design and modeling paradigm for memristor-based crossbar circuits," *IEEE Transactions on Nanotechnology*, vol. 11, no. 6, pp. 1151–1159, 2012. DOI: 10.1109/tnano.2012.2217153.

[114] G. Papandroulidakis, I. Vourkas, N. Vasileiadis, and G. C. Sirakoulis, "Boolean logic operations and computing circuits based on memristors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, no. 12, pp. 972–976, 2014. DOI: 10.1109/TCSII.2014.2357351.

[115] S. Jain, A. Ranjan, K. Roy, and A. Raghunathan, "Computing in memory with spin-transfer torque magnetic ram," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 3, pp. 470–483, 2018. DOI: 10.1109/TVLSI.2017.2776954.

[116] C. Reinbrecht, A. Aljuffri, S. Hamdioui, M. Taouil, and J. Sepúlveda, "Grinch: A cache attack against gift lightweight cipher," in *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2021, pp. 549–554. DOI: 10.23919/DATE51398.2021.9474201.

[117] N. C. Dao and D. Koch, "Memristor-based reconfigurable circuits: Challenges in implementation," in *2020 International Conference on Electronics, Information, and Communication (ICEIC)*, 2020, pp. 1–6. DOI: 10.1109/ICEIC49074.2020.9051174.

[118] J. Sun, X. Zhao, and Y. Wang, "A three input look-up-table design based on memristor-cmos," *Communications in Computer and Information Science*, pp. 275–286, 2018. DOI: 10.1007/978-981-13-2829-9_25.

[119] F. Zhang, D.-Y. Fan, Q.-P. Lin, Q. Huo, Y. Li, L. Dai, C.-Y. Chen, and H.-H. Shen, "The application of non-volatile look-up-table operations based on multilevel-cell of resistance switching random access memory," *2018 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, 2018. DOI: 10.1109/vlsi-dat.2018.8373268.

[120] A. K. Mishra, D. P. Acharya, and P. K. Patra, "Novel design technique of address decoder for sram," in *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, 2014, pp. 1032–1035. DOI: 10.1109/ICACCCT.2014.7019253.

[121] S. Singh and S. Akashe, "Low power consuming 1 kb (32 × 32) memory array using compact 7t sram cell," *Wireless Personal Communications*, vol. 96, no. 1, pp. 1099–1109, 2017. DOI: 10.1007/s11277-017-4226-z.

[122] B.-D. Yang, "Low-power and area-efficient shift register using pulsed latches," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 6, pp. 1564–1571, 2015. DOI: 10.1109/TCSI.2015.2418837.

[123] J. Galvan Hernandez and D. Vrijenhoek, "A 5-bit vernier time-to-digital converter in 90nm cmos technology, unpublished,"

[124] J. Silva, S. Lanceros-Mendez, G. Minas, and J. G. Rocha, "Cmos x-ray image sensor array," in *2007 14th IEEE International Conference on Electronics, Circuits and Systems*, 2007, pp. 1067–1070. DOI: 10.1109/ICECS.2007.4511178.

[125] C. A. Lara-Nino, A. Diaz-Perez, and M. Morales-Sandoval, *Fpga-based assessment of midori and gift lightweight block ciphers*, Cryptology ePrint Archive, Paper 2018/979, https://eprint.iacr.org/2018/979, 2018. DOI: 10.1007/978-3-030-01950-1_45. [Online]. Available: https://eprint.iacr.org/2018/979.

[126] *Synopsys spyglass products*. [Online]. Available: https://www.synopsys.com/verification/static-and-formal-verification/spyglass.html.

[127] H. Li, S. Wang, X. Zhang, W. Wang, R. Yang, Z. Sun, W. Feng, P. Lin, Z. Wang, L. Sun, and Y. Yao, "Memristive crossbar arrays for storage and computing applications," *Advanced Intelligent Systems*, vol. 3, no. 9, p. 2 100 017, 2021. DOI: https://doi.org/10.1002/aisy.202100017. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aisy.202100017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aisy.202100017.

[128] J. Wu, F. Mo, T. Saraya, T. Hiramoto, and M. Kobayashi, "A monolithic 3-d integration of rram array and oxide semiconductor fet for in-memory computing in 3-d neural network," *IEEE Transactions on Electron Devices*, vol. 67, no. 12, pp. 5322–5328, 2020. DOI: 10.1109/TED.2020.3033831.

[129] C.-H. Wang, C. McClellan, Y. Shi, X. Zheng, V. Chen, M. Lanza, E. Pop, and H.-S. Philip Wong, "3d monolithic stacked 1t1r cells using monolayer mos2 fet and hbn rram fabricated at low (150°c) temperature," in *2018 IEEE International Electron Devices Meeting (IEDM)*, 2018, pp. 22.5.1–22.5.4. DOI: 10.1109/IEDM.2018.8614495.

[130] M. Ezzadeen, D. Bosch, B. Giraud, S. Barraud, J.-P. Noel, D. Lattard, J. Lacord, J. Portal, and F. Andrieu, *Ultra-high-density 3d vertical rram with stacked junctionless nanowires for in-memory-computing applications*, Nov. 2020.

[131] B. R. Fernando, Y. Qi, C. Yakopcic, and T. M. Taha, "3d memristor crossbar architecture for a multicore neuromorphic system," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8. DOI: 10.1109/IJCNN48605.2020.9206929.

[132] H. Tawalbeh, S. Hashish, L. Tawalbeh, and A. Aldairi, "Security in wireless sensor networks using lightweight cryptography".," *Journal of Information Assurance and Security.*, vol. 12, pp. 118–123, Nov. 2017.

[133] C. Manifavas, G. Hatzivasilis, K. Fysarakis, and K. Rantos, "Lightweight cryptography for embedded systems – a comparative analysis," in *Data Privacy Management and Autonomous Spontaneous Security*, J. Garcia-Alfaro, G. Lioudakis, N. Cuppens-Boulahia, S. Foley, and W. M. Fitzgerald, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 333–349, ISBN: 978-3-642-54568-9.

[134] J. Nechvatal, E. B. Barker, L. E. Bassham, W. E. Burr, M. Dworkin, J. Foti, and E. Roback, "Report on the development of the advanced encryption standard (aes)," *Journal of Research of the National Institute of Standards and Technology*, vol. 106, pp. 511–577, 2001.

[135] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949. DOI: 10.1002/j.1538-7305.1949.tb00928.x.

A

Cryptography Background

According to [95]–[97], [132] there are three variables that define Lightweight Cryptography (LWC) for devices with limited resources: (physical) cost, performance and security. these are summarized in Table A.1. There are two cryptography styles: symmetric and asymmetric cryptography. Symmetric algorithms use a single private key for encryption and decryption, whereas asymmetric cryptography (also known as public key cryptography) uses two separate keys (private and public) for encryption and decryption. When it comes to LWC, symmetric schemes are best suited. This is because they are a lot faster and less complex [96], [97]. Asymmetric ciphers are more demanding and better suited for powerful devices. Moreover, the most efficient lightweight asymmetric scheme, i.e. *elliptic curve cryptography* [108], is still 100-1000 times slower than standard symmetric algorithms [133]. Then there's the fact that symmetric cryptography has been building a good reputation for a long time, which resulted in it becoming a widely accepted industry standard [134].

As illustrated in Figure A.1, symmetric encryption can be further categorized into *block ciphers* and *stream ciphers*. The encryption/decryption in block ciphers is done in blocks of 64 bits or more. Stream ciphers on the other hand, do this continuously, bit by bit. In 1949, mathematician Claude Shannon specified the properties of *confusion* and *diffusion* [135], which should be there in a cipher in order for it to be considered secure. To fulfill the property of confusion, each character of the ciphertext should be dependent on multiple parts of the secret key. For diffusion, changing one character in the plaintext should also change several characters in the ciphertext. When it comes to stream ciphers, they only deploy confusion. Block ciphers, however, contain both of the properties and make decryption more complex [96], [97]. This is one of the reasons to pick the block-cipher as category for the design project.

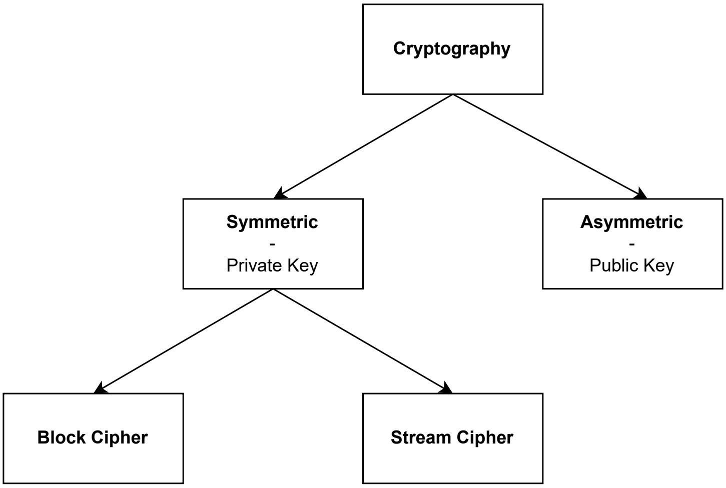


Figure A.1: Cryptography tree

Table A.1: Variables of LWC.

Variables	**Means**
(Physical) Cost	Memory space, registers Energy/Power consumption Area overhead: logic blocks, GE
Performance	Device power: throughput, latency,
Security	Number of rounds Number of bits Attack prevention: fault-injection, side-channel