ON-CHIP LEARNING IN A SPINTRONICS-BASED HARDWARE NEURAL NETWORK: A DEVICE-CIRCUIT-SYSTEM CO-STUDY

DIVYA KAUSHIK

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

January 2023

ON-CHIP LEARNING IN A SPINTRONICS-BASED HARDWARE NEURAL NETWORK: A DEVICE-CIRCUIT-SYSTEM CO-STUDY

by

DIVYA KAUSHIK

Submitted

in fulfillment of the requirements of the degree of

Doctor of Philosophy

to the

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

January 2023

Dedicated to

My Family

CERTIFICATE

This is to certify that the thesis entitled **" ON-CHIP LEARNING IN A SPINTRONICS-BASED HARDWARE NEURAL NETWORK: A DEVICE-CIRCUIT-SYSTEM CO-STUDY "** being submitted by Ms. **Divya Kaushik** for the award of the degree of Doctor of Philosophy in the Department of Electrical Engineering, Indian Institute of Technology Delhi, is a record of bonafide work done by her under my supervision and guidance. The matter embodied in this thesis has not been submitted for the award of any other degree or diploma. I consider her research work substantial enough for the award of the doctorate.

Place: New Delhi
Date:

Prof. Debanjan Bhowmik
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology, Delhi
Hauz Khas, New Delhi- 110016, India

ACKNOWLEDGEMENTS

(The acknowledgment section has been revised to address referee 2's comment 27.)

First and foremost, I am incredibly grateful to my supervisor Prof. Debanjan Bhowmik for his inspiration, invaluable advice, unfailing continuous support, and patience during my Ph.D. study. His immense knowledge and ample experience have encouraged and helped me throughout my academic research and daily life. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my supervisor, I would like to thank my thesis committee members: Prof. Shouri Chatterjee, Prof. A P Prathosh, and Prof. Pranaba Kishor Muduli, for their insightful comments and encouragement and for their insightful comments and encouragement the relevant suggestions and questions that helped me to widen the scope my research.

The completion of this project could not have been accomplished without the support of my lab-mates. Ms. Upasana Sahu has helped me enjoy my work. I will never forget our precious tea-time discussions. I also wish to thank undergraduate students who worked in my research group under the supervision of Prof. Debanjan Bhowmik: Utkarsh Saxena, Janak Sharda, Utkarsh Singh, and Varun Bhavin Desai, for their help with developing device models, coding, etc.

I am thankful to all my fellow lab mates (Mr. Manoj Kumar and Mrs. Kritika) for the stimulating discussions, the sleepless nights they worked with me before deadlines, and the fun we had during those days. Their support was incredible, particularly during the highly uncertain pandemic situation.

I would like to express my gratitude to my father, Prof. Subodh Ranjan, my mother, Mrs. Suman Lata, my brother Mr. Nikhil Kaushik, my husband, Mr. Praveen Kumar, and my mother-in-law Ms. Bimla Devi. Without their tremendous understanding and encouragement, it would not be possible to complete my study. My friend cum sister Dr.

Deepti Singh helped me in every phase of my life: from applying for Ph.D. to making balance in my life after my marriage. Her constant encouragement helped a lot. It would have been challenging to work in the COVID-19-induced lockdown situation without my family's support. Because of their love, care, and support, I have completed my research work.

Last, but not the least, I would want to praise and thank God, the Almighty, for bestowing many benefits, knowledge, and opportunities, allowing me to complete the thesis. I believe that one can only truly appreciate what one has been given; therefore, I would like to thank You, God, for life and all it contains.

Divya Kaushik

ABSTRACT

KEYWORDS: Spintronics; neuromorphic computing; non von Neumann computing ; spin-orbit torque; domain-wall device; on-chip learning; gradient descent algorithm; back-propagation algorithm; fully connected neural network (FCNN).

The abstract has been revised to address referee 2's comment 1.

This thesis focuses on the design and implementation of a neuromorphic computing scheme: using a domain-wall-synapse-based crossbar array, device-circuit-system co-design and co-simulation of on-chip learning on a fully connected neural network (FCNN) are shown here. A combination of micromagnetic-physics-based synapse-device modeling, SPICE simulation of a crossbar-array circuit using such synapse devices, and system-level coding using a high-level language (Python) is used for the purpose. While training the FCNN, the gradient descent algorithm and back-propagation algorithm (extension of the gradient descent algorithm for multi-layer FCNNs) are utilized, and analog crossbar array implementation of the same is shown for data classification applications. Matrix vector multiplication during forward computation is performed with time and energy efficiency with the help of this analog crossbar array architecture. The architecture requires a non-volatile synaptic device for storing the weights/conductance of each synapse in the neural network. On-chip learning is achieved by modulating the conductances of the synapses, corresponding to weights stored in synapses, with electrical programming pulses, at every iteration.

For this purpose, a ferromagnetic-metal-heavy-metal-based spin-orbit torque-driven domain wall-based synaptic device (a spintronic device) is proposed here as a non-volatile synaptic device in the crossbar array. This device is shown here to exhibit a linear and symmetric conductance response. On the contrary, conductance response characteristics of existing synaptic devices, Resistive Random Access Memory (RRAM) or a Phase Change

Memory (PCM) device are highly non-linear and asymmetric (between positive and negative conductance update). This leads to issues with the design of peripheral circuits for on-chip learning, and network performance also gets affected.

The ferromagnetic-metal-heavy-metal-based spintronic domain-wall device (my proposed synaptic device here) is simulated here through micromagnetics. Although the conductance variation range is much smaller for the domain-wall synapse compared to that for RRAM and PCM synapses, it is shown here through micromagnetic simulation that the conductance response of the domain-wall synapse to programming current pulses is highly linear and symmetric, unlike that of RRAM and PCM synapses. This conductance response is then incorporated in Verilog-A modules. Inserting these Verilog-A modules as synaptic devices at intersection points of crossbar architecture, SPICE circuit simulations are performed. The conductance of the domain-wall synapse has been quantized here to take the effect of domain wall pinning by defects into account. Despite the quantization, high accuracy is obtained in my circuit simulations on a popular machine learning data set: Fisher's Iris data set. Also, the time taken and energy consumed for on-chip learning of the DW synapse-based FCNN circuit are orders of magnitude lower than that for RRAM and PCM synapse-based FCNN circuits.

In the later part of the thesis, a modification to the standard gradient descent algorithm, for training neural networks, is proposed using appropriate thresholding units. This leads to optimization of the synapse cell at each intersection of the crossbar architectures, making the system scalable. For the system to approximate a wide range of functions for data classification, a hidden layer is added. The back-propagation algorithm (extension of gradient descent algorithm for multi-layered FCNN) for on-chip learning is implemented next. Through a combination of micromagnetic and SPICE circuit simulations like before, an improved accuracy is shown for the domain wall synapse-based FCNN with a hidden layer compared to that without a hidden layer, for different machine learning data sets. How the classification accuracy of the modified design is affected by device-to-device/process variations and noise (additive and multiplicative noise) is also studied.

सारांश

यह थीसिस एक न्यूरोमॉर्फिक कंप्यूटिंग योजना के डिजाइन और कार्यान्वयन पर केंद्रित है: डोमेन-वॉल-सिनैप्स-आधारित क्रॉसबार सरणी, डिवाइस-सर्किट-सिस्टम सह-डिज़ाइन और पूरी तरह से जुड़े तंत्रिका नेटवर्क पर ऑन-चिप सीखने के सह-अनुकरण का उपयोग करना। एफसीएनएन) यहां दिखाए गए हैं। माइक्रोमैग्नेटिक-फिजिक्स-आधारित सिनैप्स-डिवाइस मॉडलिंग का एक संयोजन, इस तरह के सिनैप्स उपकरणों का उपयोग करके एक क्रॉसबार-एरे सर्किट का स्पाइस सिमुलेशन, और एक उच्च-स्तरीय भाषा (पायथन) का उपयोग करके सिस्टम-लेवल कोडिंग का उपयोग इस उद्देश्य के लिए किया जाता है। एफसीएनएन को प्रशिक्षित करते समय, ग्रेडिएंट डिसेंट एल्गोरिथम और बैक-प्रॉपेगेशन एल्गोरिदम (मल्टी-लेयर एफसीएनएन के लिए ग्रेडिएंट डिसेंट एल्गोरिथम का विस्तार) का उपयोग किया जाता है, और डेटा वर्गीकरण अनुप्रयोगों के लिए उसी का एनालॉग क्रॉसबार ऐरे कार्यान्वयन दिखाया गया है। आगे की गणना के दौरान मैट्रिक्स वेक्टर गुणन इस एनालॉग क्रॉसबार ऐरे आर्किटेक्चर की मदद से समय और ऊर्जा दक्षता के साथ किया जाता है । तंत्रिका नेटवर्क में प्रत्येक अन्तर्ग्रथन के भार / चालन को संग्रहीत करने के लिए आर्किटेक्चर को एक गैर-वाष्पशील सिनैप्टिक डिवाइस की आवश्यकता होती है। प्रत्येक पुनरावृत्ति पर विद्युत प्रोग्रामिंग दालों के साथ सिनैप्स में संग्रहीत भार के अनुरूप , सिनैप्स के संचालन को संशोधित करके ऑन-चिप लर्निंग प्राप्त की जाती है।

इस उद्देश्य के लिए, एक फेरोमैग्नेटिक-मेटल-हेवी-मेटल-आधारित स्पिन-ऑर्बिट टॉर्क-चालित डोमेन वॉल-आधारित सिनैप्टिक डिवाइस (एक स्पिंट्रोनिक डिवाइस) यहां क्रॉसबार सरणी में एक गैर-वाष्पशील सिनैप्टिक डिवाइस के रूप में प्रस्तावित है। यह डिवाइस एक रैखिक और सममित चालन प्रतिक्रिया प्रदर्शित करने के लिए यहां दिखाया गया है। इसके विपरीत, मौजूदा सिनैप्टिक डिवाइस, रेसिस्टिव रैंडम एक्सेस मेमोरी (आरआरएएम) या फेज चेंज मेमोरी (पीसीएम) डिवाइस की चालन प्रतिक्रिया विशेषताएँ अत्यधिक गैर-रैखिक और असममित (सकारात्मक और नकारात्मक चालन अद्यतन के बीच) हैं। इससे ऑन-चिप सीखने के लिए परिधीय सर्किट के डिजाइन के साथ समस्याएं होती हैं, और नेटवर्क प्रदर्शन भी प्रभावित होता है।

फेरोमैग्नेटिक-मेटल-हेवी-मेटल-आधारित स्पिंट्रोनिक डोमेन-वॉल डिवाइस (यहां मेरा प्रस्तावित सिनैप्टिक डिवाइस) को माइक्रोमैग्नेटिक्स के माध्यम से यहां सिम्युलेटेड किया गया है । यद्यपि आरआरएएम और पीसीएम सिनैप्स की तुलना में डोमेन-वॉल सिनैप्स के लिए चालन भिन्नता रेंज बहुत छोटी है, लेकिन यह माइक्रोमैग्नेटिक सिमुलेशन के माध्यम से यहां दिखाया गया है कि डोमेन-वॉल सिनैप्स की प्रोग्रामिंग वर्तमान दालों की चालन प्रतिक्रिया अत्यधिक रैखिक और सममित है। आरआरएएम और पीसीएम सिंकैप्स के विपरीत। यह चालन प्रतिक्रिया तब Verilog-A मॉड्यूल में शामिल की जाती है। इन वेरिलॉग-ए मॉड्यूल को क्रॉसबार आर्किटेक्चर के इंटरसेक्शन पॉइंट्स पर सिनैप्टिक डिवाइस के रूप में सम्मिलित करते हुए, स्पाइस सर्किट सिमुलेशन का प्रदर्शन किया जाता है। दोषों द्वारा डोमेन वॉल पिनिंग के प्रभाव को ध्यान में रखते हुए डोमेन-वॉल सिनैप्स के संचालन को यहाँ परिमाणित किया गया है। परिमाणीकरण के बावजूद, एक लोकप्रिय मशीन लर्निंग डेटा सेट पर मेरे सर्किट सिमुलेशन में उच्च सटीकता प्राप्त की जाती है: फिशर आइरिस डेटा सेट। इसके अलावा, डीडब्ल्यू सिनैप्स-आधारित एफसीएनएन सर्किट के ऑन-चिप सीखने के लिए लिया गया समय और ऊर्जा आरआरएएम और पीसीएम सिनैप्स-आधारित एफसीएनएन सर्किट की तुलना में कम परिमाण के आदेश हैं।

थीसिस के बाद के भाग में, उपयुक्त थ्रेशोल्डिंग इकाइयों का उपयोग करके तंत्रिका नेटवर्क के प्रशिक्षण के लिए मानक ग्रेडिएंट डिसेंट एल्गोरिथम में संशोधन प्रस्तावित है। इससे क्रॉसबार आर्किटेक्चर के प्रत्येक चौराहे पर सिनैप्स सेल का अनुकूलन होता है, जिससे सिस्टम स्केलेबल हो जाता है। डेटा वर्गीकरण के लिए कार्यों की एक विस्तृत श्रृंखला का अनुमान लगाने के लिए सिस्टम के लिए, एक छिपी हुई परत जोड़ी जाती है। ऑन-चिप लर्निंग के लिए बैक-प्रपोगेशन एल्गोरिथम (मल्टी-लेयर्ड एफसीएनएन के लिए ग्रेडिएंट डिसेंट एल्गोरिथम का विस्तार) को आगे लागू किया गया है। पहले की तरह माइक्रोमैग्नेटिक और स्पाइस सर्किट सिमुलेशन के संयोजन के माध्यम से, डोमेन वॉल सिनैप्स-आधारित FCNN के लिए एक छिपी हुई परत के बिना एक छिपी हुई परत के साथ विभिन्न मशीन लर्निंग डेटा सेट के लिए एक बेहतर सटीकता दिखाई गई है। डिवाइस-टू-डिवाइस/प्रक्रिया विविधताओं और शोर (एडिटिव और मल्टीप्लिकेटिव शोर) से संशोधित डिज़ाइन की वर्गीकरण सटीकता कैसे प्रभावित होती है, इसका भी अध्ययन किया जाता है।

Contents

ACKNOWLEDGEMENTS **ii**

ABSTRACT **iv**

LIST OF TABLES **ix**

LIST OF FIGURES **xiii**

ABBREVIATIONS **xiv**

NOTATION **xiv**

NOTATION **xvi**

1 Introduction and Thesis Outline **1**

1.1 Artificial Intelligence/Machine Learning 1

1.2 Supervised Machine Learning 3

1.3 Neuromorphic Computing 4

1.4 Fully Connected Neural Network 6

1.5 Crossbar-Array based Analog Hardware Neural Network 8

1.6 Motivation for On-Chip Learning 10

1.7 Outline of the Thesis 11

2 *Background Material on Devices and Algorithms* **14**

2.1 Existing Non-volatile Synaptic Devices 14

2.1.1 Resistive Random Access Memory (RRAM) devices 15

2.1.2 Phase Change Memory (PCM) devices 16

2.2 *Alternative non-volatile synaptic device: Domain-wall Device* 17
2.3 Learning Algorithm for FCNNs 20
2.3.1 FCNN without a hidden layer 20
2.3.2 FCNN with a hidden layer 22
2.3.3 Effects of process/device-level variations and 'write' noise 25

3 *Comparing the Synaptic Behaviour of the Domain-Wall Device with that of RRAM and PCM Devices* 26

3.1 Proposed Spintronics-based Synaptic Device 26
3.1.1 Working Principle 26
3.2 Device Simulations 28
3.2.1 Device Specifications 29
3.2.2 Micromagnetic Simulations 30
3.3 *Experimental Calibration* 33
3.4 System Simulations 34
3.5 *Peripheral Circuitry* 35
3.6 Circuit Simulations 36
3.7 Network-Training Results based on the Domain-Wall Device 40
3.8 Comparison with alternative synaptic technology: RRAM and PCM 41
3.9 *Energy Calculation Method* 44

4 Proposed Modified Learning Algorithm 47

4.1 Design of proposed optimized 'synapse cell' 49
4.1.1 Thresholding Circuit 50
4.1.2 2-Transistor Multiplier Circuit 51
4.2 SPICE simulation of the proposed 'synapse cell.' 53
4.3 Optimized 'synapse cell'-based FCNN with circuit implementation of back-propagation algorithm. 55

4.3.1 Device Simulations 55

4.3.2 System Simulations 56

4.3.3 Circuit Simulations 58

4.4 Design of optimized 'synapse cell'-based multi-layer FCNN. 58

4.5 Circuit simulation for comparing classification accuracy of single-layer and multi-layer FCCN 61

5 Device Variations and Noise Analysis **65**

5.1 Process/ Device-to-Device Variations 65

5.2 Cycle-to-Cycle Variation/ Write Noise 68

5.2.1 Multiplicative Noise 68

5.2.2 Additive Noise 70

6 *Conclusions* **76**

List of Tables

3.1 Different simulation parameters used with their respective values while performing micromagnetic simulations on 'mumax3'. ***(The term 'PMA constant' has been replaced by 'anisotropy constant' to address referee 2's comment 15.)*** 30

3.2 Performance comparison between Domain Wall, RRAM and PCM based fully connected neural networks (single-layer FCNN) for on-chip learning.Classification accuracy, time and energy consumed for the three cases are highlighted. . . 44

4.1 Performance metrics (classification accuracy, time taken and energy consumed) for on-chip learning using the domain wall synapse based FCNN with a hidden layer (Figure 4.6) on Fisher's Iris and MNIST datasets. 64

List of Figures

1.1 Machine Learning is a sub-field of Artificial Intelligence. 2

1.2 A single-layer perceptron, or a neural network with only an input layer and an output layer (no hidden layer included). The weighted sum of inputs applied to the non-linear activation function computes forward computation. . . . 7

1.3 Schematic of a crossbar-array based neural network with an input layer and an output layer. Non-volatile memory elements are used as synapses at the junctions of the horizontal and vertical crossbars. Crossbar calculates matrix-vector multiplication in a parallel fashion mimicking the human brain. . . . 9

2.1 Schematic of the domain-wall-based synaptic device is shown. The heterostructure has a 'read' and 'write' path. R_{read} and R_{write} are the resistance of the 'read' and 'write' path, respectively, and I_{read} and I_{write} are the 'read' and 'write' current. In-plane I_{write} flowing through the heavy metal layer causes movement of the domain wall, changing the conductance of the 'read' path. The orientation of alignment of magnetization of the free layer changes from one to another as the domain wall moves. The antiferromagnetic layer at the edges over the ferromagnetic free layer pins the magnetic moments of the ferromagnetic layer, preventing the domain wall from getting destroyed at the edge. 18

2.2 FCNN without any hidden layer is known as single-layer (perceptron) neural network. Weights are updated using standard gradient descent technique using a non-linear activation function. 21

2.3 FCNN with one or more hidden layers is known as multi-layer neural network. Here weights of first sub-network are updated using backpropagation algorithm and that of second sub-network using standard gradient descent technique similar to that used in single-layer neural network. 23

3.1 Velocity vs in-plane current density, obtained from micromagnetic simulation. After the certain threshold point, the graph shows linear conductance characteristics. 31

3.2 (a) 'Write' current pulses applied on the heavy metal layer of constant magnitude in both polarities as a function of time. (b) Domain Wall motion in the ferromagnetic layer for different time instants corresponding to different current pulses of fixed magnitude as shown in (a). Domain Wall moves from its initial position when in-plane current flows through heavy metal due to spin-orbit torque on the magnetization of the ferromagnetic layer above it. The cross product of $\vec{M_{avg}}$ and $\vec{\sigma}$ decides the direction in which the domain wall will move.(The waveform for the programming current has been added (sub-figure (a)) to address referee 2's comment 16. 32

3.3 Schematic of crossbar architecture with domain wall based synapse along with transistor based analog peripheral circuits for in-chip learning of single-layer FCNN. Conductance of the DW devices takes positive values and ranges between G_{min} and G_{max}. Corresponding weights can take positive and negative values. Extra conductance ($G_{parallel}$) is added in parallel to each synapse, and the negative of the voltage applied to the synapse is applied to it (10). *The unit 'Ohm' has been added in the figure to describe the unit for the resistance R and thus address referee 1's comment 12.* 38

3.4 Transistor based implementation of quantizer circuit. Input-output characteristic of the circuit, obtained from SPICE simulation, shows quantizing behaviour. 39

3.5 Conductance response of 'read' path vs. in-plane 'write' current pulse. Conductance increases or decreases in fixed step (ΔG) pattern due to the application of fixed magnitude current pulse in either polarity. 40

3.6 (a) Quantized 'write' current pulse applied on a particular domain wall synapse while performing on-chip learning of the overall domain wall synapse based FCNN on the Fisher's Iris dataset. (b) Corresponding change in conductance of the domain wall synapse. The arrow pointing downwards shows the decrease in conductance, while the arrow pointing upwards indicates the increase in conductance corresponding to the polarity of the applied I_{write}. (This conductance change is reflected in the next epoch.) 41

3.7 Train and test accuracy of domain wall synapse-based neural network circuit as a function of epochs during on-chip learning on Fisher's Iris dataset. The proposed quantizing concept is implemented. 42

4.1 *Schematic of the modified 'synapse cell': 2-transistor multiplier circuit and the domain-wall device form the 'synapse cell.' This figure has been revised (the 'ground' symbol removed) to address referee 1's comment 10.* 48

4.2 (a) Operational amplifier based implementation of thresholding circuit. (b) Input-output characteristics of the circuit, as obtained from SPICE simulation. The range of characteristics graph is a bit different from the last case. This is explained later in the chapter. 50

4.3 *SPICE simulation of the 'synapse cell' in Figure 4.1 is carried out here. Inputs to the cell $Q_1(x_m)$ and $Q_2(\Delta w_n)$ are plotted as function of time. 'write' current flowing into the domain wall synapse in the cell and corresponding change in conductance in 'read' path of domain wall synapse are also plotted as function of time. This figure was there in the original thesis; but now the figure caption has been highlighted through italicizing to address referee 1's comment 10.* 54

4.4 Conductance response of 'read' path vs. in-plane 'write' current pulse. Conductance increases or decreases in fixed step pattern due to the application of fixed magnitude current pulse in either polarity. 56

4.5 Schematic of domain wall synapse based crossbar FCNN, without a hidden layer, along with analog peripheral circuits for on-chip learning (utilizing modified 'synapse cell.') 57

4.6 Schematic of domain wall synapse based crossbar FCNN, with a hidden layer, along with analog peripheral circuits for on-chip learning. Circuit makes use of 3 crossbars. Modified 'synapse cell' with 2-transistor-based multiplier and domain wall device is used. Thresholding is at input and common part of weight update. 59

4.7 Train and test accuracy obtained from SPICE simulations are plotted as function of epochs for Fisher's Iris dataset (a) for domain wall synapse based single-layer FCNN (Figure 4.5) and (b) for domain wall synapse based FCNN with a hidden layer (Figure 4.6). Performance of the single-layer FCNN is quite low with the modified thresholding algorithm. 62

4.8 Train and test accuracy from SPICE circuit simulations are plotted as function of epochs for MNIST dataset using modified synapse cell (a) for single-layer FCNN and (b) for FCNN with a hidden layer. 63

5.1 (a) Train accuracy (b) Test accuracy of the proposed FCNN are shown as a function of training epochs for Fisher's Iris dataset, taking different levels of process/ device-to-device variations into account. The accuracy plots clearly show that the effects of different levels of process variations on train and test accuracy are almost negligible. 66

5.2 a) Train accuracy (b) Test accuracy of the proposed FCNN are shown as a function of training epochs for MNIST dataset, taking different extents of device-to-device variations into account. Like the above Figure 5.1, train and test accuracy numbers are negligibly affected by process variations. 67

5.3 (a) Train accuracy (b) Test accuracy of the proposed FCNN is shown as a function of training epochs for Fisher's Iris dataset taking different levels of multiplicative noise into account. 0% process variation means no variation, and the case is ideal. Hardly any variation in the train and test accuracies is observed due to different extents of multiplicative noise. 69

5.4 (a) Train accuracy (b) Test accuracy of the proposed FCNN are shown as a function of training epochs for MNIST dataset taking different levels of multiplicative noise into account. Similar to the case of process variations, multiplicative noise does not create much change in the classification accuracy numbers. 70

5.5 Train accuracy of the proposed FCNN is shown as a function of training epochs for Fisher's Iris dataset, taking different levels of additive noise into account. Train accuracy numbers (a) with 5% additive noise (b) with 10% additive noise (c) with 15% additive noise (d) with 20% additive noise are plotted. It is evident from the accuracy plots that with increasing extents of the additive noise, the training of the network degrades. 72

5.6 Test accuracy of the proposed FCNN is shown as a function of training epochs for Fisher's Iris dataset with different levels of additive noise variations. Test accuracy numbers (a) with 5% additive noise (b) with 10% additive noise (c) with 15% additive noise (d) with 20% additive noise are plotted. Evidently, with increasing extents of the additive noise, test accuracy degrades drastically. 73

5.7 Train accuracy of the proposed FCNN is shown as a function of training epochs for MNIST dataset, taking different levels of additive noise into account. Train accuracy (a) with 5% additive noise (b) with 10% additive noise (c) with 15% additive noise (d) with 20% is plotted. The graphs clearly show that additive noise affects the performance of the proposed domain wall-based FCNN, and the train accuracy number falls drastically with increasing additive noise. 74

5.8 Test accuracy of the proposed FCNN is shown as a function of training epochs for MNIST dataset, taking different levels of additive noise into account. Test accuracy (a) with 5% additive noise (b) with 10% additive noise (c) with 15% additive noise (d) with 20% additive noise are plotted. The graphs clearly show that additive noise affects the performance of the proposed domain wall-based FCNN, and the test accuracy number falls drastically with increasing additive noise. 75

ABBREVIATIONS

IITD	Indian Institute of Technology, Delhi
AI	Artificial Intelligence
ML	Machine Learning
NN	Neural Network
DL	Deep Learning
SGD	Stochastic Gradient Descent
VMM	Vector-weight Matrix Multiplication
MVM	Matrix Vector Multiplication
ALU	Arithmetic and Logic Unit
FCNN	Fully Connected Neural Network
MVM	Matrix Vector Multiplication
SVM	Support Vector Machines
NVM	Non-Volatile Memory
MTJ	Magnetic Tunnel Junction
TMR	Tunnel Magnetoresistance
RRAM	Resistive Random Access Memory
PCM	Phase Change Memory
DW	Domain Wall
SOT	Spin Orbit Torque
ADC	Analog-to-Digital Converter
DAC	Digital-to-Analog Converter
CMOS	Complementary Metal-Oxide Semiconductor
FPGA	Field Programmable Gate Array
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
CPU	Central Processing Unit
GPU	Graphics processing unit
DMI	Dzyaloshinskii–Moriya Interaction

KCL	Kirchhoff's Current LAw
Pt	Platinum
ReLu	Rectified Linear Activation Unit
tanh	Hyperbolic Tangent
PMA	Perpendicular Magnetic Anisotropy
DRAM	Dynamic Random Access Memory
SRAM	Static Random Access Memory
STTMRAM	Spin-transfer torque MRAM
MNIST	Modified National Institute of Standards and Technology

NOTATION

α	Damping ratio
A	Exchange correlation constant
Δw	Change in weight of the synapse
ΔG	Change in conductance
G_{max}	Maximum conductance of the MTJ
G_{min}	Minimum conductance of the MTJ
J_s	Spin current density
J_c	Charge current density
K	Perpendicular Magnetic Anisotropy (PMA) constant
Ms	Saturation magnetization
θ_{SH}	Spin hall angle
R_{read}	Resistance of the 'read' path of the domain wall device
R_{write}	Resistance of the 'write' path of the domain wall device
I_{read}	'Read' current of the synapse
I_{write}	'Write' Current of the synapse
m_z	average magnetization
f	non-linear activation function
θ_{SH}	Spin hall angle
λ	Hyper-parameter used during learning
η	Learning rate
$\sum$	Summation