The magneto-optical Kerr effect for efficient characterization of thermal stability in dense arrays of p-MTJs

S. Van Beek, R. Carpenter, S. Kundu, S. Couet, J. Swerts, and G. Sankar Kar

Cite as: AIP Advances 9, 125236 (2019); https://doi.org/10.1063/1.5130179
Submitted: 03 October 2019. Accepted: 10 November 2019. Published Online: 26 December 2019

AIP Conference Proceedings
FLASH WINTER SALE!
50% OFF ALL PRINT PROCEEDINGS
ENTER CODE 50DECE19 AT CHECKOUT
The magneto-optical Kerr effect for efficient characterization of thermal stability in dense arrays of p-MTJs

S. Van Beek, a) R. Carpenter, b) S. Kundu, S. Couet, J. Swerts, and G. Sankar Kar

AFFILIATIONS
IMEC, Kapeldreef 75, 3001 Leuven, Belgium

Note: This paper was presented at the 64th Annual Conference on Magnetism and Magnetic Materials.

Electronic mail: simon.vanbeek@imec.be
Electronic mail: robert.carpenter@imec.be

ABSTRACT
Magnetic Random Access Memory (MRAM) targeting cache memory replacement consists of high density arrays of perpendicular Magnetic Tunnel Junctions (p-MTJs). One of the key advantages to MRAM is its non-volatility, where the measurement of retention is critical. However, evaluating the thermal stability of high density arrays is both time consuming and requires electric read out of many functional devices. It is shown that by using the Magneto-Optical Kerr Effect (MOKE) to measure the magnetic after-effect at fixed fields, efficient and in-line characterization of both the magnetic properties and data retention is possible. This is qualified through cross measurement of multiple dies using both MOKE and electrical readout of a Mbit array. Data from both techniques is then fitted using a macrospin model where the values of both \( \Delta \) and \( H_k \) are shown to come within 10% of each other.

© 2019 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/1.5130179

I. INTRODUCTION

First generation Spin Transfer Torque MRAM (STT-MRAM) products are now entering full scale production. For this, high density arrays are required. To enable successful roll out of the technology, sufficient in-line metrology is required in order to improve line yield. For electrical properties of the p-MTJ, such as the Resistance Area (RA) product and Tunneling Magnetoresistance (TMR), Current In-Plane Tunneling (CIPT) can be used. However, as CIPT is a contact measurement with limitation in the probe size and fitting models, this is only available on blanket wafers. For simple magnetic properties, such as the Free Layer (FL) coercivity (\( H_c \)) or offset field (\( H_{off} \)), a MOKE in a polar configuration may be used. As MOKE is an optical technique, the measurement can be performed not only on blanket, but also patterned wafers. This is assuming the array density is sufficient and any top layers are either transparent or thin enough to allow the incident light to interact with the layers of interest. Despite this, a number of critical parameters, such as the thermal stability (\( \Delta \)) and the effective anisotropy field (\( H_k \)), cannot be measured until full formation of the device for electrical readout.

The values of \( \Delta \) and \( H_k \) are measured typically by fitting the time dependence of the ‘magnetic after-effect’ by a macrospin model, such as that proposed by Sharrock. For technical applications, a \( \Delta > 54 \) is required for an error rate of less than \( 10^{-6} \) over 10 years. Therefore, in order to observe the time dependence in a reasonable period ‘acceleration’ is required. This can be carried out through self-heating via current, application of an external field or baking. In order to obtain sufficient statistics from each of these techniques, electrical readout from a fully selectable memory cell is required. These measurements are time consuming and the number of process steps after etching of the p-MTJ required before they can be carried out are significant. This limits the feedback time before potential faults can be found, impacting line yields.

In this work, a measurement technique is proposed in order to measure \( \Delta \) and \( H_k \) directly from a patterned array of p-MTJ pillars. This is done using the ‘field acceleration’ method and measured on a Mbit array of MRAM devices using Polar-MOKE (PMOKE).
The technique was validated on 60 nm devices with a 200 nm pitch (60P200), measured across 3 separate dies on the wafer. These results were then compared to electrically obtained values measured in the fully Complementary Metal Oxide Semiconductor (CMOS) integrated Mbit arrays from the same dies.

II. METHOD

A. Sample fabrication

The p-MTJ stacks were deposited by Physical Vapour Deposition (PVD) in a Canon-Anelva EC7800 300mm cluster tool. All wafers were post stack annealed in a TEL-MS2 MRT5000 batch furnace at 400°C for 30 mins. The stack design used a bottom pinned p-MTJ as described in previous work\textsuperscript{11,12} and shown schematically in Fig. 1(a).

Transfer of the mask pattern for the 60P200 arrays was carried out via immersion lithography. Patterning for all wafers consisted of a two-step process with the hardmask receiving a reactive ion etch and the p-MTJ pillar formation by ion beam etch. An example side-view Scanning Electron Measurement (SEM) image of the 60P200 array can be seen in Fig. 1(b). The electrically active 60P200 devices are shown in Fig. 1(c), where to enable electrical readout, 1 in 20 p-MTJs is connected to a transistor periphery.

B. Wait time measurement technique

The PMOKE setup used was a 300 mm KLA-Microsense PKM-RAM tool with a water-cooled magnet of max field 2.4 T, automated polarizer optics and diode laser of spot size ∼1 mm. The size of the laser spot with respect to the Mbit array is shown in Fig. 2(b). Field ranges and polarizer settings were chosen to increase the FL signal while minimizing spurious optical effects due to wafer orientation or topography. The procedure for the wait time method used in this work is as follows:

\[-2 \text{T} \xrightarrow{30 \text{s}} \text{P-state} \xrightarrow{3 \text{s}} \text{AP-state} \xrightarrow{600 \text{s}} \text{P-state} \xrightarrow{10 \text{s}} \text{0.4 T} \xrightarrow{10 \text{s}} \text{0 T} \xrightarrow{10 \text{s}} -2 \text{T}\]

Where the first field applied is done to initialize all devices in the parallel (P) state. The field is then swept to 0 T for two purposes: 1) to provide a reference at positive saturation for the normalization and 2) to change the Gauss range of the Hall probe from 20 kG to 4 kG to improve field resolution during the wait time measurement. When reducing the field, the Reference Layer (RL) aligns Anti-Parallel (AP) to the Hard Layer (HL) and, consequently, the FL. The time ($t_r$) to change to the waiting field ($x$) is minimized in order to limit the impact of the field sweep on the fitting of the time dependence. The measurement is then completed with the reversal of all devices to the P state at 0.4 T, a sweep to 0 T for the normalization reference at negative saturation and a change in the Gauss range followed by a final initialization at −2 T. This is shown schematically for an example measurement in Fig. 2(c) along with examples of the change in the Kerr signal and applied field with respect to time over the course of the measurement, Figures 2(d/e) respectively.

The process to normalize the Kerr signal is shown in Fig. 3. In order to account for any linear, field dependent signal a straight line is fitted at both positive and negative saturation and the slope
removed. The data is then normalized with respect to the minimum/maximum Kerr signal. An example measurement is shown in Fig. 3(a) with the resulting normalized Kerr value plotted in Fig. 3(b) as a function of time from the start of ramp up to the waiting field. This is done separately for each waiting time measurement to account for any long term laser and/or sensor drift. The data for each measurement field is then collated and fitted with the methodology discussed in Sec. II C.

As there is no continuous monitoring of magnetic field and p-MTJ resistance during the electrical measurement of the Mbit array, the procedure is different and described as follows: 1) All working devices (99%) are initialized by a magnetic field. 2) The ramp of the magnetic field from 0 to the waiting value is not controlled. However, it is in the order of milliseconds and so considered abrupt. 3) The duration of the wait time is increased in logarithmic steps in between which the field is turned off for electrical read out of the Mbit array.

C. Thermal stability fitting methodology

The p-MTJs are assumed to switch by a thermally activated mechanism as described by an Arrhenius law:[1]

$$\lambda(t) = f_0 \exp\left(-\frac{E_b(t)}{k_B T}\right),$$

(1)

where $\lambda$ is the switching rate, $f_0$ the attempt frequency, taken as 1 GHz and $E_b$ the energy barrier between the P and AP state. The value of $E_b$ is derived for the macrospin model in an out of the plane magnetic field:[14]

$$E_b(t) = \Delta \left(1 - \frac{H(t)}{H_k}\right)^2,$$

(2)

where $H(t)$ is the applied magnetic field as a function of time. The normalized magnetic moment for 1 particle will decay exponentially as follows:

$$m(t) = 2 \exp\left(\int_0^t -\lambda(t) \, dt\right) - 1,$$

(3)

where the AP (P) state is defined as 1 (-1). The integral is then split in order to separate the contribution of the reversals that occur during the time of ramp ($t_r$) from those during the waiting time. These contributions are described by:

$$m(t) = \begin{cases} 2 \exp\left(\int_0^{t_r} -\lambda(t) \, dt\right) - 1 & \text{for } t_r \leq t_r \\ 2 \exp\left(\int_0^{t_r} -\lambda(t) \, dt\right) \exp(-\lambda (t_r - t_r)) - 1 & \text{for } t_r > t_r \end{cases}$$

(4)

For the macrospin model the integral is solved analytically where the normalized magnetic moment of a single particle is given by Eq. 4.

The distribution of pillars in a 1 Mbit array are assumed to be normally distributed. As a result the normalized magnetic moment of the Mbit array is described as follows:

$$m_{\text{array}}(t) = \int_0^\infty m(t) \cdot f(x, \mu_b, \sigma_b) \, g(H_k, \mu_{H_k}, \sigma_{H_k}) \, dH_k \, dh_k,$$

with

$$f(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

(5)

where $\mu_b, \mu_{H_k}, \sigma_b, \sigma_{H_k}$ are the fitting parameters. Finally, the normalized Kerr signal of different waiting fields is fitted using a least square algorithm and Eq. 5. To reduce the number of data points, and the impact of the telegraph noise, the continuous normalized Kerr signal is sampled and averaged at logarithmic spaced timestamps.

For the fitting of the electrically measured Mbit array the rise time is considered to be negligible, $t_r \to 0$.s. Instead of the normalized Kerr value, the normalized magnetic moment is used, with an equal contribution of each p-MTJ to the total magnetic moment.

III. RESULTS AND DISCUSSION

The previously described measurement and fitting method for the PMOKE wait time technique is validated on 60P200 arrays. This is done by comparing directly with electrical measurements obtained from Mbit arrays located within the die of measurement. The normalized and fitted data for the 60P200 arrays is shown for 3 separate dies in Fig. 4. The rise time $t_r$ used for these measurements was 30 s. The calculated values for the mean $\Delta (\mu\Delta) (70/76, 67/76, 74/77)$ and mean $H_k (\mu_{H_k}) (554/541 \text{ mT}, 599/556 \text{ mT}, 532/548 \text{ mT})$ are in good agreement being within 10% for most dies. These results were obtained for AP-to-P switching, however, similar results were obtained for P-to-AP switching (not shown). This demonstrates that the wait time measurements performed by PMOKE on large p-MTJ arrays are equivalent to those obtained from electrical characterization of full CMOS connected arrays. This brings with it a number of advantages.

Firstly, for the PMOKE all magnetic material within the 1 mm area of the laser spot is measured and so information from all the p-MTJ pillars, including damaged and non-connected devices, is obtained. Meanwhile, for the electrical measurement it is only those that are fully connected that are available, ~1 M. Furthermore, any non or poorly functioning devices are rejected (~1%) meaning that only the ‘best’ pillars are measured. Therefore the distribution as measured by the PMOKE is more representative of the full array. However, this highlights a disadvantage. Due to limited signal, the measurement is not sensitive to single device reversals.
This leads to a loss in sensitivity at the tails of the distribution. Furthermore, dummy structures, which make up ~8% of the total area and have a different CD (~100 nm), are also present, ‘poisoning’ the signal. This likely contributes to the larger calculated values of $\sigma_{H_k}$ in the PMOKE measurement. However in real production wafers, the dummy p-MTJs have the same dimensions and are not a concern.

Secondly, in the PMOKE measurement the Kerr signal is measured continuously with an averaging time of 0.1 s. This allows more data to be gathered about the shape of the magnetic reversal during the field sweeps. This can be included in the fitting by modelling the sweep rate dependence of the reversal. In addition to this, the entire wait time measurement is obtained in a single measurement of ~20 mins, which is then repeated at multiple fields. In contrast, during the electrical characterization each device is measured independently, preventing continuous measurement during the wait time. Therefore the field is applied for a time $x$ and then removed. The state of each device is then measured in order to observe those which have reversed. This is a time consuming process which, for each data point, can take upwards of ~5 mins which must be repeated for each wait time and field.

Finally, as PMOKE is an optical technique, the measurement can be carried out immediately after etching of the p-MTJ. This enables rapid feedback of the impact on $\Delta$ and $H_k$ of any processes without waiting for full device formation. As such, the effect of aggressive patterning on p-MTJ properties can be investigated without depending on successful electrical connection.

IV. CONCLUSIONS

In conclusion, it has been shown that for STT-MRAM, the device densities provide sufficient signal to perform measurements following the wait time method with PMOKE. This in-line metrology technique can be performed immediately after patterning of the p-MTJ and does not require the forming of the full device. In addition to this, through efficient selection of the wait time and fields, the time to acquire $\Delta$ and $H_k$ can be greatly minimized. For example, with a wait time of 600 s carried out at 5 field values the total measurement time will be ~60 mins. With a short loop cost of ~$2.5-5k$ per wafer and a time to fabricate of ~1-2 weeks, this technique can provide significant reductions. In particular for tool-tool or site-site matching of new, or modified, processes. Finally, with further improvements in signal to noise not only can the measurement accuracy be improved, but other time dependent phenomena in p-MTJs can be studied.

ACKNOWLEDGMENTS

This work was supported by imec’s Industrial Affiliation Program on STT-MRAM devices. The authors would also like to acknowledge the support of imec’s fab, line and hardware teams.

REFERENCES

1. S. Chung, T. Kishi, J. W. Park, M. Yoshikawa, K. S. Park, T. Nagase, K. Sunouchi, H. Kanaya, G. C. Kim, K. Noma, M. S. Lee, A. Yamamoto, K. M. Rho, K. Tsuchida, S. J. Chung, J. Y. Yi, H. S. Kim, Y. S. Chun, H. Oyamatsu, and S. J. Hong, in Proc. IEDM, 27.1.1–27.1.4 (2016).
2. D. C. Worledge and P. L. Trouilloud, Applied Physics Letters 83, 84–86 (2003).
3. R. Street and J. C. Wooly, Proceedings of the Physical Society Section A 62, 562 (1949).
4. L. Neel, Journal Physique Radium 11, 49–61 (1950).
5. K. Ohashi, H. Tsuji, S. Tsunashima, and S. Uchiyama, Japanese Journal of Applied Physics 19, 1333–1338 (1980).
6. P. J. Flanders and M. P. Sharrock, Journal of Applied Physics 62, 2918–2928 (1987).
7. L. Thomas, G. Jan, S. Le, Y. J. Lee, H. Liu, J. Zhu, S. Serrano-Guisan, R. Y. Tong, K. Pi, D. Shen, R. He, J. Haq, Z. Teng, R. Annapragada, V. Lam, Y. J. Wang, T. Zhong, T. Torng, and P. K. Wang, in Proc. IEDM, 26.4.2–26.4.4 (2015).
8. D. Bedau, L. Liu, L. Z. Sun, J. A. Katine, E. E. Fullerton, S. Mangin, and A. D. Kent, Applied Physics Letters 97, 262502 (2010).
9. J. Z. Sun, R. P. Robertazzi, J. Nowak, P. L. Trouilloud, G. Hu, D. W. Abraham, M. C. Gaidis, S. L. Brown, E. J. O’Sullivan, W. J. Gallagher, and D. C. Worledge, Physical Review B 84, 064413 (2011).
10. L. Thomas, G. Jan, and P. K. Wang, Applied Physics Letters 106, 162402 (2015).
11. Y. C. Wu, W. Kim, S. Rao, K. Garello, S. Van Beek, S. Couet, E. Liu, J. Van Houdt, G. Groeseneken, and G. Sankar Kar, Applied Physics Letters 113, 142405 (2018).
12. S. Van Beek, K. Martens, P. Roussel, Y. C. Wu, W. Kim, S. Rao, J. Swerts, D. Crotti, D. Linten, G. S. Kar, and G. Groeseneken, AIP Advances 8, 055909 (2018).
13. W. F. Brown, Phys. Rev. 130, 1677–1686 (1963).
14. X. Feng and P. B. Visscher, Journal of Applied Physics 95, 7043–7045 (2004).