Voltage control of domain walls in magnetic nanowires for energy-efficient neuromorphic devices

Md Ali Azam¹, Dhritiman Bhattacharya¹, Damien Querlioz², Caroline A Ross³ and Jayasimha Atulasimha¹,⁴

¹ Department of Mechanical and Nuclear Engineering, Virginia Commonwealth Univ., Richmond, VA, United States of America
² Centre national de la recherche scientifique/Univ. of Paris Saclay, Paris, France
³ Department of Materials Science and Engineering, Massachusetts Inst. of Technology, Cambridge, MA, United States of America
⁴ Department of Electrical and Computer Engineering, Virginia Commonwealth Univ., Richmond, VA, United States of America

E-mail: jatulasimha@vcu.edu

Received 5 September 2019, revised 3 November 2019
Accepted for publication 16 December 2019
Published 16 January 2020

Abstract
An energy-efficient voltage-controlled domain wall (DW) device for implementing an artificial neuron and synapse is analyzed using micromagnetic modeling in the presence of room temperature thermal noise. By controlling the DW motion utilizing spin transfer or spin–orbit torques in association with voltage generated strain control of perpendicular magnetic anisotropy in the presence of Dzyaloshinskii–Moriya interaction, different positions of the DW are realized in the free layer of a magnetic tunnel junction to program different synaptic weights. The feasibility of scaling of such devices is assessed in the presence of thermal perturbations that compromise controllability. Additionally, an artificial neuron can be realized by combining this DW device with a CMOS buffer. This provides a possible pathway to realize energy-efficient voltage-controlled nanomagnetic deep neural networks that can learn in real time.

Supplementary material for this article is available online

Keywords: domain wall, synapse, artificial neural network, voltage control, SOT, chiral, DMI

(Some figures may appear in colour only in the online journal)

1. Introduction
There has been considerable recent progress in the development of dedicated CMOS processors for neuromorphic computing such as IBM’s TrueNorth that can implement 1 million spiking neurons and 256 million configurable synapses [1] while consuming ~70 mW power. However, these neuromorphic processors have drawbacks such as lack of onboard (real-time) learning/training. More importantly, they have poor energy efficiency in comparison to the human brain, which has ~100 billion neurons and ~500 trillion synapses and consumes a mere ~20 Watts of power [2]. Thus, a key challenge for hardware implementation of artificial neural networks lies in finding energy-efficient hardware implementations of neurons and non-volatile synapses whose weights can be changed easily and deterministically with very little energy as the network learns from data in real time. Artificial neurons and synapses have been proposed using current-controlled nanomagnets [3–8], organometallic systems [9–12] and memristors [13–18]. These artificial synapses can perform several brain-like functions such as short-time and long-time potentiation, spike-timing-dependent-plasticity etc [19]. Synaptic devices have also been shown to emulate artificial sensory systems [20]. While this new generation of artificial neural networks is potentially more energy efficient than pure CMOS implementations, there is still room for improving the energy efficiency.
We propose implementing energy efficient artificial synapse using a magnetic tunnel junction (MTJ). In biological systems, the strength of the connection between the neurons are determined by synaptic weight which is adjusted dynamically. Our DW device is inspired by this and operates analogously to this function. Different synaptic weights are achieved by arresting a domain wall (DW) at different positions in a free layer of the MTJ comprising a magnetoststrictive nanowire racetrack made of CoFe or CoFeB for example. Here we model the magnetization dynamics of the DWs in the racetrack. The wall is driven along the racetrack clocked with current passing through the racetrack exerting a spin transfer torque (STT) [21–24], or by spin–orbit torque (SOT) due to current flowing in a heavy metal layer directly underneath the racetrack [25–28]. The heavy metal layer leads to a perpendicular magnetic anisotropy (PMA) in the CoFeB layer, and a Dzyaloshinskii–Moriya interaction (DMI) which stabilizes chiral DWs. In order to achieve controlled positioning of the DW we propose current clocked DW motion in conjunction with a gradient in the PMA [29, 30]. Notches are placed at regular intervals to arrest the DW at different locations of the racetrack. For the SOT clocked DW motion [28, 31–33] instead of using a notched race track and PMA gradient, a racetrack of uniform width is used along with modulation of the PMA at regular intervals. This modulation (PMA reduction) creates a barrier to the motion of the DW, arresting it at different locations depending on the voltage applied. We also describe the manner in which this device, in combination with a CMOS buffer, can also function as a neuron and implement deep neural networks (DNNs). Such an implementation is important in applications where energy efficiency is at a premium, such as medical processors and sensor networks that need to learn from data in real time rather than be trained offline, and where synaptic weights of limited accuracy are sufficient [34].

Section 2 describes the working principle of the device and our micromagnetic modeling approach. Section 3 presents and discusses simulation of the DW dynamics in the presence of PMA gradients and STT/SOT while section 4 compares the energy efficiency of this approach with other spintronic and memristor approaches.

2. Device working principle and micromagnetic modeling approach

The working principle of the device (figure 1) is explained in terms of the DW dynamics within the magnetic free layer of a MTJ. The resistance of the MTJ, which consists of a free layer, a tunnel junction and a fixed layer pinned by a synthetic antiferromagnet, varies with the location of the DW in the free layer. Therefore, the DW position determines the non-volatile resistance states of the spintronic synapse and can be programmed by a voltage, as described below. As the DNN learns from data in real time, a backpropagation algorithm [35] implemented on a CMOS application-specific co-processor can calculate the new weights for different synapses and output these as specific programming voltages (not addressed in this paper). These voltages should be able to reprogram the resistance states of the synapses to update their resistance values during the clocking cycle of fixed time and current, as described in this work.

2.1. Clocking

Consider a PMA racetrack consisting of a heavy metal/ferromagnet bilayer that could be deposited on a piezoelectric film to realize our proposed device as shown in figure 1. Such a bilayer (e.g. Pt/CoFe) derives its PMA from interfacial effects and exhibits significant DMI that stabilizes the formation of chiral Néel DWs [28].

2.2. SOT clock

SOT acting on the magnetization is generated when current flows in the heavy metal layer. The damping like field (DL-field) thus produced is responsible for translating the Néel DW in the ferromagnetic layer [28]. Reversing the direction of the current in the Pt layer reverses the direction of DW motion, resets the DW position, and hence resets the resistance of the DW MTJ device.

2.3. STT clock

Alternatively, for current clocking by STT, electrons passing through the DW are preferentially polarized along the magnetization orientation of the region through which they pass and exert a torque on the magnetization of the subsequent region they enter [21–24]. This causes the spins within the wall to rotate thus initiating a DW motion in the direction of the electron flow.

2.4. SOT versus STT clock

When CoFe is used with a heavy metal (Pt) underlayer that leads to PMA and DMI, the ratio of the current flowing through the CoFe that leads to STT and to the current flowing through the Pt that leads to SOT depends inversely on the ratio of their resistances. Instead of considering the case of mixed STT and SOT, we consider the two extreme cases: pure SOT and pure STT in our simulations to understand the DW motion with these two clocking mechanisms. We also study the manner in which the DW can be arrested in a specific region of the racetrack by applying a voltage-induced strain under these two different clocking mechanisms.

2.5. Voltage control of DW position

Stopping the DW at a specific position along the racetrack is accomplished by applying a voltage to the side electrode (figure 1(a)) while the DW is being ‘clocked’ by SOT or STT. Consider a DW that has been ‘reset’ to one end of the racetrack and is moved along the racetrack towards the other end in figure 1(a) by SOT from a current in the adjacent Pt layer
or by STT from a charge current through the free layer. Application of a voltage between the side electrode and the bottom contact of the piezoelectric layer produces an electric field through the piezoelectric thickness, which in turn produces an in-plane stress in the manner described in \([36]\). This leads to a local strain gradient in the piezoelectric, which is transferred to the ferromagnetic layer, altering its PMA as shown in figure 1(a). A modified scheme as shown in figure 4 can also be used as discussed later.

The mechanism of generation of the strain gradient is explained in detail in figure 2. When a voltage is applied to the top electrode, a local electric field is generated across the thickness of the piezoelectric between the area directly underneath the top and the bottom electrode. This causes an out-of-plane expansion of the piezoelectric and consequently an in-plane contraction (due to Poisson’s ratio) of the area below the top electrode. This produces a tensile in-plane strain in the region of the piezoelectric immediately adjoining the electrode, with a magnitude decreasing with distance away from the electrode. This creates a strain gradient as shown in figure 2, upper schematic. While a similar strain gradient is created in the in-plane direction orthogonal to that shown in the figure, we are only concerned with the strain gradient along the DW MTJ device. Furthermore, if the piezoelectric is deposited on a stiff substrate, the bottom of the piezoelectric is clamped but the top part of the piezoelectric can experience the in-plane strain gradient.

The benefit of this scheme is that the piezoelectric film develops a strain gradient even though it is not patterned provided the in-plane dimension of the electrode is approximately equal to the thickness of the film \([36]\). The strain gradient will be most significant within a distance of one to two times the piezoelectric film thickness \([36]\). This in-plane strain in turn modulates the perpendicular anisotropy of the soft layer and provides a spatial variation of the energy landscape of the Néel DW in the racetrack.

Thus, the device relies on stress generated by the electrode to arrest the SOT/STT-induced motion of the DW, leaving the DW pinned at a notch or a specific location where there is a PMA barrier (figure 4, discussed later). The strain and therefore modulation of PMA is largest at the right end, and minimum at the left end (figure 2). This PMA gradient in conjunction with the torque on the DW and notches patterned in the wire determines the position where the DW is arrested.

### 2.6. Micromagnetic model

Mumax \([37]\) was used to perform simulations of the DW dynamics using the Landau–Lifshitz–Gilbert equation in the presence of thermal noise at room temperature. The time rate of change of magnetization in a volume element of the
magnetic material is given by:

\[
\frac{\partial \tilde{m}}{\partial t} = \gamma \left( -\gamma \cdot \frac{1}{1 + \alpha^2} \right) \left( \tilde{m} \times \mathbf{H}_{\text{eff}} \right) + \alpha (\tilde{m} \times (\tilde{m} \times \mathbf{H}_{\text{eff}})),
\]

where \(\tilde{m}\) is the reduced magnetization \((\tilde{M}/M_{\text{sat}})\), \(M_{\text{sat}}\) is the saturation magnetization, \(\gamma\) is the gyromagnetic ratio and \(\alpha\) is the Gilbert damping coefficient. The quantity \(H_{\text{eff}}\) is the effective magnetic field, which includes contribution of demagnetization, Heisenberg exchange coupling, DMI \(38, 39\) and perpendicular anisotropy. The effect of strain is incorporated by modulating the anisotropy.

Thermal noise is modeled by a random, effective magnetic field \(H_{\text{thermal}}\) applied in the manner described in \(40, 41\) within the micromagnetic framework \(37\). The case where STT dominates is simulated using the Zhang–Li type torque. As non-adiabatic torques can be small \(42\), we neglect this for simplicity. To simulate the effect of SOT, we use Slonczewski type torque as given in \(37\). To equate these two torques we use polarization \(p = 1\), Slonczewski parameter \(\Lambda = 1\), secondary spin-torque parameter \(\epsilon = 0.05\) and \(m_p = j_k \times z\) where \(j_k\) is the direction of charge current. Spin hall angle is taken to be 0.1. In our simulations, we do not consider STT and SOT at the same time. Instead, we present two cases: one in which only STT is considered and the other in which only SOT is considered. This leads to an understanding of the clocking of DW motion and its arrest at a specific position using voltage induced strain for these two different mechanisms.

The discretization cell sizes used for the simulations were 4 nm × 4 nm × 1 nm. These sizes are smaller than the exchange length \(\sqrt{(2\Lambda/M_{\text{sat}})} = 5.6\) nm. The time step was chosen to be 0.25 ps. The material parameters used for CoFe (soft layer) of the Pt/CoFe/MgO heterostructure are summarized in table 1. CoFe has sufficient magnetostriction to produce a PMA gradient that can arrest the DW at specific positions in the model. While Gilbert damping is ∼0.01–0.03 in these materials, we used a higher value (0.1) so the DW exhibits more stable dynamics. In practice, defects and edge roughness are likely to impede the DW, leading to dynamics characteristic of the higher damping. In the supplementary document, we show one such case (available online at stacks.iop.org/NANO/31/145201/mmedia).

### Table 1. Material parameters used for the CoFe soft layer in the Pt/CoFe/MgO heterostructure as compiled from previously published works \(28, 43–45\). A higher value of Gilbert damping is chosen for the simulation.

| Parameters used in simulation | Value |
|-------------------------------|-------|
| Saturation magnetization \((M_{\text{sat}})\) | \(1 \times 10^4\) A m\(^{-1}\) |
| Exchange constant \((A_{ex})\) | \(2 \times 10^{-11}\) J m\(^{-1}\) |
| Perpendicular anisotropy constant \((K_{ex})\) | \(7.5 \times 10^5\) J m\(^{-2}\) |
| Gilbert damping \((\alpha)\) | 0.1 |
| DMI constant \((D)\) | 0.001 J m\(^{-1}\) |
| Saturation magnetostriction \((\lambda)\) | 250 ppm |

### 3. Discussion of modeling results

We discuss the modeling results for nanowires of length 1000 nm and width 100 nm patterned with five notches as shown in figure 3. A gradient in PMA caused by the strain gradient drives the DW towards the lower PMA region in order to reduce the DW energy, i.e. motion is induced in the direction of the negative PMA gradient. We consider the case where the PMA gradient and STT from the charge current drive the DW in the same direction, and then we perform simulation in the presence of thermal noise to understand if these DWs can be arrested deterministically at room temperature at specific notches. We subsequently perform another SOT-driven simulation in which the DW is arrested in regions of varying PMA produced by an alternative electrode design (figure 4) and no notches are used.

#### 3.1. PMA gradient assists STT driven DW motion

In a MTJ racetrack with CoFe soft layer, the DW motion was assumed to be initiated by STT assisted by a PMA gradient (without including the effect of thermal noise). Given that the current pulse acts as the clocking signal, the ON time for current is fixed at 6 ns. This is considered the ‘write time’ for reprogramming the synaptic weight and, based on the analysis of the DW motion at different currents and PMA gradients, it is sufficient to translate the DW to any possible location within this STT clocked device.

Initially, a DW can be created at notch 1 by applying a current pulse through an overlying current wire. Table 2 shows the PMA gradients required to drive the DW from notch 1 to any of the other notches (as shown in figure 3), if the current density of the clock is kept constant at \(8.7 \times 10^{12}\) A m\(^{-2}\), which is just below the critical current needed to de-pin and initiate the DW motion. For a CoFe layer of 1 nm thickness and width 100 nm this corresponds to a current of ∼1 mA (we do not consider any current through the Pt layer and do not consider SOT for this case). We do not account for increase in current density at the notches in this simulation. Furthermore, from the point of view of programming the synapse, the voltage required to achieve a certain DW position need not be a linear function of position, as this is envisaged to be a pre-calculated analog voltage that is output by the co-processor that implements the learning algorithm.

These results show that at this current density, the current alone cannot drive the DW out of the first notch, but a combination of the current and PMA gradient can. The PMA gradient reduces the current density required to initiate and sustain the DW motion along the racetrack when compared to the case of current only. The DW motion is retarded as it moves into a region of lower PMA, helping arrest the DW at the notches. The synergistic effect of PMA gradient and current lead to a lower energy operation. Finally, we confirm through simulation (figure 3) that each of the different positions of the DW is achievable deterministically, because at a certain PMA gradient the combined strength of the current and PMA gradient is sufficient to move the DW to the desired notch during the current ON period. Therefore, for different
3.2. Thermal noise effect at room temperature and scaling issues

The simulations were repeated in the presence of thermal noise at 300 K. With thermal noise, the minimum current required to initiate the motion of the DW has a lower value when compared to the zero thermal noise case. Moreover, there is a reduction in the effectiveness of the notch in arresting the motion of the DW in presence of thermal noise. Thus, a slight reduction of current is also required to regain the effectiveness of the notches while keeping it high enough so that DW motion can be initiated. This determined our choice of a current density of

5
Table 2. PMA profile for achieving different positions of the DW. The DW starts at the 1st notch. The second column gives the change in PMA over the entire length of the device required to stop the DW at different notches. The third column shows the gradient in PMA for a 1 μm long device.

| Final position of the DW | Required ΔPMA (J m⁻³) over device length | PMA gradient that assists DW arrest (J m⁻³ nm⁻¹) |
|--------------------------|------------------------------------------|-----------------------------------------------|
| 2nd notch                | 0.2 × 10⁴                               | 2.00                                          |
| 3rd notch                | 1.6 × 10⁴                               | 16.00                                         |
| 4th notch                | 3 × 10⁴                                 | 30.00                                         |
| 5th notch                | 3.7 × 10⁴                               | 37.00                                         |

8.4 × 10¹² A m⁻² to explore the effect of thermal noise in arresting the DW motion at notch 3 with a PMA gradient of 10 Jm⁻³/µm. This value of the PMA gradient is chosen between the PMA gradient required for arresting the DW at the 2nd notch (2 J m⁻³ nm⁻¹) and the 3rd notch (16 J m⁻³ nm⁻¹). The probability distribution of the final position of the DW is shown in figure 3(b). The DW was most likely to be found in notches 3 or 4, but there was a significant probability of its being in notches 2 or 5.

This can be attributed to the relatively small change in energy due to PMA modulation through strain compared to that of thermal energy \( k_B T = 4.14 \times 10^{-21} \) J. To illustrate the energy scales, considering a 10 J m⁻³ nm⁻¹ PMA gradient (i.e. a total PMA modulation of 10 kJ m⁻³ over the length of 1 μm nanowire), this energy change, \( \Delta E = (\Delta \text{PMA}) / L \times \text{notch spacing} \times \text{Volume} = (10 \text{ J m}^{-3} \text{ nm}^{-1}) \times 167 \text{ nm} \times 16700 \text{ nm}^3 = 26 \times 10^{-21} J \approx 6.5 \times k_B T \) Here \( \Delta \text{PMA}/L \) is the PMA gradient \( \Delta \text{PMA} \) over 1 μm length, the volume corresponds to the volume of free layer between two notches, \( k_B \) is the Boltzmann constant and \( T \) is room temperature in Kelvin. This shows that the change in PMA is modest and hence the PMA gradient does not have deterministic control in positioning the DW in the presence of thermal noise. To circumvent this issue, a higher PMA gradient could be used, though this would require a greater strain gradient, or a thicker free layer could be used with its PMA derived from bulk effects (e.g. magnetoelastic) instead of interfacial anisotropy. This analysis has not considered edge roughness [46], which can provide additional pinning sites and even remove the need for lithographic patterning of notches.

3.3. SOT-driven DW motion with pinning of the DW achieved by spatial PMA modulation

We also simulated DW motion clocked by SOT without including the effect of thermal noise in a MTJ racetrack with CoFe soft layer and no notches. Here we assume no current flows in the CoFe layer and hence there is no STT. We found that discrete PMA variation as shown in figure 4 was more suitable to control DW in this case with SOT, instead of using notches and a uniform PMA gradient as in the previous case with STT. The electrode arrangement shown in figure 4 alters the PMA at specific regions of the racetrack between the electrodes enabling creation of regions of different PMA. Moreover, due to the different spacing between the pairs of electrodes, the one with smaller gap will create a larger decrease in PMA due to higher stress. However, the PMA profiles in figure 4(b) are ideal representation of the actual PMA profile based on scaling arguments presented in section IV and prior work [36]. Determination of the real stress and PMA profile with detailed finite element analysis is beyond the scope of this paper.

The SOT clock was simulated with a current density \( J_c \approx 7 \times 10^{10} \) A m⁻² through the Pt layer of length 1000 nm, width 100 nm and thickness 1 nm for a clocking period of 20 ns. Initially, the DW is located at the left end of the nanotrack. This initialization can be accomplished by small geometric modification such as fabricating a notch. For lower voltage applied to the strain electrode the PMA reduction is small at the left end. Therefore, the DW experiences smaller barriers and can be translated further to the right. With increasing voltage, barriers at the left increase, thus arresting the DW closer to the left end as shown in figure 4(b). The geometric arrangement of the electrodes will ensure that the PMA decrease due to adjacent electrodes varies in a linear fashion. In this case, the ratio between the maximum and minimum PMA change between right most and left most electrode position is kept at 3. This will ensure by only varying the voltage the DW can be arrested at different electrode locations without requiring a very large stress to be applied by the piezoelectric. The maximum PMA change for each case of the DW position is shown in table 3. The DW would eventually drift to the center of the racetrack to minimize magnetostatic energy once the applied voltage is removed. However, with realistic edge roughness the DW is pinned [46] and hence implements a non-volatile synapse.

4. Energy efficiency of DNNs with voltage control of DWs (VC-DW)

While we have discussed the DW dynamics and operation of the non-volatile voltage programmed synapse in detail, figure 5(a) shows the manner in which this device can be adapted to form a hybrid DW-CMOS neuron. The CMOS buffer implements the threshold functionality of a neuron (figure 5(a)) as well as the ability of the neuron output of one stage to drive inputs to various
neurons of the next stage (high fan-out of the CMOS stage) via synapses. In order to reset the neuron, the current/SOT clock is used with current flowing in the opposite direction. Thereafter, the clock is used to synchronize the information flow from one state of the DNN to the next. Figure 5 describes the manner in which the outputs of one set of neurons can be multiplied by the synaptic weights and input to a neuron at the next stage.

4.1. Energy efficiency and area-density of the voltage control DW DNN implementation versus other implementation schemes

The energy dissipation in the device can be divided into two parts. One part consists of charging the piezoelectric layer for stress generation, which is essentially the energy lost in charging the capacitor \((1/2)CV^2\). \(C\) = capacitance of the piezoelectric layer between the metal contacts, \(V\) = voltage applied. The other part is the \(IFR\) loss of the clocking current through the magnetic layer of the racetrack or current through the platinum layer.

For the device simulated in the previous sections, the maximum \(\Delta PMA\) was \(0.37 \times 10^3\) J m\(^{-3}\) across the length of the device. The stress required to obtain this may be estimated as \(\sigma = \frac{\Delta PMA}{3/2\lambda}\), where \(\lambda\) is the magnetostriction. For CoFeB, with \(\lambda \approx 30 \times 10^{-6}\), an unreasonably high stress of \(~800\) MPa would be required, but for CoFe with \(\lambda \approx 250 \times 10^{-6}\) a stress of order \(~100\) MPa can produce the needed \(\Delta PMA\). From the Young’s modulus of CoFe of \(\sim 2.0 \times 10^5\) MPa, a strain of \(\sim 5 \times 10^{-4}\) is required (or more specifically a strain gradient of \(5 \times 10^{-4}\) over \(1\,\mu m\) distance). Reference [36] shows that with a \(500\,\mu m\) thick lead zirconate titanate (PZT) substrate and electrodes of side \(600\,\mu m\), a strain gradient of \(\sim 10^{-3}\) over a distance of \(500\,\mu m\) is feasible with application of \(1.5\) kV. This suggests that a strain gradient of \(\sim 10^{-3}\) \((\sim 5 \times 10^{-4}\) if a single electrode is considered\) over \(1000\,nm\) distance is possible with application of \(3\) V. The effective capacitance, \(C = \sim 26\,fF\), assuming relative permittivity of \(3000\). Hence, the energy dissipated is \((1/2) CV^2 = 117\,fJ\) for the electrical control of a scaled DW nanowire device with a footprint of \(1000\,nm \times 100\,nm\) and electrode of footprint \(1000\,nm \times 1000\,nm\). In the SOT-clocked case, considering \(70\,nm\) square electrodes and \(100\,nm\) thick PZT, the total capacitance is \(C = \sim 18\,fF\) while voltage needed for strain generation is \(0.3\) V following [36]. This results in an energy dissipation of \(~1\) fJ.

For the charge current through the CoFe layer with resistivity of \(280\,\Omega\,nm\) [47] and a current density of \(J_c \sim 8.7 \times 10^{12}\,A\,m^{-2}\) with a dimension of \(1000\,nm\) length, \(100\,nm\) width and \(1\,nm\) thickness for a clocking period of \(6\) ns the energy dissipated due to \(IFR\) loss is \(~15\,pJ\). We note that, the current density needed to operate our device is large. Further, effective spin polarization ratio of CoFe interfaced with a heavy metal can decrease (e.g. spin polarization was found to be \(0.26\) in [48]). This will result in even higher current requirement and the Joule heating due to this will be unmanageable. Therefore, DW device operated with STT will be impractical if further optimizations are not made. For the SOT scheme, considering the resistivity of Pt to be \(100\,\Omega\,nm\) and charge current density \(J = \sim 7 \times 10^{10}\,Am^{-2}\) through the Pt layer of length \(1000\,nm\), width \(100\,nm\) and thickness \(1\,nm\) for a clocking period of \(20\) ns, the \(IFR\) loss is \(~100\,fJ\). Therefore, energy consumption in the device is dominated by the current that produces the SOT. This can be further reduced if low damping materials such as iron garnets are used [49], which have the further advantage of avoiding current shunting through the magnetic layer. The difficulty will lie in optimizing the edge roughness and geometrical design of the racetrack to provide controllability at such low Gilbert damping. In summary, this clocked DW device concept provides a pathway to realize novel energy-efficient DW neuromorphic devices where reprogramming of synaptic weights can be performed at \(~100\,fJ\) per synapse during the learning phase and similarly small \(~100\,fJ\) per neuron during the inference phase of the neural network. In fact, during the inference phase, a neuron implemented with only CMOS devices (not a hybrid DW-CMOS device) would only need \(~\) few fJ per neuron and the synapses would consume no energy as they are non-volatile.

It is interesting to compare these numbers with alternative implementations of artificial neurons and synapses. The most important benefit of our approach for artificial synapses is the large reduction in energy consumption. The use of voltage control in conjunction with SOT drastically reduces the energy requirements versus purely spin torque DW-based devices [50]. Non-spintronic nanodevices can also provide multilevel synapses, such as oxide-based memristors [51] and phase change memories [34]. Programming such devices requires the physical motion of atoms in order to create or dissolve

![Figure 5](image.png)

**Figure 5.** Left: schematic of the spintronic DW-neuron implemented with SOT. Right: hybrid neuron and synapse.
conductive filaments (oxide-based memristor) or to crystallize amorphous volumes of chalcogenide materials (phase change memory), which has an inherent energy cost, usually higher than picojoules even in highly scaled devices. On the other hand, these alternative technologies may provide more compact synapses than spintronic ones. Our solution therefore offers an extremely energy efficient approach to potentially implement real time learning-capable systems.

In contrast, the benefits of our approach for implementing artificial neurons is reduction in area (density). As neurons do not require non-volatility, CMOS-based solutions are typically used for neurons and have comparable energy consumption with ours. On the other hand, they typically require multiple transistors and several square micrometers of area [52].

5. Conclusion

The feasibility of an energy-efficient voltage-controlled DW implementation of an artificial neuron and synapse was demonstrated using micromagnetic simulations. In this approach, modulation of perpendicular anisotropy with stress in combination with SOT or STT is used to program different synaptic weights as well as to mimic a neuron. Scaling this device to smaller dimensions (for example, ~500 nm × 50 nm × 1 nm) could result in much lower energy dissipation as well as high densities for comparable energy dissipation (for implementing neurons) compared to competing approaches. However, to avoid loss of controllability in deterministic positioning of the DW in the presence of thermal noise, careful optimization of material and device geometry are necessary. In summary, this work provides a pathway to the realization of energy-efficient voltage-controlled artificial neuron networks with real time learning capability and could stimulate more experimental work in this direction.

Acknowledgments

JA, DB and MA acknowledges support from NSF CAREER Grant CCF-1253370 (supplement), ECCS 1609303 and Collaborative/Educational Seed Project with Manufacturing Techniques, Inc. (MTEQ) Prime Contract with the US Army CSISR Night Vision and Electronic Sensors Directorate (NVESD) and Virginia Microelectronics Consortium (VMEC), Grant #MTEQ-19-028 TO 54_FP10321. CAR acknowledges support of the National Science Foundation under Award 1639921, and the Nanoelectronics Research Corporation (NERC), a subsidiary of the Semiconductor Research Corporation (SRC), through an SRC-NRI Nanoelectronics Research Initiative award, and SMART, an nCORE Center of SRC and NIST.

ORCID iDs

Caroline A Ross Ⓛ https://orcid.org/0000-0003-2262-1249
Jayasimha Atulasimha Ⓛ https://orcid.org/0000-0002-5681-0884

References

[1] Merolla P A et al 2014 A million spiking-neuron integrated circuit with a scalable communication network and interface Science 345 668–73
[2] Drachman D A 2005 Do we have brain to spare? Neurology 64 2004–5
[3] Sharad M, Augustine C, Panagopoulos G and Roy K 2012 Spin-based neuron model with domain-wall magnets as synapse IEEE Trans. Nanotechnol. 11 843–53
[4] Chung J, Park J and Ghosh S 2016 Domain wall memory based convolutional neural networks for bit-width extendability and energy-efficiency Proc. 2016 Int. Symp. on Low Power Electronics and Design 332–7
[5] Ma X, Zhang Y, Yuan G, Ren A, Li Z, Han J, Hu J and Wang Y 2018 An area and energy efficient design of domain-wall memory-based deep convolutional neural networks using stochastic computing Proc.—Int. Symp. Qual. Electron. Des. 314–21
[6] Sengupta A, Shim Y and Roy K 2016 Proposal for an all-spin artificial neural network: Emulating neural and synaptic functionalities through domain wall motion in ferromagnets IEEE Trans. Biomed. Circuits Syst. 10 1152–60
[7] Srinivasan G, Sengupta A and Roy K 2016 Magnetic tunnel junction based long-term short-term stochastic synapse for a spiking neural network with on-chip STDP learning Sci. Rep. 6 1–13
[8] Sengupta A, Al Azim Z, Fong X and Roy K 2015 Spin–orbit torque induced spike-timing dependent plasticity Appl. Phys. Lett. 106 093704
[9] Xu W, Min S-Y, Hwang H and Lee T-W 2016 Organic core-sheath nanowire artificial synapses with femtojoule energy consumption Sci. Adv. 2 e1501326
[10] Xu W, Cho H, Kim Y-H, Kim Y-T, Wolf C, Park C-G and Lee T-W 2016 Organometal halide perovskite artificial synapses Adv. Mater. 28 5916–22
[11] Yu H, Gong J, Wei H, Huang W and Xu W 2019 Mixed-halide perovskite for ultrasensitive two-terminal artificial synaptic devices Mater. Chem. Front. 3 941–7
[12] Chen Y et al 2015 Artificial synapses based on nanomaterials Nanotechnology 30 012001
[13] Querlioz D, Bichler O and Gamrat C 2011 Simulation of a memristor-based spiking neural network immune to device variations Proc. Int. Joint Conf. on Neural Networks 1775–81
[14] Merkel C, Kudithipudi D, Suri M and Wysocki B 2017 Stochastic CBRAM-based neuromorphic time series prediction system ACM J. Emerg. Technol. Comput. Syst. 13 1–14
[15] Yakopcic C, Alom M Z and Taha T M 2016 Memristor crossbar deep network implementation based on a Convolutional neural network Proc. Int. Jt. Conf. Neural Networks 963–70
[16] Querlioz D, Bichler O, Dollfus P and Gamrat C 2013 Immunity to device variations in a spiking neural network with memristive nanodevices IEEE Trans. Nanotechnol. 12 288–95
[17] Wang Z Q, Xu H Y, Li X H, Yu H, Liu Y C and Zhu X J 2012 Synaptic learning and memory functions achieved using oxygen ion migration/diffusion in an amorphous InGaZnO memristor Adv. Funct. Mater. 22 2759–65
[18] Cheng M, Xia L, Zhu Z, Cai Y, Xie Y, Wang Y and Yang H 2017 TIME: a training-in-memory architecture for memristor-based deep neural networks Proc. 54th Annual Design Automation Conf. 2017 1–6
[19] Han H, Yu H, Wei H, Gong J and Xu W 2019 Three-terminal artificial synapses: recent progress in three-terminal artificial synapses: from device to system Small 15 1970170
[20] Kim Y et al 2018 A bioinspired flexible organic artificial afferent nerve Science 360 998–1003
[21] Yamaguchi A, Ono T, Nasu S, Miyake K, Mibu K and Shinjo T 2004 Real-space observation of current-driven domain wall motion in submicron magnetic wires Phys. Rev. Lett. 92 077205
[22] Tatara G and Kohno H 2004 Theory of current-driven domain wall motion: spin transfer versus momentum transfer Phys. Rev. Lett. 92 1–4
[23] Li Z and Zhang S 2004 Domain-wall dynamics and spin-wave excitations with spin-transfer torques Phys. Rev. Lett. 92 207203
[24] Thiaville A, Nakatani Y, Miltat J and Suzuki Y 2005 Micromagnetic understanding of current-driven domain wall motion in patterned nanowires Europhys. Lett. 69 990–6
[25] Manchon A and Zhang S 2009 Theory of spin torque due to spin–orbit coupling Phys. Rev. B 79 1–9
[26] Matos-Abiague A and Rodríguez-Suárez R L 2009 Spin–orbit coupling mediated spin torque in a single ferromagnetic layer Phys. Rev. B 80 1–6
[27] Miron I M, Gaudin G, Auffret S, Rodmacq B, Schuhl A, Li Z and Zhang S 2017 Nonadiabatic spin-torque induced magnetic domain wall motion in patterned nanowires Nano Lett. 17 5869–74
[28] Emori S, Martínez E, Lee K J, Lee H W, Bauer U, Ahn S M, Agrawal P, Bono D and Beach G S D 2014 Spin Hall torque magnetometry of Dzyaloshinskii-Moriya interactions in CoFeB films Phys. Rev. B 89 1–13
[29] Y. Zhang, X. Luo, S. Yang and Y. Yang 2013 Spin–orbit-torque induced magnetic domain wall motion in Ta/CoFe nanowires with sloped perpendicular magnetic anisotropy Sci. Rep. 3 527–33
[30] Brignen W H, Hu X, Hassan N, Bennett C H, Incorvia J A C, Garcia-Canclini F and Friedman J S 2019 Graded-anisotropy-induced magnetic domain wall drift for an artificial spintronic leaky integrate-and-fire neuron IEEE J. Explor. Solid-State Comput. Devices Circuits 5 1–9
[31] Ryu K-S, Thomas L, Yang S-H and Parkin S 2013 Chiral spin torque at magnetic domain walls Nat. Nanotechnol. 8 527–33
[32] Bhownik M, Nowakowski M, You L, Lee O, Keating D, Wong M, Bokor J and Salahuddin S 2014 Deterministic domain wall motion orthogonal to current flow due to spin–orbit torque Sci. Rep. 5 11823
[33] Yu G, Upadhyaya P, Wong K L, Jiang W, Alzate J G, Tang J, Amin P K and Wang K L 2014 Magnetization switching through spin-Hall-effect-induced chiral domain wall propagation Phys. Rev. B 89 1–6
[34] Ambrogi S et al 2018 Equivalent-accuracy accelerated neural-network training using analogue memory Nature 558 60–7
[35] Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors Nature 323 533–6
[36] Cui J, Hockel J L, Nordeen P K, Pisani D M, Liang C Y, Carman G P and Lynch C S 2013 A method to control magnetism in individual strain-mediated magnetoelectric islands Appl. Phys. Lett. 103 232905
[37] Vlassenbroeke A, Vettenbroeke J, Dvornik M, Garcia-Sanchez F and Van Waeyenberge B 2014 The design and verification of Mumax3 AIP Adv. 4 107133
[38] Dzyaloshinsky I 1958 A thermodynamic theory of ‘weak’ ferromagnetism of antiferromagnetics J. Phys. Chem. Solids 4 241–55
[39] Moriya T 1960 Anisotropic superexchange interaction and weak ferromagnetism Phys. Rev. 120 91–8
[40] Spedalieri F M, Jacob A P, Nikonov D E and Rooyenwoudhury V P 2011 Performance of magnetic quantum cellular automata and limitations due to thermal noise IEEE Trans. Nanotechnol. 10 537–46
[41] Mayugey I D, Bertotti G and Serpico C 2009 Nonlinear Magnetization Dynamics in Nanosystems (Amsterdam: Elsevier)
[42] Burrowes et al 2010 Non-adiabatic spin-torques in narrow magnetic domain walls Nat. Phys. 6 17–21
[43] Belmeguenai M, Gabor M S, Roussigné Y, Stashkevich A, Cherif S M, Zighem F and Tiusan C 2017 Brillouin light scattering investigation of the thickness dependence of Dzyaloshinskii-Moriya interaction in CoFeB films Phys. Rev. B 93 1–8
[44] Emori S, Martínez E, Lee K J, Lee H W, Bauer U, Ahn S M, Agrawal P, Bono D and Beach G S D 2014 Spin Hall torque magnetometry of Dzyaloshinskii-Moriya domain walls Phys. Rev. B 90 1–13
[45] Bilzer C, Devolder T, Von K J, Council G, Chappert C, Cardoso S and Freitas P P 2006 Study of the dynamic magnetic properties of soft CoFeB films J. Appl. Phys. 100 053903
[46] Dutta S, Siddiqui S A, Currivan-Incorvia J A, Ross C A and Baldo M A 2017 The spatial resolution limit for an individual domain wall in magnetic nanowires Nano Lett. 17 5869–74
[47] Zahnd G, Vila L, Pham T V, Marty A, Laczkowski P, Savero Torres W, Beigne C, Vergnaud C, Jamet M and Attane J-P 2016 Comparison of the use of NiFe and CoFe as electrodes for metallic lateral spin valves Nanotechnology 27 035201
[48] Zhang X et al 2019 Low spin polarization in heavy-metal-ferromagnet structures detected through domain-wall motion by synchronized magnetic field and current Phys. Rev. Appl. 11 054041
[49] Avci C O, Rosenberg E, Caretta L, Büttner F, Mann M, Marcus C, Bono D, Ross C A and Beach G S D 2019 Interface-driven chiral magnetism and current-driven domain walls in insulating magnetic garnets Nat. Nanotechnol. 14 561–6
[50] Lequeux S, Sampio J, Clos V, Yakushiji K, Fukushima A, Matsumoto R, Kubota H, Yuasa S and Grollier J 2016 A magnetic nanoscale: multilevel spin-torque memristor with perpendicular anisotropy Sci. Rep. 6 31510
[51] Przezioso M, Merrikh-Bayat F, Hoskins B D, Adam G C, Litharev K K and Strukov D B 2015 Training and operation of an integrated neuromorphic network based on metal-oxide memristors Nature 521 61–4
[52] Bennett C H, Parmar V, Calvet L E, Klein J-O, Suri M, Marinella M J and Querlioz D 2019 Contrasting advantages of learning with random weights and backpropagation in non-volatile memory neural networks IEEE Access 7 1–16