nature nanotechnology

Supplementary information

https://doi.org/10.1038/s41565-024-01788-x

Energy-efficient picosecond spin–orbit torque magnetization switching in ferro- and ferrimagnetic films

In the format provided by the authors and unedited

Contents

S1. Detection and calibration of ps-wide pulses in a coplanar waveguide

The technique used to detect and calibrate ps-wide electrical pulses propagating in a coplanar waveguide, reported in a previous work [1], is described here in the context of our present work.

The ps-pulses are generated and detected with a pump-probe experiment. First, the biased generator switch is discharged by exciting it with the laser pump. The electrical pulse is then transmitted through the waveguide and is detected by a second photoconductive switch, called the *detector*, where the probe is focused (see Fig. 1a in main text). The electric field of the passing pulse serves as a transient bias voltage in the detector and, as a consequence, when the probe laser is on, the switch will be transiently closed. As a result, a current proportional to the pulse's electric field will flow through the detector switch. Following this principle, the pulse profile can be temporally-resolved by measuring this current as a function of the delay between pump and probe.

The time-dependent current is measured using two lock-in amplifiers in a double-modulation scheme. The probe is modulated at 40 kHz by a photo-elastic modulator between two crossed polarizers, and the pump is modulated at 350 Hz by a mechanical chopper. A typical measurement is depicted in Fig. S1a, in raw reading units of current.

These ps-pulses can be calibrated into real amplitude units by sampling ns-pulses of known amplitude from a pulse generator. The ns-pulses are injected into the waveguide by moving the contact tip bypassing the generator switch (see Fig. 1b in the main text), and are sampled using the same detection principle as for the ps-pulses in a pump-probe experiment. The pump is, in this case, used to start a trigger chain consisting of a fast Si photodiode, a delay generator, and a pulse generator. The usage of a delay generator is necessary because the combined response time of the photodiode and pulse generator is larger than the time window accessible with our optical delay line (∼13 ns), thus making it impossible to synchronize pump and probe. The addition of the delay generator allows us to fall within the delay line time window and perform the pump-probe experiment by *synchronizing the **n** pump trigger with the **n+1** probe pulse.* A typical measurement of a ns-pulse is shown in Fig. S1b in raw reading units. By measuring ns-pulses of different amplitudes we find that the reading scales proportionately with the real voltage amplitude, as is shown in Fig S1c.

When measuring pulses generated at different photoswitch bias, we have consistently observed that the read currents grow proportionately with bias, whatever the laser fluence. We depict this proportionality in Fig. S2a, together with two references curves from previous measurements on previous (different) devices. If we normalize pulses obtained at different bias (keeping all other parameters identical), we find that increasing bias does not change the shape of the pulse (Fig. S2b), which means that pulse width is conserved. In order to achieve SOT switching, it is normally necessary to set the photoswitch bias at a several tens of V. By sampling such pulses, there is a high chance that the photoswitch or the magnet burns, due to the high electric fields in continuous mode. The relations shown in Fig. S2 permit us to conclude that for a given fluence, it is sufficient to sample a ps-pulse at low bias (a few V) and then scale the measured trace to a higher bias, thus safeguarding the device from breakdown.

We have also studied the dependence on the laser pump fluence (Fig. S3a). By exciting with a fluence of ∼0.4 mJ/cm^2, the resulting pulse is symmetrical, with a Gaussian profile. However, as we increase the laser pump fluence (from 1 to 15 mJ/cm^2), the pulses start exhibiting a long tail. This effect has been attributed in literature to trap saturation in LT-GaAs, which limits its spectral response to sub-THz frequencies [2] and produces longer electrical discharges. Taking advantage of this effect, we are able to cover the pulse duration range from ∼7 ps to ∼70 ps (Fig. S3b). While longer pulses should be achievable by further increasing the fluence, there is a high risk that such fluences will end up burning the photoswitch during the pump-probe measurement of the pulse. This issue might be overcome in the future by optimizing the switch capacitance and/or the photoactive layer.

Finally, when calibrating ps-pulses, some error sources must be considered:

1. The ns-pulse pump-probe measurement (Fig. S1b) presents some electronic noise picked up before lock-in amplification of the very small (~pA) signals. In the present case, by averaging at the pulse plateau, we find that the real input voltage of 1.15 V is equivalent to 17.6 ± 0.8 pA. This translates into an error of ~5% in the calibration.
2. As shown in Fig. S1c, I_{read} and V_{real} are directly proportional. This relation can be described by a linear fit: $I_{\mathrm{read}}(V_{\mathrm{real}}) = (14.8 \pm 0.2) \cdot V_{\mathrm{real}}$. By replacing the real amplitude of the ns-pulse of Fig. S1b (1.15 V), the linear fit returns a current value of 17 ± 0.2 pA. This means that scaling our measurements proportionately introduces an error of ~1% in the calibration.
3. The probe positioning over the detector switch can drift over time. In general, before any measurement, the probe is positioned by maximizing the measured current. If the probe drifts over time, the measured current for the same electrical signal will depend on the time passed since the probe was optimized. Therefore, it is advisable to do the calibration immediately after the ps-pulse sampling in order to keep the measuring conditions as similar as possible between the two measurements. Because of this, and in spite of our efforts to keep the probe optimized, we allow an error of ~10% in the calibration of our signals.

From this analysis, we conclude that the errors associated to noise limit and linear scaling are largely contained within the error induced by the probe drifting with time. Henceforth, we perform the calibration without any error propagation and then add a 10% error in the resulting voltage. An example of calibrated pulses can be seen in Fig. S3a.

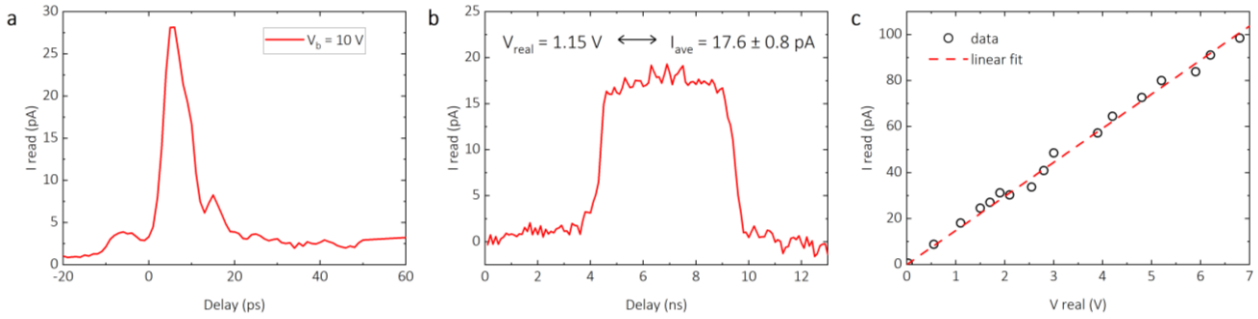


Figure S1. a. Example of detection of a ps-pulse generated with the photoswitch biased with 10 V, in raw reading units. **b.** ns-pulse of real amplitude 1.15 V, detected with the same technique, in raw current units. **c.** Calibration curve made by sampling ns-pulses of different amplitudes. The relationship between read current and real voltage is linear.

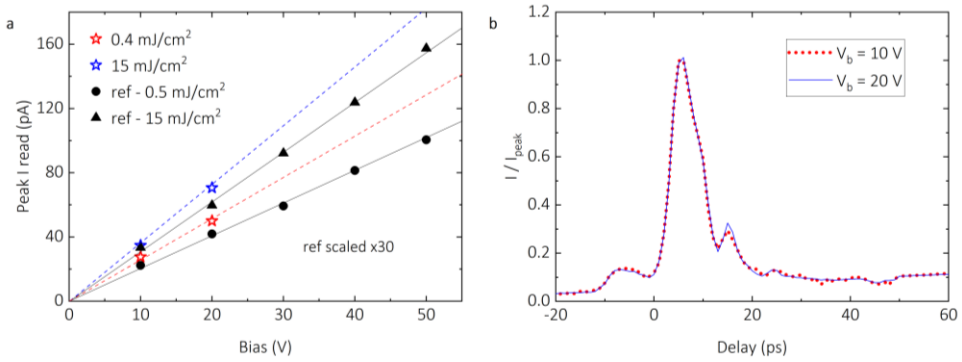


Figure S2. a. Peak intensity as a function of photoswitch bias, for the present samples and for previous (reference) ones, at different pump laser fluences. The reference samples peak values are scaled for visualization purposes. **b.** Pulses at different bias normalized by their respective peak value. The pulse shape is conserved.

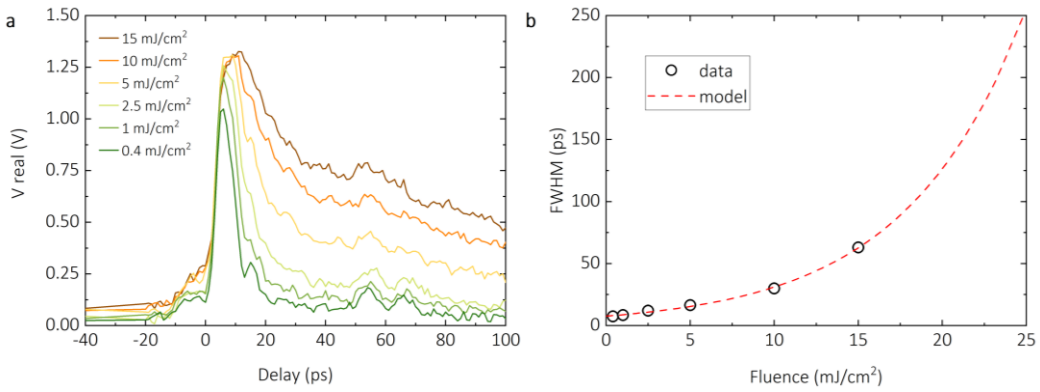


Figure S3. a. Pulses in real amplitude units as a function of pump fluence. The same bias tension (5 V) was applied for all fluences. **b.** Pulse width (FWHM) as a function of pump fluence.

S2. Determination of the reflection coefficient

The reflection coefficient (Γ) of all samples was determined using a Rhode & Schwarz® RTO2044 4-GHz oscilloscope. The device was contacted by a 40-GHz GSG probe bypassing the photoconductive switch (as shown in Fig. 1b of the main text) and connected to a pulse generator. A single 1-ns pulse was sent. Before arriving to the device, the pulse passed through the oscilloscope configured at an internal resistance of 1 MΩ. In this way, the pulse could be traced by the instrument while ensuring minimal power absorption. We note that the 4-GHz scope has a limited bandwidth when operated in the 1 MΩ input-impedance setting, which results in a smoothing of the pulses, as shown in Fig. S4. In reality, however, the pulses contain some ringing with faster frequencies.

The pulse was then injected into the magnetic sample and, due to the slight impedance mismatch with the waveguide (50 Ω), a portion of it was reflected back. The reflection was also read by the oscilloscope without significant power absorption. The incident and reflected pulses for all samples are shown in Fig. S4. The reflection coefficient Γ was calculated as the ratio between the reflected and incident amplitudes.

The reflection coefficient was also confirmed for ps pulses measured by a pump-probe experiment, yielding similar results. The incident pulse (**1st peak**) and two reflected pulses (**2nd and 3rd peak**) are depicted in Fig. S5 for all three samples. The first reflection (**2nd peak**) comes from the impedance mismatch at the magnetic sample, while the second reflection (**3rd peak**) comes from the first reflected pulse, which bounces at the photoswitch and is later detected (see inset Fig. S4a). The photoswitch acts as an infinite impedance element, hence why the 2nd and 3rd peak are almost identical in amplitude.

As a final check, we measured the resistivity of the Co sample to be $117 \cdot 10^{-8}$ Ω·m. From this value, we can estimate the resistance of our device to be $R_{Co} = 65 \pm 5\ \Omega$. The expected reflection coefficient for an impedance change from a $Z_0 = 50\ \Omega$ waveguide to a sample of R_{Co}, yields $\Gamma = (Z_0 - R_{Co})/(Z_0 + R_{Co}) \sim 0.13 \pm 0.035$, once again confirming our estimations from Fig. S4a.

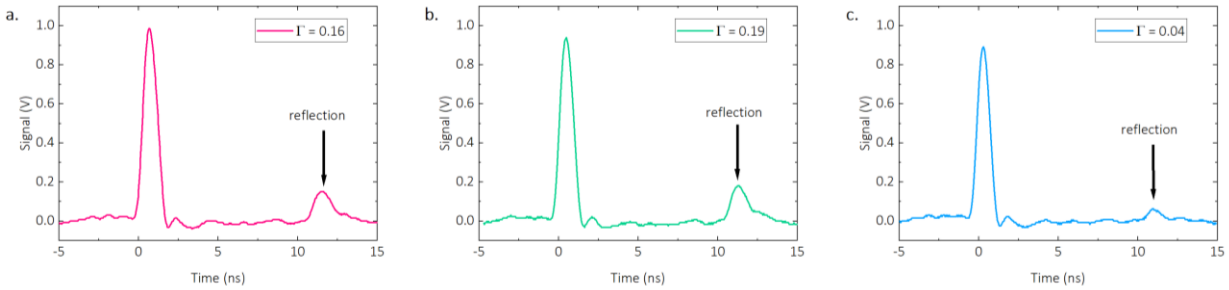


Figure S4. Incident and reflected ns pulses for **a.** Co, **b.** CoFeB, and **c.** CoGd. The reflection coefficient value is indicated in the legend.

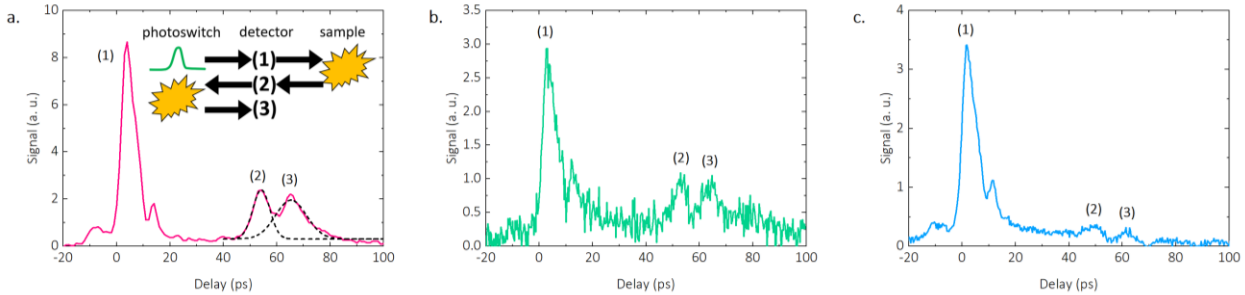


Figure S5. Incident and reflected ps pulses for **a.** Co, **b.** CoFeB, and **c.** CoGd.

S3. Determination of critical current (j_c) and energy (E)

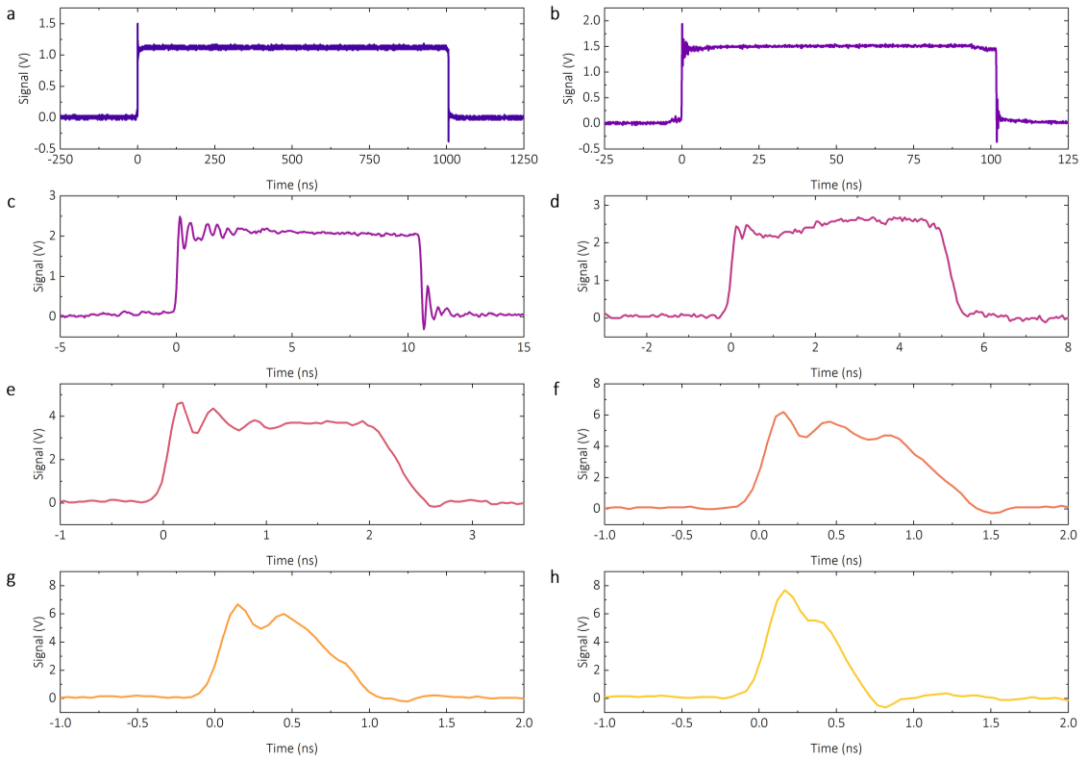


Figure S6. Shape of the electrical pulses from a commercial pulse generator, as observed in a 4-GHz oscilloscope. Nominal pulse width: **a.** 1000 ns, **b.** 100 ns, **c.** 10 ns, **d.** 5 ns, **e.** 2 ns, **f.** 1 ns, **g.** 0.75 ns, and **h.** 0.5 ns.

S3a. Pulses with duration > 0.5 ns

Pulses between 0.5 and 1000 ns were generated using an AVPP-2 series pulse generator from AVTECH® with a rise time of 200 ps, and voltages up to 20 V. In Fig. S6 we show the $V(t)$ trace of pulses of several durations detected with a 4-GHz oscilloscope. As pulse duration goes down to the ns and below, that is, to the limit of the equipment, the rising time and overshoot oscillations are made more evident. In order to determine the critical voltage and current associated to each pulse, we assume that they are square shaped. The error bar associated to each is therefore larger as we reduce the pulse duration and deviate from the square shape.

We take the trace $V(t)$ and define pulse duration (τ_p) as the full width at its half-maximum (FWHM). To define the critical voltage, we calculate the area (A) under the curve $V(t)$ and define the peak value as:

$$\mathrm{A} = \int V(t)dt \quad \rightarrow \quad V_{peak} = A/\tau_p$$

To assign the error to V_{peak} we calculate the absolute value of the deviation between $V(t)$ and V_{peak} for all the values within the pulse duration, and then average it:

$$\mathrm{diff}(t) = \mathrm{abs}(V(t) - V_{\mathrm{peak}})$$

$$\Longrightarrow \Delta V_{\mathrm{peak}} = \overline{\mathrm{diff}(t)}$$

With this, we can define the critical voltage and current at the sample as:

$$V_c = (1 + \Gamma) \cdot (V_{\text{peak}} \pm \Delta V_{\text{peak}})$$

$$I_c = \frac{(1 - \Gamma)}{Z_0} \cdot (V_{\text{peak}} \pm \Delta V_{\text{peak}})$$

Where Γ is the reflection coefficient as calculated in Suppl. Mat. S2 and Z_0 is the impedance of the line (50 Ω). With this definition, for a perfectly impedance-matched line ($R = 50\ \Omega$, $\Gamma = 0$) the current is completely absorbed by the sample, while for an open circuit ($R \to \infty$, $\Gamma = 1$) the current in the sample is zero because it is completely reflected.

The critical current density can be calculated as:

$$j_c = \frac{n \cdot I_c}{wt} = \frac{n(1 - \Gamma)}{Z_0 wt} \cdot (V_{\text{peak}} \pm \Delta V_{\text{peak}})$$

Where n is the current percentage that flows through the heavy metal layer and will experience charge-to-spin conversion (estimated in section Suppl. Mat. S8), w is the width of the sample, and t is the heavy-metal layer thickness.

To determine the energy cost of switching, we calculate the power dissipated by the sample (P) as:

$$P = V \cdot I \longrightarrow P(t) = V(t)(1 + \Gamma) \cdot \frac{V(t)(1 - \Gamma)}{Z_0} = \frac{V^2(t)(1 - \Gamma^2)}{Z_0}$$

And we determine the energy (E) as the integral over time of the dissipated power:

$$E = \int P(t)\, dt = \frac{(1 - \Gamma^2)}{Z_0} \int V^2(t) dt$$

S3b. Pulses with duration < 500 ps

For pulses shorter than 500 ps, we use the biased photoswitch excited with a laser pulse. Using the calibration method described in Suppl. Mat. S1, we are able to obtain a trace $V(t)$ of the pulse at the critical bias from a pump-probe measurement. From these traces we define pulse duration and its associated error as:

$$\tau_p = \tau_{\text{FWHM}} \pm \tau_{\text{step}}$$

Where τ_{FWHM} is the full-width at half-maximum pulse duration obtained from $V(t)$ and τ_{step} is the time step of the pump-probe measurement. To calculate the critical switching values, we take the peak value of the $V(t)$ trace:

$$V_{\text{peak}} = \max\big(V(t)\big)$$

As explained in Suppl. Mat. S1, an error of 10 % is estimated in the calibration of the $V(t)$ curves. Therefore, we define the error in V_{peak} as:

$$\Delta V_{\text{peak}} = 0.1 \cdot V_{\text{peak}}$$

The critical tension and current are, as before, defined as:

$$V_c = (1 + \Gamma) \cdot (V_{\text{peak}} \pm \Delta V_{\text{peak}})$$

$$I_c = \frac{(1 - \Gamma)}{Z_0} \cdot (V_{\text{peak}} \pm \Delta V_{\text{peak}})$$

Where we have considered the non-perfect absorption of the pulse by the sample. The energy cost of switching is also calculated as before, by integrating the $P(t)$ curve, but due to the error in the calibration, there will be a propagated error [3, p. 33] into the dissipated power:

$$P = \frac{(1-\Gamma^2)V^2}{Z_0}$$

$$\Longrightarrow P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[V^2(t) \pm V^2(t) \sqrt{\left(\frac{\Delta V(t)}{V(t)}\right)^2 + \left(\frac{\Delta V(t)}{V(t)}\right)^2} \right]$$

$$P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[V^2(t) \pm V^2(t) \sqrt{2\left(\frac{\Delta V(t)}{V(t)}\right)^2} \right]$$

$$P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[V^2(t) \pm V^2(t) \cdot \frac{\Delta V(t)}{V(t)} \cdot \sqrt{2} \right]$$

$$P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[V^2(t) \pm V(t) \cdot \Delta V(t) \cdot \sqrt{2} \right]$$

$$P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[V^2(t) \pm V(t) \cdot 0.1V(t) \cdot \sqrt{2} \right]$$

$$P(t) = \frac{(1-\Gamma^2)}{Z_0} \cdot V^2(t) \cdot \left[1 \pm 0.1\sqrt{2} \right]$$

Finally, as before, we calculate the energy as the integral of the dissipated power:

$$E = \int P(t)\, dt = \frac{(1-\Gamma^2)}{Z_0} \cdot \left[\int V^2(t) dt \right] \cdot \left[1 \pm 0.1\sqrt{2} \right]$$

S4. MOKE images for calculation of switched area

The average switched area was defined as the switched area percentage of the *visible* magnetic, after one pulse, averaged over various trials. The images were taken with a polar MOKE microscope. In Fig. S7, we show an example of such image, where only the central black region is magnetic, while the left and right sections correspond to the gold waveguide, which covers the rest of the magnetic strip (18 x 4 μm^2). The top and bottom parts correspond to the SiO_2 layer and are, therefore, non-magnetic. In order to capture more light and thus enhance magneto-optical contrast, we increased the camera's exposure time. Due to gold's high reflectivity, this results in pixel saturation, which is why the apparent noise in those regions is small. Given the small size of our samples (3 x 4 μm^2) and the additional pixel saturation in gold, which induces a brightness bleeding into it, the magnetic region is rendered smaller. Therefore, in order to obtain images with enough resolution, the switched area was measured in a slightly longer sample (4 x 4 μm^2).

The switched area was defined **manually** by comparing MOKE images of the sample at three different states: (i) before the pulse (after saturation with positive field), (ii) after the pulse, and (iii) after saturation with negative field. We observed a significant movement of our sample stage during the process due to some mechanical displacements (mostly induced by the magnetic field on some steel elements of the stage) and drift. This shift, though small, was significant enough to make the automatization of the process very difficult (for example, by using a script), hence the need to evaluate case by case.

We defined an event as the single-pulse attempt to reverse the sample magnetization. For each event we registered the 3 images mentioned above (positive saturation, after pulse, negative saturation). The statistics were made considering 10 events per each current value. Because of this, the error bars pictured in Fig. 2a (main text) correspond to a standard deviation.

We depict the qualitative difference of the switching process when induced by a 7-ps (Fig. S8a) and by a 1-ns pulse (Fig. S8b). The color scale of all images is identical. For ps switching, the final states are consistently similar for a fixed current value, and the window for which partial switching is observed is rather narrow (less than 2 % over and below threshold). No domain formation is observed. For ns switching, on the other hand, the final states can vary greatly from one another, which yields a large standard deviation. Partial switching is observed for current amplitudes even 20 % below the threshold. Multiple domain nucleation is observed in some cases.

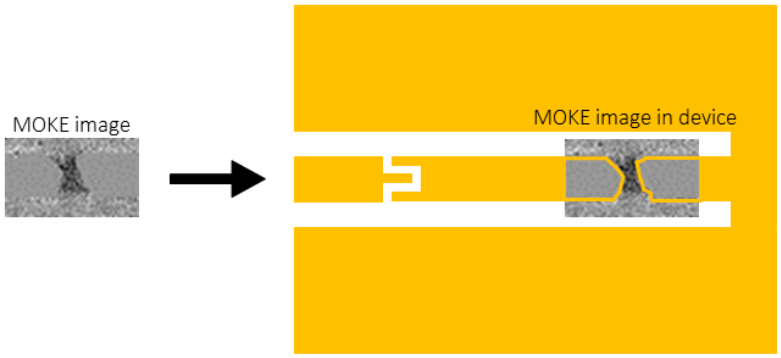
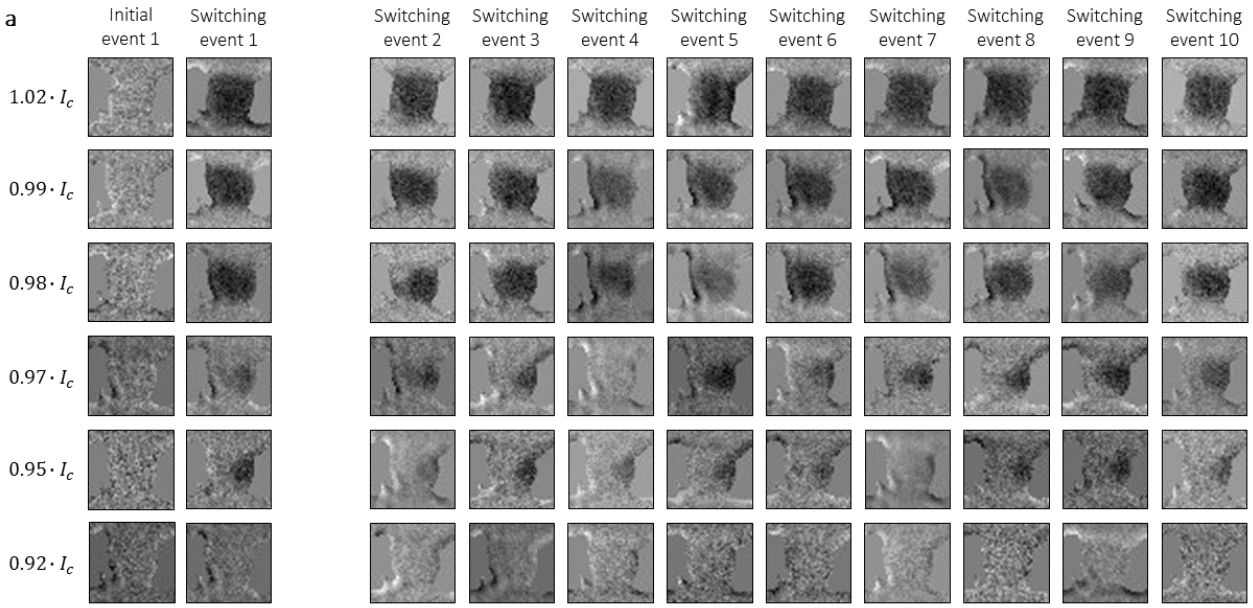


Figure S7. Left: MOKE image of a 3 x 4 μm^2 sample after SOT switching with a 7-ps pulse. **Right:** MOKE image of sample within the device for visual aid. The sample is visibly switched (black over gray) but the overflow from the electrodes distorts the sample area and leaves very few pixels available for determining switched areas.



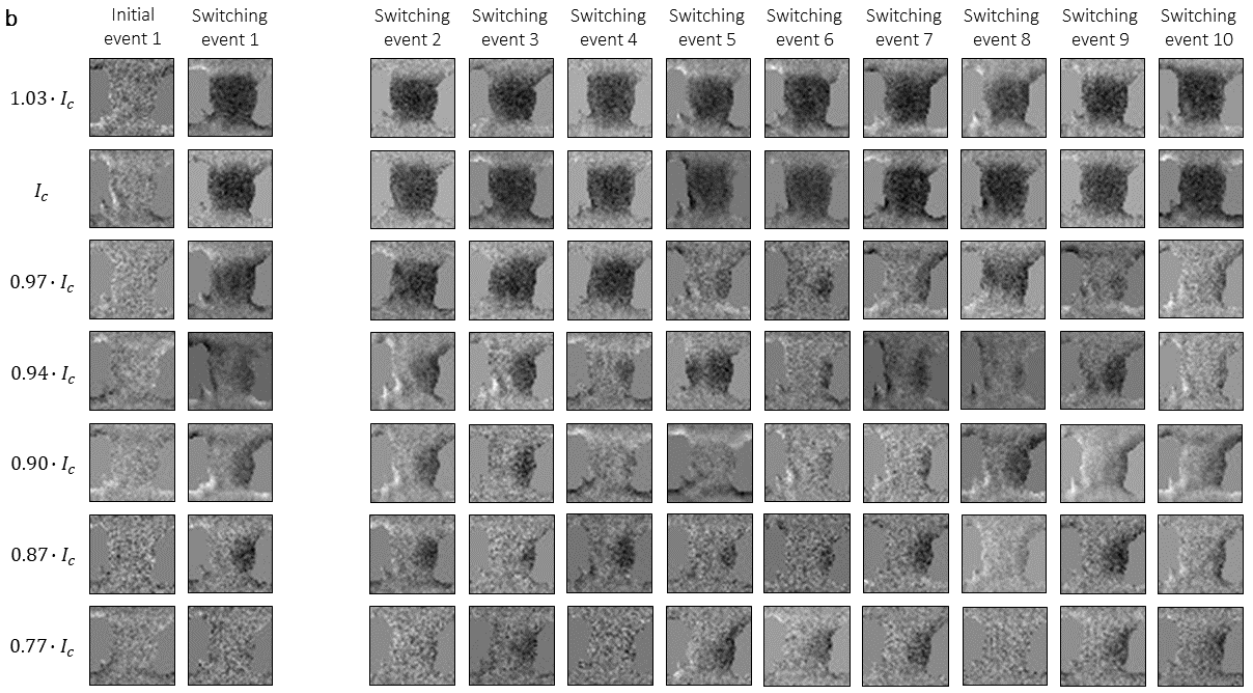


Figure S8. MOKE images of the sample after **a.** 7-ps pulse, and **b.** 1-ns wide pulse at different currents, with respect to the critical value (100 % switched area). The sample's initial state (left colon showing a gray state) is shown for the first switching attempt.

S5. In-plane field dependence of threshold current for CoFeB and CoGd

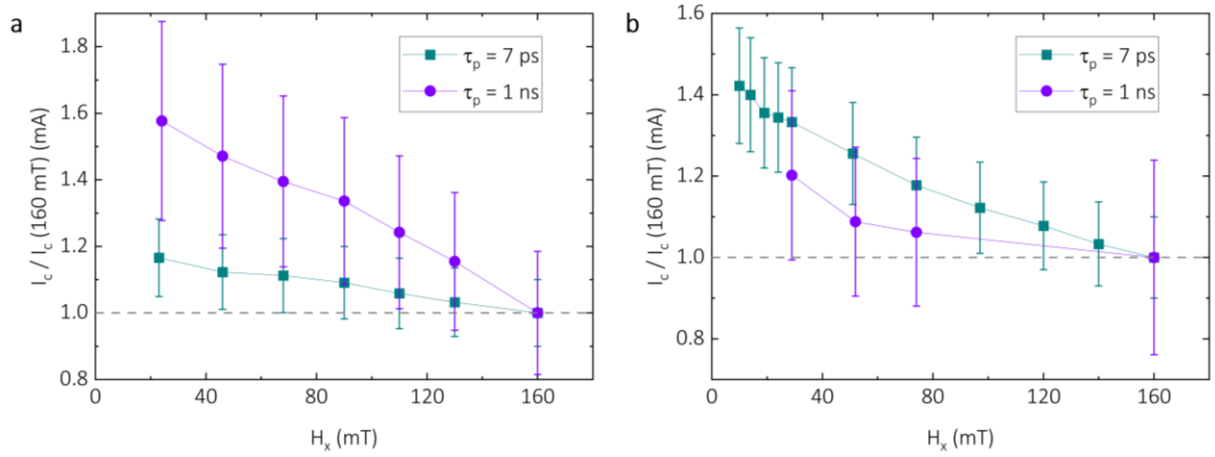


Figure S9. In-plane field dependence of critical current density for **a.** CoFeB and **b.** CoGd, for SOT switching induced by a 7-ps and a 1-ns pulse.

S6. Optical absorption calculation for all-optical switching

We performed a multilayer transfer matrix calculation to estimate the optical absorption in our magnetic heterostructure of GaAs/Ta(3)/Pt(5)/CoGd(2.4)/Ta(2)/Pt(2) for all-optical switching (see Fig. S10). The optical indices used are shown in Table S1. The total optical absorption is 18.5 %. The Co/Gd bilayer (represented as a single Co layer) absorbs about 4 %. In Fig. 3b of the main text, the reported AOS energy cost corresponds to the total absorbed power (18 %), just as for the SOT experiments, but absorption in the Co/Gd layer could potentially be improved by optical engineering.

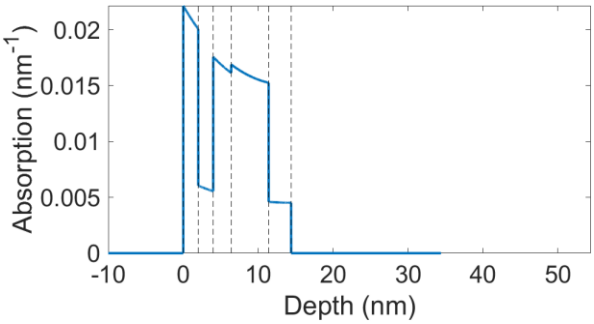


Figure S10. Absorption profile for the *CoGd* sample. Depth 0 corresponds to the top of the Pt capping layer.

Medium	Range (nm – nm)	Complex refraction index
Air	$-\infty \rightarrow 0$	1
Pt	0 – 2	2.84 – 4.95i [4]
Ta	2 – 4	1.2 – 3.53i [4]
Co (CoGd)	4 – 6.4	2.67 – 5.03i [4]
Pt	6.4 – 11.4	2.84 – 4.95i [4]
Ta	11.4 – 14.4	1.2 – 3.53i [4]
GaAs	$14.4 \rightarrow \infty$	3.68 – 0.09i [5]

Table S1. Complex refraction indices for different materials in the *CoGd* sample.

S7. Estimation of saturation magnetization by vibrating sample magnetometry (VSM)

From VSM we can estimate the saturation magnetization (M_S) by measuring the hysteresis loop in plane and out of plane. In Fig. S11 we show the out of plane hysteresis loop obtained for the Co sample. Table S2 shows the results for all the samples.

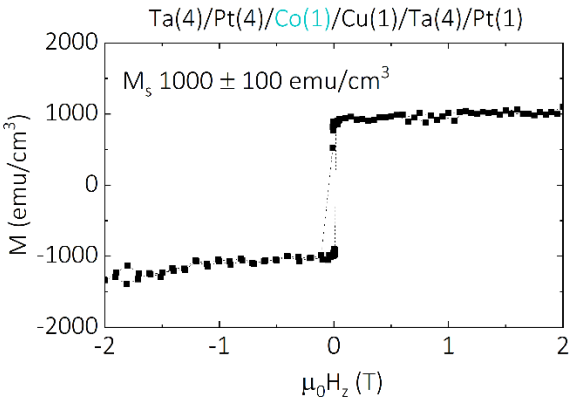


Figure S11. Hysteresis loop of the Co sample from out-of-plane VSM.

Sample	M_S (emu/cm^3)
Co	1000(100)
CoFeB	900(100)
CoGd	170(20)

Table S2. Saturation magnetization for all samples.

S8. Estimation of current density in Co stack

In Fig. 3a of the main text we report the current density for the Co sample, calculated by assuming that 40 % of the total current flows through the 4-nm Pt layer. These values were estimated using a parallel resistor model, as shown in Table S3, using estimate values from literature. We note that we did not measure the resistivity of each layer separately; the goal being only to provide an estimation of the order of magnitude of current density. These calculations do not affect any other results in the paper.

The parallel resistor model yields a total resistance value of 59 Ω, in excellent agreement with the real resistance of our sample (58.5 Ω), calculated from the measured stack resistivity of $117 \cdot 10^{-8}$ Ω m. By calculating the current percentage flowing in each layer, we see that almost 40 % of the current should flow through the high spin-Hall angle Pt layer.

Layer	Thickness (nm)	Resistivity (10^{-8} Ω m)	Source	Resistance (Ω)	Current (%)
Pt	1	200	Estimated as 2x the resistivity of Ir in Ref. [6]	1500	4
Ta	4	200	Ref. [7]	375	16
Cu	1	50	Projection from Ref. [8] as well as Ref. [6]	375	16
Co	1	100	Ref. [6]	750	8
Pt	4	80	Estimated as 2x the resistivity of Ir in Ref. [6]	150	40
Ta	4	200	Ref. [7]	375	16
			Total (summed in parallel)	59	100

Table S3. Thin film resistivities for the materials in the *Co* sample.

S9. Estimation of the anisotropy field and effective spin Hall angle by anomalous Hall effect and second harmonic measurements

From anomalous Hall effect (AHE) measurements, it is possible to extract an approximate value of the anisotropy field (H_k). This method of estimation of H_k is more reliable than VSM for ultra-thin films due to its low signal-to-noise ratio for samples with low total magnetic moment. Here we are considering a macrospin model for a sample with perpendicular magnetic anisotropy showing one easy axis and no other in-plane anisotropy. The transversal resistance measured in a Hall-bar device is usually written in this case as:

$$R_{XY} = R_H H_z + R_{AHE}\hat{m}_z + 2R_{PHE}(\hat{m}_x \cdot \hat{m}_y)$$

where R_H is the Hall resistance, R_{AHE} is the anomalous Hall resistance, R_{PHE} is the planar Hall resistance, and $\hat{\boldsymbol{m}}_{x,y,z}$ are the reduced x, y, z-components of the magnetization.

From this, and considering that in our case R_H is much smaller than R_{AHE}, we can estimate the anisotropy field from the AHE measurement by applying an in-plane field ($H_{x,y}$) smaller than H_k as:

$$R_{XY} = R_0 + R_{AHE}\sqrt{1 - (H_{x,y}/H_k)^2}$$

Where R_0 is the offset resistance. The results are illustrated in Fig. S12.

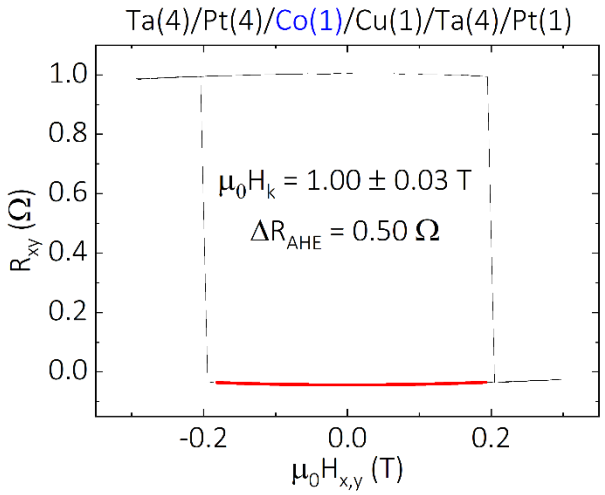


Figure S12. Estimation of H_k from transversal resistance measurements for the Co sample.

Following the well-known technique of second harmonic generation [8-11], one can estimate the spin-orbit torque (SOT) effective fields ($\boldsymbol{h}$). This is done by applying an AC current in a Hall-bar shaped device and analyzing the harmonic response of the transversal voltage. SOT is the combination of the damping-like ($h_{DL} \propto \hat{s}$) and field-like ($h_{FL} \propto \hat{m} \times \hat{s}$) effective fields, that correspond respectively to the two orthogonal components of SOT and are estimated from second harmonic measurements as a function of an in-plane field as:

$$h_{DL} = \pm 2\frac{\partial_{H_x} V_{XY}^{2\omega}}{\partial_{H_x}^2 V_{XY}^{\omega}}$$

$$h_{FL} = 2\frac{\partial_{H_y} V_{XY}^{2\omega}}{\partial_{H_y}^2 V_{XY}^{\omega}}$$

where the AC current is applied in the x-direction and V_{XY}^{ω} and $V_{XY}^{2\omega}$ are the first and second harmonic components of the transversal voltage, respectively.

We introduce an effective spin Hall angle $\theta_{SH,\text{eff}}^{DL,FL}$ which accounts for the efficiency of the spin-orbit torque effective fields with the total charge current density flowing in the device:

$$\theta_{SH,\text{eff}}^{DL,FL} = \frac{2|e|}{\hbar} M_s t_{FM} \frac{h_{DL,FL}}{j_{c,\text{TOT}}}$$

where e is the charge of the electron, $\hbar$ is the reduced Planck constant, t_{FM} is the thickness of the ferromagnetic layer and $j_{c,\text{TOT}}$ is the total charge current density flowing through the sample.

In Fig. S13 we show the first and second harmonic voltage dependence upon application of an in-plane field for the DL and FL configurations, i.e., H_x and H_y respectively, as well as the current density dependence of the effective fields for both positive and negative remanence fields.

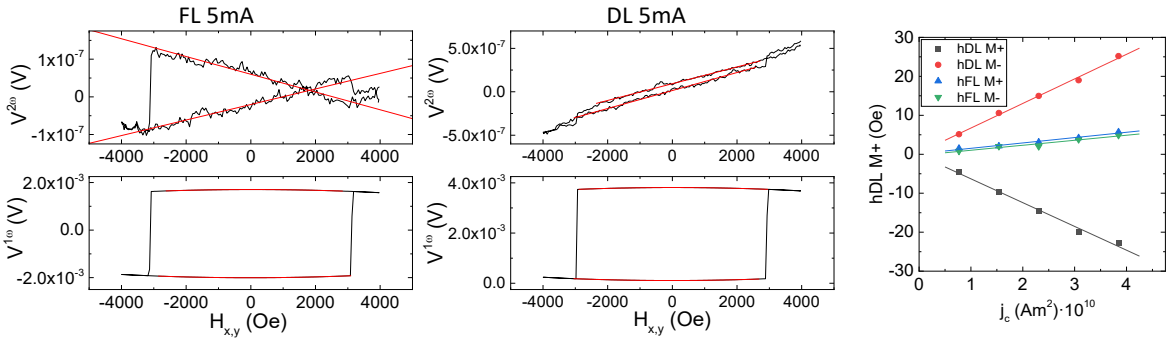


Figure S13. Estimation of the spin torque effective fields by second harmonic technique for the *Co* sample.

From these measurements, we can extract the effective spin Hall angle for the different samples, as indicated in Table S4.

	Co	CoFeB	CoGd
$\theta_{SH}^{eff}\, M^+$ DL	0.18(2)	-0.068(6)	0.37(4)
$\theta_{SH}^{eff}\, M^-$ DL	0.19(2)	-0.072(7)	0.37(4)
$\theta_{SH}^{eff}\, M^+$ FL	0.08(1)	-0.046(5)	0.046(5)
$\theta_{SH}^{eff}\, M^-$ FL	0.08(1)	-0.047(5)	0.045(5)

Table S4. Calculated effective spin Hall angles for all the samples.

S10. Estimation of Gilbert damping by ferromagnetic resonance (FMR)

Ferromagnetic resonance (FMR) measurements were conducted to determine the Gilbert damping with a NanOsc Instruments PhaseFMR 40-GHz system, using a coplanar waveguide (CPW) inside a Quantum Design PPMS system. The CPW is used to generate a RF field in the sample, which is simultaneously subjected to an external static magnetic field. As the frequency of the microwave signal is swept, resonance occurs when the frequency matches the natural precession frequency of the magnetization in the ferromagnetic sample, satisfying the Kittel formula [13]:

$$H_{\mathrm{res}}(f) = \frac{2\pi}{\gamma\mu_0} f + \mathrm{M}_{\mathrm{eff}},$$

where $\mathrm{M}_{\mathrm{eff}}$ is the effective magnetization, γ is the gyromagnetic ratio and μ_0 is the vacuum permeability. In Fig. S14 we show the resonance frequency as a function of the applied external field. From this data, we obtained a Landé factor (g) of 2.06, comparable to others in literature [14]. Gilbert damping is then obtained by estimating the change of the resonance width with frequency:

$$\Delta H(f) = \frac{4\pi\alpha}{\gamma\mu_0} f + \Delta H_0$$

where α is the Gilbert damping and ΔH_0 is the inhomogeneus linewidth. In Fig. S15, we show the measurement for the CoFeB film, where we obtained a value of $\alpha = 0.024 \pm 0.001$. Due to the low volume of the magnetic layers, these measurements could only be performed on this sample.

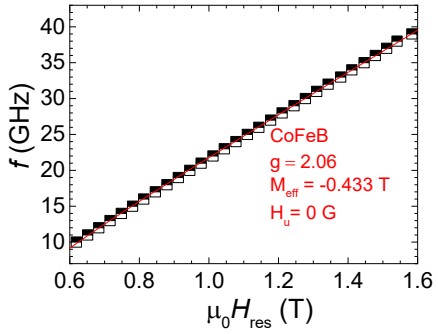


Figure S14. Dependence of the resonance frequency with the applied DC magnetic field. The effective magnetization ($\mathrm{M}_{\mathrm{eff}}$) and Landé factor ($g$) are obtained from the fit to the Kittel equation.

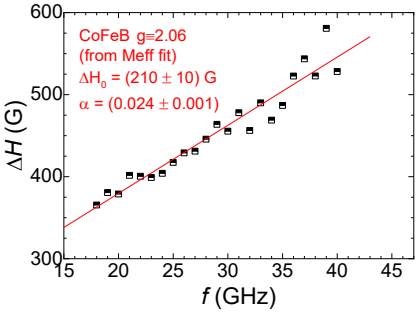


Figure S15. Dependence of the FMR linewidth with the resonance frequency. The Gilbert damping and inhomogeneous linewidth are obtained from this data.

S11. Micromagnetic simulations

Micromagnetic simulations were carried out using the Mumax[3] simulation toolkit. Magnetization dynamics was obtained by integrating the Landau-Lifshitz-Gilbert equation including the Slonczewski-like torque, which in its explicit form reads as:

$$\frac{d\vec{\boldsymbol{m}}}{dt} = \frac{\gamma}{(1+\alpha^2)}\left(\vec{\boldsymbol{m}} \times \vec{\boldsymbol{B}}_{\text{eff}} + \alpha\left(\vec{\boldsymbol{m}} \times (\vec{\boldsymbol{m}} \times \vec{\boldsymbol{B}}_{\text{eff}})\right)\right) + \gamma\beta\left(\frac{\varepsilon + \alpha\varepsilon'}{(1+\alpha^2)}(\vec{\boldsymbol{m}} \times (\vec{\boldsymbol{p}} \times \vec{\boldsymbol{m}})) - \frac{\varepsilon' - \alpha\varepsilon}{(1+\alpha^2)}(\vec{\boldsymbol{m}} \times \vec{\boldsymbol{p}})\right)$$

Where γ is the gyromagnetic ratio, α is the Gilbert's damping parameter, $\beta = \frac{j_Z \hbar}{M_S e d_0}$ with j_Z the virtual current density flowing perpendicular to plane, $\hbar$ is the reduced Planck's constant, M_S the saturation magnetization, e the electric charge of electrons and d_0 the thickness of the ferromagnetic layer. The primary spin-torque parameter $\varepsilon = P(\vec{\boldsymbol{r}}, t)\Lambda^2/((\Lambda^2+1) + (\Lambda^2-1))\vec{\boldsymbol{m}} \cdot \vec{\boldsymbol{p}}$, can be directly related to the damping-like spin-hall angle (θ_{SH}^{DL}) by assigning the polarization factor $P(\vec{\boldsymbol{r}}, t) = -\theta_{SH}^{DL}$ and the Slonczewski parameter $\Lambda = 1$. The secondary spin-torque term is directly the field-like spin-hall angle $\varepsilon' = \theta_{SH}^{FL}$.

The considered effective field $\vec{\boldsymbol{B}}_{\text{eff}}$ includes the anisotropy contribution $\vec{\boldsymbol{B}}_{\text{Anis}}$, magnetostatic $\vec{\boldsymbol{B}}_{\text{Demag}}$, direct symmetric exchange $\vec{\boldsymbol{B}}_{\text{Exch}}$, antisymmetric exchange or DMI $\vec{\boldsymbol{B}}_{\text{DMI}}$ and the external applied magnetic field or Zeeman contribution $\vec{\boldsymbol{B}}_{\text{Zee}}$. The effect of temperature is taken into account by including a Langevin stochastic field contribution ($\vec{\boldsymbol{B}}_{\text{Therm}}$) proportional to the sample temperature as well as considering a phenomenological temperature-dependence of the anisotropy following the Callen-Callen power law with fitted exponent [15]:

$$K_U(T) = K_U(0)\left(1 - \left(\frac{T}{T_C}\right)^{\alpha_M}\right)^{\beta_{\text{Anis}}}$$

We note that the Mumax[3] native stochastic fields were scaled by an artificial thermal scaling parameter η in order to tune the excitations' amplitudes to the energy barriers in the system, in order to maintain a remanent state before the electrical pulse, but to enable some stochasticity in the reversal. We also include a small spatial random deviation (1 %) of the anisotropy angle and amplitude in order to facilitate nucleation (similar to defect sites).

The temperature profile is calculated as from Ref. [16] assuming a one-temperature model in the form:

$$C_V \frac{dT}{dt} = q(t) - (T - T_{RT}) \cdot \frac{G}{d}$$

Where G is the interfacial thermal conductance and d is the total thickness of the sample. The volumetric heating source coming from the Joule heating $q(t)$ takes the form $q(t) = j_c^2(t)\rho_e$, where $j_c(t)$ is the time-dependent effective electric current density within the whole stack and ρ_e the effective electrical resistivity of the full stack. We note that in such metallic systems, electrons thermalize quickly within the nanostructure (in sub-ps time scales) due to electronic thermal diffusion. When dealing with ultrashort pulses on the order of 100 femtoseconds, electrons and phonons are out of equilibrium with each other for a few picoseconds, resulting in different electron and phonon temperatures. However, as explained in Ref. [16], when using a 6-7 ps pulse electron and phonon temperatures are pretty close to each other, particularly if heavy metals with high electron-phonon coupling constants such as Pt or Ta are involved. We can therefore reasonably assume a single temperature bath for our model in order to simplify the calculations. One can then notice that the simple heat diffusion equation, only contains 3 parameters: the heat capacity of the stack C_v, the interface conductance G/d and the Joule heating $q(t)$. Most solid metals have heat capacities on the order of 2-3 MJ/K·m^3, and as in Ref. [16], we use a weighted averaged for our stacks with values extracted from literature. Joule heating is directly determined by the current density profile. The only unknown parameter determining the temperature profile is, therefore, the interface conductance with the substrate, which we extracted via pump-probe experiments in our previous work [16].

In our simulations, the electric pulse is assumed to be a gaussian-shaped pulse in the form:

$$j_c(t) = j_c \cdot e^{\frac{-(t-t_0)^2}{2\sigma^2}}$$

With j_c the current amplitude, t_0 the pulse arrival time and σ the standard deviation which is related with the pulse width τ_p as $\sigma = \tau_p/2.35482$. We note that the assumed current density used for the torques is the one flowing through the Co (given our simplistic assumptions in S7), whereas for the thermal modelling we used the full effective current density and simulated the full thermal volume. The choice of the current density can change the ratio of heating vs torques, similarly to the choice of the thermal dependencies, and affect the thresholds. This should be further studied in future studies.

The temperature profile described by the previous differential equation was solved numerically for the different considered current densities and current widths. As Mumax[3] temperature profiles must be introduced in an analytical form, the obtained temperature profiles were fitted to the following expression:

$$T_{\text{fit}}(t) = T_{RT} + \frac{T_{\text{Max}}}{\left(e^{-\frac{(t-t_0)}{\tau_1}} + 1\right)} \cdot e^{\frac{-(t-t_0)}{\tau_2}}$$

Which is a combination of a Fermi-like step function modulated by an exponential decay.

To illustrate this point, one can get a back-of-the-envelope estimate of the maximum temperature rise for very short pulses. In that regime, all the energy is used to raise the temperature of the full stack first, and then the much slower interfacial heat-diffusion with the substrate will cool down the stack. Therefore, the maximum temperature rise can be simply approximated by the energy density absorbed by the stack, divided by the heat capacity ($\Delta T = \frac{E}{V}\frac{1}{C_v}$). For a 7 ps pulse we find an energy consumption of around 30 pJ for a device of 3 µm x 4 µm x 15 nm. The heat capacity should be close to the assumed 2.6 MJ/K·m^3. We then obtain a temperature rise of $\Delta T = 65$ K. Below, in Fig. S16, we show a temperature profile calculated for a 7 ps pulse at the critical threshold, and the simulated peak temperatures as a function of pulse duration. Our numerical simulations for 7 ps pulses result in a temperature rise of 75 K at the critical threshold, very close to the back-of-the-envelope calculation.

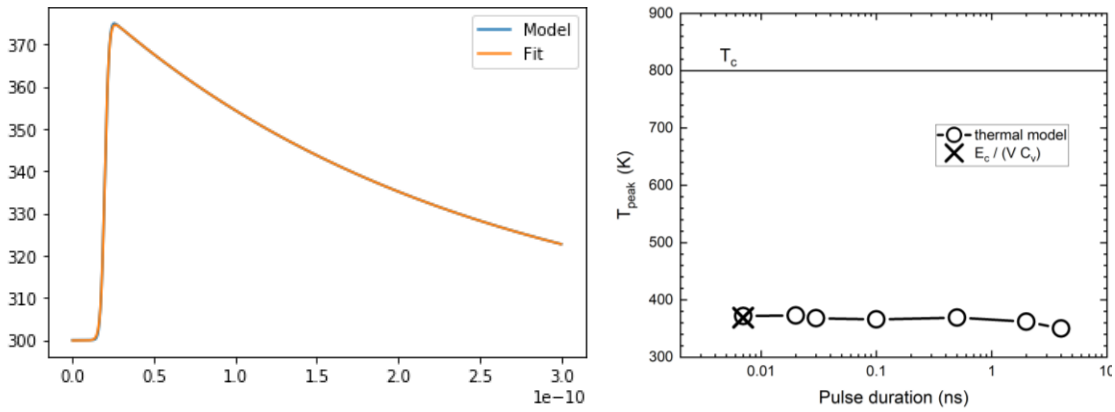


Figure S16. Left: Modelled temperatures in K (y-axis) as a function of time in seconds (x-axis), at the critical threshold for a 7 ps pulse. The fit corresponds to the analytical formula $T_{\text{fit}}(t)$. **Right:** Simulated peak temperature as a function of pulse duration at the critical threshold (circles) and simple estimation in the short pulse regime (cross). The Curie temperature is indicated with a horizontal line.

We note that all micromagnetic and macrospin simulation results have been renormalized by the negative time delay values (pre-pulse saturated state).

Finally, the parameters used in Mumax[3] code are indicated below in Table S5.

Parameter	Variable	Value	Source
Micromagnetic definitions			
Saturation magnetization	$M_{\mathrm{Sat}}\ (300\ \mathrm{K})$	$1 \cdot 10^{6}$ A m^{-1}	Measured in Ref. [16] + VSM (Supp. Mat. S6)
Uniaxial anisotropy	$K_U\ (300\ \mathrm{K})$	$1 \cdot 10^{6}$ J m^{-3}	Measured in Ref. [16] + 2nd harmonic (Supp. Mat. S8)
Exchange stiffness	$A_{\mathrm{Exch}}\ (300\ \mathrm{K})$	$20 \cdot 10^{-12}$ J m^{-1}	Order of magnitude from Ref. [17]
Interfacial DMI	$D\ (300\ \mathrm{K})$	$1 \cdot 10^{-3}$ J m^{-2}	Order of magnitude from Ref. [17]
Gilbert damping	α	0.22	Fitted in Ref. [16]
Spin-orbit torque definitions			
Damping-like spin Hall angle	θ_{SH}^{DL}	0.19	2nd harmonic (Supp. Mat. S8)
Field-like spin Hall angle	θ_{SH}^{FL}	0.08	
Structural parameters			
Cobalt thickness	d_0	$1 \cdot 10^{-9}$ m	From growth calibration
Sample thickness	d	$15 \cdot 10^{-9}$ m	
Thermal definitions			
Interfacial thermal conductance	G	$170 \cdot 10^{6}$ W m^{-2} K^{-1}	Fitted in Ref. [16]
Volumetric total heat capacity	C_V	$2.6 \cdot 10^{6}$ J m^{-3} K^{-1}	Estimated in Ref. [16]
Curie temperature	T_C	800 K	Ref. [16]
Magnetization (temperature dependent) exponent	α_M	1.7	Ref. [17]
Anisotropy (temperature dependent) exponent	β_{Anis}	2.2	Ref. [17]
Scaling parameter for thermal fluctuations	η	0.2	Phenomenological
Electrical current definitions			
Electrical resistivity	ρ_e	$117 \cdot 10^{-8}$ Ω m	Measured

Table S5. Parameters used in Mumax[3] micromagnetic simulation.

S12. Importance of edge and dimensions for reversal via domain wall propagation

When performing Mumax[3] simulations, the conditions of the edges of the area of interest are very important. It is well known that on the edges of the magnet there is often a small canting of the magnetization due to stray fields. Such canting, combined with the external field and Dzyaloshinskii-Moriya interaction (present in our films), facilitates the nucleation of a magnetic domain on the edge [18]. In our samples, the magnetic strip is a long 18 x 4 μm^2 magnet with gold electrodes contacting over it, leaving only a short 3 µm-long region uncovered (illustrated in Fig. S17). This means that the electrical current is injected within the middle of the strip and not close to the edges of the magnet. When performing simulations where the current is injected directly on the edge, we do observe such edge nucleation, and the domain wall then propagates towards the opposing end. This is depicted in Fig. S18. On the other hand, when magnetic regions with no current flow are added around the sample (which we will refer to as "contacts"), the same current densities yield no such edge-nucleation, as is depicted in Fig. S19. All studies presented in the main paper have therefore included these "magnetic contact" regions.

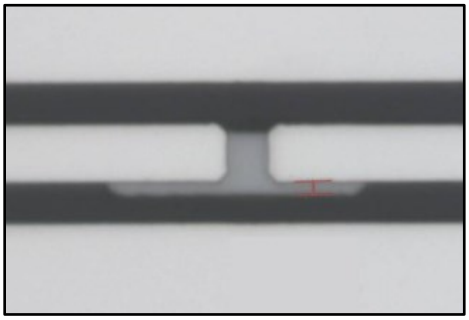


Figure S17. Micrograph of our 18 x 4 μm^2 magnetic strip (light gray) under gold electrodes (white) (red scale bar, corresponding to a ~1 µm misalignment). In this particular sample there is an important misalignment between the magnet and the contacts which makes the superposition more evident, so we chose it for illustration purposes. However, in regular samples such as the ones reported in this article, the alignment is much better.

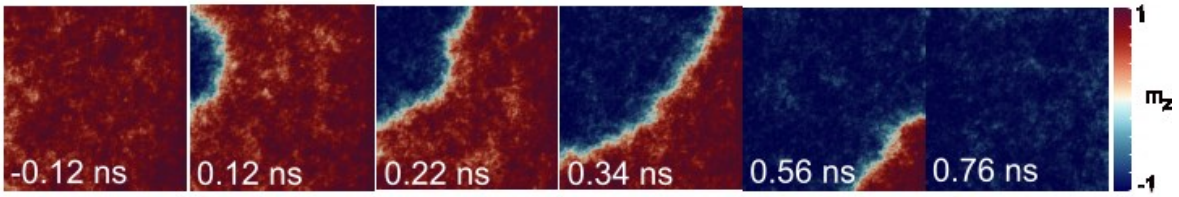


Figure S18. Micromagnetic simulation for a 128 x 128 nm^2 *Co* sample **without** contacts. This means the current is injected directly on the left/right edges of the square magnet. SOT is induced by a 2-ns pulse and current density of 0.8 x 10^{12} A m^{-2} under an in-plane field of 150 mT. Edge nucleation is observed.

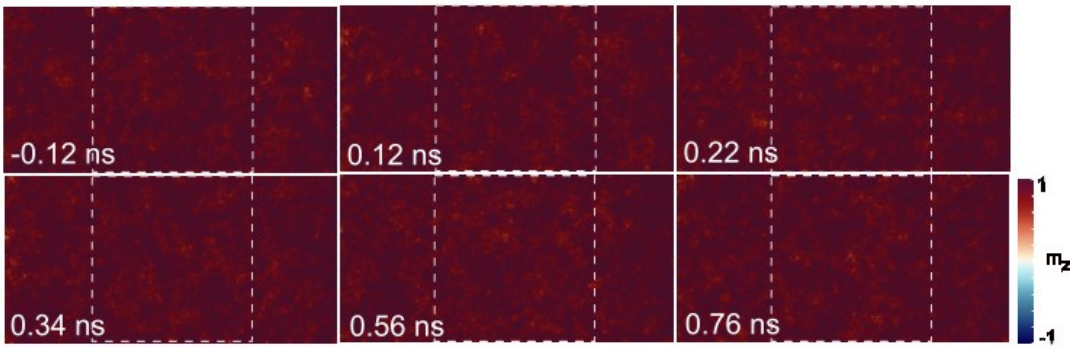


Figure S19. Micromagnetic simulation for a 128 x 128 nm^2 *Co* sample **with** contacts. This means the current only flows within the central dashed squared area. SOT is induced by a 2-ns pulse and current density of 0.8 x 10^{12} A m^{-2} under an in-plane field of 150 mT (same parameters as in Fig. S18). Edge nucleation is not observed in this case.

We now turn our attention to the importance of the magnet size for the behavior of the reversal. Many studies have shown that, when approaching the ns regime, the reversal is dominated by a single edge-to-edge domain wall propagation[18], [19]. However, most of these studies were carried out on ∼100 nm devices. In order for a domain wall to have the time to cross a 100 nm region, the current has to be sustained for a certain duration. If we now make the device 10 times larger, for the same current pulse duration, the intensity will need to increase by a factor of 10 in order for the domain wall to reach the other side by the end of the pulse. However, as the current intensity is increased, domain nucleation also becomes more likely. For large devices (micron-scale), nucleation can start to dominate over domain wall propagation [20]. In order to elucidate the importance of the dimensions in our devices, we performed the same simulations with a small (128 x 128 nm^2) and a large (1 x 1 µm^2) square device. As shown in Figs. S18 and S20, for the same current density, only the smaller device can be fully reversed. In order to reverse the larger one, the current density needs to be increased (in this case, by 12.5 %), up to a point where nucleation all over the device starts to overtake (see Fig. S21).

The combination of large devices (micron-scale) and injection of current in the middle of the magnetic device explains the likely nucleation-dominated reversal of our devices. We note that some of the parameters used in the above simulations where different from the ones presented in the main paper, but the general observations (domain wall mediated vs nucleation dominated) due to contacts (edges) and dimensions are always present.

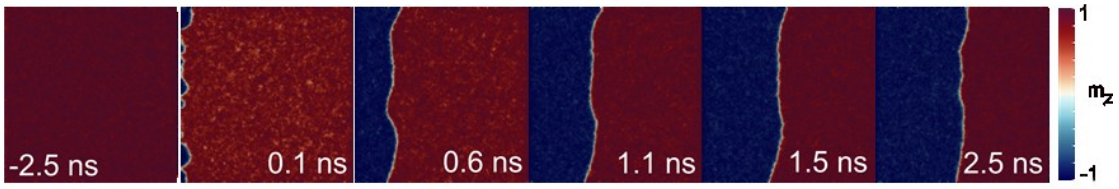


Figure S20. Micromagnetic simulation for a 1 x 1 µm^2 *Co* sample **without** contacts. SOT is induced by a 2-ns pulse and current density of 0.8 x 10^{12} A m^{-2}. Due to the size of the sample, full switching is not achieved.

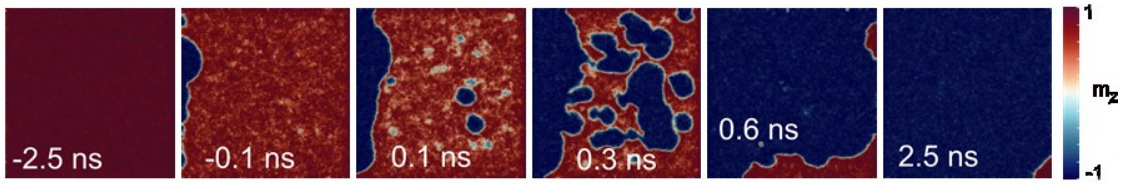


Figure S21. Micromagnetic simulation for a 1 x 1 µm^2 *Co* sample **without** contacts. SOT is induced by a 2-ns pulse and a slightly higher current density of 0.9 x 10^{12} A m^{-2}. Full switching is achieved by increasing the current by 12.5 % with respect to that of Fig. S20.

References

[1] E. Díaz, A. Anadón, M. Morassi, M. Hehn, A. Lemaître, and J. Gorchon, "Calibration of terahertz sampling detectors for intense unipolar picosecond current pulses in waveguides," *Appl. Phys. Lett.*, vol. 123, no. 14, p. 141106, Oct. 2023, doi: 10.1063/5.0169020.

[2] E. Castro-Camus, M. B. Johnston, and J. Lloyd-Hughes, "Simulation of fluence-dependent photocurrent in terahertz photoconductive receivers," *Semicond. Sci. Technol.*, vol. 27, no. 11, p. 115011, Nov. 2012, doi: 10.1088/0268-1242/27/11/115011.

[3] H. Arellano, R. Garreaud, D. Mardones, N. Mujica, Á. Núñez, and R. Soto, "Sistemas Newtonianos - Apuntes del curso." 2010. [Online]. Available: https://www.cec.uchile.cl/~mclerc/archivos/Books/Apuntes.pdf

[4] N. Bergeard *et al.*, "Hot-Electron-Induced Ultrafast Demagnetization in Co / Pt Multilayers," *Phys. Rev. Lett.*, vol. 117, no. 14, p. 147203, Sep. 2016, doi: 10.1103/PhysRevLett.117.147203.

[5] D. E. Aspnes, S. M. Kelso, R. A. Logan, and R. Bhat, "Optical properties of Al *x* Ga1− *x* As," *Journal of Applied Physics*, vol. 60, no. 2, pp. 754–767, Jul. 1986, doi: 10.1063/1.337426.

[6] M. Siniscalchi, D. Tierno, K. Moors, Z. Tőkei, and C. Adelmann, "Temperature-dependent resistivity of alternative metal thin films," *Appl. Phys. Lett.*, vol. 117, no. 4, p. 043104, Jul. 2020, doi: 10.1063/5.0015048.

[7] D. Céspedes-Berrocal *et al.*, "Current-Induced Spin Torques on Single GdFeCo Magnetic Layers," *Adv. Mat.*, vol. 33, no. 2007047, 2021, doi: 10.1002/adma.202007047.

[8] J.-W. Lim, K. Mimura, and M. Isshiki, "Thickness dependence of resistivity for Cu films deposited by ion beam deposition," *Appl. Surt. Sci.*, vol. 217, no. 1–4, pp. 95–99, Jul. 2003, doi: 10.1016/S0169-4332(03)00522-1.

[9] K. Garello *et al.*, "Symmetry and magnitude of spin–orbit torques in ferromagnetic heterostructures," *Nat. Nanotechnol.*, vol. 8, no. 8, pp. 587–593, Aug. 2013, doi: 10.1038/nnano.2013.145.

[10] M. Hayashi, J. Kim, M. Yamanouchi, and H. Ohno, "Quantitative characterization of the spin-orbit torque using harmonic Hall voltage measurements," *Phys. Rev. B*, vol. 89, no. 14, p. 144425, Apr. 2014, doi: 10.1103/PhysRevB.89.144425.

[11] J. Kim *et al.*, "Layer thickness dependence of the current-induced effective field vector in Ta|CoFeB|MgO," *Nat. Mater.*, vol. 12, no. 3, pp. 240–245, Mar. 2013, doi: 10.1038/nmat3522.

[12] A. Anadón *et al.*, "Spin-Orbit Torque from the Introduction of Cu Interlayers in Pt/Cu/Co/Pt Nanolayered Structures for Spintronic Devices," *ACS Appl. Nano Mater.*, vol. 4, no. 1, pp. 487–492, Jan. 2021, doi: 10.1021/acsanm.0c02808.

[13] R. Sbiaa *et al.*, "Ferromagnetic resonance measurements of (Co/Ni/Co/Pt) multilayers with perpendicular magnetic anisotropy," *J. Phys. D: Appl. Phys.*, vol. 49, no. 42, p. 425002, Oct. 2016, doi: 10.1088/0022-3727/49/42/425002.

[14] H. Q. Tu *et al.*, "Gilbert damping in CoFeB/GaAs(001) film with enhanced in-plane uniaxial magnetic anisotropy," *Sci Rep*, vol. 7, no. 1, p. 43971, Mar. 2017, doi: 10.1038/srep43971.

[15] H. B. Callen and E. Callen, "The present status of the temperature dependence of magnetocrystalline anisotropy, and the l(l+1)/2 power law," *J. Phys. Chem. Solids*, vol. 27, no. 8, pp. 1271–1285, Aug. 1966, doi: 10.1016/0022-3697(66)90012-6.

[16] K. Jhuria *et al.*, "Spin–orbit torque switching of a ferromagnet with picosecond electrical pulses," *Nat. Electron.*, vol. 3, no. 11, pp. 680–686, Nov. 2020, doi: 10.1038/s41928-020-00488-3.

[17] S. Schlotter, P. Agrawal, and G. S. D. Beach, "Temperature dependence of the Dzyaloshinskii-Moriya interaction in Pt/Co/Cu thin film heterostructures," *Appl. Phys. Lett.*, vol. 113, no. 9, p. 092402, Aug. 2018, doi: 10.1063/1.5038353.

[18] M. Baumgartner *et al.*, "Spatially and time-resolved magnetization dynamics driven by spin-orbit torques," *Nat. Nanotechnol.*, vol. 12, no. 10, pp. 980–986, Oct. 2017, doi: 10.1038/nnano.2017.151.

[19] E. Grimaldi *et al.*, "Single-shot dynamics of spin–orbit torque and spin transfer torque switching in three-terminal magnetic tunnel junctions," *Nat. Nanotechnol.*, vol. 15, no. 2, pp. 111–117, Feb. 2020, doi: 10.1038/s41565-019-0607-7.

[20] M. M. Decker *et al.*, "Time Resolved Measurements of the Switching Trajectory of Pt / Co Elements Induced by Spin-Orbit Torques," *Phys. Rev. Lett.*, vol. 118, no. 25, p. 257201, Jun. 2017, doi: 10.1103/PhysRevLett.118.257201.