Figure 30.2.3 shows the proposed in-situ self-termination write method. During write, hundreds of μA’s are applied to the cell, with current flowing from BL to SL when writing a 1 and from SL to BL for 0. When the free magnetic layer in the STT bitcell flips, the resistance of the cell changes abruptly, which can be detected by observing the BL voltage. Since the resistance change and current direction both flip, the sense amplifier can be easily biased on a write 1 and 0, the BL voltage drops in both cases when the write completes (Fig. 30.2.3(b)), simplifying write completion detection. The read sense amplifier is reconfigured as a continuous gate-connected voltage sense amplifier to detect the BL voltage drop as shown in Fig. 30.2.3(a). Since the write driver and detection circuit are shared with the read circuit, area overhead is negligible. Offset cancellation and precharge phases are overlapped with decoding, avoiding timing penalties. Once write completion is sensed, the ‘stop’ signal disables the write driver on that BL (Fig. 30.2.3(c)). This method auto terminates the write when complete, saving write power and improving reliability. In addition, redundant writes are seamlessly detected and avoided which conventionally require a read-before-write [5] operation and incur a full read cycle overhead.

The proposed 1Mb STT-MRAM is fabricated in 28nm embedded MRAM technology. Figure 30.2.7 shows the die photo of the MRAM chip. The 1Mb MRAM macro occupies 0.214mm². Figure 30.2.4(a) shows the read-latency vs access time for a single 128kb array, without any error correction. The array achieves a 2.8ns read-access time with an approximate error rate of 10⁻¹⁰ at 25°C, and 3.6ns at 120°C. For an access time less than 2.8ns, higher error rates are observed with the self-generated Vref because Rref takes longer to stabilize due to driving 16 sense amplifiers. However, read self-reference generation shows an error rate similar to an externally generated Vref for access times greater than 2.8ns at room temperature. However, it exhibits a lower error rate at higher temperatures since it dynamically tracks the voltage drop, while the sense amplifier is fixed and generates Vref at a constant temperature with 6σ offset cancellation tracks array-to-array variation, improving the read failure rate (Fig. 30.2.6(b)) because the write time decreases with temperature. Figure 30.2.6(c) shows that the self-reference generation tracks array-to-array variation, improving the read failure rate from 2x10⁻⁶ to 9x10⁻⁶ compared with a fixed externally generated Vref as measured across 8 arrays.

The measured Shmoo plot in Fig. 30.2.4(c) shows that the memory can be successfully read below 0.6V. Figure 30.2.4(d) shows the measured Vref,mem across 8 arrays for the proposed sense amplifier. Due to offset cancellation, the sense amplifier is robust at low supply voltages with an average Vref,mem of 0.57V (σ=19mV).

Figure 30.2.5(a) shows that the required write access time must be greater than 20ns in order to achieve a 10⁻⁹ failure rate using a conventional, fixed write time. Since the required write time varies significantly across cells, a constant write time write completes (Fig. 30.2.3(b)), simplifying write completion detection. The read completion is detected when the flip polarity between a write 1 and 0, the BL voltage drops in both cases when the free magnetic layer in the STT bitcell flips, the resistance of the cell changes abruptly, which can be detected by observing the BL voltage. Since the resistance change and current direction both flip, the sense amplifier can be easily biased on a write 1 and 0, the BL voltage drops in both cases when the write completes (Fig. 30.2.3(b)), simplifying write completion detection. The read sense amplifier is reconfigured as a continuous gate-connected voltage sense amplifier to detect the BL voltage drop as shown in Fig. 30.2.3(a). Since the write driver and detection circuit are shared with the read circuit, area overhead is negligible. Offset cancellation and precharge phases are overlapped with decoding, avoiding timing penalties. Once write completion is sensed, the ‘stop’ signal disables the write driver on that BL (Fig. 30.2.3(c)). This method auto terminates the write when complete, saving write power and improving reliability. In addition, redundant writes are seamlessly detected and avoided which conventionally require a read-before-write [5] operation and incur a full read cycle overhead.

Acknowledgements:
This work was supported by TSMC university joint development program and university shuttle program.

References:
Figure 30.2.1: Required write time and read margin vary with PVT, wasting write power and posing a challenge for sensing circuit design. Row-wise reference cells are used to generate a read-reference voltage that tracks row-wise PVT variation.

Figure 30.2.2: Constant-current based voltage sensing with a single-capacitor based offset cancellation. Compared with a conventional sense amplifier, input offset is reduced by >60%, significantly improving the sensing margin.

Figure 30.2.3: In-situ write-end detection and self-write termination using a continuous offset-cancelled sense amplifier reconfigured from the read-sense amplifier with no timing and area overhead.

Figure 30.2.4: Read operation measured results.

Figure 30.2.5: Write operation measured results.

Figure 30.2.6: Comparison table to other STT-MRAM work.
Figure 30.2.7: Die photo of 1Mb 28nm STT-MRAM.