Noise Sensitivity Analysis of Deep Belief Networks: A Monte Carlo Simulation for Memristive Crossbars

Adedoyin Adepegba, Ramtin Zand, and Ronald F. DeMara
Department of Engineering and Computer Science
University of Central Florida
Orlando, FL, USA 32816-2362

Abstract— Deep Belief Networks (DBNs) offer promising alternatives to conventional processors to conduct unsupervised learning with reduced chip area and energy consumption. Nonetheless, their reliance on analog computation warrants verification of noise-tolerance which to-date has been lacking in the literature. Herein, the noise sensitivity of a representative DBN circuit implementation using resistive weights and probabilistic spin logic devices as stochastic binary neurons is assessed. A Probabilistic Inference Network-Simulator (PIN-Sim) framework was adapted to optimize the circuit implementation of *784×10* and *784×200×10* DBN topologies. Noise-induced voltage variations ranging from ±1mV to ±20mV are applied to the input of the p-bit based neurons. Circuit simulation results indicate maximum fluctuations of 3% and 1.4% in the recognition accuracy of *784×10* and *784×200×10* DBNs, respectively. Thus, acceptable robustness from noise-induced variations are achievable for the investigated hybrid spintronic/CMOS hardware implementation of the DBNs under the conditions assessed herein.

Keywords—Deep Belief Network (DBN), Probabilistic Spin Logic Devices, Magnetic Tunnel Junction, p-bit, noise.

I. INTRODUCTION

In recent decades, the fields of machine learning (ML), artificial intelligence (AI) and artificial neural networks (ANN) have grown significantly [1]. The most commonly used technique used in ANNs rely on supervised learning, where the error rate is measured by comparing the output from the network with a known desired output. Unsupervised learning approaches, used in intelligent biological systems, allow systems to evolve as they encounter and analyze more data samples. Deep belief networks (DBNs) are a class of ML techniques that utilize an unsupervised learning approach, which demonstrates impressive learning abilities for various applications such as natural language understanding [1]. Recently, an energy-efficient hybrid spintronic/CMOS based hardware implementation of DBNs is proposed in [2], which we have leveraged herein to analyze the effect of noise on DBNs using the popular MNIST pattern recognition application [2].

II. DEEP BELIEF NETWORK

Figure 1 shows the structure of a Deep Belief Network (DBN), which is comprised of hierarchically-connected Restricted Boltzmann Machines (RBMs). RBMs consist of two fully-connected non-recurrent layers, called the *hidden layer* and *visible layer*. Neurons in the RBM structure compute a probabilistic sigmoidal activation function, and the weights can be trained using the contrastive divergence algorithm proposed in [3]. As shown in Figure 1, RBMs can be implemented using memristive crossbar architectures. Recently, stochastic spintronic devices are proposed to be utilized within the DBNs to realize the probabilistic sigmoidal activation functions [2].

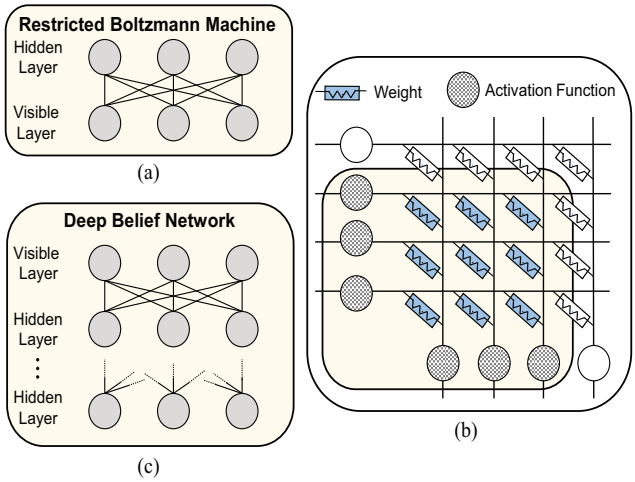

Fig. 1. (a) An RBM structure, (b) RBM implemented by crossbar architecture, (c) a DBN structure including multiple hidden layers [2].

III. SPIN-BASED PROBABILISTIC ACTIVATION FUNCTION

In [4], probabilistic spin logic devices (p-bits) are proposed as the building blocks to realize stochastic binary neurons generating a probabilistic sigmoidal activation function. In particular, Magnetic Random Access Memories (MRAMs), which are conventionally used for memory applications are modified to generate a probabilistic sigmoidal function. Magnetic Tunnel Junctions (MTJs), which are the building block for embedded MRAM-based p-bits, include two nanomagnets that are separated by an oxide layer. MTJs have a high or low resistance level based on the relation between its magnetization directions. In a conventional MTJ device, there is an energy barrier between its high and low resistive levels which makes it a non-volatile device. However, recent experiments have proposed methods to realize an unstable, near-zero energy barrier MTJ device, the resistance of which can randomly fluctuate in presence of thermal noise [5]. These MTJ devices with stochastic nanomagnets are used as the building blocks for p-bit based neurons realizing the probabilistic sigmoidal function required in DBN architectures [2, 3]. Energy efficiency gains arise by replacing the floating-point cross-product calculations with an analog resistive crossbar, although their narrow operating margins may be susceptible to noise.

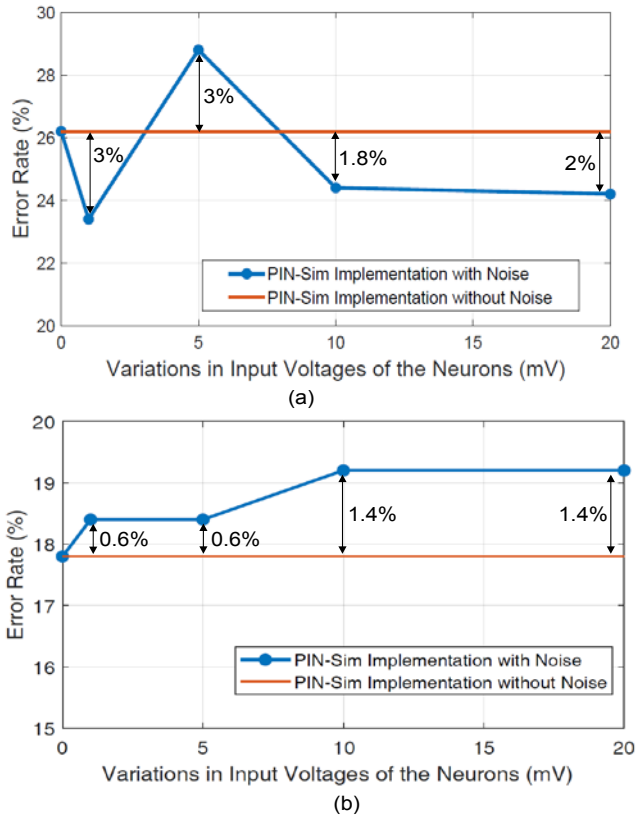

Fig. 2. Error rate versus the neuron's input voltage noise, (a) *784×10* DBN trained by 3,000 images, (b) *784×200×10* DBN trained by 3,000 images.

IV. Simulation Framework

Herein, we use the Probabilistic Inference Network-Simulator (PIN-Sim) proposed in [2] to investigate the effect of noise on the accuracy of two different DBN topologies for the MNIST pattern recognition application [6]. PIN-Sim is a hierarchical simulation framework that consists of five main modules: *(1) trainDBN:* a MATLAB-based module used for training the deep belief network architecture [7], (2) *mapWeight:* a module developed in MATLAB that converts the trained weights and biases to their corresponding resistance values, (3) *mapDBN:* a python-based module which provides a circuit-level implementation of the restricted Boltzmann machine using the obtained weight and bias resistances, (4) *neuron:* a SPICE model of the MRAM-based stochastic neuron [4], (5) *testDBN:* the main module developed in Python that executes test evaluations to assess the error rate and power consumption using the other modules in PIN-Sim. The PIN-Sim framework was utilized to conduct 1,000-trial Monte Carlo circuit-level simulations using a p-bit SPICE model for 14nm CMOS technology operating at a nominal voltage of 0.8V.

V. Noise Analysis

Herein, we have modified Module-4 in the PIN-Sim framework, i.e. the neurons, by adding random variations to the input voltage of the p-bits, which can be induced by different sources of noise [8]. In particular, we have added random noises ranging between ±1 mV, ±5 mV, ±10 mV, and ±20 mV to the input of the p-bit based neurons in two *784×10* DBN and *784×200×10* DBN topologies, as shown in Figures 2(a) and 2(b), respectively. The results exhibit that the hybrid spintronic/CMOS based implementation of the DBNs using memristive weights and p-bit based neurons is relatively robust against noise-induced variations in input voltage of the neurons. Moreover, it is shown that the maximum fluctuations in the error rate for a *784×10* network is 3%, while a *784×200×10* DBN exhibits a maximum of 1.4% increase in error rate. Thus, larger networks are seen to exhibit increased tolerance against noise. This result matches expectations as network size grows and also provides an enlightening quantification of noise impact in memristive DBNs as being tolerable as they are scaled-up.

VI. Conclusion

Modification of the PIN-Sim framework proposed in [2] to investigate the effect of noise on the hardware implementation of DBNs using a resistive crossbar as weighted arrays and probabilistic spin logic devices as stochastic binary neurons. In particular, we have applied noise-induced variation ranging from ±1mV to ±20mV to the input voltage of the p-bit based neurons using the modified PIN-Sim tool. The results exhibited the robustness of hybrid spintronic/CMOS DBN circuits to noise-induced variations. Moreover, it was shown lower fluctuation in the recognition accuracy of the DBNs can be achieved by using a sufficiently-sized network having increased numbers of layers and/or artificial neurons.

Acknowledgments

This work was supported by the Center for Probabilistic Spin Logic for Low-Energy Boolean and Non-Boolean Computing (CAPSL), one of the Nanoelectronic Computing Research (nCORE) Centers as task 2759.006, via Semiconductor Research Corporation (SRC) and by NSF CCF-1739635.

References

[1] R. Sarikaya, G. E. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE/ACM Trans. on Audio, Speech and Language Proc. (TASLP),* vol. 22, pp. 778-784, 2014.

[2] R. Zand, K. Y. Camsari, S. Datta, and R. F. DeMara, "Composable Probabilistic Inference Networks Using MRAM-based Stochastic Neurons," *ACM Journal on Emerging Technologies in Computing (JETC)*, 2019.

[3] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, "A learning algorithm for Boltzmann machines," Cog. Science, vol. 9, no. 1, pp. 147-169, 1985.

[4] K. Y. Camsari, S. Salahuddin and S. Datta, "Implementing p-bits with Embedded MTJ," in *IEEE Electron Device Letters*, vol. 38, no. 12, pp. 1767-1770, Dec. 2017.

[5] P. Debashis, R. Faria, K. Y. Camsari and Z. Chen, "Design of Stochastic Nanomagnets for Probabilistic Spin Logic," in *IEEE Magnetics Letters*, vol. 9, pp. 1-5, 2018, Art no. 4305205.

[6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE,* vol. 86, pp. 2278-2324, 1998.

[7] M. Tanaka and M. Okutomi, "A novel inference of a Restricted Boltzmann Machine," in*, 22nd International Conference on Pattern Recognition (ICPR)*, pp. 1526-1531, 2014.

[8] B. Liu *et al.*, "Digital-assisted noise-eliminating training for memristor crossbar-based analog neuromorphic computing engine," in *Design Automation Conference (DAC), 2013 50th ACM/EDAC/IEEE*, 2013, pp. 1-6.