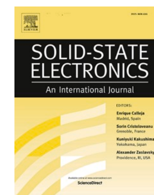
Solid-State Electronics 185 (2021) 108075

Contents lists available at ScienceDirect

Solid State Electronics

journal homepage: www.elsevier.com/locate/sse



Two-pulse switching scheme and reinforcement learning for energy efficient SOT-MRAM simulations

R.L. de Orio [a,*], J. Ender [b], S. Fiorentini [b], W. Goes [c], S. Selberherr [a], V. Sverdlov [b]

[a] Institute for Microelectronics, Gußhausstraße 27-29/E360, TU Wien, Vienna 1040, Austria
[b] Christian Doppler Laboratory for Nonvolatile Magnetoresistive Memory and Logic at the Institute for Microelectronics, Austria
[c] Silvaco Europe Ltd, Cambridge, United Kingdom

ARTICLE INFO

Handling Editor: Bogdan Cretu

Keywords:
Spin–orbit torque MRAM
Magnetic field-free switching
Two-pulse switching scheme
Reinforcement learning
Machine learning

ABSTRACT

We demonstrate by means of numerical simulations the switching of a perpendicularly magnetized free layer by spin–orbit torques based on a two-pulse switching scheme with improved writing power efficiency. In this scheme, the first pulse selects the cell, while the second pulse completes the switching deterministically. It is shown that the magnitude of the second current pulse can be reduced to about 50% of the critical current and the switching remains reliable with a switching time of 300 ps. With such a significant current reduction the writing power required for switching decreases by 40%, which results in a very energy efficient scheme. In addition, we develop a reinforcement learning approach to optimize the pulse configuration with the goal of achieving the shortest switching time. With this approach a switching time of 146 ps has been obtained, a reduction of 50% in relation to the non-optimized configuration. These research findings confirm that reinforcement learning is a promising tool to simplify and automate the search for a faster, energy efficient scheme in the two-pulse switching approach.

1. Introduction

The classical charge-based solid-state memory cells, the static random access memory (SRAM) cell and the dynamic random access memory (DRAM) cell, are intrinsically volatile, which has resulted in increasing standby power consumption as the cells have been downscaled. A solution to this issue can be obtained only with the introduction of nonvolatile memory cells, which must exhibit operation characteristics comparable to those of SRAM or DRAM cells [1].

Spin-transfer torque magnetoresistive random access memory (STT-MRAM) is currently the state of the art MRAM technology, with embedded STT-MRAM having densities in the order of gigabytes already demonstrated [2]. It is suitable for embedded nonvolatile memory applications as well as for a replacement of flash memories [3] and L4 caches [4]. The core element of the STT-MRAM cell is a magnetic tunnel junction (MTJ), which is formed by two ferromagnetic layers separated by a tunnel barrier [5]. However, fast operation with timings in the order of nanoseconds demands large switching currents flowing through the MTJ and leading to oxide reliability issues, reducing the MRAM endurance.

Spin–orbit torque magnetoresistive random access memory (SOT-MRAM) is a promising future nonvolatile memory solution beyond the STT-MRAM [6]. In particular, it is a viable option for a nonvolatile replacement of high-level caches, as it delivers high operation speed and large endurance. However, for deterministic SOT switching of a perpendicularly magnetized free layer (FL) an external magnetic field is required [7], which is cumbersome for large scale integration. In order to circumvent this issue, several field-free schemes which break the cell mirror symmetry at the physical level have been proposed [8–12]. Among the most recent works, Honjo et al. [13] showed the integration of a canted SOT cell, controlling the shape of the cell and its orientation with respect to the applied current direction. In turn, Garello et al. [14] demonstrated a successful integration of a cobalt nanomagnet into the SOT cell, which provides the required magnetic field for deterministic switching. Furthermore, the approach is suitable for full 300 mm wafer fabrication [15]. These schemes typically require precise geometrical shaping or the introduction of specific materials or layers into the cell stack, which poses several challenges for the processing steps and fabrication of the cells.

In the search for more energy-efficient SOT-MRAM, for which the write current and power are reduced, various schemes based on purely electrical field-free control of magnetization switching have been

* Corresponding author.
E-mail address: orio@iue.tuwien.ac.at (R.L. de Orio).

https://doi.org/10.1016/j.sse.2021.108075
Received 5 December 2020; Received in revised form 29 March 2021; Accepted 10 May 2021
Available online 1 June 2021
0038-1101/© 2021 Elsevier Ltd. All rights reserved.

suggested. The main idea of these schemes is to combine SOT with STT switching [16–18] and/or to combine SOT with voltage-controlled magnetic anisotropy (VCMA) [16,19,20]. In the first case, two current pulses are applied to the cell. One pulse is applied through the heavy metal layer under the magnetic FL, while another current pulse is applied through the MTJ. These currents generate SOT and STT, respectively, which act together to switch the magnetization. The amplitude and duration of both currents can be tuned to achieve low-power switching with sub-nanosecond timing and enhance the write endurance of the MTJ. In the second case, SOT is combined with VCMA. Here, a voltage is applied across the MTJ and the effect of VCMA lowers the energy barrier for switching, leading to a reduction of the write current and acceleration of the switching. In this way, a more efficient switching scheme is obtained. In fact, the interplay of these three mechanisms, SOT, STT, and VCMA leads to a high speed energy-efficient switching scheme [16].

In this work, an alternative magnetic field-free scheme based on purely SOT switching controlled by two orthogonal current pulses is considered. The scheme was initially proposed to switch an in-plane magnetized FL [21,22]. For an in-plane cell, switching can be achieved by the application of a single current pulse. However, it was shown that, by applying current pulses to two orthogonal heavy metal wires in contact with the FL (c.f. Fig. 1), the switching could be speeded up. This two-pulse scheme was later applied to switch a perpendicularly magnetized FL, where the influence of the overlap of the second heavy metal wire with the FL and the influence of the duration of the second current pulse on the switching time and robustness were investigated [22,23]. The large currents required to realize the switching are still an issue. They not only demand higher writing power and energy, but also make the integration of the cell in memory arrays more difficult, because they can disturb the magnetization of half-selected cells of the array.

The memory cell for the two-pulse scheme is depicted in Fig. 1. The writing SOT-cell is formed by a perpendicularly magnetized FL on top of a heavy metal wire (NM1) and a second orthogonal heavy metal wire (NM2) is placed on top of the FL, partially overlapping it. An MTJ is placed next to the SOT-cell for the reading operation. In this cell, the integration of the additional NM2 wire does not require a direct modification of the MTJ stack and it can be deposited and patterned using conventional processes. However, a specific challenge for its realization is to properly stop the etching on the FL, so that the thin ferromagnet is not damaged during the process.

Machine learning (ML) has been increasingly applied in the realm of physics [24]. While the predominant number of ML applications uses supervised learning approaches, which require large amounts of data beforehand to train neural networks, the subbranch of reinforcement learning (RL) [25] has gained interest in recent years. The general reinforcement learning setup consists of an agent and an environment. The agent interacts with the environment by performing certain *actions*, making the environment transition from one *state* to another one. After every transition, the environment returns the new state, as well as a *reward* to the agent. Here, the agent tries to maximize the cumulative reward it receives over time, learning how to achieve a certain objective. First RL breakthroughs were achieved using games like chess or Go [26],

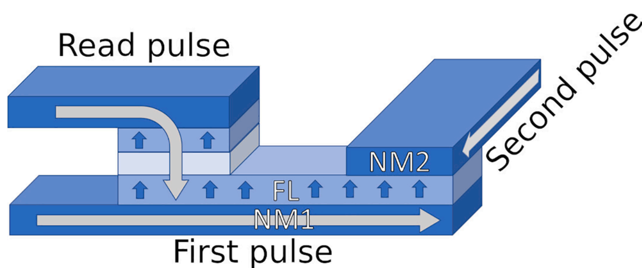


Fig. 1. SOT-MRAM cell for switching based on two orthogonal current pulses.

and these types of algorithms have also successfully been applied in physics, e.g. [27], where strategies for quantum error correction are found through RL.

This work focuses on the reduction of the applied current for switching of a perpendicularly magnetized cell based on the two-pulse scheme. After a pre-selection of the cell by the first current pulse, the switching current of the second pulse can be significantly reduced, and deterministic and fast switching of a perpendicularly magnetized FL is still guaranteed. This current reduction is accompanied by a large decrease of the writing power, improving the energy efficiency of the switching scheme. The importance of decreasing the second current below the critical value is discussed, so that the magnetization of non-selected cells is not disturbed, even if the NM2 wire is routed through several cells in an array. A particular innovation is the implementation of a reinforcement learning approach to support and automate the search for the fastest switching in the two-pulse scheme, which was confirmed by our research findings.

2. Spin–orbit torque switching scheme

2.1. Micromagnetic modeling

The writing operation of the cell is carried out by applying two current pulses to two orthogonal heavy metal wires, NM1 and NM2 (c.f. Fig. 1). It is initiated with the selection of the cell by applying the first pulse to the NM1 wire. This pulse generates the SOT which puts the magnetization in the plane of the FL orthogonal to the current. Then, the second pulse is applied to the NM2 wire. The second pulse rotates the magnetization of the FL under the NM2 wire, which dynamically creates an in-plane magnetic field to complete the magnetization switching, and thus the cell writing, deterministically [28].

The reading operation is carried out by applying a low current pulse through the MTJ and sensing the corresponding tunneling magnetoresistance ratio. Since the current through the tunnel barrier of the MTJ is small, oxide reliability is not an issue.

The magnetization dynamics is described by the Landau-Lifshitz-Gilbert equation

$$\frac{\partial \mathbf{m}}{\partial t} = -\gamma\mu_0 \mathbf{m} \times \mathbf{H}_{\text{eff}} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \frac{1}{M_S}\mathbf{T}_\mathbf{S}, \tag{1}$$

where $\mathbf{m}$ is the normalized magnetization, γ is the gyromagnetic ratio, μ_0 is the vacuum permeability, α is the Gilbert damping factor, and M_S is the saturation magnetization. $\mathbf{H}_{\text{eff}}$ is an effective magnetic field which includes the exchange field, the uniaxial perpendicular anisotropy field, the demagnetization field, the current-induced field, and the random thermal field at 300 K. $\mathbf{T}_\mathbf{S}$ is the SOT generated by the current, given by

$$\mathbf{T}_\mathbf{S} = -\gamma \frac{\hbar}{2e} \frac{\theta_{SH} j}{M_S d} [\mathbf{m} \times (\mathbf{m} \times (\mathbf{j} \times \mathbf{z}))], \tag{2}$$

where e is the elementary charge, $\hbar$ is the reduced Plank constant, θ_{SH} is an effective Hall angle, j is the applied current density, d is the FL thickness, and $\mathbf{z}$ is the unit vector perpendicular to the FL plane.

It should be pointed out that only the damping-like torque (DLT) is considered in (2). A large field-like torque (FLT) has been reported for Ta-based SOT structures [15,29], while for W-based structures a much smaller FLT is normally observed [15], up to one order of magnitude lower than the DLT [30]. In this work, we consider SOT cells based on β-tungsten heavy metal wires and neglect the FLT.

To realize the SOT switching, the applied current density must be larger than the critical current density (in the absence of an external field) [7]

$$J_C = \frac{e}{\hbar} \frac{M_S d}{\theta_{SH}} H_K, \tag{3}$$

where H_K is the effective anisotropy field. For the two-pulse switching

scheme, the selection pulse, i.e. the first pulse applied to the NM1 wire, has a current $I_1 = 130\ \mu$A and a fixed duration $T_1 = 130$ ps. This yields a current density $j_1 = 2.2 \times 10^{12}$ A/m^2, which is just above the critical one ($I_c = 120\ \mu$A, $J_c = 2.0 \times 10^{12}$ A/m^2), given the cell parameters in Table 1. In turn, the magnitude of the second current pulse (I_2) is varied and its impact on the switching is investigated.

The numerical simulations of the magnetization dynamics are carried out using an in–house tool [31] based on the finite differences method with a grid size of 1.2 nm. The parameters used in the simulations are listed in Table 1. A stability factor of 45 is computed for the cell. In order to account for the thermal distribution resulting from the random thermal field, a total of 50 realizations are considered for each simulation condition.

2.2. Reinforcement learning

The general setup of the two-pulse switching cell in an RL setting can be seen in Fig. 2. The environment contains a simulation of the two-pulse switching SOT-MRAM cell. For this purpose, our in–house micromagnetic simulation tool (as described above) was adapted in such way that it can be controlled from the outside to change the state of the pulses and that it returns the current state of the simulation together with a reward after every iteration. The agent consists of a neural network provided by a python library [32], which implements the deep Q-network (DQN) as learning algorithm [33]. The goal of the RL algorithm is to determine the pulse sequence yielding the fastest switching time, i.e. the objective of the experiments is to achieve the fastest possible transition of the average z-component of the magnetization from +1 to −0.5.

In order to direct the algorithm to take actions such that the switching time is reduced, an appropriate scheme for rewarding the agent must be chosen. The reward is an integer value returned by the environment indicating whether the actions performed by the agent were good or bad. As RL algorithms aim to maximize the cumulative reward, the rewarding scheme should be chosen such that a shorter switching time corresponds to a higher reward. For every simulation step in which the target was not yet reached, a reward of −1 is given. An upper limit of the simulation time of $t_{max} = 1$ ns was defined. Once this time is reached without the z-component reaching −0.5 the learning episode is terminated. If, however, the z-component reaches this threshold before 1 ns, a positive reward of $(t_{max} - t_{final})/dt$ is given.

The state vector returned from the environment after every iteration consists of 11 variables: the average of the three magnetization vector components (m_x, m_y, m_z), the difference of each component to the previous iteration $(\Delta m_x, \Delta m_y, \Delta m_z)$, the average component of the effective magnetic field $(H_{\mathrm{eff,x}}, H_{\mathrm{eff,y}}, H_{\mathrm{eff,z}})$, and two variables indicating whether the first and the second pulse are settable. If for example the change of the average magnetization components were not included (Δm_x, Δm_y, Δm_z), by only knowing the current value of the components, it would not be clear in which direction the magnetization is moving at this point. Based on the state information, the learning agent deduces which action to take. The current setup allows the agent to take four different actions: both pulses are off, both pulses are on, the first pulse is on, the second is off, or the first pulse is off and the second pulse is on.

Table 1
Parameters used in the simulations. β-tungsten heavy metal wires are assumed, while the magnetic FL is CoFeB on MgO [7].

Parameter	Value
Saturation magnetization, M_S	1.1×10^6 A/m
Exchange constant, A	1.0×10^{-11} J/m
Perpendicular anisotropy, K	8.4×10^5 J/m^3
Gilbert damping factor, α	0.035
Spin Hall angle, θ_{SH}	0.3
Free layer dimensions	40 nm × 20 nm × 1.2 nm
NM1: $w_1 \times l$	20 nm × 3 nm
NM2: $w_2 \times l$	16 – 36 nm × 3 nm

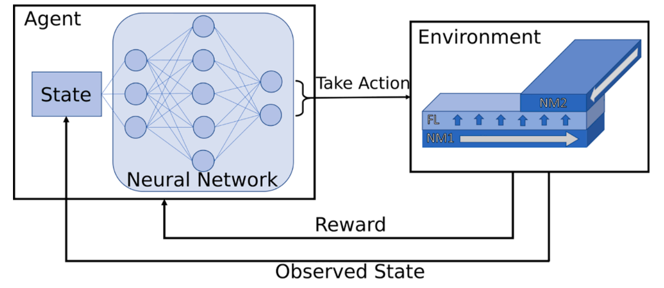


Fig. 2. Reinforcement learning setup with two-pulse switching scheme. The simulation of the SOT-MRAM cell provides the environment, which the agent interacts with to take decisions that lead to a certain goal as, for example, the fastest switching time.

3. Results and discussion

3.1. Micromagnetic simulations

Since an MTJ is located beside the writing part of the SOT-cell, the NM2 wire contacts only part of the FL depending on the wire width, as depicted in Fig. 1. Considering, initially, that the current density of the second pulse is equal to the current density of the first pulse, i.e. $j_2 = 2.2 \times 10^{12}$ A/m^2, and for a pulse width $T_2 = 150$ ps, the z-component of the magnetization dynamics for different values of the NM2 wire width is shown in Fig. 3. Deterministic switching was obtained for all values of tested w_2, where each curve represents an average of 50 realizations, due to the stochastic thermal field at 300 K. One can see that the switching characteristics remain essentially the same and the switching time stays within a range of 0.3–0.4 ns (taken at $m_z = -0.5$). These results indicate that, due to the relatively large range of suitable w_2 dimensions, a particularly tight control of the NM2 wire patterning is not necessary, which is beneficial for the cell fabrication.

The results shown in Fig. 3 were obtained for a constant current density, $j_2 = 2.2 \times 10^{12}$ A/m^2, slightly above the critical value. In order to keep a constant current density, the applied current has to be adjusted proportionally to the wire width w_2, i.e. $I_2 = j_2 w_2 l$. Thus, reducing the wire width leads to a reduction of the current, provided that the current density is kept constant. Since the power is proportional to the square of

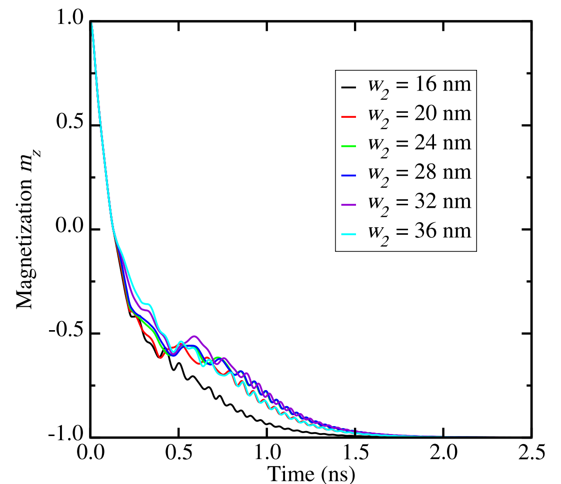


Fig. 3. Switching dynamics for different values of the NM2 wire width. The applied current is varied proportionally to w_2, while the current density is kept constant at $j_2 = 2.2 \times 10^{12}$ A/m^2 and $T_2 = 150$ ps.

the current, lowering the current is very important and advantageous, because the writing power is also reduced. The normalized switching power considering both current pulses ($P \propto (I_1^2 T_1 + I_2^2 T_2)/(T_1 + T_2)$) for different values of w_2 is shown in Table 2. The power is normalized in relation to $I_2 = 130$ μA, which is the current amplitude of the pulse applied to the NM1 wire. By adjusting the second wire width and the corresponding current, a writing power reduction factor of more than 2.0 is possible.

Although fast and efficient switching has been obtained, the second current density is still larger than the critical one. This means that the NM2 wire cannot be routed through various cells in an array, because the SOT generated by the second current pulse drives the magnetization of the non-selected cells to the FL plane. Not only is the stored information temporarily disturbed, but also the probability of an undesired bit flip increases. This prevents the use of multi-cell routing. For each bit cell an extra access transistor is required to control the current flow through the selected cell only. This increases the number of transistors per cell, the cell area, and the wiring overhead, which certainly makes the cell integration more complex. Therefore, the switching current density must be reduced. The above approach to reduce the current is bounded by the minimum wire width. Moreover, if the NM2 wire contact with the FL becomes too narrow, the switching times increase again [22]. To further reduce the second current, and thus the power dissipation, the current density has to be decreased below the critical value.

Fig. 4 shows the magnetization switching for several magnitudes of the second current below the critical one, where $w_2 = 20$ nm is fixed. In this case the current decrease also leads to a decrease of the current density. Even for $I_2 = 60$ μA, which corresponds to a reduction as large as 50% of the critical current, the switching characteristics are still preserved, i.e. the switching remains deterministic and fast. This is further demonstrated in Fig. 5, which summarizes the switching times as a function of the second current magnitude and pulse duration. The writing current density can be decreased to about 50% of the critical one, while maintaining the switching time at 0.3 ns for $T_2 \geqslant 150$ ps. If the current is decreased below 60 μA, the switching becomes non-deterministic.

As a consequence of the lower applied current, the power consumption is also significantly reduced. Table 3 shows the current, the current density, and the power consumption for a fixed wire width $w_2 = 20$ nm. Decreasing the second current by 50%, from 130 μA to 65 μA, corresponds to a reduction of 75% of the second pulse's power/energy. When both current pulses are considered, the total writing power is reduced by 40%. If we take into account this power reduction together with that from the adjustment of the NM2 wire width (c.f. Table 2), the total writing power is decreased by a factor of 3.3, which represents a power reduction of 70%. It should be pointed out that such a power reduction is obtained without loss in switching performance. Therefore, the efficiency of the switching scheme has been significantly improved.

For a current of the second pulse smaller than the critical value, the corresponding SOT is weaker and non-selected cells of a memory array are not disturbed. This is confirmed by the magnetization dynamics reported in Figs. 6 and 7. Here, the cell is not selected by the first pulse

Table 2
Magnitude of the second current pulse and power consumption as a function of the NM2 wire width. The simulation parameters are: $I_1 = 130$ μA ($j_1 = 2.2 \times 10^{12}$ A/m^2), $T_1 = 130$ ps, and $T_2 = 150$ ps. I_2 varies proportionally with w_2 in order to keep the current density fixed at $j_2 = 2.2 \times 10^{12}$ A/m^2.

w_2 (nm)	I_2 (μA)	Normalized power
36	234	2.20
32	208	1.84
28	182	1.51
24	156	1.24
20	130	1.00
16	104	0.81

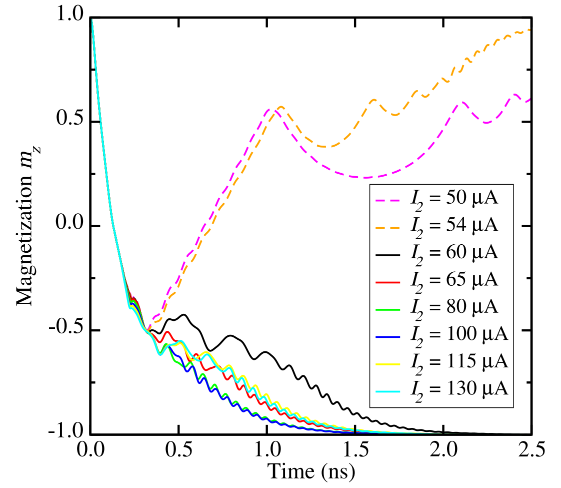


Fig. 4. z-component of the magnetization for various current pulses I_2 below the critical current and considering $w_2 = 20$ nm. The switching is deterministic and fast even for a current 50% lower than the critical one.

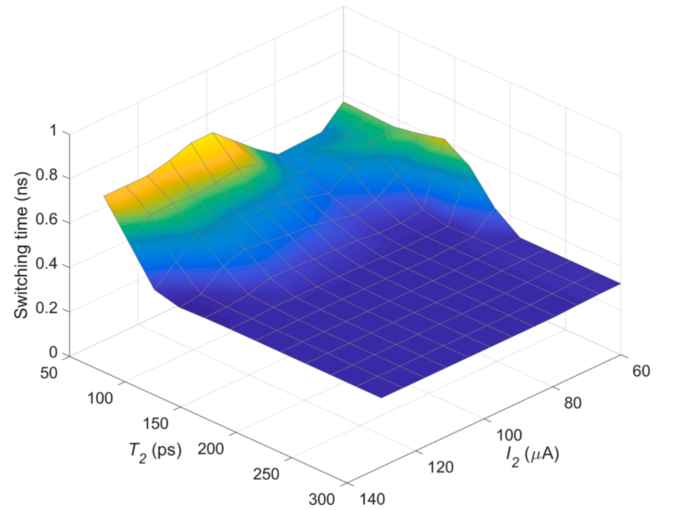


Fig. 5. Switching time as a function of the second current I_2 and as a function of the pulse duration T_2. The applied current can be significantly reduced, while the switching time remains constant.

Table 3
Second pulse current, current density, and writing power for currents below the critical value. $w_2 = 20$ nm is fixed.

I_2 (μA)	j_2 (10^{12}A/m^2)	Normalized power
130	2.2	1.00
115	1.9	0.88
100	1.7	0.78
80	1.3	0.67
65	1.1	0.60
60	1.0	0.58

and only the second current pulse is applied. The weak SOT slightly deviates the magnetization from the initial out-of-plane orientation. As Fig. 7 shows, m_z does not fall below 0.9 and the magnetization rapidly moves back to its initial position after some precession around the perpendicular anisotropy field, indicated by the oscillating m_x and m_y components in Fig. 6. This proves that non-selected cells remain undisturbed and an undesired bit flip does not occur, when a cell is subject to the second current pulse only.

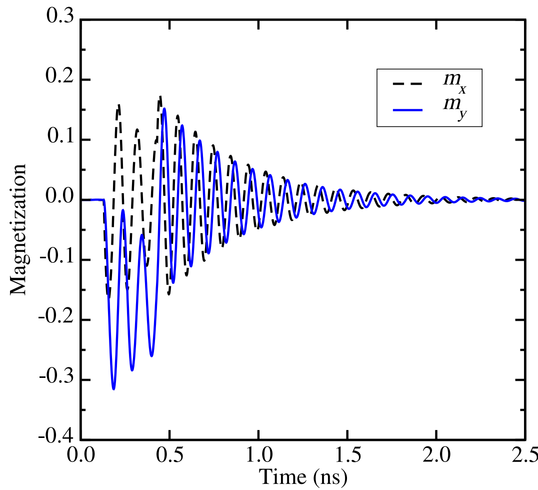


Fig. 6. Magnetization components m_x and m_y without pre-selecting the cell. The simulations parameters are: $I_1 = 0$, $T_1 = 130$ ps, $I_2 = 100$ μA, $T_2 = 150$ ps.

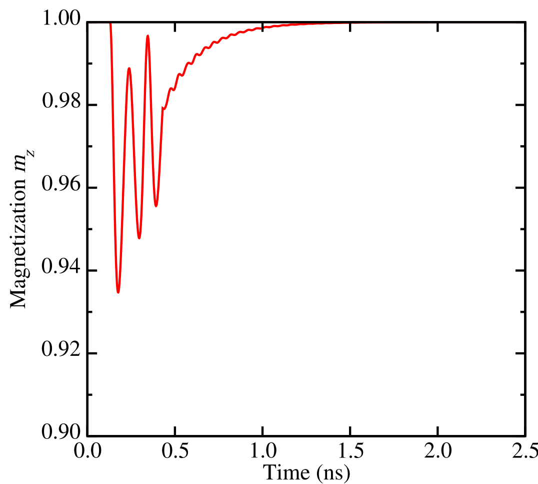


Fig. 7. Detail of the z-component of the magnetization without pre-selecting the cell.

3.2. Reinforcement learning

The previous analysis focused on the reduction of the current amplitude and its impact on the switching. It was shown that the current can be lower than the critical value and a reliable switching is still guaranteed. A minimum switching time of 300 ps was obtained.

In this section we apply the RL approach to find out the pulse configuration which leads to the fastest switching condition, defined here by the time the z-component of the magnetization reaches -0.5. The amplitude of the two pulses was fixed to $I_1 = 130$ μA for the first pulse and $I_2 = 100$ μA for the second pulse. It should be pointed out that the latter is lower than the critical current, following the results from the micromagnetic simulations of the previous section. To simplify the learning process, the amount of possible actions carried out by the agent (c.f. Fig. 2) has been restricted. The current setup of the experiment allows the agent to individually switch the two pulses on or off, with a minimum pulse width of 100 ps for each pulse. A learning episode was considered finished once the z-component of the magnetization reached -0.5. It should be emphasized that the focus of the RL experiments is to train a model for magnetization reversal which leads to the shortest switching time. These experiments have not been designed to take into account the effect of fluctuations or errors in the pulse parameters and timing.

Six individual experiments with different random seeds were performed. Fig. 8 shows results of the learning process. In Fig. 8(a), 20 independent results of the switching time over the course of the learning period are shown and one can see that learning is not yet fully deterministic. Fig. 8(b) presents the mean switching time (episode length) over the six best learning runs. Over the course of 10^6 learning steps, first an increase in the episode length, due to the initial focus on exploration of the state–action space, can be observed, which is followed by a decrease resulting in a final mean switching time of around 240 ps.

Fig. 9 shows the pulse sequence learned by the DQN algorithm. Both pulses are turned on right in the beginning and the first pulse is turned off again after 100 ps, after which the z-component of the magnetization drops below the threshold of -0.5. We observe that, although the target value is reached, since the second pulse is left on for the rest of the simulation, the magnetization does not fully switch to -1, but rather converges to a value of around -0.8. This demonstrates the importance of the rewarding scheme and general setup of the RL experiment. As the RL agent was rewarded for finishing the episode as fast as possible and the episode was considered finished as soon as the threshold of -0.5 was reached, the agent learned how to get to the threshold and did not care what happened afterwards.

Fig. 10 shows the dynamics of the z-component of the magnetization for pulse sequences which are slight variations of the learned one. The blue curves show that, when the second pulse is turned off once -0.5 is reached, the magnetization moves towards -1. This confirms that turning both pulses on right in the beginning indeed leads to fast switching. Further experiments with small timing differences were performed and Fig. 10 shows that extending the first pulse and/or delaying the second pulse (orange and greens curves) leads to longer switching times compared to the learned sequence with the second pulse switched off. In one scenario (purple curve) it even leads to no switching at all. As previously mentioned, the current RL setup does not account for possible errors in the pulse timing, so a relatively small variation of the ideal pulse sequence can lead to an unreliable scheme. Nevertheless, the results yield a proof of concept. The proposed RL method has been able to achieve its goal and find a pulse sequence producing the fastest magnetization switching by the two-pulse scheme.

In order to demonstrate that the learned scheme with the second pulse switched off (blue curve in Fig. 10) results in deterministic switching, 50 switching realizations of this pulse sequence were performed for which the results are given in Fig. 11. The variation between the different realizations is very small and all of them switch. It should

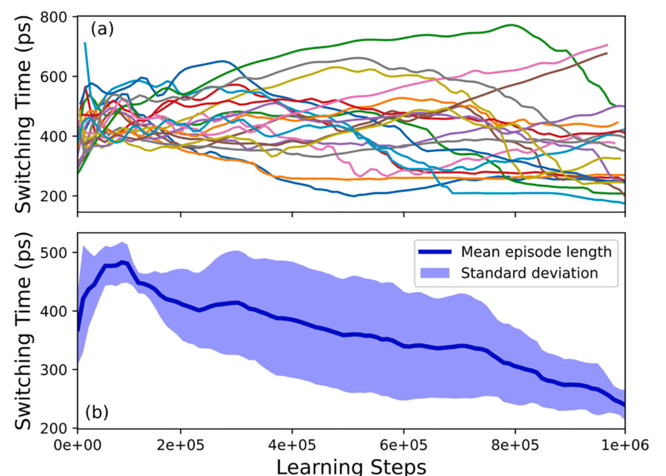


Fig. 8. Learning curve showing the mean episode length over 10^6 time steps. (a) Switching time over the course of the learning period of 20 independent runs. (b) Mean and standard deviation of the switching time over the six best runs.

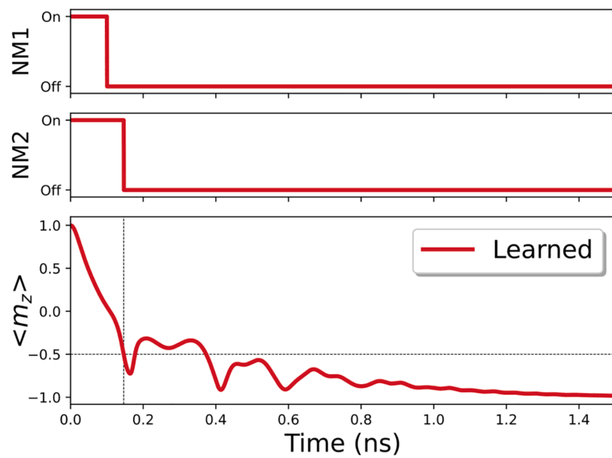


Fig. 9. Pulse sequence learned by the DQN agent. Dashed lines indicate the threshold, when the magnetization was considered switched (−0.5) and the corresponding time (146 ps).

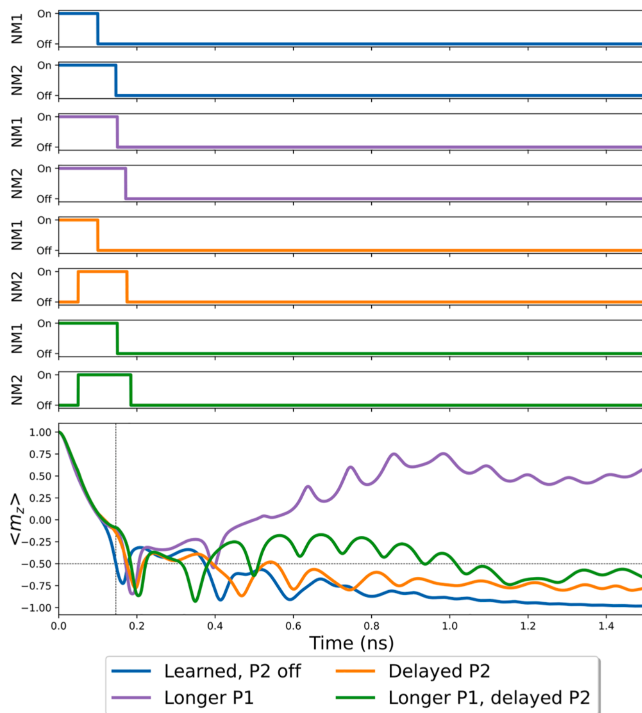


Fig. 10. Comparison of different pulse sequences. The learned sequence is compared with slightly modified ones. The learned sequence exhibits fastest switching of the z-component to −0.5.

be pointed out that based on the learned pulse sequence, a switching time of 146 ps is obtained. This is about one-half of the switching time obtained by the manual analysis of the previous section, 300 ps, which demonstrates the potential of the RL approach in combination with micromagnetic modeling.

4. Conclusion

Deterministic switching of a perpendicularly magnetized free layer by spin–orbit torques can be accomplished using a two-pulse approach with increased power efficiency. The first pulse selects the cell, while the second, low current pulse completes the switching. The switching characteristics remained practically unchanged, even when the second current was 50% lower than the critical one. In this case, the switching power was reduced by 40%. Thus the scheme became more efficient. A reinforcement learning approach was applied to determine the pulse sequence which leads to the shortest switching time. With this approach a switching time as short as 146 ps was obtained in comparison to the 300 ps of the manually optimized configuration. Our results confirm that reinforcement learning is a promising tool to automate the search for a faster, energy efficient switching scheme for SOT-MRAM cells.

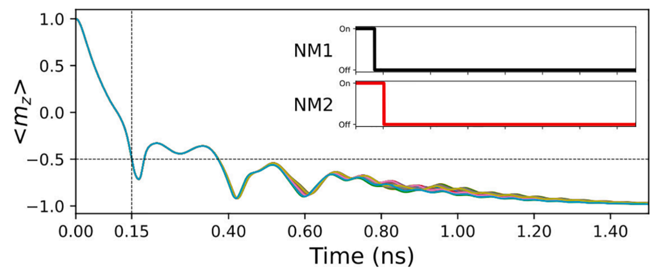


Fig. 11. z-component of the magnetization of 50 switching realizations using the switching scheme found by the RL algorithm, which is shown in the inset. The horizontal dashed line indicates the value of the average magnetization at which the cell is considered to be switched. The vertical dashed line indicates the corresponding time at which switching has occurred.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The financial support by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

[1] Lee SW, Lee KJ. Emerging three-terminal magnetic memory devices. Proc IEEE 2016;104:1831–43. https://doi.org/10.1109/JPROC.2016.2543782.

[2] Lee K, Bak JH, Kim YJ, Kim CK, Antonyan A, Chang DH, et al., 1 Gbit high density embedded STT-MRAM in 28 nm FDSOI technology. In: Proc IEDM Conf; 2019. p. 2.2.1–4. https://doi.org/10.1109/IEDM19573.2019.8993551.

[3] Golonzka O, Alzate JG, Arslan U, Bohr M, Bai P, Brockman J, et al., MRAM as embedded non-volatile memory solution for 22FFL FinFet technology. In: Proc IEDM Conf; 2018. p. 18.1.1–4. https://doi.org/10.1109/IEDM.2018.8614620.

[4] Alzate JG, Arslan U, Bai P, Brockman J, Chen YJ, Das N, et al., 2 MB array-level demonstration of STT-MRAM process and performance towards L4 cache applications. In: Proc IEDM Conf; 2019. p. 2.4.1–4. https://doi.org/10.1109/IEDM19573.2019.8993474.

[5] Wolf SA, Awschalom DD, Buhrman RA, Daughton JM, von Molnár S, Roukes ML, Chtchelkanova AY, Treger DM. Spintronics: A spin-based electronics vision for the future. Science 2001;294:1488–95. https://doi.org/10.1126/science.1065389.

[6] Miron IM, Gaudin G, Auffret S, Rodmacq B, Schuhl A, Pizzini S, Vogel J, Gambardella P. Current-driven spin torque induced by the Rashba effect in a ferromagnetic metal layer. Nat Mater 2010;9:230–4. https://doi.org/10.1038/nmat2613.

[7] Fukami S, Anekawa T, Zhang C, Ohno H. A spin-orbit torque switching scheme with collinear magnetic easy axis and current configuration. Nat Nanotechnol 2016;11:621–6. https://doi.org/10.1038/nnano.2016.29.

[8] Fukami S, Zhang C, DuttaGupta S, Kurenkov A, Ohno H. Magnetization switching by spin-orbit torque in an antiferromagnet-ferromagnet bilayer system. Nat Mater 2016;15:535–41. https://doi.org/10.1038/nmat4566.

[9] Oh YW, Baek SHC, Kim YM, Lee HY, Lee KD, Yang CG, et al. Field-free switching of perpendicular magnetization through spin-orbit torque in antiferromagnet/ferromagnet/oxide structures. Nat Nanotechnol 2016;11:878–84. https://doi.org/10.1038/nnano.2016.109.

[10] Wu H, Razavi SA, Shao Q, Li X, Wong KL, Liu Y, Yin G, Wang KL. Spin-orbit torque from a ferromagnetic metal. Phys Rev B 2019;99:184403. https://doi.org/10.1103/PhysRevB.99.184403.

[11] MacNeill D, Stiehl GM, Guimaraes MHD, Buhrman RA, Park J, Ralph DC. Control of spin-orbit torques through crystal symmetry in WTe2/ferromagnet bilayers. Nat Phys 2016;13:300–5. https://doi.org/10.1038/nphys3933.

[12] Yu G, Upadhyaya P, Fan Y, Alzate JG, Jiang W, Wong KL, Takei S, Bender SA, Chang L-T, Jiang Y, Lang M, Tang J, Wang Y, Tserkovnyak Y, Amiri PK, Wang KL. Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields. Nat Nanotechnol 2014;9:548–54. https://doi.org/10.1038/nnano.2014.94.

[13] Honjo H, Nguyen TVA, Watanabe T, Nasuno T, Zhang C, Tanigawa T, et al., First demonstration of field-free SOT-MRAM with 0.35ns write speed and 70 thermal stability under 400 °C thermal tolerance by canted SOT structure and its advanced patterning/SOT channel technology. In: Proc IEDM Conf; 2019. p. 28.5.1–4. https://doi.org/10.1109/IEDM19573.2019.8993443.

[14] Garello K, Yasin F, Hody H, Couet S, Souriau L, Sharifi SH, et al., Manufacturable 300mm platform solution for field-free switching SOT-MRAM. In: Proc IEEE Symp VLSI Circ; 2019. p. T194–5. https://doi.org/10.23919/VLSIC.2019.8778100.

[15] Garello K, Yasin F, Couet S, Souriau L, Swerts J, Rao S, et al., SOT-MRAM 300mm integration for low power and ultrafast embedded memories. In: Proc IEEE Symp VLSI Circ; 2018. p. 81–2. https://doi.org/10.1109/VLSIC.2018.8502269.

[16] Grimaldi E, Krizakova V, Sala G, Yasin F, Couet S, Sankar Kar G, Garello K, Gambardella P. Single-shot dynamics of spin-orbit torque and spin transfer torque switching in three-terminal magnetic tunnel junctions. Nat Nanotechnol 2020;15: 111–7. https://doi.org/10.1038/s41565-019-0607-7.

[17] Wang Z, Cai W, Zhu D, Wang Z, Kan J, Zhao Z, Cao K, Wang Z, Zhang Y, Zhang T, Park C, Wang J-P, Fert A, Zhao W. Field-free switching of a perpendicular magnetic tunnel junction through the interplay of spin-orbit and spin-transfer torques. Nat Electron 2019;1:582–8. https://doi.org/10.1038/s41928-018-0160-7.

[18] Wang Z, Zhou H, Wang M, Cai W, Zhu D, Klein J, Zhao W. Proposal of toggle spin torques magnetic RAM for ultrafast computing. IEEE Electron Dev Lett 2019;40: 726–9. https://doi.org/10.1109/LED.2019.2907063.

[19] Inokuchi T, Yoda H, Koi K, Shimomura N, Ohsawa Y, Kato Y, Shirotori S, Shimizu M, Sugiyama H, Oikawa S, Altansargai B, Kurobe A. Real-time observation of fast and highly reliable magnetization switching in Voltage-Control Spintronics Memory (VoCSM). Appl Phys Lett 2019;114:192404. https://doi.org/10.1063/1.5097063.

[20] Kato Y, Saito Y, Yoda H, Inokuchi T, Shirotori S, Shimomura N, et al. Improvement of write efficiency in voltage-controlled spintronic memory by development of a Ta −B spin hall electrode. Phys Rev Appl 2018;10:044011. https://doi.org/10.1103/PhysRevApplied.10.044011.

[21] Makarov A, Windbacher T, Sverdlov V, Selberherr S. SOT-MRAM based on 1Transistor-1MTJ-cell structure. Technical digest of the non-volatile memory technology symposium. NVMTS; 2015. p. 105–6.

[22] Sverdlov V, Makarov A, Selberherr S. Two-pulse sub-ns switching scheme for advanced spin-orbit torque MRAM. Solid-State Electron 2019;155:49–56. https://doi.org/10.1016/j.sse.2019.03.010.

[23] de Orio RL, Makarov A, Selberherr S, Goes W, Ender J, Fiorentini S, Sverdlov V. Robust magnetic field-free switching of a perpendicularly magnetized free layer for SOT-MRAM. Solid-State Electron 2019:107730. https://doi.org/10.1016/j.sse.2019.107730.

[24] Mehta P, Bukov M, Wang CH, Day AG, Richardson C, Fisher CK, et al. A high-bias, low-variance introduction to machine learning for physicists. Phys Rep 2019;810: 1–124. https://doi.org/10.1016/j.physrep.2019.03.001.

[25] Sutton RS, Barto AG. Reinforcement learning: An introduction. 2nd ed. The MIT Press; 2018.

[26] Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 2018;362:1140–4. https://doi.org/10.1126/science.aar6404.

[27] Fösel T, Tighineanu P, Weiss T, Marquardt F. Reinforcement learning with neural networks for quantum feedback. Phys Rev X 2018;8:031084. https://doi.org/10.1103/PhysRevX.8.031084.

[28] de Orio RL, Makarov A, Goes W, Ender J, Fiorentini S, Sverdlov V. Two-pulse magnetic field-free switching scheme for perpendicular SOT-MRAM with a symmetric square free layer. Phys B 2020;578:411743. https://doi.org/10.1016/j.physb.2019.411743.

[29] Lee JM, Kwon JH, Ramaswamy R, Yoon J, Son J, Qiu X, Mishra R, Srivastava S, Cai K, Yang H. Oscillatory spin-orbit torque switching induced by field-like torques. Commun Phys 2018;1:1–7. https://doi.org/10.1038/s42005-017-0002-3.

[30] Takeuchi Y, Zhang C, Okada A, Sato H, Fukami S, Ohno H. Spin-orbit torques in high-resistivity-W/CoFeB/MgO. Appl Phys Lett 2018;112:192408. https://doi.org/10.1063/1.5027855.

[31] Makarov A. Modeling of emerging resistive switching based memory cells, Ph.D. thesis, Institute for Microelectronics, TU Wien, Vienna; 2014. https://www.iue.tuwien.ac.at/phd/makarov/.

[32] Raffin A, Hill A, Ernestus M, Gleave A, Kanervisto A, Dormann N. Stable Baselines3, https://github.com/DLR-RM/stable-baselines3; 2019.

[33] Mnih V, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. Nature 2015;518:529–33. https://doi.org/10.1038/nature14236.

Roberto Lacerda de Orio studied electrical engineering at the University of Campinas(UNICAMP), where he received a Bachelor degree in 2004 and a Master degree with emphasis on Electronics, Microelectronics and Optoelectronics in 2006. He joined the Institute for Microelectronics, Technische Universität Wien (TU Wien) in October 2006, where he received his doctoral degree in 2010 and worked as a post-doctoral researcher until 2013. From 2014 to 2018 he was Assistant Professor at the Faculty of Electrical and Computer Engineering of UNICAMP. In October 2018 he rejoined the Institute for Microelectronics, TU Wien as a research scientist on non-volatile magnetic memories and logic.