New Approaches to MRAM Switching

Neil Robertson

Many thanks to team members in WD Research
Energy-speed performance of emerging memories

Based on reported experimental results

- STT-MRAM has an advantage in terms of speed and energy efficiency over other emerging memory technologies.
- Endurance and retention can also be excellent depending on how the STT-MRAM is operated.
- Integrating MRAM as an embedded memory requires only 2 additional masks (versus 10 for eFlash), making it a good fit for the embedded space.

Recent developments of STT-MRAM memories

- Demonstration of stand-alone 4 Gb STT-MRAM memory by SK Hynix/Toshiba.
  S.-W. Chung et al., IEDM16 -659 (2016)

- Demonstration of 8Mb 1T-1MTJ STT-MRAM embedded in 28nm CMOS logic by Samsung.
  Y. J. Song et al., IEDM16 - 663 (2016)

- Significant technological development of STT-MRAM for both stand alone and embedded memory platforms.
The next step - MRAM as embedded SRAM replacement

- For some high value applications, last level cache takes up the majority of the die layout.

- In order to reach the required speed and WER to replace eSRAM, the STT-MRAM cell will require a large overdrive current. Density will be reduced to accommodate the large transistor required to provide the write current.

- Endurance is a significant concern here, as large write currents are needed to meet SRAM speed targets.

- STT-MRAM power is higher than SRAM.

- Innovations that reduce the current requirements for high-speed STT-MRAM are an active area of investigation (but won’t be covered in this talk).

- Alternatively, one could consider other fast MRAM technologies like SOT and VCMA switching for high performance, low-power embedded memory.
Spin orbit torque (SOT)-MRAM

Alternative MRAM cell design for high speed/endurance and power efficiency

- **Basic SOT-MRAM cell:**
  - Consists of spin-orbit torque layer (typically a heavy metal like Pt, Ta) in contact with MTJ
  - Write current is passed through SOT layer, while read current it is passed through MTJ
  - Cell is 3-terminal and requires 2 transistors per cell to controllably write and read

- **SOT-MRAM potential advantages:**
  - Higher endurance (as write voltage is never applied across tunnel barrier)
  - Lower write-energy (as write current is applied through low resistance metal instead of high-resistance MTJ)
  - Higher writing speed (as higher overdrive current can be applied due to the above)
  - Lower write error rate (for the same reason as the above)

Figure from: Y. Kim et al., IEEE Trans. Electron Devices 62, 561 (2015)
Physics of the SOT switching

*Spin currents generated via spin orbit coupling*

- For electrical current applied in x-direction, spin Hall effect (SHE) generates spin current of y-polarized spins in the z-direction $J_{yz}$ and z-polarized spins in the y-direction $J_{zy}$.
- $J_{yz}$ can exert spin torque on the FM layer deposited on top.
- The magnitude of spin current depends on the spin Hall angle $\theta_{SH}$ which is materials parameter in the range of 1 – 40%.

\[ J_s = \frac{\hbar}{2e} \theta_{SH} J_e \]

- SOT is most conveniently described physically in terms of the effective SOT field it generates on the FM layer:

\[ H_{SOT} = \frac{\hbar}{2e} \frac{\theta_{SH}^{eff}}{(M_s t)_FL} J_e (\hat{y} \times \textbf{m}_{FL}) \]

- Difference from STT-MRAM: Spin current is polarized in y-direction (unit y-vector in the cross product), even if $\textbf{m}_{FL}$ is perpendicular.
SOT MRAM Flavors

**PMA SOT**
- In-plane field required
- High anisotropy (No shielding)
- **Highest density**
- Potentially faster switching
- Write current increases as the bit shrinks
- Highest write current
- **No solution yet for field-free deterministic write**

**In-plane SOT (Type-y)**
- Deterministic w/o a field
- Lowest write current
- Low anisotropy (field immunity issues)
- Shape anisotropy may limit bit size

**In-plane SOT (Type x)**
- Out-of-plane field required
- Potentially faster switching
- Middle write current
- Requires z-field
- Low anisotropy (field immunity issues)
- Shape anisotropy may limit bit size

Critical write power per cell area

- **STT-MRAM:**
  \[ p = \frac{P}{A} = \frac{V_{c0}^2}{RA} = (RA)J_{c0}^2 \]
  for \( V_{c0} = 0.25 \text{ V} \) and \( RA = 10 \ \Omega \mu \text{m}^2 \)
  \( \Rightarrow p \sim 0.6 \text{ MW/cm}^2 \)

- **SOT-MRAM:**
  \[ p = \frac{P}{A} = \frac{RI_{c0}^2}{Lw} = \frac{\rho L^2 J_{c0}^2}{wtLw} \cdot \frac{t}{t} = (\rho t)J_{c0}^2 \]
  for \( J_{c0} = 2 \times 10^7 \text{ A/cm}^2 \), \( \rho = 100 - 200 \ \mu \Omega \text{cm} \), \( t \sim 5 \text{ nm} \)
  \( \Rightarrow p = 0.02 - 0.04 \text{ MW/cm}^2 \)

So for bits the same size, Pt or β-W SOT devices require 20x times less power, and the energy difference is even greater, since the switching time for SOT is faster.
3-terminal SOT Devices

- Pt based, top pinned, SAF reference layer

10 Ta/60 Pt/10 CoTiB/16.5 CoFeB/2 CoFe/700sec rf-MgO/2 CoFe50/20.5 CoFeB/3 CoFe50/4 Co/8 Ru/4 Co/13 CoFe50/10 Co/100 IrMn/50 NiCr/30 Ru/30 Ta/30 Ru

- SOT pillar is nominally in the center of a 1.05 µm x 1.05 µm Hall cross defined in the Pt SHE channel

![Diagram of SOT Devices with labels and structure.](image-url)
Nanosecond pulsed switching test results

\[ I_c = I_{c0} \left(1 + \frac{\tau_0}{t_p}\right) \]

- W-based - average of 5 devices:
  \[ I_{c0} = (0.51 \pm 0.02) \text{ mA and } \tau_0 = (1.67 \pm 0.21) \text{ ns from ns switching} \]
  \[ I_{c0} = (0.49 \pm 0.02) \text{ mA from thermal activation model } t_p = 100 \text{ ms} \]

- Pt-based - average of 5 devices:
  \[ I_{c0} = (1.13 \pm 0.07) \text{ mA and } \tau_0 = (1.74 \pm 0.30) \text{ ns from ns switching} \]
  \[ I_{c0} = (0.95 \pm 0.03) \text{ mA from thermal activation model } t_p = 100 \text{ ms} \]

- We find no difference in High -> Low vs Low -> High switching for \( I_{c0} \) and \( \tau_0 \)

- SOT switching with 50% probability can be achieved down to \(~1.7\) ns with overdrive currents of \(~2I_{c0}\).
Extension to sub-nanosecond switching times

\[ I_c = I_{c0} \left( 1 + \frac{\tau_0}{t_p} \right) \]

- Red data points were measured using a different pulse generator to extend switching times to the sub-nanosecond regime.
- The black line is fit to the equation above with:
  - \( I_{c0} = 1.15 \text{ mA} \) and \( \tau_0 = 1.8 \text{ ns} \) (Low -> High switching)
  - \( I_{c0} = -1.15 \text{ mA} \) and \( \tau_0 = 1.6 \text{ ns} \) (High -> Low switching)
- The results show that that simple formula \( I_c = I_{c0} \left( 1 + \frac{\tau_0}{t_p} \right) \) describes the results empirically well in the sub-10ns range down to 0.5 ns, with an average \( \tau_0 \approx 1.7 \text{ ns} \) as found above.

- SOT switching with 50% probability can be achieved down to \(~0.5\) ns with overdrive currents of \(\sim 4.5I_{c0}\).
We have demonstrated SOT switching in in-plane (type-y) with sub-nanosecond pulses in 3-T devices down to 30x90 nm with pinning fields > 1 kOe, TMR > 120%, and coercivity up to 250 Oe.

- Both Pt and β-W are potential SOT materials
- The switching between high- and low-resistance states is quite symmetric
- Key future challenges involve further lowering $I_{c0}$ and increasing bit coercivity
- There are opportunities for materials innovations to address these challenges

The most likely path to high density SOT-MRAM involves PMA bits.

- Similar SOT materials may be used
- One needs reasonable switching currents at <40 nm pillar diameters without an external field
- Still need a resolution of the deterministic issue
CoFeB-MgO MTJs exhibit voltage-controlled magnetic anisotropy, where an electric field changes the density of electrons at the CoFeB/MgO interface, and affects the perpendicular anisotropy. A decrease in the electron density at the interface increases perpendicular anisotropy. (Maruyama, et al. Nat. Nano 4, 158 (2009)).

Since this magnetoelectric coupling is not strain-mediated, it is not endurance limited, making it compatible with logic and memory applications.

Unlike an external magnetic field that tilts the energy barrier between P and AP orientation, the voltage controlled anisotropy either reduces or increases the energy barrier, depending on the polarity.

\[
E_i(V) = E_i(0) - \varepsilon V/d, \quad \text{where} \quad d \text{ is free layer thickness,}
\]
\[
\text{and} \quad \varepsilon \text{ is the VCMA coefficient.}
\]
For standard PMA CoFeB/MgO tunnel junctions, \(E_i(0) \sim 1-2 \text{ mJ/m}^2\) and \(\varepsilon \sim 30-50 \text{ fJ/Vm}\).
We demonstrated VCMA-driven writing in 80 nm x 80 nm perpendicular bits with TMR readout. Effect can be scaled to smaller bits if higher VCMA coefficient is achieved.

- Achieved ~100% switching in a time window ~ 700 ± 250 ps.
- Device optimization to achieve a wide write window.
- Very little current flows thru the barrier, so endurance is unlimited, and power is lower than STT-MRAM (power per unit area scales with $V_c^2/RA$)
- Switching energy is ~ 10 fJ/bit, write time is < 1 ns

High RA MgO barrier is used, so negligible current flows. This is an E-Field driven effect.

VCMA effect eliminates perp anisotropy, allowing free layer to precess about $H_{\text{eff}}$.
Write Error Rate vs. Total Write Time

- Note this is a toggle memory
- Write is unipolar, so applying read in opposite direction eliminates read-disturbs
- Write verification is assumed after each VCMA pulse
- Each cycle takes 2.5 ns
- An increased VCMA coefficient is required to reduce the amplitude of the write pulse (dashed lines)

Since the precession period is dependent on the effective field, bit-to-bit variations in $H_{eff}$ are a concern, as is the magnetic field environment of the chip.

Variations in the shape/duration of sub-ns voltage pulses will impact WER across wafer.
MeRAM Requires Large VCMA Coefficients

Regardless of the writing mechanism, maintaining thermal stability as the technology node scales requires larger values of $K_i$.

The present CoFeB/MgO-based systems provide typical values of 30 fJ/Vm, while newer materials have been reported to provide values potentially suitable for 14 nm and below; e.g. Y. Kato, et al. *Applied Physics Express* 11(5):053007 (2018).

Note that the requirements on both $K_i$ and $\varepsilon$ may be relaxed for applications where non-volatility is not required, i.e., where $\Delta < 40$. 

STT + Precessional Switching

- Intermediate RA-product film with high VCMA coefficient is used. Proper polarity is required.
- Initial pulse at positive 1.2V cancels interfacial PMA and magnetization begins to precess.
- Relatively short pulse at lower voltage generates enough current for STT effect to create deterministic switching. Polarity determines switching direction.

- Unlike purely precessional MeRAM, this combination is far less sensitive to the field at the bit and the pulse duration.
- A high VCMA coefficient is required to avoid endurance issues with the initial VCMA pulse.
- The endurance issues are less severe than those associated with overdriving standard STT-MRAM, since the VCMA pulse is quite short. Very short pulse rise times are required for negative values of \( V_{b2} \) to switch reliably.

VCMA MRAM Summary and Outlook

• VCMA provides another “knob” we can use to switch MRAM bits
  – Temporarily lowers the energy barrier, allowing very low-power switching.
  – Can enable a precessional toggle memory
  – Can be combined with other effects to enhance performance

• Large VCMA coefficients are needed for all of these implementations to be practical
  – Important area of MRAM materials research
  – Preliminary results are promising

• This is the most immature of potential MRAM technologies discussed in this talk, and a lot of work is still required to evaluate the feasibility of these technologies.

• Note, though, that VCMA effects are present even in standard STT-MRAM devices, and it may be possible to exploit them to, for example, make switching more symmetric.
Questions ?
What is an STT-MRAM cell

**CMOS transistor + magnetic tunnel junction (MTJ)**

**Physical sketch:**

- Relative orientation between magnetizations of the reference (pinned) layer and free (storage) layer can be set by passing electric current through MTJ via spin transfer torque effect. This enables memory **writing** operation.

- If no current is passed through MTJ, relative orientation between magnetizations of the reference (pinned) layer and free (storage) layer remains unchanged — **nonvolatile storage**.

**Symbolic description**

- Resistance of MTJ depends on relative orientation between magnetizations of the reference (pinned) layer and free (storage) layer. When they are parallel resistance is low (0) and when they are antiparallel resistance is high (1). This enables memory **reading** operation.

Figure from: D. Apalkov et al., Proceedings of IEEE 104, 1796 (2016)
Spin-orbit Torque (SOT) switching due to SHE

- Magnetization in plane
- Current orthogonal to \( \mathbf{m} \)

\[
H_{SOT} \parallel \frac{d\mathbf{m}}{dt}
\]

very efficient antidamping torque similar to STT

\[
H_{SOT}^{ip-y} \sim \alpha \left( H_k + \frac{M_{eff}}{2} \right) \quad \quad \quad H_{SOT}^{ip-x} \sim \left( \frac{M_{eff}}{2} \right)
\]

\( \alpha \sim 0.01 \)

Current orthogonal to \( \mathbf{m} \) \( \rightarrow \) “Type Y”

Current parallel to \( \mathbf{m} \) \( \rightarrow \) “Type X”

\[
H_{SOT} \approx \frac{\hbar \theta_{SH}^{\text{eff}}}{2e (M_{st})_{FL}} J_e (\hat{y} \times \mathbf{m}_{FL})
\]

- PMA Free Layer

\[
\left( H_{SOT}, \frac{d\mathbf{m}}{dt} \right) \approx 0
\]

- inefficient field-like torque
- **not deterministic**
- requires high effective field (i.e. high \( J_e \)) to switch FL magnetization

Current orthogonal to \( \mathbf{m} \) \( \rightarrow \) “Type Y”

Current parallel to \( \mathbf{m} \) \( \rightarrow \) “Type X”

\[
H_{SOT}^{\text{perp}} \approx \frac{1}{2} H_k \perp
\]
Spin-orbit Torque (SOT) Device

- MRAM SOT switching versus STT does not flow current across the barrier so unlimited endurance is possible.
- Switching is extremely fast (< 2 ns) at reasonable current levels.
- Unlike STT-MRAM, in a 3-terminal SOT device, we need 2 transistors (or a transistor and a diode).
- Although the footprint is larger than STT-MRAM, it is still more compact than SRAM.
- Significantly lower power than STT-MRAM as shown on next slide.

TEM Image of 3-T device with top-pinned AP-reference layer and Pt SOT layer.

Stopping the ion mill precisely in a 6 nm Pt layer without producing milling “tails” is a challenge