Solid State Electronics 155 (2019) 49–56

Contents lists available at ScienceDirect

Solid State Electronics

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

Two-pulse sub-ns switching scheme for advanced spin-orbit torque MRAM

Viktor Sverdlov[a,*], Alexander Makarov[b], Siegfried Selberherr[b]

[a] Christian Doppler Laboratory for Nonvolatile Magnetoresistive Memory and Logic at the Institute for Microelectronics, TU Wien, Gußhausstraße 27-29, A-1040 Wien, Austria

[b] Institute for Microelectronics, TU Wien, Gußhausstraße 27-29, A-1040 Wien, Austria

ARTICLE INFO

The review of this paper was arranged by Profs. S. Luryi, J. M. Xu, and A. Zaslavsky

Keywords:
Spin-orbit torque
MRAM
Switching
In-plane MTJ
Perpendicular magnetic anisotropy
Sub-500ps switching

ABSTRACT

The steady increase in performance and speed of modern integrated circuits is continuously supported by constant miniaturization of complementary metal-oxide semiconductor (CMOS) devices. However, a rapid growth of the dynamic and stand-by power due to transistor leakages becomes a pressing issue. A promising way to slow down this trend is to introduce non-volatility in circuits. The development of an electrically addressable non-volatile memory combining high speed and high endurance is essential to achieve these goals. To further reduce the energy consumption, it is essential to replace SRAM in modern hierarchical multi-level processor memory structures with a non-volatile memory technology. The spin-orbit torque magnetic random access memory (SOT-MRAM) combines non-volatility, high speed, and high endurance and is thus suitable for applications in caches. However, its development is still hindered by relatively high switching currents and the need of an external magnetic field for deterministic switching of perpendicularly magnetized layers. The switching by means of two orthogonal current pulses allows achieving deterministic sub-500 ps and magnetic field-free switching in perpendicularly magnetized rectangular structures. Complementing the two-pulse switching scheme with weak perpendicular interface-induced magnetic anisotropy reduces the switching current significantly for achieving sub-500 ps switching in in-plane structures.

1. Introduction

Continuous miniaturization of complementary metal-oxide semiconductor (CMOS) devices is behind the unprecedented increase of speed and performance of modern integrated circuits. Numerous outstanding technological challenges have been resolved on this exciting path. Among the most crucial technological achievements implemented by the semiconductor industry within the last 15 years to boosting CMOS performance while maintaining the gate control over the semiconductor channel are the introduction of strain [1], high-k gate dielectrics and metal gates [2], and three-dimensional (3D) tri-gate transistor architecture [3–5]. While chips with 5 nm technology based on nano-sheets are already nearing production [6], the semiconductor industry is now focusing on a 3 nm technology node. Although setting limits for scaling has proven to be a mere meaningless task in the past, it is obvious that the conventional transistor scaling is showing signs of saturation. To sustain the growing demand for high performance small area central processing units (CPUs) and high-capacity memory needed to handle an increasing information flow, an introduction of a disruptive technology employing new computing principles is anticipated. Most importantly, any emerging technology must be energy efficient as a harmful active power penalty already prevents the clock frequency from being increased in CMOS circuits. Indeed, clock frequencies for high-end consumer-level workstation CPUs have now saturated at approximately 3.7 GHz with the possibility to be boosted for a short time up to 4.2 GHz under heavy load. Although the transistor size is scaled down, the load capacitance value per unit area remained approximately unchanged. This prevents the on-current from decreasing while maintaining appropriate high speed operation due to an unavoidable charging of this capacitance.

Small transistor dimensions lead to rapidly increasing leakages. A rapid increase of the stand-by power due to transistor leakages at small transistor dimensions, as well as of the dynamic power and the need to refresh the data in dynamic random access memory (DRAM), is becoming a pressing issue. The microelectronics industry is facing major challenges related to power dissipation and energy consumption, and the microprocessors' scaling will soon hit a power wall.

An attractive path to mitigate the unfavorable trend of increasing power at stand-by is to introduce non-volatility in the circuits. The development of an electrically addressable non-volatile memory, which combines fast operation, simple structure, and high endurance, is essential to mitigate the increase of the stand-by power and to introduce

* Corresponding author.
E-mail addresses: Sverdlov@iue.tuwien.ac.at (V. Sverdlov), Makarov@iue.tuwien.ac.at (A. Makarov), Selberherr@iue.tuwien.ac.at (S. Selberherr).

https://doi.org/10.1016/j.sse.2019.03.010

Available online 04 March 2019

instant-on architectures without the need of data initialization when going from a stand-by to an operation regime. Oxide-based resistive RAM (RRAM) possesses filamentary switching between the on/off states and is thus intrinsically prone to significant resistance fluctuations in both states. In addition, the endurance reported is only slightly better than that of flash memory. Although RRAM possesses a simple structure and a large on/off current ratio, it is premature to consider RRAM at its current stage of development for digital applications. As continuous conductance modulation is suitable for implementing analog synaptic weights, both filamentary and non-filamentary switching RRAM types are currently intensively investigated, particularly for neuromorphic applications [7].

Spin-transfer torque magnetic RAM (STT-MRAM) is fast (10 ns), possesses high endurance (10^{12}), and has a simple structure. It is compatible with CMOS and can be straightforwardly embedded in circuits, e.g. [8]. It is particularly promising to employ non-volatility in internet of things (IoT) and automotive applications, as well as a replacement of conventional volatile CMOS-based DRAM and non-volatile flash. High-density STT-MRAM arrays with 4Gbit capacities have been already demonstrated [9]. The availability of high-capacity non-volatile memory close to high-performance CMOS circuits allows exploring conceptually new logic-in-memory [10] and computing-in-memory [11,12] architectures for future artificial intelligence and cognitive computing [13].

Although the use of STT-MRAM in last-level caches is conceivable [14], the switching current for operating at a speed faster than 10 ns is growing rapidly. This leads to large dimensions of the access transistor. The need of even higher switching currents for faster operation in higher-level caches potentially prevents STT-MRAM from entering in L2 and L1 caches currently mastered by static RAM (SRAM). In addition, rapidly increasing critical currents required for operating STT-MRAM at 5 ns result in large current densities running through magnetic tunnel junctions. This leads to oxide reliability issues, which in turn reduces the MRAM endurance to that of the flash memory, thus negating one of the important MRAM advantages over flash.

The engineering of an electrically addressable non-volatile memory combining high speed (sub-ns operation) and high endurance suitable for replacing SRAM in higher-level caches of hierarchical multi-level processor memory structures with a non-volatile memory [10] cannot be based on STT, and the use of new physical principles is required. Among the newly discovered physical phenomena suitable for next-generation MRAM is the spin-orbit torque (SOT) assisted switching at room temperature in heavy metal/ferromagnetic [15–22] or topological insulator/ferromagnetic [23–26] bilayers. In this memory cell the magnetic tunnel junction's (MTJ's) free layer is grown on a material with a large spin Hall angle. The SOT acting on the adjacent magnetic layer is generated by passing the current through this material. The relatively large switching current is injected in-plane along the heavy metal/ferromagnetic bilayer and does not flow through the MTJ, while a much smaller read current is applied through the MTJ. This results in a three-terminal configuration where the read and write current paths are decoupled. Since the large write current does not flow through the oxide in the MTJ, this prevents the tunnel barrier from damage. It improves device reliability by eliminating correlations between the switching current and the retention time. Therefore, three-terminal MRAM cells are promising candidates for future generations of non-volatile memory for fast sub-ns switching [21]. SOT-MRAM is an electrically addressable non-volatile memory combining high speed and high endurance and is thus suitable for applications in caches [21]. Although the high switching current is not flowing through a magnetic tunnel junction but rather through a heavy metal wire under it, the current is still high, and its reduction is the pressing issue in the field of SOT-MRAM development.

Topological insulators (TIs) are promising materials for reducing the switching current as they are characterized by the high spin Hall angle and the efficiency Θ_{SH} of charge to spin conversion due to the peculiar perpendicular spin-momentum locking in the interface states. In addition, the strong spin-orbit interfacial Rashba field helps generating spin density in TIs boosting the charge to spin conversion efficiency above 100%.

Although high charge to spin conversion efficiency in TIs has been reported, the electrical conductivity of TIs needed for practical applications to build a high-density, ultra-low power, and ultra-fast non-volatile memory was not sufficiently high because of then insulating bulk. Recent developments introduce BiSe [25] and BiSb [26] based TIs as suitable candidates for emerging SOT-MRAM as they possess a charge to spin conversion efficiency of 18.8 and 52 times, respectively. This allows to reduce the switching current by two orders of magnitude as compared to W-based SOT-MRAM. In addition, BiSb samples [26] possess very high electrical conductivity making thin BiSb films leading candidates for emerging fast and low-power SOT-MRAM, and the process integration of BiSb into a realistic MTJ stack is currently under scrutiny.

A proper integration of SOT-MRAM represents a significant challenge as the memory cells must withhold at least the back-end-of-line thermal budget. As IMEC presented recently a technology to integrate a perpendicular beta-phase W/CoFeB/MgO/CoFeB/synthetic antiferromagnetic stack based SOT-MRAM on a 300 mm CMOS wafer by only CMOS compatible processes [27], there is a cautious confidence that fast low-power non-volatile magnetoresistive memory suitable for processor caches will be developed soon.

However, despite an undisputable progress in developing SOT-MRAM, there is one important issue which has not been convincingly resolved so far. Namely, a static magnetic field is still required to guarantee deterministic switching [28] or a perpendicularly magnetized free layer. Several paths to achieve the deterministic switching without magnetic fields were suggested. They require unusual solutions to break the mirror symmetry either by means of the shape of the dielectric [29] or the free layer [30], or by controlling the crystal symmetry of the metal line at the microscopic layer [20]. Biasing the free layer by employing an exchange coupling to an antiferromagnet [31–34] as well as the use of a peculiar tapered sample shape, which controls the switching [35], were recently reported.

However, even if in most of these studies a field-free switching was reported, these methods either require a local intrusion into the fabrication process, or are based on solutions whose scalability is questionable (antiferromagnets, shapes), which makes further large scale integration of the fabricated memory cells problematic.

In this work we explore a switching scheme based on the use of two consecutive orthogonal sub-nanosecond current pulses. This scheme is suitable for integration in a cross-bar architecture [11]. Applied to in-plane structures [11], the switching scheme has been shown to accelerate the switching of the in-plane structure by a factor of three [36].

First, we investigate by means of extensive micromagnetic simulations the potential of the two-pulse switching scheme to reduce the switching currents while keeping the switching fast in in-plane magnetized layers. The introduction of relatively weak interface-induced perpendicular magnetic anisotropy further helps reducing the current for sub-0.5 ns switching.

Next, we apply the two-pulse switching scheme to application relevant free magnetic layers with perpendicular magnetization. Our simulations suggest that deterministic, fast, and external magnetic field-free switching of the perpendicularly magnetized layers of a rectangular shape is reliably achieved.

2. Method

The memory cell is shown in Fig. 1. It consists of an in-plane magnetized MTJ with its free layer (FL) lying on top of a heavy metal wire NM1. The magnetization of the fixed magnetic reference layer (RL) is pinned by the strong exchange coupling to a synthetic antiferromagnetic (SAF) stack (not shown). The SAF stack is usually

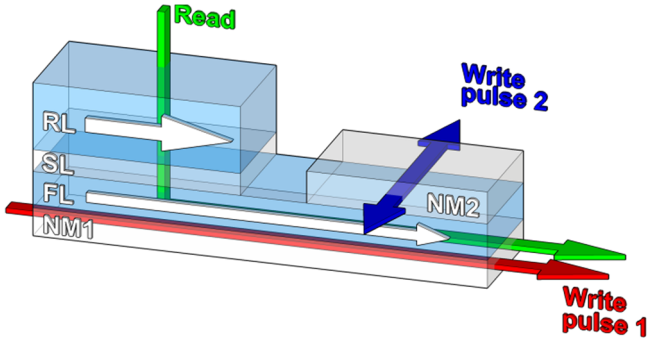

Fig 1. Scheme of the in-plane SOT-MRAM cell. The free layer of an MTJ is grown above the heavy metal wire NM1, through which the current pulse is applied. The wire NM2 serves to conduct the perpendicular current "Write pulse 2" of the in-plane SOT-MRAM cell.

designed to also compensate the influence of the unwanted stray magnetic field from RL on the switching dynamics of the FL. The RL is needed to read the MTJ resistance $R(\varphi)$, which depends on the relative angle φ between the RL and FL magnetization as

$$R(\varphi) = \frac{1}{2}(R_p(1 - \mathrm{TMR}\cos\varphi) + R_{AP}), \tag{1}$$

where $R_{P(AP)}$ is the MTJ resistance in the parallel (antiparallel) magnetization configuration, and TMR is the tunneling magnetoresistance ratio. The parameters and the dimensions of the free magnetic layer are chosen the same as for the perpendicular magnetic structure discussed in the next section. The free layer of dimension $a \times b \times d = 52.5 \times 12.5 \times 2\,\mathrm{nm}^3$ overlaps fully with the heavy metal line NM1 of 3 nm thickness, with the effective Hall angle Θ_{SH} describing the efficiency of the charge-to-spin conversion. Another heavy metal wire NM2 with an overlap from the right side lesser than the total free layer width serves to apply the second perpendicular current pulse and the spin-orbit torque associated with it. As the resistance of the MTJ depends on the FL magnetization orientation, the primary goal is to find the FL magnetization dynamics due to the spin-orbit torques generated by the currents through NM1 and NM2. The magnetization dynamics of the magnetic system due to the spin current densities and the spin accumulations is well described by the Landau-Lifshitz-Gilbert equation [21,37]

$$\frac{\partial \mathbf{m}}{\partial t} = -\gamma \mathbf{m} \times \mathbf{H}_{\mathrm{eff}} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \gamma \frac{\hbar}{2e} \frac{\Theta_{SH} J_1}{M_S d} [\mathbf{m} \times (\mathbf{m} \times \mathbf{y})] \Theta(t) \Theta(t_1 - t) - \gamma \frac{\hbar}{2e} \frac{\Theta_{SH} J_2}{M_S d} [\mathbf{m} \times (\mathbf{m} \times \mathbf{x})] \Theta(t - t_1) \Theta(t_2 + t_1 - t), \tag{2}$$

where $\mathbf{m}$ is the position-dependent magnetization $\mathbf{M}$ normalized by the saturation magnetization M_S: $\mathbf{m} = \mathbf{M}/M_S$, γ is the gyromagnetic ratio, α is the Gilbert damping parameter, e is the electron charge, ħ is the reduced Plank constant, and Θ_{SH} is an effective Hall angle relating the strength of the spin current density to the charge current density J_i (i = 1,2) passing through NM1 and NM2, $t_{1,(2)}$ is the duration of the first (second) pulse. The spin-orbit torques due to the charge currents through NM1 and NM2 represented by the third and fourth terms on the r.h.s. of (2) are acting on the free layer at the places of the overlaps of the FL with NM1 and NM2 (Fig. 1).

The spin current entering the free ferromagnetic layer is becoming quickly aligned with the local magnetization $\mathbf{M}$ generating the torque [38], which acts on the magnetization, as described by the last term in (2). This spin-orbit torque together with the magnetic field $\mathbf{H}_{\mathrm{eff}}$ describes the damped magnetization dynamics (2). The magnetic field $\mathbf{H}_{\mathrm{eff}}$ includes the external field as well as the contributions due to bulk and interface-induced magnetic anisotropies, exchange field, and demagnetization field, and the Ampere field due to the electric current. Thermal effects on the magnetization dynamics are incorporated by means of a random magnetic field added to $\mathbf{H}_{\mathrm{eff}}$. The strength of the thermal field fluctuations is related to temperature [39] set to 300 K. To describe the magnetization dynamics, we employ our in-house open-source tool [40] based on the finite difference discretization method. The tool has been carefully checked [41] to reproduce known examples and data from other simulators.

3. Two-pulse scheme for current reduction in in-plane magnetized free layer

We apply the two-pulse switching scheme previously proposed for fast switching of an in-plane structure [11] to demonstrate the switching current reduction. We assume $M_S = 4 \times 10^5$ A/m, and the exchange strength $A = 2 \times 10^{-11}$ J/m. The SOT due to the first pulse through the nonmagnetic heavy metal wire NM1 tilts the magnetization of the free layer in-plane perpendicular to the direction of the "Write pulse 1" (Fig. 1). We note that the "Write pulse 1" does not guarantee the deterministic magnetization switching as an external magnetic field is required to break the mirror symmetry and complete the switching [28].

In contrast, the SOT of the second pulse through the wire NM2 is able to switch the magnetization deterministically even if it is applied alone [28]. However, at switching by the "Write pulse 2" the SOT is not efficient at the beginning and the switching process possesses a long initial incubation period before the precessions begin. When the precessions start, the switching can be completed, if the torque overcomes the damping term. The damping value is about a hundredth and is thus small. This results in relatively small critical current density values. However, the switching takes long as many precessions are required before the magnetization flips. If one wants to switch faster, the current density increases rapidly inversely proportional to the switching time [8], and the schemes to reduce the current should be considered in case of fast switching by the "Write pulse 2" alone.

Fig. 2 shows the reduction of the current, which is achieved by supplying the "Write pulse 2" through the NM2 wire with a reduced overlap to the free layer, provided the current density is fixed. The pulse duration is set at 1.2 ns. A simple reduction of the overlap results in a smaller area of the free layer, where the SOT is acting. Surprisingly, the increase of the switching time is only a factor of two in the case

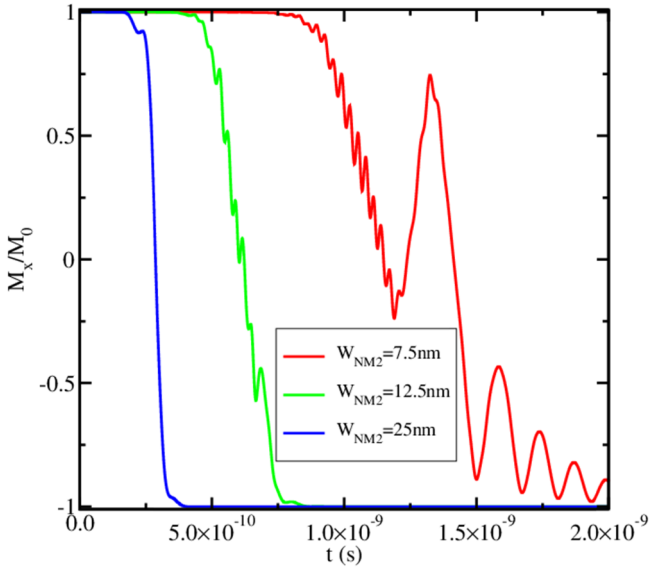

Fig 2. Time dependent magnetization evolution for several values of the second wire NM2 overlap. The current is scaled with NM2 and is equal to 180 μA for NM2 = 25 nm, 90 μA for NM2 = 12.5 nm, and 54 μA for NM2 = 7.5 nm. The current pulse duration through NM2 is 1.2 ns. The current through wire NM1 is zero.

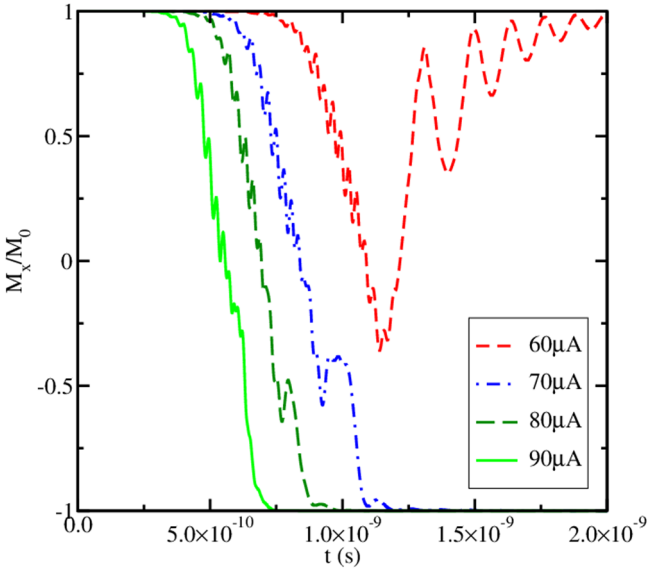

Fig 3. Time evolution of the magnetization for several values of the current through wire NM2 with the width 12.5 nm. The current pulse duration through NM2 is 1.2 ns. The current through wire NM1 is zero.

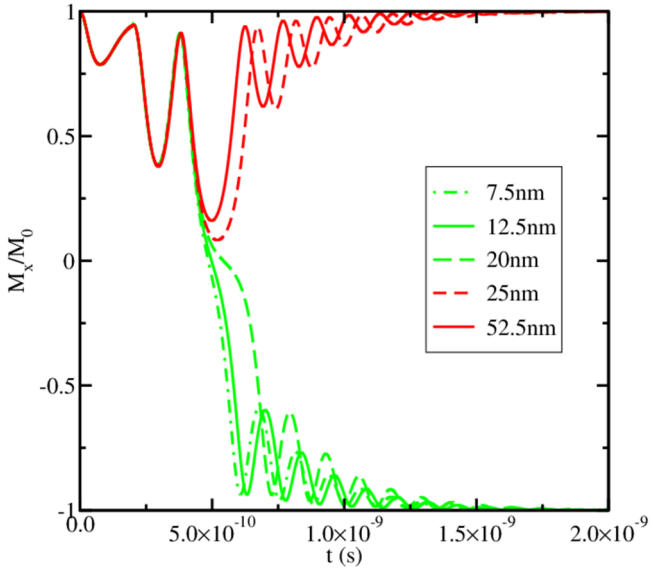

Fig 4. The magnetization dynamics for several overlap lengths of wire NM2. Two consecutive perpendicular pulses with 30 μA current and 200 ps duration are applied through NM1 and NM2.

when the NM2 is reduced from 25 nm to 12.5 nm [42], comparable to the current scaling with the switching times in STT-MRAM in fast precessional regime [8]. However, when the overlap is further decreased to 7.5 nm, the switching time increases more rapidly compared to the expected increase dictated by a geometrical scaling. This is due to the fact that the torque magnitude acting only on a small part of the free layer becomes insufficient to switch the whole layer efficiently, if the current pulse is 1.2 ns. A close inspection of the switching dynamics indicates the creation of a soliton-like spin wave which propagates through the FL further and back causing the oscillations of the average magnetization (Fig. 2, $W_{NM2} = 7.5$ nm).

The overlap of 12.5 nm of the NM2 wire, which is equal to the width of the NM1 wire, is sufficient to provide sub-ns switching at a current value of 90 μA. Fig. 3 shows the magnetization switching if the current value is further decreased, while the pulse duration is fixed at 1 ns. It appears that a switching current of 70 μA still switches the magnetization, while 60 μA is not sufficient to invert the magnetization within 1 ns. We also note that the current of 30 μA applied within 1 ns, which is about a factor of two weaker than the critical switching current, does not manage to develop any visible dynamics.

One of the reasons why the magnetization does not show any substantial dynamics at a current value of 30 μA is that the SOT due to the "Write pulse 2" is not efficient at the beginning when the magnetization is along the **x** direction parallel to the spin accumulation. Indeed, the corresponding term in (2) is close to zero. Therefore, although the maximal torque amplitude described by the coefficient in front of the cross product in the last term in (2) is sufficient to overcome the Gilbert damping and to invert the magnetization; the switching is characterized by a very long incubation time [28] necessary to develop a substantial deviation of the magnetization from the **x** axis. The initial spatial magnetization distribution is created during a 200 ps thermalization process under the influence of the random thermal magnetic field (Figs. 2, 3). Although at every point the magnetization is not perfectly aligned with the **x** axis and contains projections on the other axes **y** and **z**, for which the SOT is efficient, the projections on these axes averaged over the whole sample are close to zero due to the random character of the thermal field. As a development of a substantial deviation of the averaged projections from zero is more probable in a smaller sample area, the SOT for NM2 = 12.5 nm, albeit weaker than that for NM2 = 25 nm, is more efficient from the beginning due to this deviation. This explains the better scaling of the switching time with the current in Fig. 1 as compared to the one at STT switching and demonstrates the importance of making the SOT efficient from the beginning of switching for the reduction of the switching time.

In order to create an initial deviation of the average magnetization, we apply the two-pulse switching with the first current pulse running through the wire NM1 just before the second, switching, pulse through the wire NM2 [11]. We consider equivalent pulses of 200 ps duration [42]. The first pulse through NM1 creates an initial deviation of the magnetization from its equilibrium direction and tends to put the magnetization along the **y** axis. This makes the torque from the second pulse efficient from the beginning thus removing the incubation period. The two-pulse scheme allows reducing the switching current by a factor of 3 as compared to single pulse switching of a similar duration. Fig. 4 demonstrates that the optimal width of NM2 is around 12.5 nm as its further reduction results in a high current density. In Fig. 4 the magnitude of the currents of both pulses is fixed at 30 μA.

The purpose of the first pulse is to create an initial deviation of the magnetization to make the torque due to the second pulse running through NM2 effective. Therefore, applying the first pulse allows shortening or removing the long incubation period at switching by the pulse through NM2. The scheme is not sensitive to the current amplitude of the first current pulse. Indeed, if the current density in the first pulse is larger than the critical one, it will align the magnetization in-plane along the *y* axis [28]. However, the first pulse cannot complete the switching, and the external perpendicular magnetic field is required to make the switching deterministic [28]. In the suggested two-pulse scheme the switching is completed by applying the second pulse through NM2 of a corresponding polarity. The SOT due to the second pulse is efficient and completes the switching even if the current amplitude in the second pulse is small.

To further decrease the switching current, a perpendicular magnetic anisotropy typically developed at the MgO/CoFeB interface is introduced [42,43]. The perpendicular anisotropy compensates the large demagnetizing contribution when the magnetization is out of plane. However, the perpendicular anisotropy may compromise the temperature stability of the in-plane structure, when its value K is too large. Fig. 5 shows the time evolution of the magnetization projection on the **z** direction perpendicular to the structure, for several K, for a current fixed at 30 μA, while Fig. 6 displays a typical dependence of the in-plane magnetization along the easy axis. Our simulations predict that the effect of K on the switching process is not monotonic as the

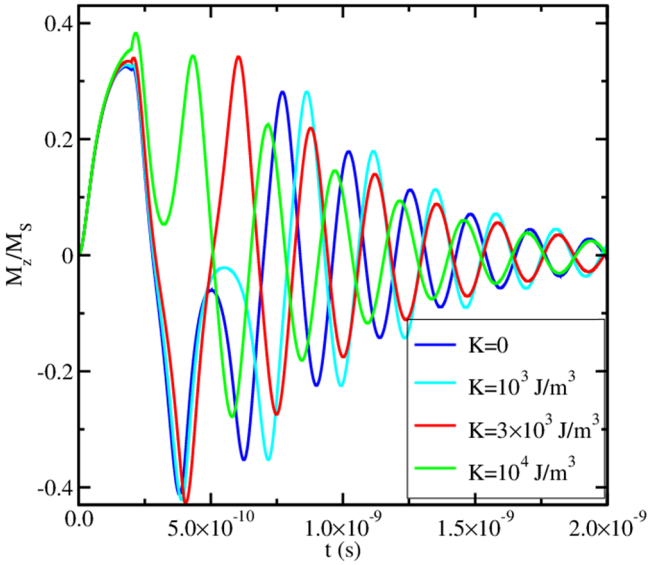

Fig 5. Time dependence of the magnetization projection on the perpendicular **z** axis, for several values of the interface-induced perpendicular anisotropy.

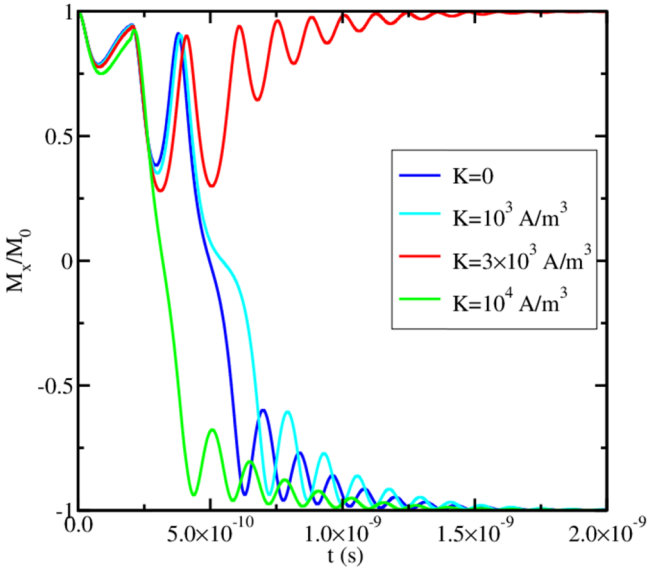

Fig 6. Adding perpendicular magnetization in a two-pulse scheme allows to reduce the current by a factor of 6 for sub-0.5 ns switching.

switching becomes longer at $K = 10^3 J/m^3$, while no switching is observed at an intermediate value $K = 3 \times 10^3 J/m^3$. Importantly, increasing the value of the interface-induced perpendicular anisotropy to $K = 10^4 J/m^3$ results in a fast, sub-500 ps switching. This value of the anisotropy is very close to a value at which the perpendicular magnetic anisotropy reduces the demagnetization energy of the magnetization to be out of plane to such an extend that it becomes equal to the demagnetization energy of the magnetization in-plane but along the **y** axis. Thus, this value is the maximum allowed for the particular geometry and dimensions of the free layer in Fig. 1 as its further increase compromises the cell thermal stability factor.

Summarizing this section, properly combining the two-pulse switching scheme [11] with the carefully tailored interface-induced magnetic anisotropy results in a six-fold current reduction as compared to the switching by a single current pulse, while still yielding fast sub 0.5 ns switching. The addition of the perpendicular anisotropy in an in-plane structure does not compromise the thermal stability, if the minimum energy barrier separating the two stable states is determined

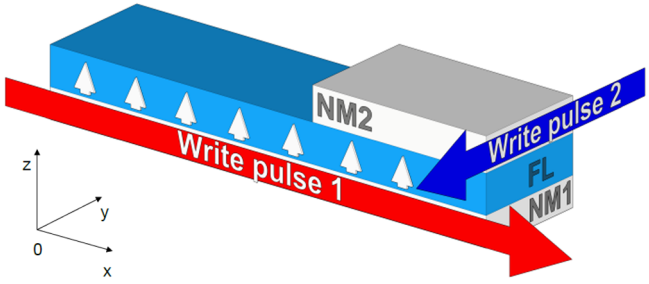

Fig 7. Perpendicular SOT-MRAM memory cell with a 52.5 nm × 12.5 nm × 2 nm free layer. After the 100 ps/100 μA current pulse the second perpendicular pulse is applied.

by the sample shape anisotropy. This implies that there exists a maximum value of the perpendicular anisotropy for a particular shape of the free in-plane magnetic layer, which should not be exceeded. However, when the perpendicular anisotropy becomes so large that it can bring the magnetization out-of-plane to be perpendicular to the free layer interface, the free magnetic layer becomes again promising for application [44]. The applicability and benefits of applying the two-pulse switching scheme in the latter case are described in the next section.

4. Two-pulse scheme for magnetic field free switching in perpendicularly magnetized free layer

We apply the two-pulse switching scheme to a perpendicularly magnetized structure (Fig. 7). It consists of a perpendicularly magnetized MTJ with its free layer lying on top of a heavy metal wire NM1 of 3 nm thickness. The dimensions of the free layer are $52.5 \times 12.5 \times 2\,nm^3$, the same as in the case of the in-plane structure [45]. The dimensions of NM1 are 3 nm thickness × 12.5 nm. A heavy metal wire NM2 with an overlap from the right side lesser or equal to the total free layer width of 52.5 nm serves to apply the second perpendicular current pulse and the spin-orbit torque associated with it.

We employ the parameters listed in Table 1. These parameters are relevant for the perpendicular CoFeB free layer on MgO tunnel barrier and result in a thermal stability factor defined as the ratio of the minimum energy barrier separating the two memory states at a temperature of around 40. Although this factor is too small for stand-alone memory applications, it is sufficient for employing such a memory cell in high level cashes, like SRAM, where the requirements for TMR and the stability are somewhat relaxed in favor of fast operation.

To realize the switching scheme, a 100 ps short and 100 μA strong current pulse is applied through NM1, Fig. 7, after a 100 ps initial waiting period. This current density equals $2.6 \times 10^{12}\,A/m^2$, which is similar to the critical current density $2 \times 10^{12}\,A/m^2$ reported in [28]. The consecutive perpendicular current pulse of varying strength and duration is applied through NM2. The magnetization dynamics of the magnetic system appears due to the charge current running through the heavy metal wire along the x axis and is described by the Landau-Lifshitz-Gilbert equation (2) with typical values for $M_S = 4 \times 10^5\,A/m$ and $\alpha = 0.05$ for perpendicular structures. Thermal effects at 300 K on the magnetization dynamics are described by a fluctuating magnetic

Table 1
Parameters used in simulations.

Name	Value
Saturation magnetization M_S	4×10^5 A/m
Exchange constant A	2×10^{-11} J/m
Perpendicular anisotropy K	2×10^5 J/m^3
Damping α	0.05
Spin Hall angle Θ_{SH}	0.3

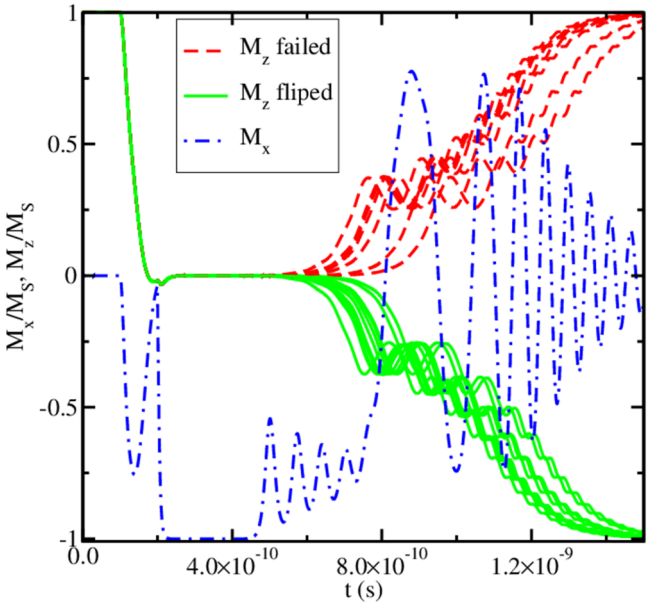

Fig 8. Time evolution of the magnetization in two-pulse switching (second pulse 1 mA/200 ps). 20 different switching realizations are obtained after 100 ps initial thermalization.

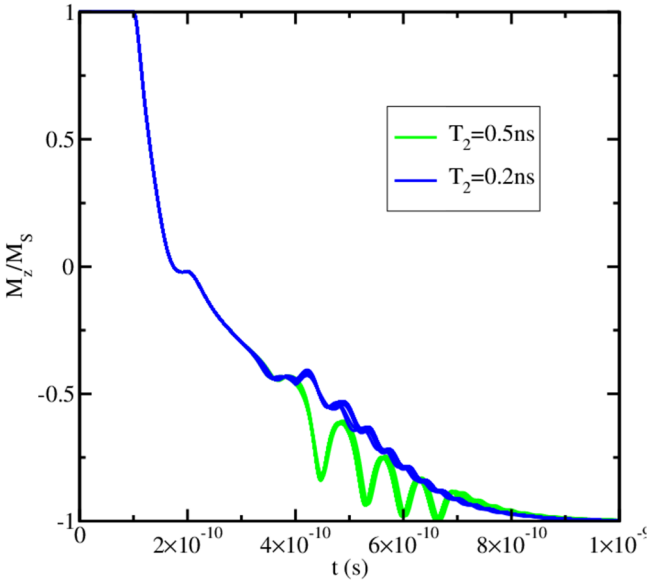

Fig 9. Same as in Fig. 8, but for the second pulse current of 200 μA, for two different durations of the second pulse. The 20 realizations switch evenly, almost without dispersion. The switching is deterministic.

field. Different realizations of the field produce diverse realizations of the switching process. Importantly, no external magnetic field is applied.

The time dependent magnetization dynamics for 20 realizations is shown in Fig. 8, where the current density in "Write pulse 2" of 200 ps is larger than the critical current applied to NM2 with the width 52.5 nm (equal to the free layer width). Although the current is high (1 mA), it does not provide a deterministic switching. It orients the magnetization in-plane against the **x** axis, but there is no torque along the **z** axis. Therefore, after the current is turned off, the magnetization randomly relaxes either in the same or in the opposite direction with equal probability because of the perpendicular magnetic anisotropy.

Unexpectedly, when the current is reduced to 200 μA and the current density to 1.25×10^{12} A/m^2 drops below its critical value, the switching becomes deterministic, for all 20 realizations (Fig. 9). In this case the torque due to "Write pulse 2" is not sufficiently strong to align the magnetization along **x** direction and it only tilts the magnetization away from the **y** direction. In this case the switching becomes deterministic as the tilted magnetization precesses about the shape effective anisotropy field along the **x** direction, which plays the role of the external field required for making the switching of perpendicular structures deterministic [28]. The polarity of the "Write pulse 2" therefore defines whether the magnetization relaxes up or down relative to the **z** axis. If the current of the "Write pulse 2" is reduced to the value of that in the "Write pulse 1", which results in a current density of 0.62×10^{12} A/m^2, switching becomes unreliable, as shown in Fig. 10 for several pulse durations.

However, if the width of NM2 is reduced to 12.5 nm (width of NM1), the switching becomes deterministic again (Fig. 11). In this case the magnetization below NM2 is quickly aligned along the **x** axis. Due to the exchange interaction, the rest of the free layer tilts towards the **x** direction and the anisotropy field acting on it completes the switching deterministically.

We note that the scheme works perfectly provided the "Write pulse 2" is short. As the current density in the second pulse is larger than the critical current density, we expect that the scheme may become less robust for longer pulses. Indeed, the magnetization of the free layer remains fixed along the **x** axis under NM2, while the magnetization of

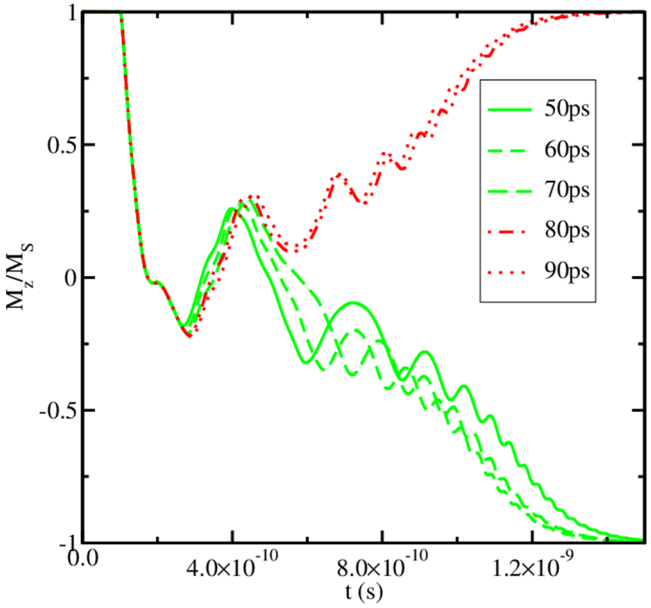

Fig 10. Randomly chosen switching realizations for the second pulse current of 100 μA, for several durations. The switching fails for longer pulses.

the rest of the free layer evolves, subject to the shape and perpendicular anisotropies, which may result in a complicated precessional motion resulting in occasional switching failures. We stress however, that this scenario may become realistic only, when the duration "Write pulse 2" is longer than that considered here.

Fig. 12 shows the dependence of the switching time on the NM2 width, for several durations of "Write pulse 2". It appears that the fastest reliable deterministic switching is achieved when the NM2 width is about the width of the free layer which coincides with the NM1 width. Most importantly, for these widths the switching is not only fast, but also not very sensitive to the width and pulse duration fluctuations, which relaxes the restrictions for layout and pulse timing.

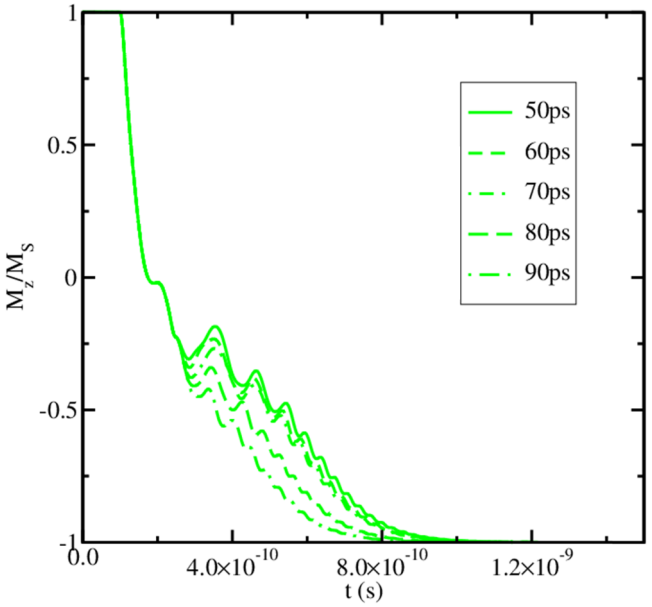

Fig 11. Same as in Fig. 10, but the width of NM2 is 12.5 nm (partial overlap). For equal wires NM1 and NM2 and equal currents the switching becomes fast and deterministic.

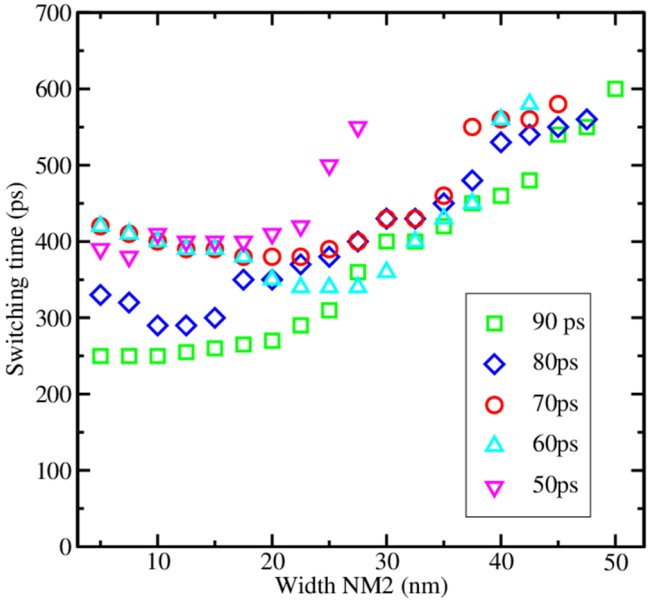

Fig 12. The switching time as a function of the NM2 width. The NM2 width of 12.5 nm is optimal as it guarantees fast, robust, and deterministic switching which is insensitive to small variations of the pulse duration or NM2 width.

5. Conclusion

The two-pulse switching scheme is demonstrated suitable for switching both in-plane and perpendicularly magnetized free flat magnetic layers of rectangular shape by means of spin-orbit torques. A partial overlap of the second wire is efficient in reducing the switching current in the in-plane structure keeping it fast. Adding an interface-induced perpendicular magnetization further improves the speed of switching. As compared to the single pulse switching, the introduction of a perpendicular anisotropy into the two-pulse switching scheme allows reducing the switching current by a factor six, while preserving fast sub-500 ps writing suitable for cache applications. A proper engineering of the interface-induced anisotropy is required as its addition does not necessarily improve the switching characteristics. In addition, a free layer shape-dependent maximum value of the perpendicular anisotropy is existing as the increase of the anisotropy beyond this value compromises the layer's thermal stability.

For perpendicularly magnetized structures, the spin-orbit torque due to the first 100 ps pulse tilts the magnetization of the free layer in-plane perpendicular to the "Write pulse 1" direction. The spin-orbit torque of the second consecutive current "Write pulse 2" results in an additional precession of the magnetization in the part of the free layer under it, which is transferred to the remaining part of the free layer through the exchange interaction. Depending on the direction of this precession, the magnetization of the remaining part tilts up or down with respect to its in-plane orientation set after the "Write pulse 1". The part under the NM2 wire follows the precession after the current is turned off, thus completing the switching. Results of the switching time calculations for several pulse durations as a function of the width of the second pulse wire demonstrate that the fastest, sub-300 ps, and 100% reliable, and magnetic field-free switching is achieved at around 30% overlap of the second pulse wire NM2 with the free layer.

Acknowledgement

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sse.2019.03.010.

References

[1] Thompson S-E, Armstrong M, Auth C, et al. A 90-nm Logic technology featuring strained-silicon. IEEE Trans Electron Devices 2004;51:1790–7.
[2] Mistry K, Allen C, Auth C, et al. A 45nm Logic Technology with High-k + Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-Free Packaging. IEDM Techn. Digest; 2007. p. 247–50.
[3] Natarajan S, Armstrong M, Bost M, et al. A 32nm logic technology featuring 2nd-generation high-k + metal-gate transistors, enhanced channel strain and 0.171μm2 SRAM cell size in a 291Mb array. IEDM Techn. Digest; 2008. p. 941–3.
[4] Xie R, Montanini P, Akarvardar K, et al. A 7nm FinFET technology featuring EUV patterning and dual strained high mobility channels. IEDM Techn. Digest; 2016. p. 47–50.
[5] Wu S-Y, Lin CY, Chiang MC. et al. 7nm CMOS platform technology featuring 4th generation FinFET transistors with a 0.027um2 high density 6-T SRAM cell for mobile SoC applications, IEDM Techn. Digest; 2016. p. 43–6.
[6] Loubet N, Hook T, Montanini P, et al. Stacked Nanosheet gate-all-around transistor to enable scaling beyond FinFET. 2017 Symp. VLSI Technology and Circuits: 2017. T230.
[7] Burr JGW, Shelby RM, Sebastian A, et al. Neuromorphic computing using non-volatile memory. Adv Phys: X 2017;2:89.
[8] Apalkov D, Dieny B, Slaughter JM. Magnetoresistive random access memory. Proc IEEE 2016;104:1796.
[9] Chung S-W, Kishi T, Park JW, et al.4Gbit Density STT-MRAM using perpendicular MTJ realized with compact cell structure, IEDM 2016 Techn. Digest; 2016. p. 660.
[10] Hanyu T, Endoh T, Suzuki D, et al. Standby-power-free integrated circuits using MTJ-based VLSI computing. Proc IEEE 2016;104:1844.
[11] Makarov A, Windbacher T, Sverdlov V, Selberherr S. CMOS-compatible spintronic devices: a review. Semicond Sci Technol 2016;31:113006.
[12] Jaiswal A, Agrawal A, Roy K. In-situ, in-memory stateful vector logic operations based on voltage controlled magnetic anisotropy. Sci Rep 2018;8:5738.
[13] Ielmini D, Wong H-SP. In-memory computing with resistive switching devices. Nat Electron 2018;1:333.
[14] Jan G, Thomas L, Le S, et al. Achieving sub-ns switching of STT-MRAM for future embedded LLC applications through improvement of nucleation and propagation switching mechanisms. In: Proc. of the 2016 Symp. VLSI Technology and Circuits; 2016. p. 18.
[15] Miron IM, Garello K, Gaudin G, et al. Perpendicular switching of a single ferromagnetic layer induced by in-plane current injection. Nature 2011;476:189.
[16] Liu L, Lee J, Gudmundsen TJ, et al. Current-induced switching of perpendicularly magnetized magnetic layers using spin torque from the spin hall effect. Phys Rev Lett 2012;109:096602.
[17] Liu L, Pai C-F, Li Y, et al. Spin-torque switching with the giant spin Hall effect of tantalum. Science 2012;336:555.

[18] Brataas A, Hals KMD. Spin–orbit torques in action. Nat Nanotechnol 2014;9:86.
[19] Taniguchi T, Grollier J, Stiles MD. Spin-transfer torques generated by the anomalous hall effect and anisotropic magnetoresistance. Phys Rev Appl 2015;3:044001.
[20] MacNeil D, Stiehl GM, Guimaraes MHD, et al. Control of spin–orbit torques through crystal symmetry in WTe_2/ferromagnet bilayers. Nat Phys 2017;13:300.
[21] Lee S-W, Lee K-J. Emerging three-terminal magnetic memory devices. Proc IEEE 2016;104:1831.
[22] Demasius KU, Phung T, Zhang W, et al. Enhanced spin–orbit torques by oxygen incorporation in tungsten films. Nat Commun 2016;7:10644.
[23] Han J, Richardella A, Siddiqui SA, et al. Room-temperature spin-orbit torque switching induced by a topological insulator. Phys Rev Lett 2017;119:077702.
[24] Wang Y, Zhu D, Wu Y, et al. Room temperature magnetization switching in topological insulator-ferromagnet heterostructures by spin-orbit torques. Nat Commun 2017;8:1364.
[25] Mahendr DC, Grassi R, Chen J-Y, et al. Room-temperature high spin–orbit torque due to quantum confinement in sputtered $Bi_xSe_{(1-x)}$ films. Nat Mater 2018. https://doi.org/10.1038/s41563-018-0136-z.
[26] Huynh N, Khang D, Ueda Y, Hai PN. A conductive topological insulator with large spin hall effect for ultralow power spin–orbit torque switching. Nat Mater 2018. https://doi.org/10.1038/s41563-018-0137-y.
[27] Garello K, Yasin F, Couet S, et al. SOT-MRAM 300mm integration for low power and ultrafast embedded memories. In: 2018 Symposia on VLSI Technology and Circuits; 2018. C8-2.
[28] Fukami S, Anekawa T, Zhan C, Ohno H. A spin–orbit torque switching scheme with collinear magnetic easy axis and current configuration. Nat Nanotechnol 2016;11:621.
[29] Yu G, Upadhyaya P, Fanet Y, et al. Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields. Nat Nanotechnol 2014;9:548–54.
[30] Yu G, Chang L-T, Akyol M, et al. Current-driven perpendicular magnetization switching in Ta/CoFeB/[TaO_x or MgO/TaO_x] films with lateral structural asymmetry. Appl Phys Lett 2014;105:102411.
[31] Fukami S, Zhang C, DuttaGupta S, et al. Magnetization switching by spin–orbit torque in an antiferromagnet–ferromagnet bilayer system. Nat Mater 2016;15:535–41.
[32] van den Brink A, Vermijs G, Solignac A, et al. Field-free magnetization reversal by spin-hall effect and exchange bias. Nat Commun 2016;7:10854.
[33] Lau Y-C, Betto D, Rode K, et al. Spin–orbit torque switching without an external field using interlayer exchange coupling. Nat Nanotechnol 2016;11:758–62.
[34] Oh Y-W, Baek S-HC, Kim YM, et al. Field-free switching of perpendicular magnetization through spin–orbit torque in antiferromagnet/ferromagnet/oxide structures. Nat Nanotechnol 2016;11:878–84.
[35] Safeer CK, Jué E, Lopez A, et al. Spin-orbit torque magnetization switching controlled by geometry. Nat Nanotechnol 2016;11:143–6.
[36] Makarov A, Windbacher T, Sverdlov V, Selberherr S. SOT-MRAM based on 1transistor-1MTJ-cell structure. 2015 15th Non-Volatile Memory Technology Symposium (NVMTS); 2015. p. 1–4.
[37] Gilbert T. A phenomenological theory of damping in ferromagnetic materials. IEEE Trans Magn 2004;40:3443.
[38] Kent AD, Ozylimaz B, del Barco E. Spin-transfer-induced precessional magnetization reversal. Appl Phys Lett 2004;84:3897.
[39] Finocchio G, Carpentieri M, Azzerboni B, et al. Micromagnetic simulations of nanosecond magnetization reversal processes in magnetic nanopillar. J Appl Phys 2006;99:08G522.
[40] VIENNAMAG 2016. www.iue.tuwien.ac.at/index.php?id=24.
[41] Makarov A. Modeling of emerging resistive switching based memory cells. PhD Thesis; 2014. www.iue.tuwien.ac.at/phd/makarov/.
[42] Sverdlov V, Makarov A, Selberherr S. Switching current reduction in advanced spin-orbit torque MRAM. Book of abstracts of the 2018 joint international EUROSOI workshop and international conference on ultimate integration on silicon; 2018. p. 57–8.
[43] Liu L, Moriyama T, Ralph DC, Buhrman RA. Reduction of the spin-torque critical current by partially canceling the free layer demagnetization field. Appl Phys Lett 2009;94:122508.
[44] Ikeda S, Miura K, Yamamoto H, et al. A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction. Nat Mater 2010;9:721–4.
[45] Makarov A, Sverdlov V, Selberherr S. Two-pulse sub-ns switching of a perpendicular spin-orbit torque MRAM cell without external magnetic field. In: Abstracts Advanced Research Workshop Future Trends in Microelectronics: Vingt Ans Après; 2018. p. 51.

Viktor Sverdlov received his Master of Science and PhD degrees in physics from the State University of St.Petersburg, Russia, in 1985 and 1989, respectively. From 1989 to 1999 he worked as a staff research scientist at the V.A. Fock Institute of Physics, St. Petersburg State University. During this time, he visited ICTP (Italy, 1993), the University of Geneva (Switzerland, 1993–1994), the University of Oulu (Finland, 1995), the Helsinki University of Technology (Finland, 1996, 1998), the Free University of Berlin (Germany, 1997), and NORDITA (Denmark, 1998). In 1999, he became a staff research scientist at the State University of New York at Stony Brook. He joined the Institute for Microelectronics, *TU Wien*, in 2004. His scientific interests include device simulations, computational physics, solid-state physics, and nanoelectronics.