Robustness of Binary Stochastic Neurons Implemented With Low Barrier Nanomagnets Made of Dilute Magnetic Semiconductors

Rahnuma Rahman and Supriyo Bandyopadhyay
Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA 23284, USA
∗ Fellow, IEEE

Received 5 Aug 2022, revised 16 Aug 2022, accepted 17 Aug 2022, published 26 Aug 2022, current version 10 Oct 2022.

Abstract—Binary stochastic neurons (BSNs) are excellent hardware accelerators for machine learning. A popular platform for implementing them is low- or zero-energy barrier nanomagnets possessing in-plane magnetic anisotropy (e.g., circular disks or quasi-elliptical disks with very small eccentricity). Unfortunately, small geometric variations in the lateral shapes of such nanomagnets can produce large changes in the BSN response times if the nanomagnets are made of common metallic ferromagnets (Co, Ni, Fe) with large saturation magnetization. In addition, the response times become very sensitive to initial conditions, i.e., the initial magnetization orientation. In this letter, we show that if the nanomagnets are made of dilute magnetic semiconductors with much smaller saturation magnetization than common metallic ferromagnets, then the variability in their response times (due to shape variations and variation in the initial condition) is drastically suppressed. This significantly reduces the device-to-device variation, which is a serious challenge for large-scale neuromorphic systems. A simple material choice can, therefore, alleviate one of the most aggravating problems in probabilistic computing with nanomagnets.

Index Terms—Spin electronics, low barrier magnets, binary stochastic neurons, correlation time, dilute magnetic semiconductors.

I. INTRODUCTION

Artificial intelligence platforms often employ spiking neural networks (SNN) activated by binary stochastic impulses to execute machine learning tasks [Pfeiffer and Pfeil 2018, Roy 2019]. The common building block of SNNs is a binary stochastic neuron (BSN) that has two distinct output states; the BSN will output either state with a probability determined by a specific function of an input impulse. BSN states can also represent Ising spins and are useful hardware platforms for Ising machines that can solve combinatorial optimization problems.

Recently, it has become popular to implement BSNs with low (or zero) energy barrier nanomagnets (LBMs) possessing in-plane magnetic anisotropy because of their superior energy efficiency and relatively fast response times [Faria 2017, Hassan 2019, 2021]. The magnetization vectors of these LBMs, which are circular or nearly circular disks, fluctuate randomly at room temperature owing to thermal perturbations. To implement a BSN, one would employ a magnetic tunnel junction (MTJ) with an LBM as the soft layer, and the MTJ resistance will fluctuate in time between the high and low values (corresponding to the antiparallel and parallel configurations). A threshold can be set midway between the two resistance values, and during a sampling interval, if the MTJ resistance is above this threshold, it will be interpreted as the binary bit 1; when it is below, it will be interpreted as the binary bit −1. Normally, the probabilities of −1 and 1 will be equal, but if a spin-polarized current is injected into the MTJ, it will produce a spin transfer torque and make the probabilities unequal. Thus, with this construct, one can produce a binary state $m_i$ (−1 or +1) at time step $(n+1)$ given by the analytical expression

$$m_i(n+1) = \text{sgn}[\tanh(I_i(n)) - r_i]$$  (1)

where $I_i$ is the dimensionless input spin current that biases the output either toward −1 or toward 1 (depending on its sign) and $r_i$ is a random number uniformly distributed between −1 and +1. Each BSN described by (1) receives its input from a weighted sum of other BSNs obtained from a synapse $I_i(n) = \sum_j W_{ij} m_j(n)$. A wide variety of problems can be solved by properly designing or learning the weights $W_{ij}$, e.g., classification problems [Cilingiroglu 1991], constraint satisfaction problems [Fonseca Guerra and Furber 2017], generation of cursive letters [Mizrahi 2018], etc.

A critical parameter for a BSN implemented with LBMs is the “correlation time” $\tau$, which is the full-width-at-half maximum of the temporal decay characteristic of the autocorrelation function of the magnetization fluctuations [Hassan 2019]. This quantity determines the BSN’s response speed. LBMs with in-plane anisotropy typically have a value of $\tau$, that is about two orders of magnitude smaller than LBMs with perpendicular magnetic anisotropy [Hassan 2019] and are therefore favored because they lend themselves to faster circuit operation. The correlation time, however, is very sensitive to small geometric variations if LBMs with in-plane magnetic anisotropy are realized with magnetic materials possessing large saturation magnetization, e.g., Co, Ni, or Fe [Abed 2019, Rahman 2021a]. The reason for this is that the in-plane shape anisotropy barrier $E_b$ in a nanomagnetic disk that is not a perfect circle but is slightly elliptical, as shown in Fig. 1, is given (within the macrospin approximation) as

$$E_b = \frac{\mu_0}{2} M_i^2 \Omega (N_{d,xx} - N_{d,yy})$$  (2)

where $\mu_0$ is the permeability of free space, $\Omega$ is the nanomagnet volume, $N_{d,xx}$ and $N_{d,yy}$ are the demagnetization factors along the minor and major axes of the ellipse, respectively, and $M_i$ is the saturation magnetization. The expressions for the demagnetization factors are
II. RESULTS

We simulate the fluctuating room temperature magneto-dynamics of two sets of LBM, one made of Co and the other made of GaMnAs, using the stochastic Landau–Lifshitz–Gilbert equations and different initial conditions (i.e., different initial magnetization orientations) [Rahman 2021a]. This yields the magnetization vector as a function of time and therefore the autocorrelation plots of any component of the magnetization as a function of the delay [Rahman 2021a], from which we can extract the correlation time $\tau_c$ for that magnetization component. In each set, we consider nanomagnets of two different shapes—a slightly elliptical one of major axis 100 nm and minor axis 99 nm, and a perfectly circular one of diameter 100 nm. The magnet thickness is 6 nm. The in-plane shape anisotropy energy barrier in the case of the slightly elliptical Co nanomagnets is $\sim$4 kT, while in the case of GaMnAs nanomagnets, it is $10^{-4}$ kT at room temperature. The Gilbert damping factor in GaMnAs is 0.01 [Khazen 2008], similar to that of Co.

We simulate the magneto-dynamics in the presence of thermal noise at room temperature (300 K), taking into account spin inertia, which causes nutational dynamics over short time spans [Wegrowe 2012]. The nutational dynamics last for a duration of $\tau$ (which depends on the magnetic vector’s moment of inertia and damping), and we consider three values of $\tau = 1, 10,$ and 100 ps. The detailed procedures for the calculation can be found in Rahman [2021a] and not repeated here to avoid redundancy.

We calculate the autocorrelation functions of the fluctuating magnetization components along the minor and major axes of the nanomagnets ($x$- and $y$-axes, respectively) at room temperature, $C_x(t') = \int_0^\infty m_x(t)m_x(t+t')dt$ and $C_y(t') = \int_0^\infty m_y(t)m_y(t+t')dt$ with two sets of initial conditions: $m_x(0) = 0.995; m_y(0) = 0.095; m_y(0) = 0.031$ (initial magnetization orientation along the minor axis) and $m_y(0) = 0.095; m_x(0) = 0.995; m_x(0) = 0.031$ (initial orientation along major axis). We then plot the autocorrelation functions as a function of the delay $\tau$ for the Co and GaMnAs nanomagnets for both slightly elliptical and circular nanomagnets and for both initial conditions in Fig. 2. These characteristics are obtained by averaging over 1000 simulation runs, which is equivalent to ensemble averaging over 1000 identical nanomagnets for each plot owing to ergodicity (the averaging is necessary since thermal fluctuations are random). We have ascertained that a larger ensemble size does not perceptibly change any result.

We see no significant dependence on spin inertia under any situation since the plots do not change appreciably for the three different values of $\tau$, but this is expected since the nutational duration $\tau$ is much smaller than the correlation time $\tau_c$ in all cases. There are situations when the nutational dynamics can affect outcomes that emerge long after nutation has ceased [Rahman 2021b], but this is not such a case. Clearly, the fluctuation magneto-dynamics does not retain any memory of the nutation long after it ceases to exist.

The interesting observation is that in the case of GaMnAs, unlike in the case of Co, the autocorrelation decay plots are not very sensitive to small deviations in shape (perfectly circular versus slightly elliptical), nor are they particularly sensitive to the initial orientation of the magnetization for the case of the slightly elliptical nanomagnet. The latter feature can be understood easily. Consider the case when the energy barrier is large, as in the case of Co. If the initial orientation of the magnetization is along the major axis of the ellipse, which

\begin{align*}
N_{a-xx} &= \frac{\pi}{4} \left( \frac{t}{a} \right) \left[ 1 + \frac{5}{4} \left( \frac{a-b}{a} \right) + \frac{21}{16} \left( \frac{a-b}{a} \right)^2 \right] \\
N_{a-yy} &= \frac{\pi}{4} \left( \frac{t}{a} \right) \left[ 1 + \frac{1}{4} \left( \frac{a-b}{a} \right) - \frac{3}{16} \left( \frac{a-b}{a} \right)^2 \right]
\end{align*}

where $a$ = major axis dimension, $b$ = minor axis dimension, and $t$ = thickness of the elliptical nanomagnet. For a perfect circle, $a = b$, and hence $E_b = 0$. For an ellipse, $a > b$, and hence $E_b > 0$.

It is clear from (2) and (3) that a slight change in the nanomagnet dimension $a$ or $b$ can cause a large change in the energy barrier $E_b$ if the saturation magnetization of the nanomagnet $M_s$ is large. A material like Co has a saturation magnetization of $10^6$ A/m, whereas a dilute magnetic semiconductor like GaMnAs has a saturation magnetization of only $5 \times 10^3$ A/m [Lawniczak-Jablonska 2011] at room temperature, which is more than two orders of magnitude smaller. Hence, everything else being the same, a GaMnAs nanomagnet will have an energy barrier that is roughly four orders of magnitude smaller than that of a Co nanomagnet. Consequently, the same change in a nanomagnet dimension ($a$ or $b$) will cause a much smaller change in the energy barrier if the nanomagnet is made of GaMnAs than if it is made of Co. Thus, a BSN implemented with a GaMnAs LBM is expected to be much more robust against small geometric variations than one implemented with a Co nanomagnet because the BSN characteristics are sensitive to the energy barrier.

In this connection, we mention that the use of a lower saturation magnetization material like GaMnAs does not introduce any problem of its own. It is ferromagnetic at room temperature with Mn concentration close to 20% [Lawniczak-Jablonska 2011]. There is no problem with read-out; the magnetization states are usually read with an MTJ whose reading ability depends on the tunneling magnetoresistance (TMR). Large TMR has been predicted for GaMnAs/GaAlAs/GaMnAs MTJs [Brey 2004]. Hence, the readout is not significantly affected if Co is replaced by GaMnAs. Furthermore, we do not have to make the devices larger if we use GaMnAs; we consider nanomagnets with lateral dimensions of $\sim 100$ nm, which will remain ferromagnetic at room temperature (not turn superparamagnetic) despite the lower saturation magnetization.

Lastly, GaMnAs is a well-known and well-researched dilute magnetic semiconductor. It is not significantly different from GaAs, and there is a large body of work on integrating GaAs with Si technology [Bietti 2013]. Hence, integrating GaMnAs-based BSN with complementary metal-oxide semiconductor (CMOS) in a Si platform will not pose insurmountable challenges.

Fig. 1. Nanomagnet shaped like a slightly elliptical disk.
is the easy axis, then the magnetization is in a stable state, and it will take a long time for thermal perturbations to destabilize the magnetization and make it veer away from the major axis because the energy barrier is large. On the other hand, if the magnetization is initially along the minor axis (hard axis), it is at the maximally unstable location, and thermal perturbations will quickly destabilize and make it move away from the minor axis. That is why in the case of Co $\tau_c$ is $\sim$20 ns if the initial orientation is along the major axis and only $\sim$1.5 ns if it is along the minor axis. The difference is more than an order of magnitude. The difference is much smaller for GaMnAs because the energy barrier is much smaller for the same nanomagnet dimensions. The large difference in the correlation times in the case of nanomagnets with large saturation magnetization will introduce a large variability in BSN response times if their initial magnetization orientations are different. This would mandate an initialization step where the initial magnetizations of all BSNs in a circuit will have to be forcibly aligned in the same direction with a magnetic field, which is an inconvenience. This inconvenience is alleviated

Fig. 2. Left column (a)–(e) are for Co nanomagnets of slightly elliptical and circular shapes for different initial orientations of the magnetizations, and right column (f)–(j) are the corresponding plots for GaMnAs nanomagnets. Figures (a)–(e) are reproduced from Rahman [2021a].
if we use a material with lower saturation magnetization, such as GaMnAs.

Second, we notice that in the case of Co, even a 1% change in a lateral dimension can change the correlation time $\tau_c$ by a factor of 4! Given that 11 nm magnetic structures have been produced for spin transfer torque random-access memory [Nowak 2016], a 1% size control in a 100 nm diameter object (which corresponds to 1 nm variation) may be achievable, but any tighter control than 1 nm will demand unrealistic fabrication tolerance. On the other hand, in the case of the GaMnAs nanomagnets, $\tau_c$ varies between 5.95 and 6.23 ns for all cases, showing that neither is the correlation time sensitive to small geometric variations nor is it sensitive to initial conditions. This is, of course, a consequence of the much smaller energy barrier in the case of GaMnAs compared to Co. Therefore, replacing high-saturation magnetization materials like Co with low-saturation magnetization materials like dilute magnetic semiconductors (e.g., GaMnAs) is beneficial for BSNs.

### III. CONCLUSION

Device-to-device variability in BSNs, and neuromorphic circuits in general, is a serious problem, and countermeasures have been proposed to alleviate it. One recent suggestion was to use hardware-aware in situ learning [Kaiser 2022]. Here, we propose a simpler solution: replacement of material magnetic possessing a large saturation magnetization with one possessing smaller saturation magnetization, such as a dilute magnetic semiconductor like GaMnAs. This reduces the sensitivity of the correlation time (response time) of BSNs to small geometric variations and also initial conditions, which aids ultra-large-scale integration.

### ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grants CCF-2006843 and CCF-2001255.

### REFERENCES

Abied M A and Bandyopadhyay S (2019), "Low energy barrier nanomagnet design for binary stochastic neurons: Design challenges for real nanomagnets with fabrication defects," *IEEE Magn. Lett.*, vol. 10, 4504405, doi: 10.1109/LMAG.2019.2929484.

Bietti S, Scaccabarozzi A, Frigeri C, Bollani M, Bonera E, Falub C V, von Kanel H, Miglio L, and Sanguinetti S (2019), “Monolithic integration of optical grade GaAs on Si (001) substrates deeply patterned at a micron scale,” *Appl. Phys. Lett.*, vol. 103, 262106, doi: 10.1063/1.4987835.

Brey L, Tejedor C, and Fernández-Rossier J (2004), “Tunnel magnetoresistance in GaMnAs: Going beyond Julliére formula,” *Appl. Phys. Lett.*, vol. 85, 1996, doi: 10.1063/1.1789241.

Chikarami S (1964), *Physics of Magnetism*. New York, NY, USA: Wiley.

Cilingiroglu U (1991), “A purely capacitive synaptic matrix for fixed-weight neural networks,” *IEEE Trans. Circuits Syst.*, vol. 38, pp. 210–217, doi: 10.1109/81.68299.

Faria R, Camsari K Y, and Datta S (2017), “Low-barrier nanomagnets as p-bits for spin logic,” *IEEE Magn. Lett.*, vol. 8, 410505, doi: 10.1109/LMAG.2017.2685358.

Fonseca Guerra G A and Furber S B (2017), “Using stochastic spiking neural networks on SpinNaker to solve constraint satisfaction problems,” *Front. Neurosci.*, vol. 11, 714, doi: 10.3389/fnins.2017.00714.

Hassan O, Datta S, and Camsari K Y (2021), “Quantitative evaluation of hardware binary stochastic neurons,” *Phys. Rev. Appl.*, vol. 15, 064046, doi: 10.1103/PhysRevApplied.15.064046.

Hassan O, Faria R, Camsari K Y, Sun JZ, and Datta S (2019), “Low-barrier magnet design for efficient hardware binary stochastic neurons,” *IEEE Magn. Lett.*, vol. 10, 4502085, doi: 10.1109/LMAG.2019.2910787.

Kaiser J, Borders W A, Camsari K Y, Fukami S, Ohno H, and Datta S (2022), “Hardware-aware in situ learning based on stochastic magnetic tunnel junctions,” *Phys. Rev. Appl.*, vol. 17, 014016, doi: 10.1103/PhysRevApplied.17.014016.

Khazen Kh, Bardeleben H J V, Cubukcu M, Canlin J T, Novak V, Olejnik K, Cuk M, Thevenard L, and Lemaître A (2008) “Damping and magnetic anisotropy of GaMnAs thin films,” *arXiv*:0809.4644.

Lawniczak-Jablonska K, Libera J, Wolska A, Klepka M T, Dłużewski P, Dudowski J, Wasik D, Twardowski A, Kwiatkowski A, and Sato K (2011), “The source of room temperature ferromagnetism in granular GaMnAs layers with zinc blende clusters,” *Physica Status Solidi B*, vol. 5, pp. 62–64, doi: 10.1002/pssr.201004503.

Mizrahi A, Hirtzel T, Fukushima A, Kubota H, Yuasa S, Grollier J, and Querlioz D (2018), “Neural like computing with populations of superparamagnetic basis functions,” *Nat. Commun.*, vol. 9, 1533, doi: 10.1038/s41467-018-03963-w.

Nowak J J, Robertazzi R P, Sun J Z, Hu G, Park J H, Lee J, Annunziata A J, Lauer G P, Kothandaraman R, Sullivan E J O, Trouilloud P L, Kim Y, and Worledge D C (2016), “Dependence of voltage and size on write error rates in spin-transfer torque magnetic random-access memory,” *IEEE Magn. Lett.*, vol. 7, 3102604, doi: 10.1109/LMAG.2016.2539256.

Pfeiffer M and Pfeil T (2018), “Deep learning with spiking neurons: Opportunities and challenges,” *Front. Neurosci.*, vol. 12, 774, doi: 10.3389/fnins.2018.00774.

Rahman R and Bandyopadhyay S (2021a), “The strong sensitivity of the time response and pinning current of binary stochastic neurons employing low energy barrier nanomagnets to small geometrical variations,” *arXiv*:2108.04319v2, doi: 10.48550/arXiv.2108.04319.

Rahman R and Bandyopadhyay S (2021b), “An observable effect of spin inertia in slow magneto-dynamics: Increase of the switching error rates in nanoscale ferromagnets,” *J. Phys.: Condens. Matter*, vol. 33, 355801, doi: 10.1088/1361-648X/ac09eb.

Roy D, Chakraborty I, and Roy K (2019), “Scaling deep spiking neural networks with inertial effects,” *Amer. J. Phys.*, vol. 80, 607, doi: 10.1119/1.4709188.