Special MRAM poster session

IEDM (Dec 10-12, 2019, Hilton Union Square, San-Francisco)

Wednesday morning Dec 11, 9:00am-12:00 noon

Yosemite room

For the 4\textsuperscript{th} consecutive year, a special poster session entirely dedicated to MRAM is organized during IEDM. This session is technically organized by the IEEE Magnetics Society and embedded in the IEDM 2018 conference. This event will be a great opportunity to foster closer interactions between the microelectronics and magnetism communities. The posters will cover a number of topics including MRAM materials, phenomena, technology (STT, SOT, E-field control), testing, hybrid CMOS/MTJ technology and circuits, and MRAM applications. This year, 30 posters were accepted for presentation.

The list is shown below.

Bernard DIENY and Kaizhong GAO
IEEE Magnetics Society

Presented posters:

1. Novel multifunctional RKKY coupling layer for ultrathin perpendicular synthetic antiferromagnet for easy integration of STT-MRAM cells

Jyotirmoy Chatterjee*, Stephane Auffret, Ricardo Sousa, Paulo Coelho, Ioan-Lucian Prejbeanu and Bernard Dieny
Univ. Grenoble Alpes, CEA, CNRS, IRIG-SPINTEC, Grenoble France
*At present: Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, 94720, USA

A novel multi-functional antiferromagnetic coupling layer (MF-AFC) combining Ru and W was developed to realize an extremely thin (3.8 nm), back-end-of-line (BEOL) compatible as well as magnetically and electrically stable perpendicular synthetic antiferromagnetic layer (pSAF), essential for spintronic memory and logic device applications. This thin pSAF layer, comprising Si/Ta$_3$/Pt$_5$/[Co$_{0.5}$/Pt$_{0.25}$]$_3$/Co$_{0.5}$/Ru$_{0.4}$/W$_{0.2}$/FeCoB$_{1.2}$/MgO/cap layer (thicknesses are in nm) with various thicknesses of Ru and W were deposited and studied.

In addition to achieving antiferromagnetic RKKY coupling, this MF-AFC also acts as a Boron sink and texture-breaking layer. A detailed optimization of the thickness of the various involved layers has been carried out to obtain extremely thin-pSAF reference layer with stable magnetic properties, which enables the realization of sub-20 nm STT-MRAM cells.

The maximum RKKY coupling energy obtained using Ru/W is 0.86 erg/cm$^2$ after 340\degree C annealing which is slightly larger than the values obtained at the second peak of Ru (0.8-0.9 nm). The optimized thickness of Ru and W are 4 Å and 2 Å respectively to obtain BEOL compatible thin-pSAF. These optimizations enable to obtain an extremely thin, 3.8 nm thick pSAF layer ([Co$_{0.5}$/Pt$_{0.25}$]$_2$/Co$_{0.5}$/Ru$_{0.4}$/W$_{0.2}$/FeCoB 1.2 nm) stable up to 3.3 kOe. Functional STT-MRAM cells were patterned, demonstrating the stability of the
reference layer against applied write voltage pulses up to 1 V and exhibiting very low dipolar field on the storage layer (< 200 Oe) for sub-20nm electrical diameters.

Two important advantages are provided by this ultrathin reference layer: the easing of the reference layer etching thanks to its reduced thickness and also the minimization of the dipolar field acting on the storage layer’s magnetization.

**Figure 1.** RKKY coupling energy density ($J_{RKKY}$) of the pSAF versus thicknesses of Ru and W layers of MF-AFC after annealing at 340°C.

**Figure 2.** Descending branch of M(H) loops of thin-pSAF of composition: Si/SiO$_2$ 500/Ta 3/Pt (t)/[Co 0.5/Pt 0.25]$_3$/Co 0.5 /Ru 0.4/W 0.2/ FeCoB 1.2/ MgO/cap layers (thicknesses are in nm) with different Pt thicknesses, t = 5, 10, 20, 30 nm after annealing at 340°C. The arrows indicate the magnetization reversal process along the field-sweep from +6 kOe to -6 kOe. Inset of this figure displays the minor loops of FeCoB PLs.

**Figure 3.** Voltage-field stability diagram of a patterned magnetic tunnel junction with 53 nm electrical diameter. This diagram exhibits enough stability of the ultrathin-pSAF with Ru0.4nm/W0.2nm multifunctional AF coupling layer.
2. Novel seed-less multilayers for BEOL compatible spintronic memory devices

Jyotirmoy Chatterjee*, Paulo Coelho, Antoine Chavant, Stephane Auffret, Ricardo Sousa, Claire Baraduc, Ioan-Lucian Prejbeanu and Bernard Dieny
Univ. Grenoble Alpes, CEA, CNRS, IRIG-SPINTEC, Grenoble France
*At present: Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, 94720, USA

We report a novel seedless multilayer (SL-ML) of the form [Co/Insertion layer/Pt]$_n$ grown on top of half-magnetic tunnel junction (half-MTJ) (MgO/FeCoB/texture breaking layer). Among the various metals investigated, the Ta insertion produced high effective perpendicular anisotropy ($K_{\text{eff}}$) and almost 100% square out-of-plane M(H) loops after annealing at 400°C as shown in Fig. 1. Thicknesses of different components of the SL-ML, e.g. texture breaking layer, Co, Ta-insertion layer and Pt layer, have been optimized. A back-end-of-line (BEOL) compatible hard top reference layer (RL) using this SL-MLs is achieved, which exhibits more than twice effective perpendicular anisotropy compared to conventional top reference layer.

Using this RL, we demonstrate (a) top-pinned MTJ stacks able to withstand annealing at 425°C, (b) back-end-of-line (BEOL) compatible double-MTJ memory cell exhibiting high-voltage reading and low voltage writing and (c) 2-bit MTJ stack combining STT and SOT writings.

Hence, the proposed SL-MLs is promising for different spintronic memory or logic applications requiring top hard reference layer.

Figure 1. (a) Out-of-plane, (b) in-plane magnetic cycle when the field is varied from positive to negative, (c) $M_s$ and (d) $K_{\text{eff}}$ of “Ta/FeCoB 0.3/MgO/FeCoB 1.1/ W 0.6/[Co 0.6/ I]/Pt 1.7/Ru 5 nm” MLs with different insertion (I) layers after annealing at 400°C for 10mins. $K_{\text{eff}}$ and saturation magnetization ($M_s$) were calculated considering the total thickness of Co, Pt and FeCoB layers.

Figure 2. (a) R(H) loops in read and write modes of patterned 50 nm memory cell fabricated out of double-MTJ stack. Read mode exhibit four times larger TMR than that of write mode as expected. Voltage-field phase diagram of a double-MTJ memory cell in (b) read and (c) write mode respectively exhibiting the average critical voltage for switching in read mode is significantly larger than that in write mode.
3. Dielectric breakdown of MgAl2O4 based magnetic tunnel junctions using DC and 10 ns voltage pulses

Hiroaki Sukegawa1,*, Zhenchao Wen1, Shinya Kasai1, Seiji Mitani1, Shinto Ichikawa2, Katsuyuki Nakada1,∗∗
1 Research Center for Magnetic and Spintronic Materials, National Institute for Materials Science, 1-2-1 Sengen, Tsukuba 305-0047, Japan
2 Advanced Products Development Center, Technology & Intellectual Property HQ, TDK Corporation, 2-15-7 Higashi-Ohwada, Ichikawa 272-8558, Japan,
e-mail: * SUKEGAWA.Hiroaki@nims.go.jp, ** knakada@jp.tdk.com

Improving breakdown voltages of magnetic tunnel junctions (MTJs) is one of the most important issues for various spintronic applications such as spin-transfer torque type magnetoresistive random access memories (STT-MRAMs). Although most applications use MgO-based MTJs, it is very difficult to maintain a high breakdown voltage when the MgO thickness is less than 1 nm. Also, the effect of lattice mismatch with CoFeB electrodes is not negligible since introduced misfit dislocations work as electron trapping sites that significantly reduce the breakdown voltage. Use of a spinel MgAl$_2$O$_4$ (MAO) barrier provides perfectly lattice-matched interfaces with CoFeB [1], leading to elimination of misfit-driven breakdown by electric stress [2]. Therefore, in this study, we focused on breakdown voltages of lattice-matched MAO-MTJs with resistance area ($R_A$) values between 3 and 10 $\Omega \cdot \mu$m$^2$ using DC and 10-ns voltage pulses. We prepared epitaxial CoFe/MAO/CoFe(001) MTJ stacks on an MgO(001) substrate using magnetron sputtering. An MAO barrier was formed using natural oxidation of Mg-Al. A tunnel magnetoresistance (TMR) ratio and $R_A$ were evaluated by current-in-plane technique (CIPT) to be 155% (220%) and 3.3 $\Omega \cdot \mu$m$^2$ ($7 \Omega \cdot \mu$m$^2$), respectively. Breakdown voltages were evaluated using ramping DC voltages and 10-ns voltage pulses after patterning into 200×100 nm$^2$ scale junctions. We observed breakdown voltages of ~1.1 V by DC and ~1.5 V by 10-ns pulses for MTJs with 3.3 $\Omega \cdot \mu$m$^2$. In the poster session, we discuss about the TMR and breakdown characteristics of our MAO-based MTJs in detail.


4. Dielectric breakdown of MgO in MRAM

A. Teramoto1,2, J. Tsuchimoto2,3, H. Park2, M. Hayashi1,2, K. Tsunekawa2, T. Suwa2, R. Kuroda2, S. Sugawa2
1 Hiroshima University, 2 Tohoku University, 3 CANON ANELVA CORPORATION
The dielectric breakdown of MgO film is one of the most crucial issues for the reliability of MRAMs. We have investigated the dependence of formation method of MgO, roughness of the bottom electrode, and edge length on the breakdown characteristics of MgO by using the Metal-Insulator-Metal (MIM: CoFeB/MgO/CoFeB) capacitors [1-3]. The crystallization of the MgO film itself and top and bottom electrode show the large impact of the lifetime of MgO films [1]. The microroughness of bottom electrode induces the excess current of MIM capacitors and lowering breakdown voltage because of the electric field concentration [2]. When the metal is adsorbed on the side wall of MgO film at the etching process of bottom electrode for formation of magnetic tunnel junction cells the excess leakage current is increased. We have reported the edge defect can be suppressed by the accurate control of the bottom electrode etching condition [3]. However, if this leakage current cannot be suppressed, this current impacts on the lifetime of MIM capacitor strongly [3] although the magnetoresistance (MR) ratio is not affected by this leakage current. Because the MR ratio can be affected only in a cell, in contrast, we must take the all edge length in the memory device into the lifetime expectation. Fig. 1 shows the Failure rate of MgO films with metal adsorption on the sidewall as a function of the edge length. Our experimental results include the edge length of less than 500 μm, however, the edge length varied from several meters to several hundred meters for Mbits to Gbits MRAMs, and then the reliability of real MRAM becomes very poor. The suppression of metal adsorption on the MgO sidewall at the bottom etching is very important for maintaining the MRAM reliability.


5. Thermal and Reliability Modeling of FinFET-Driven STT-pMTJ Array Considering Mutual Coupling, 3D Heat Flow, and BEOL Effects

Ya-Jui Tsou, Chia-Che Chung, Jih-Chao Chiu, Huan-Chi Shih, and C. W. Liu*
Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan  *cliu@ntu.edu.tw

Self-heating of vertically-stacked magnetic tunnel junctions (MTJs) and scaled FinFETs causes the increase of MTJ temperature ($T_{MTJ}$), which degrades time-dependent dielectric breakdown (TDBB) of MgO and read disturbance rate (RDR). To simulate the temperature evolution of MTJ, the multi-step pseudo isothermal plane model is used to calculate the equivalent thermal resistance ($R_{th}$) of BEOL plane by plane (Fig. 1(a)~(c)). A non-flat isothermal plane is obtained (Fig. 1(b)) due to the asymmetric BEOL structure and the low thermal conductivity of MTJ [1]. The modeled $R_{th, BEOL}$ is further verified by 3D thermal TCAD (Fig. 1(d)). The transient thermal behavior of FinFETs is accurately modeled by our distributed $R_{th,Ct}$ network [2,3]. The Joule heat waveform is determined by the hybrid FinFET/MTJ circuit, including incubation delay and transition of MTJ switching by the LLG-based model. The heat flow along vertical and lateral directions in an array with flip-chip package is considered. The $T_{MTJ}$ decreases with the increasing number of rows in an array and reaches the saturation. The mutual thermal coupling between MTJ and FinFET through BEOL leads to an additional increase of $T_{MTJ}$ (Fig. 2(a)) and FinFET junction temperature ($T_j$) (Fig. 2(b)). The self-heating can be mitigated by increasing wait time, increasing via density, and relocating MTJ to higher metal levels in BEOL. The TDBB time-to-failure (TTF) of MgO and RDR (Neel-Brown model [4]) can therefore be improved (Fig. 2(c)).
Acknowledgement – This work is supported in part by MOST, Taiwan (108-2622-8-002-016-, 108-2218-E-002-027-, and 108-2218-E-002-016-) and MOE, Taiwan (NTU-CC-108L891701).


Figure 1. Envelope of (a) transient TMTJ and (b) transient FinFET Tj with and without mutual thermal coupling. (c) Improvement of TDDB TTF and RDR.

Figure 2. (a) Structure of BEOL and MTJ in Rth simulation by TCAD. (b) Constant temperature contour on A-A' plane and B-B' plane. (c) Multi-step pseudo isothermal plane modeling of Rth, BEOL. (d) Rth, M1~M5 values obtained from different methods.
Critical current ($I_C$) distributions of STT-pMTJ considering thermal fluctuations and process variations are investigated using the LLG-based physical model with Monte Carlo simulation. The thermal fluctuations (intrinsinc variation of MTJ) lead to the different initial angle ($\theta_0$) of magnetization in each switching process [1]. The smaller $\theta_0$ requires the larger current to complete the STT switching due to the smaller initial spin torque ($\tau \propto \sin \theta$). The decrease of write time ($t_w$) increases the mean of $I_C$ and broadens the $I_C$ distribution (Fig. 1(a)). The extrinsic variations originate from the process variations such as free layer (FL) thickness and MTJ diameter. The Gaussian distributions are used to depict the process variations with the variability of 1% ~ 5%. The FL volume dependence of damping constant due to spin pumping [2] is further taken into account, where the damping constant increases with the decreasing FL thickness and the decreasing MTJ diameter. The $I_C$ of STT-pMTJ is more sensitive to the variation of FL thickness (Fig. 1(b)), since the FL thickness highly affects the interfacial perpendicular magnetic anisotropy. Considering the total variation with the process variability of 1%, the $I_C$ variability is dominated by the variation of initial angle (Fig. 2(a)). For the process variability of 5%, the variation of FL thickness dominates the $I_C$ variability. The write current to reach the write error rate (WER) of $10^{-6}$ is extrapolated from the simulated WER data (Fig. 2(b)). The $t_w$ strongly affects the write current with the WER of $10^{-6}$.

Acknowledgement This work is supported in part by MOST, Taiwan (108-2622-8-002-016-, 108-2218-E-002-027-, and 108-2218-E-002-016-) and MOE, Taiwan (NTU-CC-108L891701).

7. Demonstration of high write endurance and low write bit error rate with spin current type magnetic memory array

Advanced Products Development Center, TDK Corporation, Chiba, Japan
E-mail: shiokawa@jp.tdk.com

Recently, three-terminal spin current type magnetic memory that use spin orbit torque (SOT) has attracted much attention because of its high speed magnetization switching (<1 ns) [1]. In previous work, we showed high write endurance of this spin current type memory in 64 elements array with type-Y magnetic tunnel junction (MTJ) [2]. However, an issue of thermal attack to W-SOT line by Joule heating due to high resistivity of $\beta$-W was clarified. In this study, we demonstrate the characteristic that satisfied both high write endurance and low write bit error rate (WER). Furthermore, we propose new multi-layer structure of SOT line for lower resistivity in order to achieve higher write endurance. Figure 1 (a) shows the low WER of under $10^{-7}$ with write pulse width of 10 ns in 64 elements array. Figure 1 (b) shows the result of the write endurance test in one MTJ element of the 64 elements array under the write condition of low WER. In the presentation, more detail and new multi-layer structure of SOT line will be discussed.

Reference

![Figure 1](image)

Fig 1 (a) WER in 64 elements array, (b) Write endurance result in one MTJ element of the 64 elements array under the write condition of low WER.

8. Magnetization switching and thermal stability in Perpendicular Magnetic Tunnel Junctions with surface defects

Iana Volvach$^{1,3}$ and Vitaliy Lomakin$^{2,3}$

$^1$Material Science and Engineering Department, University of California at San Diego, La Jolla, California 92093, USA
$^2$Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA
$^3$Center for Memory and Recording Research, La Jolla, CA 92093, USA

Magnetic devices, such as perpendicular magnetic tunnel junctions (p-MTJ) or magnetic random access memories (MRAM) experimentally have an arbitrary shapes with various edge and surface defects. Accurate models are needed to enable the simulation of novel magnetic devices with the arbitrary shapes and defects. We developed the LLG simulator, which self-consistently models the dynamics of the magnetic device and present results of using this simulator for accurate modeling of p-MTJ switching to understand the effect of the surface defects. The simulator uses FastMag micromagnetic solver [1] for
considering complex dynamics in p-MTJ, accounting for magnetoresistance, non-uniform magnetization, finite temperature, and complex material and geometry composition. In the presented approach a magnetic device is represented as regular p-MTJ stack: free layer (FL) and reference layer (RL) separated by nonmagnetic spacer. It has been noticed that the switching dynamics has noticeable effects on non-uniform switching dynamics, thermal stability, critical current density and figure of merit when the p-MTJ stack has surface defects. Fig. 1. (a) Shows the minimal energy path for the devices with and without any surface defect respectively. Fig. 1. (b) shows the difference in slopes for the probability of not switching for the devices with and without surface defect. The switching current density for the device with the surface defects is increasing compared to the device without any surface defect.

Fig. 1. Energy barrier path for the device with and without surface defect (b) Probability of not switching for the pulse tpw = 5 ns for a device with D = 40 nm and tR = 1.2 nm. The magnetic properties include the magnetic exchange constant \( A_s = 1.0 \times 10^{-8} \text{ erg/cm} \), saturation magnetization \( M_s = 1200 \text{emu/cm}^3 \), surface anisotropy \( K_s = 1.088 \text{erg/cm}^2 \), and damping constant \( \alpha = 0.008 \).

References:

9. Recent Development of Magnetic Tunnel Junction with an MgO Tunnel Barrier Formed by Post-oxidation Process for High Volume Manufacturing of STT-MRAM

H. Tomita1,2, K. Nakamura1, K. Nagasaki1,2, S. Bosu1,2, K. Ando1, T. Kitada1, A. Gomi1, H. Maehara1,2, and N. Watanabe1
1Tokyo Electron Technology Solutions Limited
2Tokyo Electron Limited
3National Institute of Advanced Industrial Science and Technology (AIST), Spintronics Research Center
E-mail: hiroyuki.tomita@tel.com

Spin-transfer torque magnetic random access memory (STT-MRAM) has been developed for emerging non-volatile memory as a replacement of flash memory and SRAM. In order to realize high volume manufacturing (HVM) of STT-MRAM, PVD tool for deposition of magnetic tunnel junctions (MTJ) with an MgO tunnel barrier must be able to meet the requirements such as device performance, cost and yield. In particular, reduction of particle defect originating from MgO fabrication process can bring in high yield. Basically, fabrication of the MgO using post-oxidation process is superior to RF sputtering method in terms of the particles. We have developed the fabrication process using post-oxidation which can achieve not only low particle level but also high quality MgO tunnel barrier. Device performance of p-MTJs with the MgO formed by the post-oxidation process will be discussed in the presentation.

10. From Spin Transfer Torque (STT) to Toggle Spin Torques (TST) for Ultra-fast Computing with MRAM
A combination of spin–orbit and spin-transfer torques can be used to create a versatile magnetic random access memory. Spin-orbit torque and spin-transfer torque are leading the pathway to the future of spintronic memories. However, both of the mechanisms are suffering from intrinsic limitations. In particular, an external magnetic field is required for spin-orbit torque to execute deterministic switching in perpendicular magnetic tunnel junctions; the demand for reduced spin-transfer torque switching current to realize low power and high endurance is still remaining. Thus, a more advanced switching mechanism is urgently needed to move forward spintronics for wide applications. Here, we experimentally demonstrate the field-free switching of three-terminal perpendicular-anisotropy nanopillar devices through the interaction between spin-orbit and spin-transfer torques. Increasing reduces the threshold current density of spin-orbit torque switching that of spin-transfer torque (see Fig. 1), hence a compromise for low-power p-MTJ switching can be achieved by tuning the two current densities (1). The Toggle Spin Torque switching methods (see Fig. 2) allow the ultra-fast switching and GHz computing in memory (2).

Due to the advantages of non-volatility, unlimited erasure and fast writing, magnetic random-access memory (MRAM) now is emerging as the next generation of low power general memory. The magnetic tunnel junction (MTJ) composed of ferromagnetic layer, dielectric tunneling barrier layer and ferromagnetic layer, is critical for STT-MRAM devices performance. However, for now, STT-MRAM is still hard in production for it has challenges to overcome. Most MTJ metal materials do not form volatile reactants, resulting in a large number of reaction residues by conventional etch process. To obtain excellent MRAM device performance, processing MTJ stacks requires specially developed nano-patterning solution. In industry, both reactive ion etch (RIE) and ion beam etch (IBE) processes were under research to overcome the MTJ etch bottleneck. Although both of them work for relatively isolated and large devices, none of them works for sub-30nm dense array structures.

In this paper, a combined etch solution was proposed for the patterning of perpendicular MTJ stacks. The system includes a reactive ion etch (RIE) chamber and an ion beam etch (IBE) chamber, which is designed for metal physical etch. By combining RIE & IBE systems, it can solve most of the challenges such as sidewall re-deposition and damage. And the throughput is able to be much higher. One possible etch scenario of the system is: (1) RIE is used as fast main etch, (2) IBE is used to clean MTJ sidewall, and (3) in-situ encapsulation is used to protect MTJ to be exposed to ambient. This approach largely limits IBE shading effect and provides a high throughput. In addition, due to the fact that the main etch is performed in RIE, most metal is etched in RIE. Thus, chamber contamination is mainly in RIE chamber, whose chamber clean process is well developed. Under optimized etch process, sub-30nm dense-array MTJ devices were demonstrated to be well patterned, without sidewall re-deposition, with minimum sidewall damage, and leaving a flat etch front. Electrical analysis on devices demonstrated comparable TMR after etch with the TMR on blanket MTJ film, and no shorts in all measured devices, indicating a good device performance. This approach leads to solutions to high density dense array devices, for both embedded and standalone MRAM.

12. An optimized method for critical cleaning of sidewall residues in the fabrication of 8-inch CMOS compatible STT-MRAM

Meiyin Yang\textsuperscript{1}, Yan Cui\textsuperscript{1}, Tengzhi Yang\textsuperscript{1,2}, Kaihua Cao\textsuperscript{3}, Zongxia Guo\textsuperscript{3}, Jianfeng Gao\textsuperscript{1}, Xiaobin He\textsuperscript{1}, Junfeng Li\textsuperscript{1}, Wenwu Wang\textsuperscript{1,2}, Dongchen Che\textsuperscript{3}, Kaidong Xu\textsuperscript{5}, Chao Zhao\textsuperscript{1,2}, Weisheng Zhao\textsuperscript{1,4}, and Jun Luo\textsuperscript{1,2* (present)}

\textsuperscript{1} Key Laboratory of Microelectronic Devices & Integrated Technology, Institute of Microelectronics, Chinese Academy of Sciences (IMECAS), Beijing, P. R. China
\textsuperscript{2} University of Chinese of Academy Sciences (UCAS), Beijing, P. R. China
\textsuperscript{3} Fert Beijing Institute, BDBC, and School of Electronic and Information Engineering, Beihang University, Beijing, P. R. China
\textsuperscript{4} Beihang-Geortek Joint Microelectronics Institute, Qingdao Research Institute, Beihang University, Qingdao, P.R. China
\textsuperscript{5} Jiangsu Leuven Instruments Co., Ltd. Jiangsu Province, P.R. China
E-mail: luojun@ime.ac.cn; chaochao@ime.ac.cn

As one of the promising non-volatile memories, STT-MRAM has gained extensive attention in recent years. Although prototype chips have been realized in the last few decades there are still several challenges to be resolved for producing the STT-MRAM with smaller size and larger capacity. One of challenges is the magnetic tunneling junction (MTJ) side-wall cleaning process. As the MTJ is mainly composed of ferromagnetic materials such as Co, Fe, Pt, Ru and Mn which will yield low-volatility by-products using conventional plasma etch chemistries during etching process, and the by-products will re-deposit on the side-walls of the MTJ, especially for high-density MTJ cell arrays. Side-wall residues may cause many problems such as shorts and cone-shaped MTJ which are undesired in production. Thus, high-precision side-wall cleaning technology is needed. Ion Beam Etching (IBE) technology is a suitable solution. The principle of IBE is different from Inductively Coupled Plasma Etching (ICP), the substrate of IBE can be tilted and rotated, which is helpful to remove residues on the side-wall of MTJ.
In this paper, we have successfully developed the optimal MTJ side-wall cleaning method with ion beam etching (IBE) for the sub 100 nm MTJ array etching process compatible with 8-inch CMOS technology. We experimentally compare the morphology characteristics of the MTJ array by transmission electron microscope (TEM) under different side-wall cleaning methods and eventually find the optimal method to remove the by-products. The influences of IBE incident angle, etching power and etching time on side-wall residue and bottom electrode of MTJ are investigated. By the optimization of etching condition, we obtain the high density MTJ array without side-wall residue. Our results are helpful for 8-inch high density STT-MRAM Back-End-of-Line integrate process.

13. Proposal of quad-interface MTJ technology with high performance and its demonstration

K. Nishioka1, H. Honjo1, S. Ikeda1-4, T. Watanabe1, S. Miura1, H. Inoue1, T. Tanigawa1, Y. Noguchi1, M. Yasuhi1, H. Sato1-4, and T. Endoh1-5
1Center for Innovative Integrated Electronic Systems, Tohoku University, Sendai 980-8572, Japan, 2Center for Science and Innovation Spintronics (Core Research Cluster), Tohoku University, 3Center for Spintronics Research Network, Tohoku University, 4Research Institute of Electrical Communication, Tohoku University, 5Graduate School of Engineering, Tohoku University

Double CoFeB/MgO interface magnetic tunnel junction (double MTJ) technology utilizing interfacial perpendicular magnetic anisotropy [1-3] has been applied to embedded STT MRAM [4-6] for SRAM/e-Flash replacement. To overcome the scaling issue of the STT-MRAM beyond 20 nm, it is required to develop a novel technology which can offer both high thermal stability factor $\Delta$ and high switching efficiency $\Delta/V_{CO}$. We have proposed novel quad CoFeB/MgO interface MTJ (quad MTJ) technology which brought forth an increase of both $\Delta$ and $\Delta/V_{CO}$ compared with conventional double MTJ technology. By developing the quad MTJs using 300 mm process based on a novel low damage integration process including PVD, RIE and so on [7], we have achieved about two times larger $\Delta$ and $\Delta/V_{CO}$ compared with referred double MTJs [8]. Moreover, by the stack development specific for the quad MTJ technology, we have achieved low $RA$ while keeping high TMR ratio despite increasing the number of MgO layers in the MTJ. The developed quad MTJ technology will become an essential technology for the scaling of the STT-MRAM beyond 20 nm.

This work is supported by CIES’s Industrial Affiliation on STT MRAM program, CAO-SIP, and JST-OPERA.

References

14. Demonstration of hybrid free layer concept for spintronic devices in large scale integration

E. Raymenants1,2, D. Wan1, S. Couet1, L. Souriau1, A. Thiam1, O. Bulynck1,2, D. Lin1, I. Asselberghs1, M. Heyns1,2, I.P. Radu1 and V.D. Nguyen1
1imec, Leuven, Belgium, email: van.dai.nguyen@imec.be
2KU Leuven, 3001 Leuven, Belgium

A perpendicular magnetic tunnel junction (pMTJ) based on interfacial perpendicular magnetic anisotropy (PMA) CoFeB/MgO free layer is not only at the core of emerging magnetic memories but is also the key
component of integrated spin logic devices thanks to its fast and reliable reading and low writing current. In domain-wall (DW) based spin logic devices, information is encoded and is carried by the position of a DW in the nanotrack shared by multiple MTJs. Thanks to its non-volatility, potentially low power and scalability, such devices are currently explored as an alternative path to circumvent the scaling limitations of CMOS [1]. However, recent experimental results have shown that the use of a conventional pMTJs design for spin logic devices poses some challenges which are mainly related to the slow DW speed in CoFeB/MgO free layer and difficulties of fabricating such devices using industrial integration platforms [2]. Here, we demonstrate that a full functional DW-based device can be realized using a novel engineered pMTJ which incorporates a synthetic antiferromagnet (SAF) [3] as a free layer (FL) without compromising tunneling magnetoresistance (TMR) readout signal. The DW conduit in such device can be decoupled from MgO-induced interfacial PMA, which allows a significantly wider process window and opens the door for large-scale integration of spintronic devices. Moreover, high DW speed driven by spin orbit torque (SOT) which is desired for fast logic operation, is proved. Therefore, the demonstration of fully integrated DW-based devices in imec’s 300 nm CMOS fab flow using a SAF-HFL design leads the way towards the future development of ultra-fast and more efficient spintronic devices, such as racetrack memory, artificial neural networks and spin torque majority gates.

Acknowledgments:
This work was performed as part of the imec IIAP core CMOS and the Beyond CMOS program of Intel Corporation. The authors gratefully acknowledge V. De Heyn and the P-line for operational support, and MCA team for TEM images. E. Raymenants gratefully acknowledges FWO Flanders

References

15. Materials Requirements of High-Speed and Low-Power Spin-Orbit-Torque Magnetic Random-Access Memory

Xiang Li1,2*, Shy-Jay Lin1, Mahendra DC1, Yu-Ching Liao4, Chengyang Yao2, Azad Naeemi4, Wilman Tsai1, Shan X. Wang1,2*
1Materials Science and Engineering, Stanford University, Stanford, CA, USA
2Electrical Engineering, Stanford University, Stanford, CA, USA
3Taiwan Semiconductor Manufacturing Company (TSMC), Hsinchu, Taiwan
4Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA
Email: xiangsli@stanford.edu ; sxwang@stanford.edu

As spin-orbit-torque magnetic random-access memory (SOT-MRAM) is being intensively researched as the next-generation low-power and high-speed on-chip cache memory applications, it is critical to analyze the magnetic tunnel junction (MTJ) properties needed to achieve sub-ns and sub-fJ write operation when integrated with CMOS access transistors. Recently, abundant materials research has demonstrated SOT material that can generate close and larger than 1 charge-to-spin conversion efficiency ($\xi_{ST}$). Among the various SOT materials, heavy metal such as Pt and W, [1, 2] and topological materials such as Bi$_2$Se$_3$ and WTe$_2$ are of particular interest [3, 4]. Nevertheless, no work has integrated these materials into a practical CMOS transistor-MTJ cell to study the energy and speed performance. Thus, the write energy benchmarking conducted thus far [5, 6] remain rather qualitative without considering the impact from the access transistor such as limited current drive and nonzero transistor resistance.

In this paper, we first introduce the framework used to model the 2T-1MTJ cell, SOT-induced switching current, and current distribution along the write path. Then, the SOT material requirements for SOT-MRAM to achieve sub-ns and sub-fJ write operation based on the above framework are discussed, namely
the resistivity, charge-to-spin conversion efficiency, spin Hall conductivity, and thickness. We also benchmark write energy and speed performances of SOT-MRAM cells using several existing SOT materials based on the established framework, including Pt, W, WTe$_2$, and Bi$_{(1-x)}$Se$_x$. This work will provide important guidelines for SOT-MRAM materials and device research in the future to achieve low write current and power.


16. Enhanced Spin-Orbit Torques and Multi-Level Current-Induced Switching in W/CoTb/Pt Heterostructure

Zhenyi Zheng$^{1,2}$, Yue Zhang$^{1,2}$, Xueqiang Feng$^{1}$, Kun Zhang$^{1}$, Jiang Nan$^{1}$, Zhizhong Zhang$^{1,2}$, Guanda Wang$^{1}$, Jinkai Wang$^{1}$, Na Lei$^{1}$, Dijun Liu$^{1}$, Youguang Zhang$^{1,2}$ and Weisheng Zhao$^{1,*}$

$^1$Fert Beijing Research Institute, BDBC, School of Microelectronics, Beihang University, 100191, Beijing, P. R. China
$^2$School of Electronics and Information Engineering, Beihang University, 100191, Beijing, P. R. China
*email: yz@buaa.edu.cn, weisheng.zhao@buaa.edu.cn

Spin-orbit Torque (SOT) effect based Magnetic Random-Access Memory (MRAM) is a promising next-generation electronic device with low power consumption and non-volatility [1-2]. However, order of gigahertz is a bottleneck for the working frequency of MRAM, owing to the limited natural magnetic dynamics of ferromagnets. Thus, investigating novel material and new structure in SOT system is attached of great importance and can certainly contribute to the upgrading of MRAM. Recently, rare-earth-transition-metal (RE-TM) ferrimagnets (FIMs) have attracted enormous attentions, since its working frequency can exceed terahertz near compensation point while still being detection-friendly [3-4]. However, the limited effective spin Hall angle remains a problem in near-compensated heavy metal/FIM system, owing to the net magnetization reduction and small spin mixing-conductance [5-6]. This reality degrades FIM’s future application potential.

Here, we investigate enhanced SOT effect in a heterostructure composed of a near-compensated CoTb layer sandwiched by W and Pt layers (schematic in Fig1.a), where CoTb is a typical RE-TM FIM with bulk perpendicular magnetic anisotropy and W has the opposite spin Hall angle against Pt [7-8]. As shown in Fig 1.b, beyond a large SOT efficiency (58.6 $\times$ 10$^{-10}$ Oe m$^2$/A), 1.9 times effective spin Hall angle and 20% switching power consumption can also be achieved simultaneously by optimizing the thickness of Pt layer, compared with SOT performance in conventional W/CoTb structure.

To introduce this tri-layer structure into potential application, we further investigated current-induced multi-level magnetization switching in Pt(3nm) / CoTb(2nm) / W(5nm) / CoTb(4nm) / Pt(3nm) stack (see image in Fig2.a). Anomalous Hall effect curve in Fig2.b confirms that zero coupling effect exists between two CoTb layers. By sweeping current pulse under a certain in-plane magnetic field along the current direction, four magnetization states ($\uparrow$CoTb$^{\text{thin}}$, CoTb$^{\text{thick}}$, ↓$\downarrow$), ($\downarrow$CoTb$^{\text{thin}}$, CoTb$^{\text{thick}}$, ↑$\uparrow$), ($\uparrow$CoTb$^{\text{thin}}$, CoTb$^{\text{thick}}$, ↓$\downarrow$) and ($\downarrow$CoTb$^{\text{thin}}$, CoTb$^{\text{thick}}$, ↑$\uparrow$) can be achieved in sequence (see Fig2.c), which doubles the capacity of normal SOT devices.

Our work combining efficient and multi-bit magnetization switching provides an alternative route to leverage ferrimagnetic material in future SOT-based magnetic memory devices.
Fig. 1 (a) Schematic of the heterostructure. (b) Effective spin Hall angle and power consumption tendencies versus Pt thickness.

Fig. 2 (a) Cross-section image of the stack. (b) Anomalous Hall effect measurement of the fabricated Hall bar. (c) Four-state current-induced magnetization switching phenomena.


17. Engineering of heavy metal/ferromagnetic metal interface for high-performance MRAM

Shouzhong Peng1,2, Daoqian Zhu1, Jiaqi Zhou1, Boyu Zhang1, Anni Cao1, Mengxing Wang1, Wenlong Cai1, Kaishua Cao1, and Weisheng Zhao1,2,*
Magnetic random-access memory (MRAM) attracts considerable attention due to features like nonvolatility, high scalability, low power, and high speed. Over the past few years, innovative materials and new structures in this field have stimulated emerging phenomena and led to exciting performances. Here we demonstrate that the heavy metal (HM)/ferromagnetic metal (FM) interface is playing an essential role in MRAM devices and the device properties can be significantly improved by properly engineering this interface.[1-6] Firstly, we investigated the effect of HM/FM interface on the perpendicular magnetic anisotropy (PMA) of several structures.[4-6] Then experimental and theoretical investigations about the modulation of HM/FM interface for high tunnel magnetoresistance (TMR) ratio were presented.[2] Next, magnetic damping constant and spin-orbit torque (SOT) switching were discussed with emphasis on the low-power SOT switching.[3] Subsequently, the investigations of Dzyaloshinskii-Moriya interaction (DMI) in different HM/FM systems were presented. Finally, the remaining problems and future directions were discussed. This work may provide guidelines for realizing high-performance MRAM devices as well as an outlook of future research and potential applications.

Fig. 1. A schematic illustration of the HM/FM interface (inner ellipse) and the related properties of spintronic devices (outer ellipse).


18. A statistical study of the reliability of step-structure SOT MRAM cells by thermal baking

Ding-Yeong Wang1,*, I-Jung Wang1, Jeng-Hua Wei1, Shakh Ziaur Rahaman1, Shan-Yi Yang1, Yao-Jen Chang1, Yu-Chen Hsin1, Hsin-Han Lee1, Yi-Ching Kuo1, Yi-Hui Su1, Guan-Long Chen1, Fang-Ming Chen1, Shyh-Shuyuan Sheu1, Wei-Chung Lo1, Chih-I Wu1,2, Duan-Li Deng1

1Electronic and Optoelectronic System Research Laboratories, Industrial Technology Research Institute (ITRI), Hsinchu 310, Taiwan, R.O.C. 2Graduate Institute of Photonics and Optoelectronics, National Taiwan University (NTU), Taipei 106, Taiwan, R.O.C.

E-mail: *DY.Wang@itri.org.tw
This paper presents a reliability study of SOT MRAM cells. Thermal baking reveals two types of bit fail: the usual retention failure and another failure caused by the transformation of MTJ from binary state to multiple state. Judging from the statistics, the latter type of failure may be suppressed by design.

Introduction: Despite of the continuing debate of switching mechanisms [1, 2], Spin-orbit-torque (SOT) MRAM research has made solid progress in recent years [3], solving key engineering issues of MTJ integration into CMOS process [4], optimizing cell material and structure [5, 6] and understanding the bit cell behavior in actual operation [4]. This paper presents a statistical study of the reliability of SOT MRAM cells.

Experiments and results: Thin (5nm) of Ta is chosen as SOT spin current source. Ta line is placed over recessed Cu-pads, which reduces cell write-port input (Ta line) resistance to match the low Vdd of advanced CMOS requirement. Top-pin in-plane MTJs are deposited over Ta. An etching recipe is selected to etch the hard mask and pinned layer at a fast rate, while it etches MgO very slowly. The etching is stopped on MgO and the CoFeB free layer underneath is not removed. Fig. 1 shows the SOT cell structure and Table 1 shows the key cell dimensions. The SOT MRAM wafers are field-annealed at 265C for 30 minutes to set the pinning prior to testing. This etch-stop-on-MgO process eases the stringent etch uniformity requirement of the etch-stop-on-Ta process, widens the process window and improves the cell yield.

The CoFeB is damaged by the etching ion, creates dead-layer. A post-etch anneal further thickens the dead-layer [6]. While operating the cell, the write current flows through the Ta/CoFeB bilayer. As reported in ref. [4], the write characteristics of such SOT cell, such as write current pulse width dependency and temperature effect, greatly resemble those of STT cell.

Thousands SOT-MRAM bit cells are subjected to thermal baking to study the reliability of this cell structure. The thermal baking accelerates the data retention fail and reveals a new failure, namely some bits transform from binary bits into multiple-state (MS) bits. Fig. 2(a) and (b) illustrate the post-bake R-H loop of the normal binary bits and MS bits, respectively. Fig 2(c) shows the post-bake $R_{AP}$ of of binary bits and 2(d) $R_{AP}$ of the MS bits, where $R_{AP}$ drops partially, remains greater than initial $R_P$.

Fig. 3 shows the statistics of retention fail and MS fail after 4-, 12-, and 36-Hr baking at 150C: total sample count, retention fail bit count, the MS bit count. Notice that the retention fail bits are far less than the MS fail bits. While smaller size MTJs fail retention, the larger ones fail MS.

Fig. 4 shows that MTJ (shape) anisotropy is neither sensitive to MTJ size nor Ta width. Extracted based on from switching probability $P_{sw} = 1 - \exp\left[\frac{1}{\gamma_0}\exp(-\Delta)\right]$. $\Delta$ is in the range 49.5-51.5, from the smallest to the largest MTJ. On the contrary, MS fail is quite sensitive to Ta dimension and is more pronounce in Wide Ta. Statistically, smaller MTJ on narrow-Ta tends to exhibit least MS failure (Fig.5). More study is required to understand its relation with the device structure.

Conclusion and Implications: Statistical study shows that data retention time is MTJ size dependent for Step-structure SOT-MTJ. The un-etched CoFeB outside of MTJ does not participate in the data retention. We also learn that binary bit cell may transform into multiple-state bit cell after long thermal baking. Judging from the occurrence frequency, this transformation may be minimized or eliminated by scaling down MTJ size and Ta line width.

Acknowledges: The authors would like to thank MOEA, Taiwan, for their financial support under grant 108-EC-17-A-24-1519 and MOST, Taiwan, for their financial support under grant 107-2622-8-002-018.

References:
Fig. 1 (a) Schematics of SOT device cross-section, (b) top view, and (c) RIE damaged free layer.

Table 1 Dimension of MTJ and Ta line under study

<table>
<thead>
<tr>
<th>Device label</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ta width category</td>
<td>narrow Ta</td>
<td></td>
<td></td>
<td></td>
<td>wid Ta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu-pad spacing (nm)</td>
<td>400</td>
<td>600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L_Ta (nm)</td>
<td>800</td>
<td>960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W-Ta (nm)</td>
<td>1080</td>
<td>1290</td>
<td>1500</td>
<td>1710</td>
<td>1710</td>
<td>1710</td>
<td>1710</td>
<td>1710</td>
</tr>
<tr>
<td>L-mtj (nm)</td>
<td>240</td>
<td>300</td>
<td>360</td>
<td>420</td>
<td>240</td>
<td>300</td>
<td>360</td>
<td>420</td>
</tr>
<tr>
<td>W-mtj (nm)</td>
<td>840</td>
<td>1050</td>
<td>1240</td>
<td>1740</td>
<td>840</td>
<td>1050</td>
<td>1240</td>
<td>1470</td>
</tr>
</tbody>
</table>

Fig. 2 R-H loop (a) with binary bit, (b) with multiple-state bit. Initial $R_p$ (x) and post-bake $R_{up}$ (c) are plotted against initial $R_{up}$ of binary state bits (c) and multiple-state bits (d).
19. Spin-Orbit Torque MRAM Thin Films With High Thermal Stability

Meiyin Yang\textsuperscript{1}(present), Jianfeng Gao\textsuperscript{1}, Yan Cui\textsuperscript{1}, Tengzhi Yang\textsuperscript{1,2}, Chao Zhao\textsuperscript{1,2} and Jun Luo\textsuperscript{1,2*}

\textsuperscript{1}Key Laboratory of Microelectronic Devices & Integrated Technology, Institute of Microelectronics, Chinese Academy of Sciences (IMECAS), Beijing, P. R. China. \textsuperscript{2}University of Chinese of Academy Sciences (UCAS), Beijing, P. R. China

E-mail: luojun@ime.ac.cn

SOT-MRAM (spin-orbit torque MRAM) has the potential to be the next generation MRAM, as it is a faster and much more efficient memory technology. Furthermore, with the separation of reading and writing channels in magnetic tunneling junctions, the endurance of the SOT-MRAM greatly increases which is favorable in spin logics. However, the thin film process is much more difficult than that of STT-MRAM. Compared with STT-MRAM thin films, the bottom electrodes are limited by heavy metal (W, Ta and Pt)
with small thickness (3~5 nm), which brings great challenges for the magnetic properties and thermal stability for the upper tunnel junction and synthetic antiferromagnets (SAF).

In this work, we developed a thin film process for SOT-MRAM on 8-inch wafers with high thermal stability. Perpendicular magnetized CoFeB/MgO/CoFeB tunneling structure were fabricated on W (5 nm) bottom electrodes. The reference CoFeB layer is exchange coupled with the upper CoPt multilayer SAF with W (0.6 nm). However, the film stack cannot maintain perpendicular anisotropy after 350 °C annealing, which will not survive during the back-end process when integrated with CMOS. We modified the fabrication process by changing the sputtering atmosphere, sputter pressure and other parameters to increase the thermal stability of the SOT-MRAM thin film, and maintained the perpendicular anisotropy after annealing. Our work provides a solution for the thermal stability problem of SOT-MRAM, which is vital for the MRAM integration with CMOS technology.

20. Stability of Spin-Orbit-Torque Magnetic Random Access Memory with Voltage Controlled Magnetic Anisotropy Effect

S. Z. Rahaman1,*, I-J. Wang1, H. H. Lee1, D. Y. Wang1, C. F. Pai1,2, S. Y. Yang1, Y. C. Hsin1, Y. J. Chang1, Y. C. Ku1, Y. H. Su1, S. G. Chen1, F. M. Chen1, J. H. Wei1, S. S. Sheu1, W. C. Lo1, C. I Wu1,2, and D. L. Deng1
1Electronic and Optoelectronic System Research Laboratories, Industrial Technology Research Institute, Hsinchu 31040, Taiwan
2Department of Materials Science and Engineering, National Taiwan University, Taipei 10617, Taiwan
3Graduate Institute of Photonics and Optoelectronics, National Taiwan University, Taipei 10617, Taiwan
*E-mail: skzia@itri.org.tw

To mitigate the reliability issues of STT-MRAM, the SOT-MRAM with separated read and write paths is regarded as a possible solution for nonvolatile cache memory applications in advanced computing systems [1]-[5]. Here, we report a low resistance Ta HM-based SOT-MRAM and VCMA-MRAM with an etch-stop-on-MgO-barrier layer process technology to solve the etch non-uniformity issues and we place Cu pads under the high resistivity Ta layer to reduce its resistance, allowing low voltage operation of SOT cell. Furthermore, a comprehensive study of the switching properties of such SOT cells: pulse-width, temperature, endurance, and the VCMA modulation of our SOT cell switching threshold. A schematic diagram of the three-terminal SOT-MRAM bit cell structure and multilayer film stacks are shown in Fig. 1. The optimal VCMA effect occurs at ±0.5V in our SOT cell (Fig. 2(a)). Thus, in the presence of VCMA effect, an external electric field can modulate the energy barrier of MTJ (Fig. 2(b)). Fig. 2(c) shows the VCMA-assisted multibit SOT cell structure for future high-density nonvolatile memory applications. A micromagnetic simulation program MUMAX3 [6], based on the LLG equation was employed to obtain 5ns switching of SOT cell (Fig. 3), which is in good agreement with the experimental results. Inset shows the electron current density vs reduced magnetization switching curve at 5ns pulse-width. No degradation was obtained with continuing >10⁴ cycles (Fig. 4). Fig. 5 shows the temperature accelerated testing. The memory devices were baked at various temperatures for 24 hrs and checked the corresponding resistances at each temperature. More than 35 cells for each state were tested to compare. Only two bits were flipped after 24 hours at 230 °C. Although no bit flips were observed at 220 °C, bit resistance of this device structure holds well over the bake cycle and only few bits show resistance drop. We will explain many of the observations in the poster.

Fig. 1 Schematic diagram of the three-terminal SOT-MRAM bit cell structure, and schematic view of the multilayer film stacks. The values in parentheses refer to layer thicknesses expressed in nanometers.

Fig. 2 (a) SOT switching threshold under 0V, and ±0.5V bias on MTJ, the VCMA effect, (b) illustration of the impacts of positive/negative voltages to reduce/enhance the energy barrier of MTJ in presence of VCMA effect, and (c) illustration of VCMA-assisted SOT multibit RAM cell. CoFeB between MTJ is turned non-magnetic by process treatment.

Fig. 3 SOT switching at different pulse-widths, which are obtained by MUMAX3 program [6]. Inset shows the electron current density vs reduced magnetization at 5ns.

Fig. 4 Cycling test of five randomly selected SOT-MRAM cells. No degradation were obtained with continuing >10⁴ cycles.
Spin-orbit torque (SOT) magnetoresistive random access memory (MRAM) is a viable candidate for a non-volatile replacement of high-level caches, as it delivers high operation speed complemented with large endurance. However, for a deterministic SOT switching of a perpendicularly magnetized free layer (FL) an external magnetic field is required. In a recent SOT-MRAM implementation on a 300mm wafer the field was created by a cobalt nanomagnet added to each memory cell [1]. The magnetic field free switching solutions are based on breaking the cell mirror symmetry, which can be performed by coupling the FL to a ferromagnet [2] or an antiferromagnet (AF) [3], by designing the crystal structure [4], or purely geometrically by shaping the device appropriately [5]. Recently, interesting field-free schemes based on stacking of ferromagnetic layers [6] and heavy metals with opposite Hall angles [7] have been demonstrated. However, dynamical approaches to induce the effective magnetic field were not explored.

By means of extensive micromagnetic simulations, we demonstrate that an application of the two perpendicular current pulses in a Hall bar configuration results in fast, reliable, and magnetic field free perpendicular magnetization reversal of a rectangular FL. While the first current pulse forces the magnetization in-plane perpendicular to the long rectangle edge, the second pulse deviates the magnetization towards one of the elongated rectangle sides. Then, the magnetization experiences the shape anisotropy field which plays the role of an external field and makes the switching deterministic. The speed of switching is boosted, when the second current is applied not to the whole FL, but to a part of it. In this case, the FL part subject to SOT of the second current is quickly rotated in-plane along the long edges of the rectangle. The field from this part makes the rest of the FL to rotate about it, moving the magnetization deterministically from its in-plane orientation, thus completing the switching. We demonstrate that the dynamic scheme does neither require a perfect current pulse synchronization nor precise patterning of the second current line.
22. On the patterning of Spin-orbit torque (SOT) MRAM cell on Tungsten

Y. J. Chang¹*, S. Y. Yang¹, Y. C. Hsin¹, I. J. Wang¹, Y. C. Kuo¹, J. H. Wei¹, S. Z. Rahaman¹, D. Y. Wang¹, H. H. Lee¹, Y. H. Su¹, G. L. Chen¹, F. M. Chen¹, S. S. Sheu¹, W. C. Lo¹, C. I. Wu¹,², L. X. Ye¹, T. H. Wu¹, D. L. Deng¹

¹Electronic and Optoelectronic System Research Laboratories, Industrial Technology Research Institute (ITRI), Hsinchu, Taiwan, R.O.C. ²Graduate Institute of Photonics and Optoelectronics, National Taiwan University (NTU), Taipei, Taiwan, R.O.C. ³Graduate School of Materials Science, National Yunlin University of Science and Technology (NYUST), Yunlin, Taiwan, R.O.C.

*E-mail: Yao-Jen Chang, yjchang1689@itri.org.tw

One of most demanding tasks in developing SOT MRAM is patterning the MTJ device. A step-SOT device structure has been developed [1, 2] (Fig.1). By stopping the etching on MgO and the rendering the remaining free layer under the MgO non-magnetic, one creates a step SOT cell structure. This process relaxes the demand of etching uniformity and lowers the cell input impedance when WRITE.

This study presents the etching and the annealing process employed to pattern this step SOT structure. Table 1 summarizes the etching properties. The etch rate of MgO is very slow on MgO while >200 times faster on metal. The film stack under study is Si/SiOx/W3/CoFeB1.5/MgO0.7/Ta5. All units are in nm. It takes 6.1 sec to etch away the Ta Hardmask. The total etch time is 8 sec. Thus, the etch practically stops on MgO. The etch ions partially damage the CoFeB underneath MgO [1]. We found that annealing in air can completely turn CoFeB into non-magnetic.

Fig. 2 compares the Edex profile and M-H loop of the post-etch samples, one with ~2.5nm remaining un-etched Ta Hardmask and the other 0. Both samples are annealed. Edex shows that the former is not oxidized during anneal, and its M-H loop is unchanged. The latter is oxidized, its CoFeB Ms~0. Without the Ta Hardmask capping, oxygen diffuses through MgO and oxidizes CoFeB below the MgO, and Fe and Co diffuse up. This process can selective remove post-etch CoFeB outside of MTJ while not affecting the magnetic material covered by the MTJ etching Hardmask.

This process is deployed to build SOT MRAM cell on W nanowire. The magnetic dead CoFeB/W-nanowire exhibits higher resistivity (Fig.3) than Ta. We will discuss the cell properties.

23. A self-matching complementary-reference sensing scheme for high speed and reliable toggle spin torque MRAM

Jinkai Wang, Yue Zhang*, Chenyu Lian, Guangda Wang, Yining Bai, Zhe Huang and Weisheng Zhao*
Fert Beijing Research Institute, School of Microelectronics, BDBC, Beihang University, 100191, Beijing, P. R. China
(yz@buaa.edu.cn, weisheng.zhao@buaa.edu.cn)

Toggle spin torque MRAM (TST-MRAM) has been proposed for the perpendicular anisotropy magnetic tunnel junction so as to improve the write performance [1]-[3]. Experiments and simulations have proven that TST-MRAM can provide fast (<2.5ns) and low energy write operation [3]. However, read operation performance has not been improved for TST-MRAM due to the limitation of sensing scheme. The read operation time was greater than 2.5ns for many sensing schemes [4]-[6]. In addition, with the scale down, process deterioration causes the decline of sensing margin, which seriously affect the reliability of read operation [7]-[9].

Here, we propose a self-matching complementary-reference (SMCR) sensing scheme to effectively reduce the read operation latency with favourable reliability, which adds dual-reference bit-cells and bit-lines to generate dual-reference voltages for read data, as shown in Fig.1 (a). Data voltage will select a complementation voltage from dual-reference voltage. And then, the voltage difference between data...
voltage and complementation voltage is amplified by the SMCR sense amplifier (SMCR-SA) to read the data in TST device. Waveforms and timing sequences of SMCR scheme are illustrated in Fig. 1 (b). A 128x64 TST-MRAM is built to evaluate the performance of the SMCR scheme. One thousand Monte Carlo simulations simulate the swing voltage caused by the process of manufacturing to analyse the reliability of SMCR scheme. Simulations show that SMCR scheme can provide a read access time of 0.8ns and a read bit error rate of $1.02 \times 10^{-13}$, as shown in Fig. 2 (a). These results indicate that SMCR has high reliability while reading data fast. From Fig.2 (b), with 0.8V supply voltage and 100mV voltage difference, read access energy of a bit can save by 7.5% and 20% compared with conventional single-ended voltage sensing (CSVS) and dynamic reference (DR) schemes, respectively.

Fig.2. (a) Read bit error rate and read access time per bit at different supply voltage. (b) Read access energy per bit at different supply voltage.


24. SpinCIM: Spintronic Computing in Memory with magnetic tunnel junction

He Zhang, Wang Kang*, Youguang Zhang and Weisheng Zhao
1 School of Microelectronics, Beihang University, Beijing, China
*Corresponding Author: Wang Kang; Email: wang.kang@buaa.edu.cn

Over the past decades, the overhead of data communication between the processor and the memory units results in huge performance degradation and energy consumption, called von-Neumann bottleneck [1]. Computing in memory (CIM) is a promising solution to deal with this issue, by moving the computing
task into memory, therefore avoiding the power consumption and delay of data movement. Among them, CIM with non-volatile memory is particularly concerned [2, 3], due to its ability to perform Boolean logic functions with one/several cells. In this abstract, we report a stateful spintronics CIM (SpinCIM) logic paradigm, which can perform different Boolean logic functions with typical memory like write/read operations in a single memory cell. The spintronic device can be either spin transfer torque (STT) magnetic tunnel junction (MTJ) or voltage-gated spin Hall effect (VG-SHE) MTJ [4, 5]. Fig. 1(a) shows the structure of STT-MTJ and VG-SHE-MTJ and defines the operands in logic functions. By analyzing the truth table of the three operands, an equation is obtained to describe the relationship. Different logic functions can be achieved by setting operand C to different values. To support the proposed logic paradigm, Fig. 1(b) shows the parallel structure of our memory array, which is the basic computing unit in SpinCIM. The proposed memory array works just like a typical memory cell and all the logic functions are achieved via regular memory-like write and read operations with minimal modifications in write driver and word line driver. The memory array can work either in a typical memory mode or in logic mode, which can be switched between each other freely. Fig. 1(c) shows the transient simulation waveforms of XOR operation in 40nm technology node. In order to efficiently perform logic operations in our memory array, we proposed a reconfigurable architecture (SpinCIM) as shown in Fig. 1(d). The SpinCIM is consists of multiple arrays that can perform logic operation simultaneously, and the digital part is responsible for controlling memory arrays. By utilizing the proposed SpinCIM, many memory intensive applications can be performed, such as secure hash algorithms, convolutional neural network etc. Fig. 1(e) shows an example of an implementing binary neural network in the proposed memory array. As MRAM has recently been a commercialized product, our proposed SPU is expected to be a practical PIM platform.

References

Fig. 1. (a) Basic logic paradigm of the proposed SpinCIM, and the definition of logic operands in STT MTJ cell and VG-SHE MTJ cell; (b) Reconfigurable array structure between typical memory mode and CIM mode. (c) Transient waveforms of the logic paradigm within 1T1MTJ structure; (d) Architecture of the proposed SpinCIM. The digital part is responsible for controlling the memory arrays; (e) Case study of a binary neural network in the proposed memory array.
Thermally activated leakage currents of transistors cause the fundamentally unavoidable standby power problem for CMOS logic systems such as microprocessors and SoCs, which degrades/restricts the energy efficiency of these logic systems [1]. In particular, the constituent caches account for much fraction of standby power in logic systems, and thus its reduction is considerably indispensable. Application of STT-MRAM to caches can achieve extremely low standby power owing to its nonvolatile retention ability. However, the high store energy and long store latency restrict the application to the last-level cache or main memory. Power gating (PG) architectures using the dedicated nonvolatile SRAM (NV-SRAM) using MTJs can avoid the store latency problem, since the store operation to the MTJs is carried out only before shut-down period (the MTJs are not used during the normal SRAM operation mode) [2]. Therefore, the NV-SRAM is applicable to higher level caches. However, the high store energy prolongs the break-even time (BET), i.e., temporal granularity of PG cannot be shortened. To overcome the store energy problem, several architectures, e.g., early write termination, efficient data coding for reducing store bits, and so on, have been proposed for STT-MRAM caches [3-13]. These architectures can be also applied to NV-SRAM caches. Recently, simple store skipping (SSS) [14] and hierarchical store-free (HSF) [15] architectures are also proposed, which are more suitable for NV-SRAM caches. In this paper, we propose a new store energy and latency reduction architecture using proactive flush of useless data for NV-SRAM using MTJs, which is analogous to cache line replacement techniques used in cache systems of microprocessors and SoCs. Performances of the proposed proactive useless data flush (PUF) architecture are computationally analyzed and experimentally verified using circuit parameters extracted from a fabricated NV-SRAM TEG. The PUF architecture is based on the previously proposed HSF architecture [15]. In the HSF architecture, store-unneeded blocks and subarrays in NV-SRAM are memorized as store-free flags in the purpose-built latches, in which these store-unneeded blocks and subarrays are encoded by the presence or absence of write access during the normal SRAM operation mode. Before executing the store operation to the MTJs in the NV-SRAM, the store-unneeded blocks and subarrays in the NV-SRAM are shut down in advance in accordance with the store-free flags. On the other hand, in the PUF architecture, the prior shut-down is carried out based on prediction of useless data that are judged in similar manner to cache line replacement techniques. In general, all the data memorized in caches need not be used again, i.e., there exist useless data in caches (although data coherency is required between the intended and its lower-level caches/memories). Data that have a high possibility for being useless can be predicted by several techniques similar to FIFO and LRU architectures used in ordinary caches. The proposed PUF architecture is achieved by the prior shut-down of blocks and subarrays having useless data. Energy and latency performances of the PUF architecture are investigated using 32kB, 256kB, and 2MB NV-SRAM arrays, where the cell proposed by [2], which is suitable for the PG architecture using nonvolatile retention, is employed. The NV-SRAM array is comprised of 8kB subarrays and each subarray is divided into small blocks. The power supply of the blocks is controlled by a power management unit (PMU) using power switches. The PMU has a register file to memorize useless-data flags for the blocks. Using the PUF architecture, the store energy and latency can be effectively reduced depending on the proportion of useless data blocks. Therefore, the BET can also be shortened significantly, leading to highly energy-efficient PG. The energy reduction efficiencies of the arrays are further enhanced by combining the PUF architecture with the HSF architecture. These simulation results are good agreement with results analyzed by measured circuit parameters extracted from a fabricated NV-SRAM TEG.
26. Physics-Based Stochastic Learning in MRAM-Based Binarized Neural Networks

T. Hirtzlin, M. Bocquet, J.-O Klein, J.-M Portal, and D. Querlioz

One of the most exciting applications of Spin Torque Magnetoresistive Random Access Memory (ST-MRAM) is the in-memory implementation of deep neural networks, which could allow improving the energy efficiency of Artificial Intelligence by orders of magnitude with regards to its implementation on computers and graphics cards. A particularly stimulating vision is using ST-MRAM for implementing Binarized Neural Networks (BNNs), a class of deep neural networks discovered in 2016, which can achieve state-of-the-art performance with a highly reduced memory footprint with regards to conventional artificial intelligence approaches. However, training BNNs is a difficult and computationally-expensive task and, until now, hardware implementations of BNNs have focused on inference (i.e. non-learning capable hardware) [1]. In this work, encouraged by previous works on neuroscience-inspired networks [2], we introduce an original method for training ST-MRAM-based BNNs that relies on the stochastic switching nature of magnetic tunnel junctions and has highly reduced computational and memory requirements with regards to the conventional training technique of BNNs. We validate the technique by simulations of vision tasks, and highlight the impressive extent to which imperfection of ST-MRAM cells may be tolerated. We also investigate the limitations of our approach in the case of very deep networks, and conclude on the potential applications to edge computing.


27. Binarized Neural Network based on Versatile Nonvolatile Hybrid Spin/CMOS Logic-in-memory for IoT Applications

Biao Pan, Erya Deng, Wang Kang, and Weisheng Zhao, Fellow, IEEE

School of Electronic and Information Engineering, Beihang University, Beijing, China
Fert Beijing Institute, BDBC, Beihang University, Beijing, China
School of Microelectronics, Beihang University, Beijing, China
Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing, China

Corresponding Author: Wang Kang; Email: wang.kang@buaa.edu.cn
Convolutional neural networks (CNN) are biologically inspired classes of algorithms that have recently demonstrated the state-of-the-art accuracy in large-scale classification and recognition tasks. Hardware acceleration of CNN is of paramount importance to ensure their ubiquitous presence in future computing platforms [1]. However, high demands on memory capacity and computational power make it unsuitable to implement the full precision CNN on resource-limited devices such as embedded systems and mobile devices in IoT [2]. Recent advances have shown that XNOR-network can provide a satisfying accuracy on various image datasets with significant reduction in computation and memory costs [3]. In this paper, we demonstrate how XNOR-network can be implemented in a modified von Neumann accelerator, by enabling binary convolutions within the magnetic random-access memory (MRAM) arrays, as shown in Fig.1. In general, XNOR-network is a kind of Binary Neural Networks (BNNs), whose binary convolutions consist of bit-wise XNOR operations followed by a population count operation. In order to conduct the bit-wise XNOR operations, we utilize hybrid spin/CMOS logic circuits with the spintronic-based memory to store the non-volatile logic data (weights ‘0’ and ‘1’) and CMOS logic tree to activate the volatile logic data (inputs ‘0’ and ‘1’). The 1-bit data stored in a pair of MTJs can be sensed by a sensing circuit and written by the 4T or 6T writing circuit. In order to realize the writing operation without disturbing the outputs, separating transistors contribute to insulating the MTJ cells from the sensing part and thus preventing the writing current from passing through the sensing part during this phase. The proposed BNN design can effectively solve the conflict between area and energy consumption that exists in conventional SRAM based designs. By using our developed MTJ electrical models and a CMOS design-kit, circuit operation and performance are demonstrated at the 40 nm technology node. At a system-level, we use a neural network hardware platform NeuroSim+ [4] to conduct the modified von-Neumann based system architecture. By utilizing these in-memory convolutions, we demonstrate the benefits in the overall system energy consumption and latency per inference. More details will be presented in the poster.

References
Neuromorphic computing, i.e. modeling computer architecture after the human brain, is a promising approach to the ever-increasing demands for real-time processing of massive amounts of data. But, the building blocks of the brain, i.e. neurons and synapses, have more detailed behavior than simply how they are connected, and this behavior is central to their computing efficiency. If we want to capture the energy efficiency of such a neuromorphic architecture, we need devices themselves that can capture the behavior of the biological elements.

MRAM-based materials and devices have many properties similar to the brain that make them attractive candidates as artificial neurons and synapses. I will present our recent work [1] on designing and modeling three-terminal magnetic tunnel junction circuits that can act as synapses with online learning and spike-timing-dependent plasticity behavior. The circuit consists of a three-terminal magnetic tunnel junction with a mobile domain wall between two low-pass filters and has been modeled in SPICE. The results show that the current flowing through the synapse is highly correlated to the timing delay between the pre-synaptic and post-synaptic neurons. Using micromagnetic simulations, we show that introducing notches along the length of the domain wall track pins the domain wall at each successive notch to properly respond to the timing between the input and output current pulses of the circuit, producing a multi-state resistance representing synaptic weights. We show in SPICE that a notch-free ideal magnetic device also shows spike-timing dependent plasticity in response to the circuit current. This work is key progress towards making more bio-realistic artificial synapses with multiple weights, which can be trained online with a promise of CMOS compatibility and energy efficiency.
29. A new path to magnetization switching: low dimensional chaos driven by ac spin transfer torque

Eric Arturo Montoya, Salvatore Perna, Yu-Jin Chen, Jordan A. Katine, Massimiliano d’Aquino, Claudio Serpico & Ilya N. Krivorotov

Energy-efficient magnetization switching is an essential problem in the realization of practical nonvolatile magnetic storage. Recently, several efficient methods of magnetic switching were demonstrated including spin torque, magneto-electric, and microwave-assisted switching mechanisms. In this poster, we experimentally demonstrate that low dimensional magnetic chaos induced by alternating spin torque can strongly increase the rate of thermally-activated magnetic switching of the free layer in a magnetic tunnel junction, the core component of STT-MRAM. This mechanism exhibits a well-pronounced threshold character in spin torque amplitude and its efficiency increases with decreasing spin torque frequency. We present analytical and numerical calculations that quantitatively explain these experimental findings and reveal the crucial role played by low dimensional magnetic chaos near saddle equilibria in enhancement of the switching rate. Our work unveils an important interplay between chaos and stochasticity in the energy assisted switching of magnetic nanosystems and paves the way towards improved energy efficiency of spin torque memory and logic.


30. Ultrafast Optically Switchable Co/Tb multilayers based p-MTJ for the development of photo-spintronic devices


This work reports the development of perpendicular magnetic tunnel junctions, incorporating a multilayered stack of Tb/Co that can be all-optical addressed via helicity independent single-shot switching of the magnetization. In order to fabricate optically switchable electrodes, [Tb/Co] multilayers were coupled to a FeCoB layer through a Ta ultra-thin layer. The all-optical helicity independent switching was achieved for Co-rich compositions of the multilayer, either alone or embedded in a tunnel junction stack coupled to a CoFeB electrode on an MgO barrier. Toggling of the magnetization in the CoFeB-[Tb/Co] half-MTJ stack was observed using both 60 femtosecond- and 5 picosecond-long laser pulses with fluences down to 4.7 mJ/cm². This result is obtained even after annealing at 250 °C leading to a TMR signal of 41% and RxA value of 152 μm in nanopatterned junctions. These results are very promising for the development of p-MTJ whose storage layer could be switched with single-shot ultrafast laser pulses that target applications of hybrid spintronic photonic systems with THz switching speeds and fJ switching energies.

The authors acknowledge the funding support from project SPICE (EU ITC 2020 grant agreement No. 713481)