Neuromorphic Logic Networks and Robust Stochastic Computing Under Large Perturbations and Uncertainties

A. H. Tran[1], S. N. Yanushkevich[1], S. E. Lyshevski[2] and V. P. Shmerko[1]
[1]Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada
[2]Department of Electrical and Microelectronic Engineering
Rochester Institute of Technology, Rochester, NY 14623,USA
E-mail: Sergey.Lyshevski@mail.rit.edu Web: http://people.rit.edu/seleee

ABSTRACT

This paper studies robust fault-tolerant neuromorphic computing to support enabling application-specific design and enable emerging *nanoscale* microelectronics. We develop an energy-centric probabilistic design concept and propose a library of neuromorphic networks for logic functions. These developments enable robust failure tolerance, adaptation and reconfiguration of complex large-scale networks. The innovative methods and tools in design of neuromorphic networks are verified for unreliable, defective, faulty and failed interconnect and cells which may operate under large perturbations.

***Keywords*:** networks, neuromorphic computing, processing

1. INTRODUCTION

Our interest in neuromorphic computing is motivated by large uncertainties, noise, quantum-effects variations, interferences, disturbances and other perturbations in nanoscale processing primitives [1, 2]. New methods and tools must be developed for processing platforms which are implemented using nanoscaled *macroscopic* processing and logic gates because they exhibit stochastic dynamics, probabilistic variations, oscillations and multi-stable switchings. We focus on analysis of behavior, robustness and stability of networks comprised from computing cells. The proposed neuromorphic networks can be viewed as a biologically inspired computing and memory scheme. The researched solutions may typify to some extent processing by living organisms. We utilize the most advanced design concepts and solutions such as:

1. Massive parallelism;
2. Distributed, robust and fault-tolerant processing of probabilistic, noisy and inconsistent information;
3. Robust computing under large perturbations and uncertainties.

The Hopfield networks for computing and implementation of logic functions result in small number of neurons and iterations needed to achieve *training*. The following undesirable effects are associated in redundant Hopfield networks:

1. The overall network and interconnect complexity (from 3 links in an optimal topology to 9 links in the case of 2-neuron redundancy);
2. The large number of iterations (instead of several iterations in an *optimal* network, it may require many iterations to achieve *training*);
3. The complexity of decoding (assuming an optimal topology, 2^2 logic functions must be selected from 2^3 alternatives, while in the case of 2-neuron redundancy, 2^2 values of a logic function must be selected from 2^5 alternatives).

It is imperative to enable the existing methods and approaches for emerging nanoscale microelectronics, quantum-mechanical devices and computing *fabrics*. The Hopfield and other networks can be characterized and designed by using physically-relevant quantities departing from *abstract* concepts. In this paper, we successfully utilize and demonstrated the energy- and probabilistic-centric design of complex networks. The proposed approach results in efficient and redundant fault-tolerant parallel commuting under large perturbations and uncertainties. It is documented that a robust processing of fuzzy, probabilistic, noisy or inconsistent data can be accomplished.

2. INTRODUCTION TO STOCHASTIC COMPUTING

Neuromorphic networks, which use the most promising design concepts and focus on front-end technologies, have being researched extensively [1-3]. These networks, in general, cannot be viewed as biologically inspired processing due to:

1. A lack of coherent understanding of processing in living organisms at device and system levels;
2. Fundamental shortcomings of information theory and computer science on information processing and information measures; etc.

While *engineered* systems perform data processing, vertebrates and invertebrates exhibit information processing. It is illustrated in [1, 2] that neuromorphic networks may exhibit the following properties, capabilities and attributes:

(i) Robustness and failure tolerance;
(ii) Flexibility and plasticity;
(iii) Processing of fuzzy, probabilistic, noisy, or inconsistent data;
(iv) Massively parallel and distributed data processing.

The Hopfield network [4], in contrast to the so-called *Boltzmann machine* [5, 6], leads to a deterministic computing model using a *parallel relaxation* due to the properties of neurons in a distributed system. In general, the design of networks can be accomplished by minimizing the energy in stochastically perturbed systems. The computing in cellular networks of arbitrary configurations can be realized by using the random change of states of cells with respect to the objective function until stable states of cells are achieved. These stable states encode and manifest the final result. In the proposed neuromorphic networks, this yields values of elementary logic functions. In general, the *relaxation* is a dynamic evolution. The various states of the network form a search space. A randomly chosen state will transform itself into one of the local minima at the nearest stable state. Even if the initial state contains inconsistencies, uncertainties or variations, a logic network will converge to a solution, which ensures the fewest constraints minimizing the *system* energy.

3. DEVICE PHYSICS AND PHENOMENA AT NANOSCALE

Conventional semiconductor *nanoscale* devices exhibit quantum phenomena which could significantly degrade the overall performance, capabilities and functionality [7]. Processing cells, which are comprised from the aforementioned interconnected devices, as well as design concepts, must be examined. In the modeling of cells, it is feasible to utilize the distribution laws. For example, the *Boltzmann machine* applied a probabilistic *training*. A cell has an active state with a probability, which associates with the *energy function*. The energy reflects and may explicitly characterize the state change of cells.

The average kinetic energy of a system, as given by $3k_BT/2$, depends on the absolute temperature T. Here, k_B is the Boltzmann constant. For *macroscopic* electronic devices, the Fermi-Dirac and Maxwell-Boltzmann distributions are

$$\phi(E)=\left(e^{\frac{E-E_F}{k_BT}}+1\right)^{-1},\ \phi(E)=e^{-\frac{E-E_F}{k_BT}}. \qquad (1)$$

where E_F is the Fermi energy. For intrinsic semiconductors, the Fermi energy depends on the energy gap E_{gap} and effective masses. In particular, $E_F=\frac{1}{2}E_{gap}+\frac{3}{4}k_BT\ln\left(\frac{m_{h\,eff}}{m_{e\,eff}}\right)$.

The quantization of the radiation (electromagnetic) field can be performed. The Hamiltonian operator $H=\sum_j\hbar\omega\left(a_j^+a_j+\frac{1}{2}\right)$ is derived by using the creation and annihilation operators

$$a_j^+(t)=\sqrt{\frac{1}{2\hbar\omega_j}}\left[\omega_jq_j(t)-ip_j(t)\right],\ a_j(t)=\sqrt{\frac{1}{2\hbar\omega_j}}\left[\omega_jq_j(t)+ip_j(t)\right],$$

were p_j and q_j are the momentum and coordinate operators.

The energy of the lattice vibration mode at frequency ω is given as $E_n=(n+1)\hbar\omega$. If the lattice is in the thermal equilibrium, the probability, that the mode is exited to the state n is $p(n)=\frac{e^{-\frac{E_n}{k_BT}}}{\sum_{m=0}^{\infty}e^{-\frac{E_m}{k_BT}}}$. The average excitation energy of the mode is

$$\overline{E}=\frac{1}{2}\hbar\omega+\frac{\sum_{n=0}^{\infty}n\hbar\omega e^{-\frac{n\hbar\omega}{k_BT}}}{\sum_{m=0}^{\infty}e^{-\frac{m\hbar\omega}{k_BT}}}=\frac{1}{2}\hbar\omega+\frac{\hbar\omega}{e^{\frac{\hbar\omega}{k_BT}}-1}. \qquad (2)$$

For the black body radiation, the average energy per mode in the thermal equilibrium is $\overline{E}=\frac{1}{2}\hbar\nu+\frac{\hbar\nu}{e^{\frac{\hbar\nu}{k_BT}}-1}$.

The average excitation is characterized by the average number of quanta

$$\overline{n}=\frac{\overline{E}}{\hbar\omega}=\frac{1}{2}+\frac{1}{e^{\frac{\hbar\omega}{k_BT}}-1}. \qquad (3)$$

4. NETWORKS DESIGN

We emphasized some essential features of *microscopic* and nanoscale (*macroscopic*) semiconductor electronic devices. One may use energy and examine energetic quantities designing rudimentary, complex and large-scale networks. It is possible to define the network energy and accomplish design tasks evaluating the energy changes. The cells and network energies depend on the device switching function, activity, losses, etc. Performance, characteristics and abilities of semiconductor devices, which implement logic gates, are affected by k_BT. Thus, the cells and networks can be energetically characterized and evaluated.

At the network equilibrium, for a network with $N=(N_1,N_2,\ldots)$ cells, the average energy of a cell and a *total* network energy are

$$E_{\text{cell}}=E/N=k_NS,\ E=Nk_NS, \qquad (4)$$

where k_N is the energy-relevant network-dependent quantity which is a function of parallelism, redundancy, complexity and other features [J/switching]; S is the switching-activity functions [switching].

One can define the probability that the cell is exited to the specific state, the average energy per switching, as well as other descriptive features. As statistical premises are applied to *microscopic* and *macroscopic* systems, the statistical concepts can be also used for cells and networks. For example, the Maxwell-Boltzmann and Fermi-Dirac distributions (1) are applied to distinguishable and identical particles. The probabilities for retaining of a particular energy for distinguishable and identical processing cells are

$$\phi(E)=e^{-\frac{E-E_0}{k_NS}},\ \phi(E)=\frac{1}{e^{\frac{E-E_0}{k_NS}}+1}=\left(e^{\frac{E-E_0}{k_NS}}+1\right)^{-1}, \qquad (5)$$

where E_0 is the energy of the most-energetic cell in the *relaxed* (*ground*) state in non-interacting network.

Equations (5) provide the probability that a cell is in a state characterized by the energy E. Consider the distributions (5). In equilibrium, the average (probable) number of cells in a network, which will be found in state 1 (with state energy E_1) is $N_{E1}=\phi(E_1)$.

In the threshold models of logic gates, the output of a binary cell is either logical 0 or 1, but the probability distribution is a *sigmoid function,* on average.

Two-input AND, NAND, OR and NOR elementary logic functions can be generated and realized by a single cell. Control over the type of logic function is achieved by varying the threshold θ and weights $w_i\in[1, -1]$ of two inputs. The output is the arithmetic sum. The recursive stochastic models at the gate level is

$$f=w_1\times x_1+w_2\times x_2-\theta. \quad (6)$$

Recursive stochastic models are based on the property of *relaxation*, e.g., an ability to *relax* (evolve) to a stable state. Within the proposed approach, a value of a Boolean function *f*, given an assignment of its Boolean variables, is computed through evolution of cells in the network.

In [5, 6], the *relaxation* is controlled by using the "pseudotemperature" which is not a physical quantity. In particular, the *training* is accomplished by assuming the following probability [6] $p_{x_k\to -x_k}=\frac{1}{1+e^{-\Delta E_k/T}}$.

We characterize, examine and evaluate cells and networks by using a well-defined *energy* (4). Furthermore, physically- and logic-consistent distributions (5) are applied. The *system* (network) may evolve to a higher energy configuration or state, evolving from a local minimum. While the Hopfield network uses local minima (as the memory of a network), the probabilistic concept enables the Hopfield approach by using the stimulated excitations and *relaxations* to reach a global minimum.

5. NEUROMORPHIC MODEL

The stochastic nature of various phenomena (interference, thermal noise, electron tunneling, quantization, delays, oscillations and others) were emphasized. In addition, fabricated ICs and logic networks will have defects which affect the functionality and capabilities of components, modules and systems. Stochastic methods should be applied to study the processing circuit failures, yield, etc. The aforementioned phenomena and considerations significantly affect the reliability, performance and functionality of logic cells. Reliable and robust computing in the presence of faults is addressed by applying the premise of *fault-tolerant computing*. Though problems of reliable computing by networks with faulty nodes and interconnect was addressed [8], the existing approaches are not effective or not applicable at interconnect, device and system levels. The reported techniques may be unacceptable or ineffective for improving the reliability of single and interconnected logic cells. These facts were demonstrated by von Neumann examining the NAND cells [9]. Von Neumann demonstrated a significant complexity of correction and redundancy at the gate level. This complexity reduces effectiveness of redundant-centric techniques and diminishes the reliability of logic cells which compute rudimentary logic functions.

Complexity of interconnect and redundant logic networks significantly decrease the overall effectiveness of fault-tolerant computing. Many concepts [1, 2, 10, 11] were proposed and verified to accomplish *fault-tolerant computing* thereby enabling robustness, accommodating failures, accomplishing error corrections, etc. The following features of the proposed neuromorphic models enable fault-tolerant and robust implementation of logic functions *f*:

1. The cells should be in one of two *computing* states, e.g., *active* (encoded by 1) and *inactive* (encoded by 0);
2. The network of interacting cells is fully interconnected and symmetrical, e.g., each cell is connected to other cells. Every pair of cells has a connection assigned by a weight C_{ij}. A positive weighted connection indicates that two cells tend to *activate* each other triggering mutually supportive activities. A negative weight allows an active cell to *deactivate* a neighboring cell triggering incompatible activities;
3. A *recursive computing*: The outputs of each cell feed into the inputs of other cells. If *asynchronous* control is chosen, and, each cell makes a decision based only on its state;
4. Any elementary logic function *f* of *n* Boolean variables can be implemented by a network with at least (n+1) cells.

The aforementioned new features and controlled evolutions enable robust processing and uniquely suit parallel asynchronous computing. For a network with $N=(N_1,N_2,\ldots)$ cells, there are various possible cell configurations which are characterized by E, and, $E_{\text{cell}i}=(E_{\text{cell}1},E_{\text{cell}2},\ldots)$. To ensure network evolutions through *relaxations*, we define the physically- and logic-consistent quantity

$$E=-\tfrac{1}{2}\sum_i^n\sum_{j,j\neq i}^n C_{ij}v_iv_j+\sum_i^n H_iv_i\ , \quad (7)$$

where v_i is the state of the *i*th cell; C_{ij} is the connection weight; H_i is the threshold value; *n* is the total number of cells excluding *bias* cell.

Let a cell be randomly chosen. If any of its neighbors are active, the cell computes E_j to the active neighbors. If $E_j>0$, the cell becomes active, otherwise, it becomes inactive. For the second cell, chosen at random, the process repeats until the network reaches a stable state within a local minimum. That is, the process continues until no more cells change their states. The state reflects the assignment of *truth* and *falsity* to the various hypotheses under the constraints. The statistical updating ensures probabilistic updating. Instead of setting the state of the *i*th cell deterministically, the probability p_i that a cell takes the specified state is used. The algorithm is:

1. Select initial E_i for a randomly-chosen cell *i*;

2. Calculate ΔE_i, $\Delta E = E - E_0$. If $\Delta E_i > 0$, calculate the probability $p_i = \left(1 + e^{-\Delta E_i / k_N S}\right)^{-1}$, relevant to (4), that cell i takes a specified state (for example, 1).
3. Else set the state of cell i to –1;
4. Repeat steps 1 through 3 until $|\Delta E_i| \leq \varepsilon$, $\varepsilon > 0$.

6. FAULT-TOLERANCE

A Hopfield network of logic cells is parallel and inherently fault-tolerant in its *optimal* configuration. The *optimal* configuration means that there is no other network exists which represents this function using fewer cells. In the proposed fault-tolerant design, each logic cell can contain more than three cells. Three cells network is optimal for AND, NAND, OR and NOR, while, EXOR and XNOR need more cells. We examine robustness, stability and functionality of the elementary cells to large perturbbations and uncertainties in connections (links) and cells.

The superposition of inherent noise, interference and other disturbances are considered as stochastic perturbations ξ. We use the perturbed-to-unperturbed ratio R_p which characterizes the *strength* of the ξ. In general, ξ is a mapping of all ξ_i which are hardware-dependent. Continuous and discrete distributions of random ξ_i result in probability functions $\gamma(\xi_i)$ as well as distribution functions. Normal, binomial, hypergeometric, Poisson and other distributions affect ξ.

We simplify the analysis by using $R_p \in [0\ 1]$.

For an ideal unperturbed model $R_p = 0$, while for a completely faulty cell, $R_p = 1$.

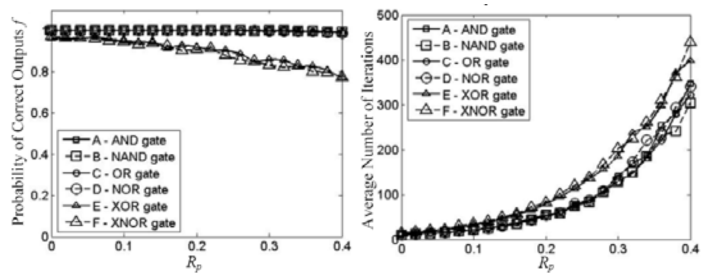
The cell states are function of ξ. The energy of the network is

$$E = -\tfrac{1}{2}\sum_i^n \sum_{j, j \neq i}^n C_{ij} v_i(\xi_i) v_j(\xi_j) + \sum_i^n H_i v_i(\xi_i)\,. \qquad (8)$$

Figures 1 illustrate the quantitative data for the neuromorphic model of AND, NAND, OR, NOR, XOR and XNOR gates. The graph shows the relationship between R_p and probability (assurance) of reaching the correct solution, which is the supposed output of logic functions f. The stochastic perturbations ξ are modeled as uniformly distributed discrete random noise. For $R_p \in [0\ 0.4]$, the correct output f for AND, NAND, NOR and OR gates is achieved within ~100% requiring 400 iterations. In contrast, gates XOR and XNOR achieve the correct f from 80% to 100% requiring ~450 iterations. As R_p increases, the required number of iterations to achieve stability increases. Figure 1.b illustrates the number of iterations for the network to evolve to a stable state for different R_p.

7. CONCLUSIONS

We proposed a new approach which allows one to design complex large-scale networks. These networks may realize arbitrary complexity logic functions. The reported procedure, design concept, network configurations and schemes lead to highly robust, reliable and fault-tolerant computing under large perturbations and uncertainties. If in the proposed redundant networks, a few cells misbehave or completely fail, as well as interconnect fails, the network may still ensure overall functionality. The fault-tolerance of networks can be further improved using redundancy. Illustrative results validated our findings and documented the effectiveness of the proposed approaches and schemes.



(a) (b)

Figure 1. Fault-tolerance of the neuromorphic model of elementary logic gates in the presence of disturbances for $R_p \in [0\ 0.4]$ (5000 runs)

REFERENCES

1. S. Lyshevski, S. N. Yanushkevich, V. P. Shmerko and V. Geurkov, "Computing paradigms for logic nanocells," *J. Computational and Theoretical Nanoscience*, vol. 5, no. 12, pp. 2377-2395, 2008.
2. V. P. Shmerko, S. N. Yanushkevich and S. E. Lyshevski, *Computer Arithmetics for Nanoelectronics*, CRC Press, Boca Raton, FL, 2009.
3. C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, no 10, pp. 1629-1639, 1990.
4. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of National Academy of Sciences*, USA, vol. 79, pp. 2554-2558, 1982.
5. D. H. Ackley, G. E. Hinton, T. J. Sejnowski, "A learning algorithm for Boltzmann machines," *Cognitive Science,* no. 9, pp. 147-169, 1985.
6. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, NJ, 1999.
7. *International Technology Roadmap for Semiconductors*, 2009 Edition, Semiconductor Industry Association, Austin, TX, 2009.
8. P. J. Kuekes, W. Robinett, G. Seroussi and R. S. Williams, "Defect-tolerant interconnect to nanoelectronic circuits: Internally redundant demultiplexers based on error-correcting codes," *Nanotechnology*, no. 16, pp. 869, 2005.
9. J. von Neumann, "Probabilistic Logics," *Automata Studies*, no. 34, Princeton University Press, 1956
10. D. Crawley, K. Nikoli´c and M. Forshaw, Eds., *3D Nanoelectronic, Computer Architecture and Implementation*, Institute of Physics Publishing, UK, 2005.
11. S. N. Yanushkevich and V. P. Shmerko, "Nature-inspired computing for predictable nanoelectronics," *J. Computational and Theoretical Nanoscience*, vol. 7, pp. 1-22, 2010.