Abstract

Development of memristor-based device in 2008 has encouraged scientists and engineers to perform brain-inspired cognitive tasks on the hardware level. Memristor is a resistor with memory which has some similarity like a synapse in our brain and can be fabricated in a 3-dimensional array with high density. This allows memristors to be used in parallel processing applications like pattern recognition.

Memristor-based neuromorphic systems have neuron and memristor integrated together as the fundamental unit which works like neuron and synapses in our brain. This unit is repeated in large number and interacts with each other to perform the complex functionalities of the brain. Many researchers have done pattern recognition and other cognitive tasks on simulation level and partially on hardware implementation. Development of full neuromorphic system requires the development of memristor and CMOS neuron circuits with spiking circuit for implementing the fundamental properties of synapses that is Long-Term Potentiation (LTP) and Long-Term Depression (LTD).

Since the continuous switching memristor is in its infant stage, in order to progress the research in the neuromorphic circuit, an On-chip externally programmable emulator has been designed which shows memory effect (pinched hysteresis curve), LTP and LTD on the hardware level. This emulator can be used for implementing any memristive behavior on the hardware level. A double barrier memristor device (DBMD) has been developed by Institute of Electrical Engineering and Information Technology, University of Kiel which shows a continuous change in resistance when the consecutive write voltage is applied. This device has to be integrated with CMOS neuron to develop potentiation and depression. An integrate & fire neuron chip with spiking circuit has been designed and fabricated in AMS350 nm process.

DBMD has been integrated with Neuron ASIC and the measurement shows LTP on hardware level which can be used for a larger network to mimic functionalities of the brain. A LTspice based simulation has been performed with DBMD model and neuron model to train 4 images, each of 4X3 pixels. The simulation shows the unsupervised learning of a memristor-based system. This paves the way toward development of neuromorphic computers.

Chapter 1

Introduction

Von Neumann Architecture (VNA) of modern day computer was given by great visionary scientist Von Neumann in 1945. VNA architecture has separate memory and processing unit as shown in figure 1.1. All programs and data in this architecture are stored in memory and communicate through a bus. Each time a new program has to be processed, the processor fetches the data from memory, process it and store it in the memory. This requires a high bandwidth of bus and separate memory to store each program and data. This architecture has a simple, fixed structure but capable of performing any computation without hardware modification. This paved the way towards the development of modern day computers and helped us in developing a modern day digital world. With the evolution of computers, broad application areas emerged which adopted digital computers for faster and efficient computing. Consequently, high-performance microprocessors and data storage were required. Technology scaling resulted in improving the speed of microprocessor by reducing the size of transistor. However, semiconductor scaling has already reached 7 nanometer (nm) technology and scaling of 3 nm is limited by tunneling. Hence, further scaling is not possible due to physical and practical limits [1]. Therefore, digital processors have limitations in improving it's further speed. Moreover, VNA suffers from communication bottleneck due to bus bandwidth and memory wall resulting from CMOS downscaling [1]. These are the challenges of VNA and limits its usability in applications where huge data and high processing is required like in Artificial Neural Network (ANN).

Challenges in VNA and need for highly efficient computing has encouraged scientists and semiconductor industry to develop a brain-inspired neuromorphic computer which is considered to be highly efficient like a biological brain. Human brain architecture is

adaptive in nature and has memory and processing at the same location, unlike VNA. Hence, brain-inspired computing on hardware requires neuromorphic architecture and cognitive device which works similar to synapses in the brain. Neuromorphic architecture can be implemented in CMOS circuit but the bottleneck remains in developing a cognitive device (synapse) which can change its property with time. In 2008, HP lab developed a two terminal device which has adaptive resistance and called memristors. This device has a property like synapses of the brain and attracted scientist and engineers for developing neuromorphic computing with this device. Since its development, multiple researchers have developed memristor with different material and architecture and this device has become a key component in neuromorphic computing [2], [3].

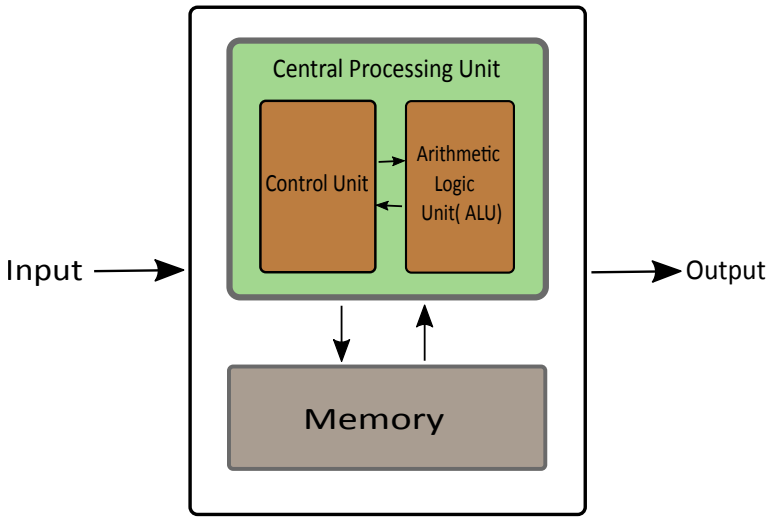

FIGURE 1.1: Von Neumann Architecture

Memristor

The memristor was first postulated by Leon Chua in 1971 [4] where he described memristor as a new two-terminal device which relates two fundamental circuit variables; charge and flux linkage. Chua presents the electromagnetic interpretation of this relationship and shows that the new device has peculiar behavior not found in basic circuit elements like resistor, capacitor, and inductor. In electrical circuits, all the four circuit quantities are linked with each other either by physical law or a physical device as shown in figure 1.2.

Resistor relates voltage & current, inductor relates current & flux and capacitor relates voltage & charge. Chua mentions the missing link in the group which links magnetic flux and introduces a new device which relates these two physical qualities. He termed

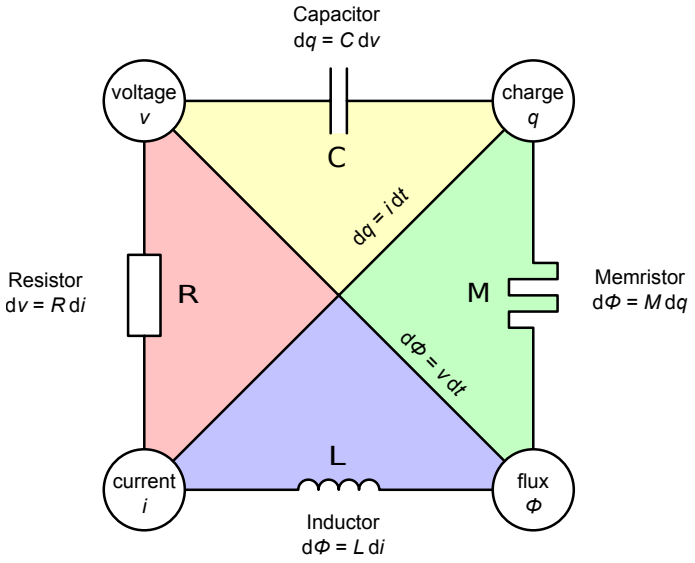

FIGURE 1.2: Two terminal circuit elements

this device name as memristor because it shows the behavior of a resistor with memory. Emulation of this model was done and a non-linear current-voltage relationship was shown in simulation and measurements. Since then, a lot of research has been done to develop such a device which has a non-volatile memory like memristor. A device with flux-charge relation could not be developed till date and it is not expected to be in foreseeable future. However, In 2008, Hewlett-Packard (HP) team under R. Stanley Williams developed a Titanium oxide TiO_2 based two terminal device which has the non-linear characteristic and voltage-current relationship like the device explained by Leon Chua in [4]. This device has similar properties like memristor but was not based on a flux-charge relation. The TiO_2 based device worked on the principle of ion movement when a voltage is applied which eventually changes the resistance of the device [5].

1.1 State of the art and applications

The memristor got attention in scientific community because of it's predicted potential application in non-volatile memory [6], [7], [8], [9] programmable analog circuits [10], [11], [12] neuromorphic circuits [13], [14], [2], [3] and many more. It has opened new possibilities for neuromorphic computers where distributed processing are supposed to be more power efficient and compact.

Non-volatile Memory

In last decades silicon based Flash memory has been widely used for non-volatile compact data storage. Flash memory has challenges like high voltage operation and cost effective-production to store ever growing data in present day situation. Research has been going on to develop next-generation flash but it faces technical challenges. Therefore, other type of non-volatile memory like resistive random access memory (ReRAM) [15], [16], [17], phase change memory (PCM) [18], [19], [20], spin-transfer torque random access memory (STT-RAM) [21], [22], [23], magnetic random access memory (MRAM) [24], [25] and Ferroelectric random access memory (FeRAM) [26], [27] have attracted researchers and semiconductor industry as possible alternative for Flash memory. Table 1.1 shows summary of all next-generation non-volatile memory. All possible next-generation non-volatile devices has low read and write voltage compared to convention Flash memory where still more than 10 V is required.

RRAM is getting much attention in industry because it is compatible with conventional semiconductor process [28], [29], [30]. In [30] authors have developed memristor on the silicon substrate and it can be integrated with other CMOS devices like a transistor, resistor, and capacitor on same ASIC. Beside silicon, there is a wide possibility of material which can be used for memristor development. In [31], authors have shown a list of materials which can be used for resistive switching as well as materials which can be used for electrodes. The device can be fabricated in 3-D architecture providing a high density of devices [32], [33]. ReRAM based devices work on the same principle as a memristor. When a positive voltage in applied device change to ON state and when a negative voltage is applied the device change to OFF state as shown in figure 1.3. The devices have different read and write voltage.

Programmable analog circuits

Memristor has attracted analog designer to use this device in programmable analog circuits because it reduces the total area of the circuit compared to the conventional approach. In [10], authors have shown applications of memristor in programmable threshold comparator, programmable gain amplifier, Schmitt trigger and analog filters. The authors have simulated analog circuits with memristor emulator and indicated the successful application in analog design. In [11] author indicates pulse coded precision programmable resistor and simulates the precise change in resistance by changing the pulse width of the clock. Application of memristor has been tested on hardware level in analog applications. In, [12] author uses Titanium oxide-based memristor device to design programmable high pass filter. It indicates the possible use of memristor in analog circuit as it is area efficient, power efficient and less complex compared to the conventional

Feature	FeRAM	MRAM	STT-RAM	PCM	ReRAM
Cell size (F^2)	Large, approximately 40 to 20	Large, approximately 25	Small, approximately 6 to 20	small, approximately 8	20 [34]
Storage mechanism	Permanent polarization of a ferroelectric material (PZT or SBT)	Permanent magnetization of a ferromagnetic material in a MTJ	Spin-polarized current applies torque on the magnetic moment	Amorphous/polycrystal phases of chalcogenide alloy	Resistive switching
Read time (ns)	20 to 80	3 to 20	2 to 20	20 to 50	10 [35]
Write/erase time (ns)	50/50	3 to 20	2 to 20	20/30	10 [35]
Endurance	10^{12}	$> 10^{15}$	$> 10^{16}$	10^{12}	10^6 [35]
Nonvolatility	Yes	Yes	Yes	Yes	Yes
Maturity	Limited production	Test chips	Test chips	Test chips	Limited production
Applications	Low density	Low density	High density	High density	High density

TABLE 1.1: Emerging Non-volatile memory summary [28]

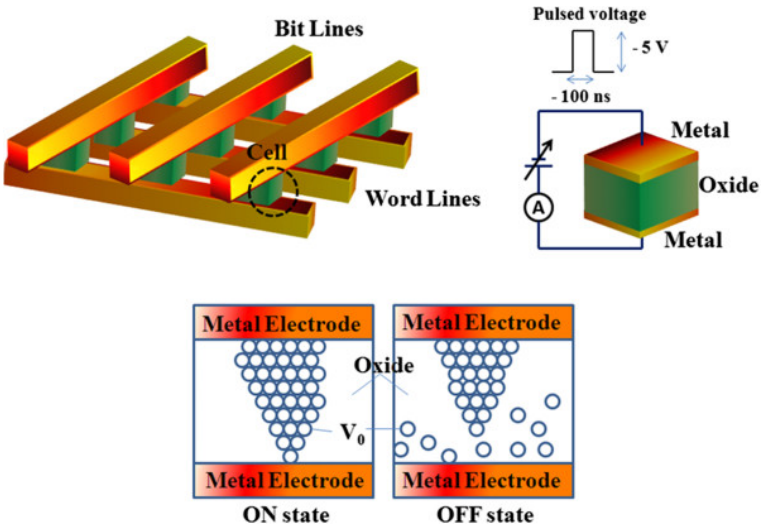

FIGURE 1.3: ReRAM working principle [28]

approach where to program the circuit resistance switching need to be implemented with the digital control circuit.

Neuromorphic circuits

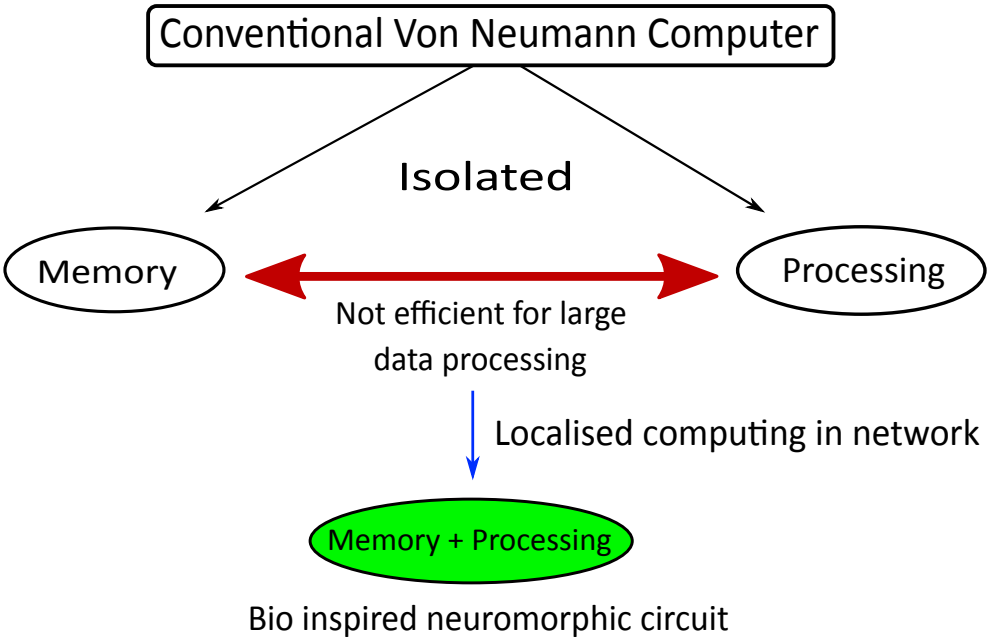

FIGURE 1.4: Conventional Von Neumann computers versus Neuromorphic circuits

Conventional Von Neumann computers have separate memory and processing as shown in figure 1.4. The ever-growing data in present day situation is an issue for even state of the art digital processors to fetch data from memory, process it and store it back in

memory. Semiconductor technology has limitation to shrink further because of fundamental and practical limits. Neuromorphic processors may be the next alternative which is bio-inspired and is supposed to be as efficient as a brain. The human brain works at 20 Watts and capable of doing a highly complex cognitive task which is impossible for the modern day computer. Whereas, implementing human-scale intelligence in digital computer will require "more than 10 MW" of power [36] and is therefore impractical [37], [38], [39]. The human brain works much more efficient than digital computers (VNA) because it works on a completely different approach than a human brain. It has a highly dense structure of neurons and synapse which adaptable in nature and remains one of the key components of the biological brain. Figure 1.5 shows biological neuron-synapse and its functional description on hardware. The basic electrically important components of a neuron are cell body, axon, dendrites, and synapse. Every neuron is connected with other neurons through dendrites and each neuron is connected with approximately 10000 other neurons via these dendrites. The interconnect between dendrite and neuron is synapse. It is self-adaptive resistance in an electrical sense which allows neurons to learn. Cell body integrates all incoming currents through these synapses and fires (generates action potential) when it reaches a threshold. Action potential travel along the axon to communicate with other neurons at distance. Implementing bio-logically inspired neuromorphic circuit on hardware requires the development of a cognitive device which is adaptive in nature can be integrated with high density and is adaptive in nature.

Memristor is an alternate device which can be fabricated in 3-D dense structure and it can be part of processing as well as memory. It has opened new possibilities for neuromorphic computers which are supposed to be more power efficient and compact.

This device may give the opportunity to emulate the functionality of the brain, instead of simulating on computers which requires a large amount of area and power. Many research groups are working on brain projects like IBM's Blue brain project, Howard Hughes Medical Institute's Janelia Farm, Harvard's Center for Brain Science etc. Even today not even a rat brain can be simulated because it requires huge power and area.

When the amount of data is huge, which is the present day situation, separate processing, and memory unit approach of von Neumann digital computer is inefficient. A new approach is needed where processing can be done in the distributed manner and memory storage can be localized in a network instead of separate memory.

"Memristor can be made extremely small, and they function like synapse. Using them, we will be able to build analog electronics circuits that could fit in a shoebox and function

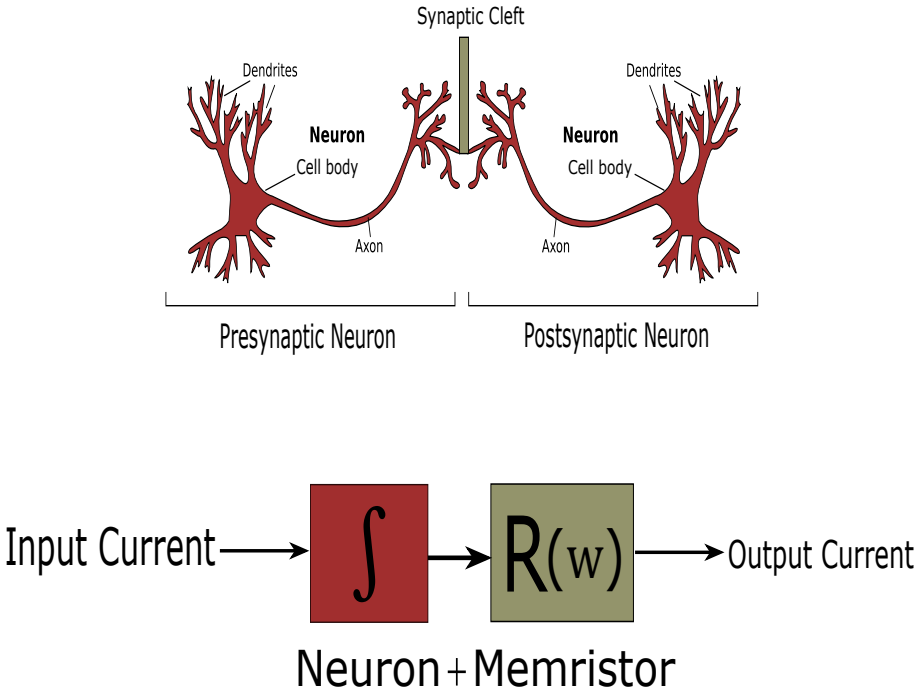

FIGURE 1.5: Synapses and Neuron

according to the same physical principle as a brain. A hybrid circuit containing many connected memristors and transistors could help us research actual brain function and disorders. Such a circuit might even lead to machine that can recognize patterns the way humans can, in those critical ways computers can't - for example, picking a particular face out of a crowd even if it has changed significantly since our last memory of it." [40]

Many researchers have shown successful results of distributed processing and learning behavior with memristors like pattern recognition. In [41] author proposes a design methodology for unsupervised learning of handwritten digits. In this work, pattern recognition has been done on simulation level which shows the pattern recognition even in presence of device variation due to adapting nature of memristor and power of homeostasis. It also shows the self-adjustment capability of the network with different input encoding schemes. Additionally, many researchers have indicated memristor based pattern recognition in simulation [42], [43], [44]. Some authors have developed memristor devices, extracted the measurement data and simulated the device data on hardware for pattern recognition. In [45], Nb_xO_y memristor measurement-based model has been used to train 10 handwritten numbers and system shows recognition rate more than $60\ \%$ even with device-to-device variation more than $40\ \%$. Different authors have used different material and its measurement results to simulate the same handwritten data from the MNIST database to achieve pattern recognition. In [46] authors have used $Pd/WO_x/W$

device to for pattern recognition and shown the change in memristor value before and after learning. Similarly [14] used TiO_x based device, [47] used $Ag/AgInSb/Ta$-based memristor, [48] used HfO_2 based memristor to simulate memristor device measurement data to perform successful pattern recognition in simulation. Memristor has also been implemented on full hardware with memristor device and FPGA [49]. Neurons, spiking circuits, and homeostasis have been implemented on FPGA and the measurements show successful pattern matching on hardware without any external processing. Memristor-based pattern recognition is less susceptible to noise due to self-adaptation of memristor after training. In [50] author implements $Pr_{0.7}Ca_{0.3}MnO_3$ (PCMO) device based pattern recognition and the system is able to learn 5X6 pixel image with $85\ \%$ recognition rate even at $10\ \%$ of input noise level.

1.2 Purpose of the work

Challenges of VNA, the requirement of high-efficiency computers, and development of memristor have encouraged the scientific community to develop memristor based neuromorphic processing. Development of neuromorphic processing includes: understanding of memristive device specification for neuromorphic applications, development of memristive devices, it's integration with CMOS circuit, it's evaluation and development of neuromorphic processing model.

Memristor has non-volatile adaptive memory and it has attracted scientific community for developing memristor based neuromorphic circuits. However, there is no commercially available memristor available till now and there is still ambiguity about the memristor behavior best suited for neuromorphic applications. Multiple researchers have developed many different memristive devices with different memristive functions [51], [52], [53], [54], [55] etc. Therefore, a memristor emulator is needed which is compact and externally programmable to implement any memristive function on hardware. Memristor emulator developed by other researchers are bulky and non-programmable ([56], [10], [57], [58]) which limits its usability in neuromorphic circuits. Development of Application Specific Integrated Circuit (ASIC) based emulator will provide freedom to implement a wide variety of memristive function and analyze the impact on learning behavior like the implementation of LTP, LTD and synaptic plasticity. Moreover, two on-chip neurons will provide an opportunity to mimic the combined learning activity of neuron and synapse on hardware.

Brain-inspired neuromorphic computing is considered to be a more efficient approach than digital approach than digital in terms of power and area. Moreover, digital computers have isolated processing and memory which faces technological challenges in dealing high volume of data due to communication bandwidth and the memory wall. Memristor-based neuromorphic processing is capable of handling huge data application like pattern recognition and audio processing due to its distributed computing in the network itself. In memristor based approach, a memristor is part of processing as well as memory. It obviates the need for separate memory and processing unit and emerges as a promising candidate for human-level cognitive tasks applications. Memristor and CMOS neuron based simulation of the neuromorphic circuit shows successful pattern recognition which is a challenge for digital computers when high definition image has to be recognized and processed. Since memristor has distributed memory and processing in the network, architecture can be extended in 3D architecture to recognize and process any high resolution of an image. It shows a wide application area where high-efficiency processors are required and lead towards the development of neuromorphic processors.

Neuromorphic computers are biologically inspired and mimic the functionalities of neuron and synapse in the brain as shown in figure 1.5. Synapse-neuron is repeated in a dense structure in our brain to perform all high-efficiency cognitive task. Therefore, for developing neuromorphic computer, the first step is to develop synapse-neuron unit which can be connected in bigger architecture to perform the human-level cognitive task. Development of NPU requires the development of memristor device, CMOS neuron circuit and memristor-CMOS integration on hardware as demonstrated in figure 1.6.

Development of memristor is beyond the scope of this work and has been developed by our project partner, Institute of Electrical Engineering and Information Technology, University of Kiel. Double barrier memristive devices based on a 4-inch wafer technology have been fabricated which consist of an ultra-thin memristive layer (Nb_xO_y) sandwiched between a Al_2O_3 tunnel barrier and a Schottky-like contact [51]. The presented device indicates a non-filamentary based continuous resistance change. The device shows some basic similarity to synapse in our brain and can be used in neuromorphic applications with CMOS circuits. The neuron ASIC mimics the fundamental properties of a biological neuron; it integrates all incoming current and generates a spike when the threshold of the neuron is reached. Since DBMD requires more than 3 V for changing its conductance, AMS350 nm process has been used for implementing ASIC. Integration of designed ASIC and memristor die develops NPU which is the basic unit of neuromorphic processing capable of learning and storing the memory in a single device. This

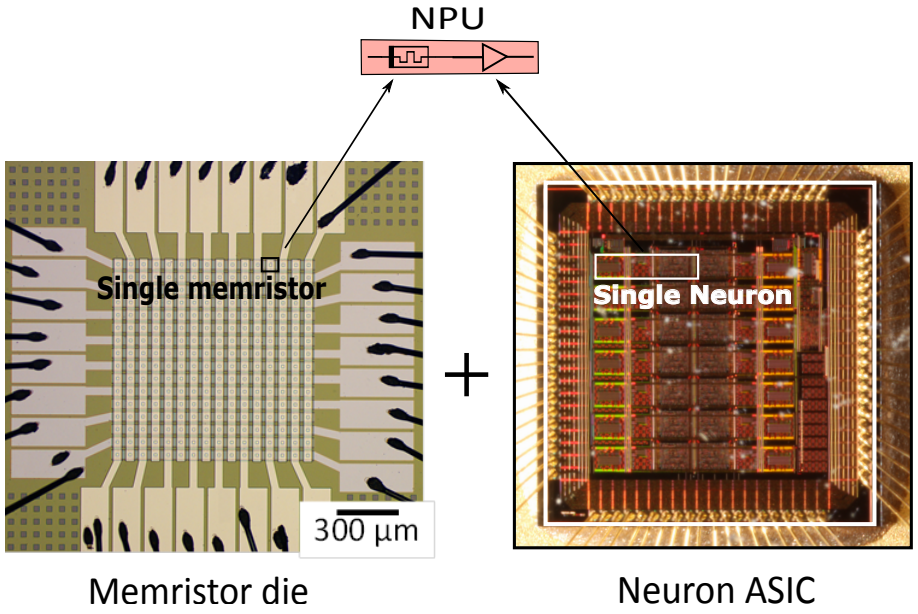

FIGURE 1.6: Memristor die and neuron ASIC integration and development of NPU

NPU paves the way towards the development of neuromorphic computers when arranged in a 3D array.

The author has to thank the German Research Foundation (DFG) for funding the research group FOR2093 (Memristive Components for Neural Systems), Project C2 (Neural Circuits), under which this research activity has been conducted.

1.3 Thesis outline

Chapter two presents the memristor working principle, and its development at nanoscale level. This chapter explains the memristor device developed in HP lab and discusses its measurement results. DBMD has been used in this work to develop memristive learning. It's architecture and advantages in neuromorphic application has been discussed in this chapter.

Chapter three describes the memristor emulator ASIC, its circuit design and measurement results. Since memristor is in its infant stage of development, it is not commercially available till now. Many memristor structure and voltage-current functions have been proposed and each one has different impact on the circuit learning [5], [51], [59], [60] etc. At this stage of research, there is still an ambiguity about the memristive function.

Every memristive function has their pros and cons. In order to understand the specification of memristor for neuromorphic application, an externally programmable memristor emulator hardware is needed. An on-chip memristor emulator has been designed which is compact and can be used in emulating any type of memristive function and can be used for bigger networks as well. The memristor emulator chip has been designed in AMS 350nm process and processing of the memristive function has been done off-chip on a FPGA. This selection allows the implementation of any memristive function on hardware. The emulator has shown a pinched hysteresis curve which is a fundamental behavior of the memristor. It shows two different approaches to implement the emulator on-chip, resistive based and switch capacitive based. It discusses the algorithm to implement the emulator on-chip together with the off-chip processing algorithm and its measurement results.

Chapter four presents the simulation of neuromorphic pattern recognition using the DBMD memristor device model and a CMOS neuron circuit. Pattern recognition simulation of four different patterns in LTSpice simulation software has been shown. It explains the neuron learning algorithm and necessary elements as competition, homeostasis, and variability. This chapter explain the limitations of VNA in implemnting ANN in digital computers and discusses the usability of neuromorphic circuits in high data application like pattern recognition.

Chapter five explains the design of a neuron ASIC on CMOS AMS350 nm process and its measurement results. CMOS neuron circuit had to be designed in such a way that it works like a neuron in the human brain and can be integrated with DBMD to perform pattern recognition which has been done only at simulation level in [45], [14]. This chapter also explains the details of neuron architecture and design considerations of analog building blocks like operational amplifier, comparator, integrator and spike generator circuit used in this neuron ASIC implementation.

Chapter six explains the integration of the CMOS neuron circuit with a real memristor (DBMS). Single memristor device has to be integrated with a CMOS designed neuron chip to develop NPU which can be repeated in 3D array for high effiecient processing like biological neural network. It show the fundamental behavior of learning like Long-Term Potentiation (LTP) which is one of the key components of learning.. The designed neuron chip has an integrate-and-fire circuit (I & F) with an on-chip spike generator circuit. The neuron circuit integrates the current that flows through the memristor at low voltages (read voltage) and generates the potentiation or depression pulses (firing) across the memristor. These pulses result in long-term potentiation or depression of the

memristor [61], [62]. The DBMD memristor shows 100 times change in conductance (100 times stronger synaptic contact) with consecutive pulses resulting in learning of memristor (increased firing rate of the neuron). This explains the integration approach and hardware measurement results with memristor retention issue. This chapter shows the learning behavior on hardware level.

Chapter seven presents conclusions and outlook of this work. It proposes neuromorphic processor model based on NPU.

Chapter 2

Memristor

Memristor is a two terminal device which changes its resistance when a voltage is applied across its terminals. The change in resistance depends upon the voltage polarity, voltage level and time duration for which voltage is applied. Memristor resistance is plastic in nature which retains its resistance when the voltage is turned off until next voltage is applied. Memristor has a positive and negative terminal as shown in figure 2.1.

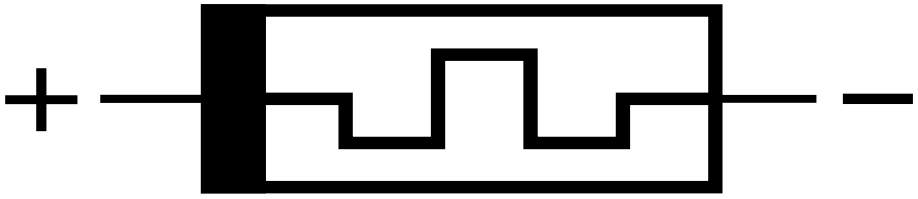

FIGURE 2.1: Memristor Symbol

When the voltage at the positive terminal is higher than the voltage at its negative terminal, the memristor becomes more conducting, while it becomes less conducting when the voltage at its negative terminal is higher. The amount of change in conductance depends on the duration for which a positive voltage is applied. When the voltage is turned off, the memristor freezes its resistance state. The change in resistance is due to the change in position of ions which migrate from one location to other in the device when exposed to a high electric field. The applied electric field has to be high to move ions and to generate a high electric field. To generate a high electric field with low voltage, the interface layer must be very thin in the nanometer scale. So the memristive effect can only be observed in the nanoscale devices.

2.1 How memristor works

The memristor has been built with different materials and different architectures which shows different behaviors. Some examples are the niobium-based (Nb_xO_y) memristor [51], and the Titanium based in (TiO_2) [5] etc. There are two basic principles for memristive devices, oxide drift based [51], [52], [53] and conductive filament based [54], [55].

In the case of the oxide-based memristor when a voltage is applied, oxygen ions change their position and makes the overall conductance change. Titanium dioxide (TiO_2) based device is a case of an oxide-based memristor.

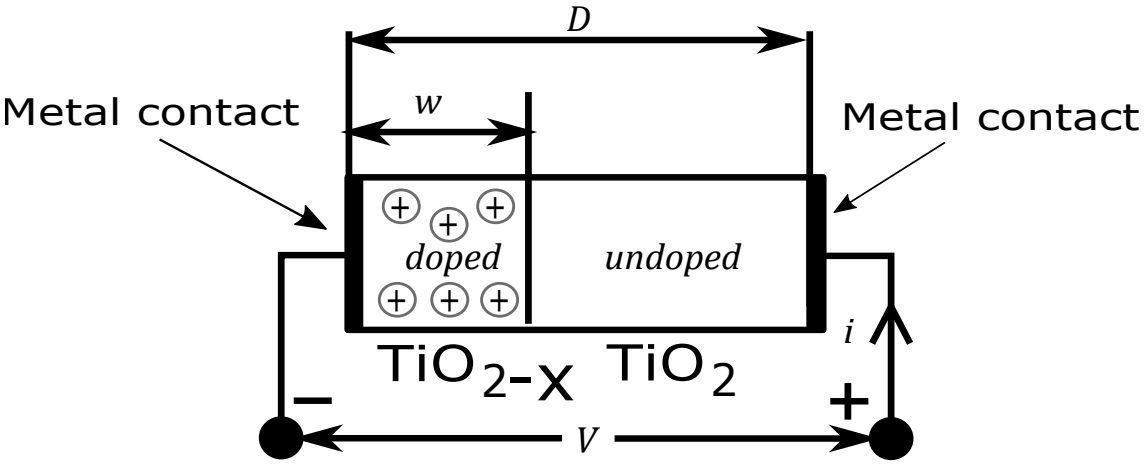

FIGURE 2.2: Memristor device model

The Titanium dioxide (TiO_2) memristor device is a nanometer cube of titanium dioxide (TiO_2) in two layers between two electrodes (metal contact) as shown in figure 2.2. The right layer (TiO_2) has 1:2 titanium to oxygen ratio which acts as an insulator, the left layer has missing oxygen ions which makes this layer oxygen deficient and therefore conductive. When a positive voltage is applied, oxygen ions move from left to right which increases the width of the doped region. Increase in the doped region width (w) eventually rises the overall conductance of memristor. When a negative voltage is applied, some oxygen ions move towards the left side reducing the width of the doped region which then reduces the overall conductance of the memristor as shown in figure 2.3.

This device also shows the special memristive effect that when the voltage is turned off, the oxygen ions do not migrate. The oxygen ions stay where they were and freeze the boundary between the two titanium oxide layers. This phenomenon makes memristor