# Chapter 5
# Characterization and Modeling of Spin-Transfer Torque (STT) Magnetic Memory for Computing Applications

**Roberto Carboni**

## 5.1 Introduction

The ubiquitous widespread of mobile devices marked the beginning of the Internet of Things (IoT) era, where the information is acquired, elaborated and transmitted by billions of interconnected smart devices. IoT demanded also a paradigm shift, from a centralized model, where acquired data was simply transmitted to a central mainframe to be processed, to a distributed model, requiring real-time data elaboration right where data is collected. For example, emerging applications such as active health monitoring, drone/robot navigation and autonomous car driving, require online elaboration of massive data. In this scenario, there is an ever increasing need for faster computing and larger/faster storage available on the IoT devices themselves.

During the last 50 years information technology achieved tremendous advancements in terms of computing power. This trend was made possible by the continuous miniaturization of the metal-oxide-semiconductor field-effect transistor (MOSFET). To describe such scaling pace, in 1965, Gordon Moore observed that the number of transistors on a silicon chip doubled every 18 months [1], speculating that such trend would continue in order to sustain the economics of electronics (Moore's law). Unfortunately, this trend is starting to slow down and is currently facing severe challenges. The so-called heat wall is one of the challenges faced by present day electronics [2], and one of the main aspects hindering Moore's law. The main reason for this high power dissipation is the increased leakage power typical of scaled transistor, which is directly connected to the Boltzmann statistics-limited sub-threshold swing of 60 mV/dec [3]. The growing difficulty in keeping up with Moore's law is one of the critical aspects in hindering next generation computing. Moreover, some other challenges are not related to technology, but they are rooted in the structure

R. Carboni (✉)
Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB) and IU.NET, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
e-mail: roberto.carboni@polimi.it

© The Author(s) 2021

A. Geraci (ed.), *Special Topics in Information Technology*,
PoliMI SpringerBriefs, https://doi.org/10.1007/978-3-030-62476-7_5

of computer systems. In fact, the great majority of computer systems are based on the von Neumann architecture, which is characterized by a rigid separation of logic and memory circuits requiring a continuous movement of data between them. This condition is usually referred as memory wall or von Neumann bottleneck [4].

Aiming at the mitigation of such effect, an improved storage solution was proposed by IBM with the Storage Class Memory (SCM) concept [5]. For this purpose it should have high read/write speeds, below 100 ns like DRAM, low cost per bit, high density and non-volatility like Flash memory. Emerging memory technologies are considered a prominent candidate for SCM implementation thanks to their non-volatility, low power/fast operation and better scalability [6]. Figure 5.1 shows various emerging memory concepts, including resistive switching memory (RRAM), phase-change memory (PCM), ferroelectric memory (FERAM) and spin-transfer torque magnetic memory (STT-MRAM) [7]. Table 5.1 reports a comparison of emerging memory technologies performances. They generally depend on material-based storage, which relies on the physics of the constituent active materials. Each of them is based on its peculiar transport and switching mechanisms, although sharing the two-terminal structure, where the application of suitable voltage pulses can change one or more properties of the active material. The remarkable properties of such switching materials enables various approaches to overcome the von Neumann bottleneck,

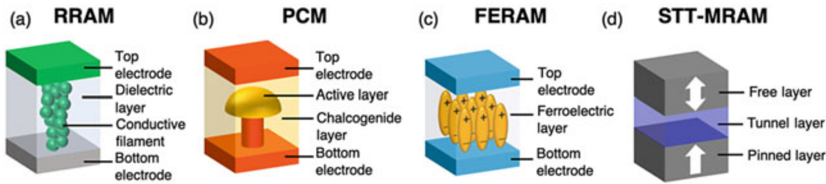



**Fig. 5.1** Emerging non-volatile memory technologies. Although sharing the same two-terminal structure, the controlled modulation of the device resistance is allowed by different materials and physical phenomena. **a** In resistive-switching memory (RRAM) resistance modulation is due to the formation/disruption of a conductive filament (CF) from the top electrode to the bottom electrode. **b** In phase-change memory (PCM) amorphous-crystalline phase transition in the active layer allows for different resistance states. **c** In ferroelectric random-access memory (FERAM), the orientation of the electric dipoles in the ferroelectric layer causes a permanent polarization, resulting in different resistive states. **d** The magnetic tunnel junction (MTJ) metal-insulator-metal structure is at the core of the spin-transfer torque magnetic memory (STT-MRAM). Here, two resistive states correspond to the two relative magnetic orientation of free and pinned layer

**Table 5.1** Comparative table of emerging memory technologies performances

| Technology | RRAM | PCM | FERAM | STT-MRAM |
|---|---|---|---|---|
| Cell size[a] ($F^2$) | 4–12 | 4–30 | 15–34 | 6–50 |
| Write time (ns) | $<10$ | $\approx 50$ | $\approx 30$ | 1–10 |
| Write current (μA) | 10–100 | 80–200 | $<100$ | $>50$ |
| Endurance | $>10^6$–$10^{12}$ | $>10^9$ | $\approx 10^{10}$ | $>10^{15}$ |

[a]F represents the minimum feature size for a given microelectronics technological node

such as stochastic and neuromorphic computing, which are currently under intense scrutiny by both academia and industry [7–9].

This chapter is focused on the electrical characterization and physical modeling of STT-MRAM technology, with emphasis on its reliability and computing applications. Firstly, the dielectric breakdown-limited cycling endurance is experimentally characterized and understood with a semi-empirical model. Then, the cycle-to-cycle variability occurring in STT-MRAM is studied thanks to a physics-based model of the stochastic switching. Finally, the stochastic switching phenomenon is exploited towards the design of true-random number generator (T-RNG) and spiking neuron for stochastic/neuromorphic computing.

## 5.2 Spin-Transfer Torque Magnetic Memory (STT-MRAM)

Among the various emerging memory technology described in Sect. 5.1, STT-MRAM is attracting a strong interest as storage-class memory (SCM) [5, 10], DRAM replacement [11], and embedded nonvolatile memory [12], due to its fast switching, non-volatility, high endurance, CMOS compatibility and low current operation [13]. In addition, STT-RAM and spintronic devices in general can be implemented in novel non-von Neumann concepts of computing, e.g., as electronic synapse in neural networks [14], nonvolatile logic [15], and random number generator (RNG) [16].

STT magnetic memory has at its core the magnetic tunnel junction (MTJ), which consists of a metal-insulator-metal tri-layered structure comprising a thin MgO tunnel barrier ($t_{MgO} \approx 1$ nm) separating two CoFeB ferromagnetic electrodes (FMs). One of these two electrodes, called pinned layer (PL), has a fixed magnetic polarization, whereas the free layer (FL) polarization can change between parallel (P) and anti-parallel (AP) with respect to the PL. The relative orientation of the magnetic polarization of the FL and PL determines two stable MTJ resistance states as a result of the tunnel magnetoresistance effect [13].

Specifically, the P state has a relatively low resistance $R_P$, while the AP state has a relatively high resistance $R_{AP}$. Electronic switching between the two stable resistive states takes place by spin-transfer torque (STT) effect, where the spin-polarized electrons flowing across the MTJ induce a change in the FL magnetic polarization by angular momentum conservation [17]. Perpendicular spin-transfer torque (p-STT), where the polarization of the two FMs is perpendicular to the MTJ plane, demonstrated lower switching current for the same retention time, thus enabling low-power operation and improved area scalability [18].

Figure 5.2a shows the structure of STT magnetic memory devices used for the experiments described throughout this chapter. It comprises a CoFeB PL (bottom electrode, BE) and FL (top electrode, TE) with a crystalline MgO dielectric layer. Note that the device shows an out-of-plane magnetized easy axis with two stable resistive states, namely a P-state with low resistance and an AP-state with high resistance. Figure 5.2b shows the measured current-voltage (I-V) characteristics under quasi-static ramped voltage (DC) conditions, where the set event, i.e. AP→P transition, occurred at a pos-

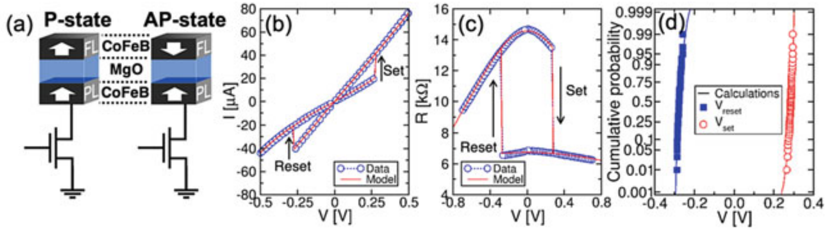



**Fig. 5.2** Perpendicular STT memory device: parallel state and antiparallel state (**a**). Measured and calculated I-V curve in DC conditions (**b**) and corresponding R-V characteristics (**c**), evidencing set (AP→P) and reset (P→AP) transitions. **d** Distributions of stochastic switching voltages $V_{set}$ and $V_{reset}$ for 50 DC cycles. Reproduced under the terms of the CC BY 4.0 Creative Commons license from [20]. Published 2019 IEEE

itive voltage $V_{set} = 0.27$ V. The reset event, i.e. P→AP transition, occurred at a negative voltage $|V_{reset}| = 0.27$ V, underlining the symmetric switching behavior of our samples. Figure 5.2c shows the resistance-voltage (R-V) curve. The figures also show the calculated conduction characteristics by an analytical model [19].

Note that in order to drive the switching current across the MTJ, bipolar voltage pulses are applied, hence a very large electric field develops in the nanometer-thick MgO layer. As a consequence, this might induce degradation and time-dependent dielectric breakdown (TDDB) in the long term. This topic will be described in Sect. 5.3. Cycle-to-cycle repetition of the switching characteristics shows statistical variation of $V_{set}$ and $V_{reset}$. Figure 5.2d shows the distributions of stochastic switching voltages $V_{set}$ and $V_{reset}$ for 50 DC cycles. In the thermal regime of switching, STT-induced switching takes place by random thermal fluctuations, thus featuring an intrinsically stochastic behavior. Sections 5.4–5.5 will study stochastic switching and describe some of its applications towards computing, respectively.

## 5.3 Understanding Dielectric Breakdown-Limited Cycling Endurance

Although the cycling endurance of STT-MRAM is sometimes referred to as virtually infinite [21], the repeated electrical stress during switching operation leads to a breakdown-limited endurance lifetime. This poses a limitation on the applicability of STT-MRAM as working memory or in-memory computing element, where extended cycling endurance is often a paramount requirement. Despite such relevant need for high endurance, the characterization methodology, the physical understanding and simulation models for breakdown-limited endurance are not yet well established. Therefore, here are presented an experimental study and a semi-empirical model of endurance failure in p-STT for the prediction of STT-MRAM lifetime.

Figure 5.3a shows the experimental set-up for the pulsed characterization of STT devices, including a waveform generator to apply triangular pulses for set (transi-

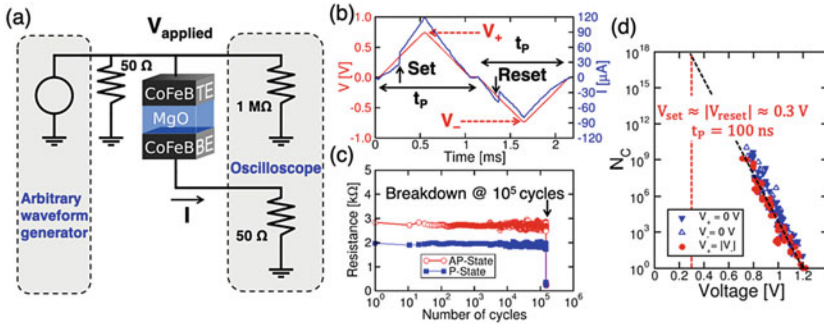



**Fig. 5.3** Experimental setup for real-time monitoring of the switching characteristics during AC cycling of the p-STT devices (**a**) and measured waveforms of voltage and current (**b**). The typical voltage waveform for switching characterization and endurance experiments comprises a positive and a negative triangular pulses, which induce set and reset transitions, respectively. P and AP measured resistances during cycling (**c**), showing TMR $\approx$ 50% and endurance failure after $1.5*10^5$ cycles; median values over 10 reads are shown. After the MgO breakdown the device showed a resistance of 300 Ω, corresponding to the contact resistance. Number of cycles at endurance failure $N_C$ as a function of the applied stress voltage for symmetric bipolar and for positive/negative unipolar cycling (**d**). Readapted under the terms of the CC BY 4.0 Creative Commons license from [22]. Published 2018 IEEE

tion from AP to P under positive voltage) and reset (transition from P to AP under negative voltage) processes, while the applied $V_{TE}$ voltage and current I across the MTJ were monitored by an oscilloscope. Figure 5.3b shows a typical sequence of positive set, and negative reset, showing the two switching events. By monitoring the switching characteristics at each cycle, the observation of degradation phenomena and the exact event of endurance failure is possible. This event is shown in Fig. 5.3b, which reports the measured resistance during a typical pulsed experiment under symmetric switching ($V_+ = |V_-|$), as a function of the number of cycles. Data evidence clearly separate P and AP states with a TMR $= \Delta R/R_P \approx 50\%$, where $\Delta R = R_{AP} - R_P$. Cycling endurance failure is marked by an abrupt drop of read resistance, corresponding to a hard breakdown of the MgO dielectric layer, after a number $N_C$ of cycles. Such destructive event can be explained by defect generation in MgO inducing a percolative path and thermal runaway [23]. After breakdown, the device shows a TMR of 0% and a constant resistance $R \approx 300\,\Omega$ which can be attributed to the metal contacts and interfaces.

Figure 5.3c shows the measured cycling endurance $N_C$ as a function of the applied voltage with a pulse-width $t_P = 100$ ns. Three cycling conditions are compared in the figure, i.e., symmetric bipolar stress with $V_+ = |V_-|$, positive unipolar stress with $V_- = 0$ V and negative unipolar stress with $V_+ = 0$ V. $N_C$ data for positive and negative unipolar stress show similar behaviors, evidencing a steep exponential voltage dependence with a slope $\approx$50 mV/dec for the three regimes in the figure. A simple extrapolation to the switching voltage indicates an estimated $N_C \approx 10^{18}$ at $V = 0.3$ V and $t_P = 100$ ns, which is high enough to comply with most SCM and DRAM applications. Figure 5.4a shows cycling endurance for asymmetric bipolar

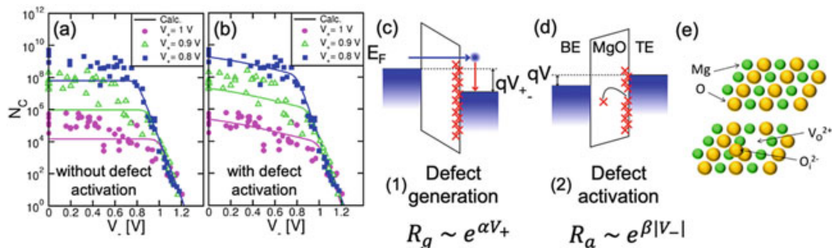



**Fig. 5.4** Schematic illustration of the semi-empirical model of MgO breakdown, comprising a defect generation phase and activation. The defects could be considered to be $O_i^{2-}-V_O^{2+}$ Frenkel pairs (**e**). Measured and calculated $N_C$ taking into account only defect generation process (**a**). **b** Calculated cycling endurance considering also defect activation process, demonstrating good agreement with experimental data. Reproduced under the terms of the CC BY 4.0 Creative Commons license from [22]. Published 2018 IEEE

stress, with variable $V_-$ and constant $V_+ = 1$, 0.9 and 0.8 V. The voltage dependence of $N_C$ data shows two distinct regions, namely (i) region A for $|V_-| > V_+$ where data show a steep slope ≈50 mV/dec, and (ii) region B for $|V_-| < V_+$ with reduced slope ≈600 mV/dec. In order to describe the dependence of $N_C$ on the voltage amplitude of the applied signal, a semi-empirical model of cycling endurance was developed [19, 22]. In this model, $N_C$ is inversely proportional to the defect concentration within the MgO layer, namely $N_C = N_{C0}(n_D/n_{D0})^{-1}$, where $N_{C0}$ and $n_{D0}$ are constant and $n_D$ was calculated as $n_D = n_{D,TE} + n_{D,BE}$, where $n_{D,TE}$ and $n_{D,BE}$ are the defect concentrations originating from the TE interface and the BE interface, respectively. Defect concentrations are given by $n_{D,TE} = n_{D0}*R_{TE}/R_0$ and $n_{D,BE} = n_{D0}*R_{BE}/R_0$, where $R_{TE}$ and $R_{BE}$ are the generation rates describing the cycling-induced degradation at the TE and BE interfaces, respectively, while $R_0$ is a constant. In crystalline MgO layer, defects might be attributed, e.g., to Frenkel pairs of O vacancies $V_O^{2+}$ and O interstitials $O_i^{2-}$ as shown Fig. 5.4e. As depicted in Fig. 5.4c, d, tunneling electrons are considered to have a primary role in MgO degradation according to a 2-stage mechanism, including (1) defect generation (Fig. 5.4c) and (2) defect activation (Fig. 5.4d).

In this model, defects are firstly generated by bond breaking due to tunneling electrons releasing their kinetic energy E to the lattice. Here, defect generation probability is assumed to increase exponentially with the energy E, thus the generation rate is given by $R_{TE} = R_0\exp(\alpha V_+)$, where $\alpha$ is a constant. Similarly, the generation rate at the BE interface can be written as $R_{BE} = R_0\exp(\alpha|V_-|)$. As demonstrated in Fig. 5.4a, the model correctly describes the steep decay of $N_C$ in region A, however the model fails to predict the weak voltage dependence in region B. To account for the impact of the smaller voltage in the MgO degradation, we considered the defect activation mechanism displayed in Fig. 5.4d. After a positive pulse of voltage $V_+$, the application of a negative pulse with amplitude $|V_-| < V_+$ can activate the defects generated by the positive semi-cycle, e.g. by displacing an interstitial oxygen ion away from the corresponding O vacancy in the newly formed Frenkel pair.

Figure 5.4b reports calculated cycling endurance with both defect generation and activation, indicating a better agreement with data in both regions A and B.

To complete the endurance model, defect generation and activation at both interfaces is considered. Moreover, an explicit dependence on the pulse-widths $t_+$ and $t_-$ of the positive and negative pulses, respectively. The total defect density due to generation and activation is thus written as:

$$n_D = n_{D0}\left[\frac{t_+}{t_0}\,e^{\alpha V_+}e^{\beta|V_-|} + \frac{t_-}{t_0}\,e^{\alpha|V_-|}e^{\beta V_+}\right], \tag{5.1}$$

## 5.4 Modeling Stochastic Switching in STT-MRAM

STT-based circuits for both memory and computing applications require accurate compact model for physics-based simulations. Various STT switching models rely on numerical simulations from the Landau–Lifshitz–Gilbert (LLG) equation [24], which are typically too computing-intensive for electronic circuit simulators. Thus, simple analytical compact models are the ideal candidates for such tasks [25]. Most analytical STT switching models are limited to the thermal regime (>200 ns) and the precession regime (<1 ns), even though practical STT-RAM applications mostly work in the intermediate regime [19, 20].

To better address the cycle-to-cycle statistical variation of STT switching, the write error rate (WER) was measured, i.e., the failure rate of the switching transition according to the experimental technique shown in Fig. 5.5. Figure 5.5a shows the voltage waveform applied to characterize the WER of set transition, consisting of: (1) a negative-voltage triangular pulse at $V_- = -0.7$ V to deterministically initialize the cell in the AP state; (2) a negative-voltage square pulse for reading the cell state; (3) a positive-voltage square pulse, with amplitude $V_A$ and duration $t_P$, to induce the stochastic set transition (AP→P); and (4) a positive-voltage square pulse for a final reading to verify the cell state. Pulses (1), (2), and (4) have the same pulsewidth of 1 µs, while the set pulse has a variable pulsewidth $t_P$ ranging from 40 ns to 10 µs. The final state of the cell indicates the success or failure of the transition to P state. The WER is then defined as the ratio between the number of write failures and the total number of applied cycles. Figure 5.5b shows a similar waveform for the dual experiment, namely the evaluation of reset WER. Figure 5.5c, d shows the measured WER as a function of the voltage V applied during the stochastic square pulse for (c) reset transition and (d) set transition at increasing pulsewidth $t_P$. The WER drops almost exponentially at increasing voltage. As $t_P$ decreases, the transition to low WER occurs at increasing voltage and with decreasing slope. Data in Fig. 5.5c, d are summarized in Fig. 5.5e, showing the Weibull scale parameter $V_{63\%}$, defined as the voltage for WER = 63%, as a function of $t_P$. The time-voltage relationship in Fig. 5.5e is usually explained by the thermal model depicted in Fig. 5.5f, g. Considering AP→P transition, the FL polarization experiences thermal fluctuations within the AP-state well, while the current-induced spin torque causes lowering of the PMA barrier

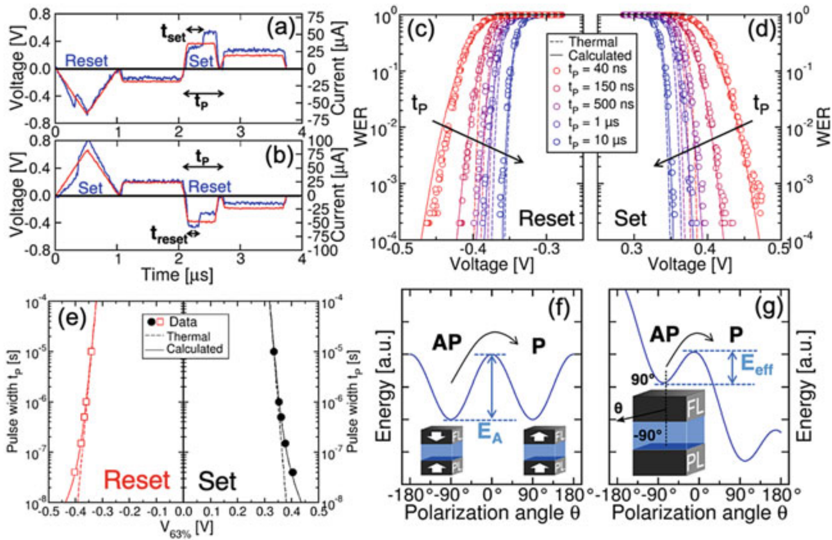



**Fig. 5.5** Applied voltage waveform and corresponding read current during a WER experiment for **(a)** set and **(b)** reset transitions. **c–d** Experimental results and calculations for WER as a function of the applied voltage for different pulse widths $t_P$. The slope reduction for shorter $t_P$ is correctly predicted by the new compact model. **e** Measured and calculated Weibull scale parameter $V_{63\%}$ for set and reset. PMA energy profile as a function of the FL magnetic polarization angle $\theta$. Thermal fluctuations induce AP→P transition across the energy barrier **(f)** with no applied voltage or **(g)** with positive applied voltage. The energy unbalance between AP and P states originates from the current-induced STT. Reproduced under the terms of the CC BY 4.0 Creative Commons license from [20]. Published 2019 IEEE

$E_A$ and the consequent transition to the P-state. Assuming a linear voltage-induced barrier lowering, the characteristic switching time is given by:

$$\tau_{th} = \tau_0 e^{\Delta\left(1-\frac{V}{V_{c0}}\right)}, \tag{5.2}$$

where $\Delta = E_A/kT$, while $\tau_0$ and $V_{c0}$ are constant [26]. The WER can thus be obtained by a Poissonian switching probability P given by:

$$\frac{dP}{dt} = \frac{(1-P)}{\tau}, \tag{5.3}$$

where $\tau$ is the characteristic switching time equal to $\tau_{th}$.

Calculations with the thermal model are reported in Figs. 5.5c, d and 5.6a, b, where the Weibull plot, i.e., log(−log(WER)) is shown. Note that the thermal model correctly predicts the linear behavior for slower $t_P$. However, it cannot explain the deviation from the linear behavior at $t_P < 200$ ns [20, 27]. To this purpose, Fig. 5.6c, d reports the Weibull shape factor dlog(−log(WER))/dV as a function of voltage, compared with the ideal value $\Delta/V_{c0}$ (i.e., the barrier lowering coefficient) from

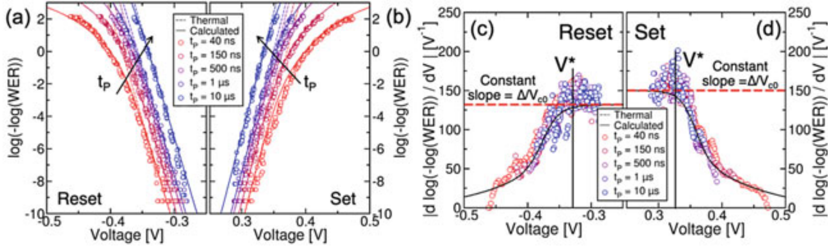



**Fig. 5.6** **a**–**b** Voltage distributions in a Weibull plot, i.e. log(−log(WER)). Data show a deviation from the expected linear dependence in thermal regime to an anomalous non-linear shape for $t_P < 200$ ns, correctly described by the compact model. **c**–**d** Experimental and calculated Weibull shape factor dlog(−log(WER))/dV as a function of applied voltage V, showing a drop of the barrier lowering coefficient after a critical V*, instead of the constant value predicted by thermal model. Reproduced under the terms of the CC BY 4.0 Creative Commons license from [20]. Published 2019 IEEE

Eqs. 5.2–5.3. Data show a drop of the Weibull shape factor beyond a critical voltage V* ≈ 0.33 V. To account for such anomalous barrier lowering at high voltage/short times, the characteristic switching time in Eq. 5.2 was rewritten as $\tau = \tau_{th} + \tau'_{th}$, where $\tau'_{th}$ is an additional time given by:

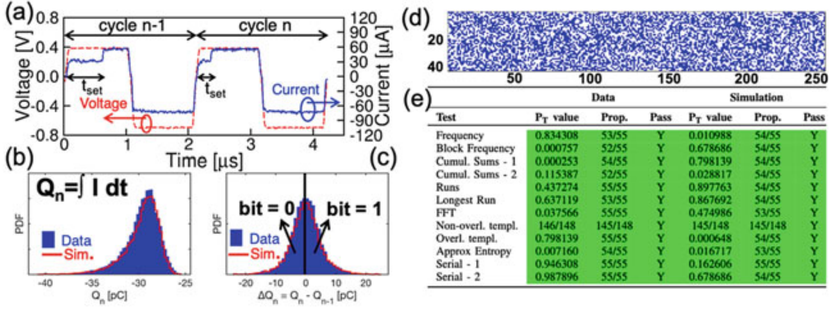
$$\tau'_{th} = \tau_0 e^{\Delta'\left(1-erf\left(\frac{V}{V'_{c0}}\right)\right)}. \tag{5.4}$$

Figures 5.5 and 5.6 show the calculation obtained from Eqs. 5.3–5.4, demonstrating the strength of the compact model to account for WER with V-dependent Weibull shape factor in both thermal (>200 ns) and intermediate regimes (<200 ns) [20].

## 5.5 Stochastic STT Switching for Security and Computing

Although stochastic switching variability is harmful to the operation of STT-based magnetic memory, it is considered beneficial for emerging concept such as true-random number generator (TRNG) [16], stochastic computing [28] and brain-inspired computing [29].

On-chip generation of true random numbers is a key feature for hardware and data security for IoT. STT-based RNG can be designed as follows [16]: repeated square set/reset pulses are applied to STT cell, resulting in the stochastic set/reset events in Fig. 5.7a. Due to the stochastic $t_{set}$, integration of the current along the n-th set/reset cycle leads to a broadly-distributed charge $Q_n = \int I dt$ (Fig. 5.7b). The difference $Q_n - Q_{n-1}$ over two consecutive cycles (Fig. 5.7c) can have either positive or negative values with 50% probability, which are then associated to random bit values 1 or 0, respectively. Figure 5.7d, e shows a sample of generated random bits and the NIST statistical test of random bits obtained by experiments and simulation with the compact model. All tests are passed, thus supporting the feasibility of STT-based TRNG.




| Test | Data P_T value | Data Prop. | Data Pass | Simulation P_T value | Simulation Prop. | Simulation Pass |
|---|---|---|---|---|---|---|
| Frequency | 0.834308 | 53/55 | Y | 0.010988 | 54/55 | Y |
| Block Frequency | 0.000757 | 52/55 | Y | 0.678686 | 54/55 | Y |
| Cumul. Sums - 1 | 0.000253 | 54/55 | Y | 0.798139 | 54/55 | Y |
| Cumul. Sums - 2 | 0.115387 | 52/55 | Y | 0.028817 | 54/55 | Y |
| Runs | 0.437274 | 55/55 | Y | 0.897763 | 54/55 | Y |
| Longest Run | 0.637119 | 53/55 | Y | 0.867692 | 54/55 | Y |
| FFT | 0.037566 | 55/55 | Y | 0.474986 | 53/55 | Y |
| Non-overl. templ. | 146/148 | 145/148 | Y | 145/148 | 145/148 | Y |
| Overl. templ. | 0.798139 | 55/55 | Y | 0.000648 | 54/55 | Y |
| Approx Entropy | 0.007160 | 54/55 | Y | 0.016717 | 53/55 | Y |
| Serial - 1 | 0.946308 | 55/55 | Y | 0.162606 | 55/55 | Y |
| Serial - 2 | 0.987896 | 55/55 | Y | 0.678686 | 54/55 | Y |

**Fig. 5.7** STT-based TRNG concept. **a** Measured square voltage pulses and current response for 2 consecutive cycles n−1 and n, (**b**) PDF of the integrated current $Q_n$ and (**c**) PDF of differential charge $\Delta Q_n = Q_n - Q_{n-1}$. **d** Representation of 10 kb from the random bitstream(blue dot = 0, white dot = 1). **e** NIST test on 55 sequences from 1 Mb output bitstream is passed, demonstrating good randomness quality. Copyright 2019, ACM/IEEE. Reprinted with permission from [27]

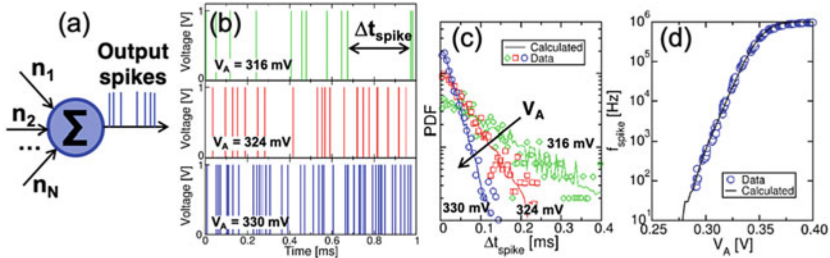



**Fig. 5.8** Schematic illustration of the stochastic spiking neuron (**a**) and measured output spiking activity for increasing applied voltage $V_A$ (**b**). Experimental and simulated exponential distributions for spike-to-spike time $\Delta t_{spike}$ (**c**), demonstrating exponential spiking statistics. **d** Experimental and calculated average voltage-dependent spiking frequency Copyright 2019, ACM/IEEE. Reprinted with permission from [27]

Set/reset switching variability in STT-MRAM can also be exploited towards stochastic spiking neurons (Fig. 5.8a) [27]. In the proposed spiking neuron, a STT-MRAM cell is biased at constant voltage $V_A$ to induce stochastic set transition from the AP state. As the switching occurs, (1) an output spike is generated and (2) the cell is reinitialized in the AP state. Figure 5.8b shows the experimental output spiking activity at increasing $V_A$. Spikes obey a Poissonian statistics as indicated by the exponential distributions of spike-to-spike time $\Delta t_{spike}$ in Fig. 5.8c. Figure 5.8d shows the average spiking frequency $f_{spike} = 1/<\Delta t_{spike}>$ at increasing voltage, demonstrating a voltage-controlled spiking neuron.

Note that the compact model for stochastic switching (Sect. 5.4) fully supports the design and simulation of STT devices for security and computing, as demonstrated by the calculations shown in Figs. 5.7 and 5.8.

The applicability of the model to computing with spiking neurons was demonstrated for analogue multiplication (stochastic computing) and SRDP-based [30] spiking neural network [27].

## 5.6 Conclusions

In this chapter, STT-MRAM are described as a promising candidate for storage-class memory concept and emerging computing applications. To fully understand the applicability and limitations of this emerging memory technology, various reliability aspects were thoroughly studied. Firstly, a comprehensive study of breakdown-limited cycling endurance was presented. Cycling endurance was experimentally monitored as a function of the pulse amplitude, polarity and timing. A semi-empirical model based on defect generation and activation in the MgO tunnel barrier allowed for endurance understanding and prediction. Then, random switching in variability in STT-MRAM was characterized and described through a physics-based stochastic switching model. The compact model is capable of computing switching probability with $10^{-4}$ accuracy in the thermal regime ($>$200 ns) and intermediate regime ($<$200 ns), accounting for WER data of STT-MRAM as a function of the applied voltage for various pulse-widths. Finally, STT switching variability is exploited towards the implementation of true RNG, spiking neurons, analogue stochastic multiplication and SNNs. The stochastic switching compact model is demonstrated as an useful tool for design and simulation of future STT-based hardware security and computing primitives.

## References

1. Moore GE (1965) Cramming more components onto integrated circuits. Electronics 38(8)
2. Ball P (2012) Computer engineering: feeling the heat. Nat News 492(7428):174
3. Salahuddin S, Ni K, Datta S (2018) The era of hyper-scaling in electronics. Nat Electron 1(8):442
4. Backus J (1978) Can programming be liberated from the von Neumann style? a functional style and its algebra of programs. Commun ACM 21(8):613–641
5. Freitas RF, Wilcke WW (2008) Storage-class memory: the next storage system technology. IBM J Res Dev 52(4/5):439
6. Burr GW et al (2008) Overview of candidate device technologies for storage-class memory. IBM J Res Dev 52(4.5):449–464
7. Ielmini D, Wong H-SP (2018) In-memory computing with resistive switching devices. Nat Electron 1(6):333
8. Carboni R, Ielmini D (2019) Stochastic memory devices for security and computing. Adv Electron Mater 1900198
9. Carboni R, Ielmini D (2020) Applications of resistive switching memory as hardware security primitive. In: Applications of emerging memory technology. Springer, Berlin, pp 93–131
10. Wong H-SP, Salahuddin S (2015) Memory leads the way to better computing. Nat Nanotechnol 10(3):191
11. Ikegami K et al (2014) Low power and high density STT-MRAM for embedded cache memory using advanced perpendicular MTJ integrations and asymmetric compensation techniques. In: IEEE international electron devices meeting
12. Shum D et al (2017) CMOS-embedded STT- MRAM arrays in 2x nm nodes for GP-MCU applications. In: Symposium on VLSI technology, pp T208–T209
13. Apalkov D, Dieny B, Slaughter J (2016) Magnetoresistive random access memory. Proc IEEE 104(10):1796–1830

14. Lequeux S et al (2016) A magnetic synapse: multilevel spin-torque memristor with perpendicular anisotropy. Sci Rep 6:31510
15. Mahmoudi H et al (2013) Implication logic gates using spin-transfer-torque-operated magnetic tunnel junctions for intrinsic logic-in-memory. Solid-State Electron 84
16. Carboni R et al (2018) Random number generation by differential read of stochastic switching in spin-transfer torque memory. IEEE Electron Device Lett 39:951–954
17. Slonczewski JC (1996) Current-driven excitation of magnetic multilayers. J Magn Magn Mater 159(1–2):L1–L7
18. Ikeda S (2010) A perpendicular-anisotropy CoFeB-MgO magnetic tunnel junction. Nat Mater 9(9):721
19. Carboni R et al (2016) Understanding cycling endurance in perpendicular spin-transfer torque (p-STT) magnetic memory. In: IEEE international electron devices meeting (IEDM)
20. Carboni R, Vernocchi E et al (2019) A physics-based compact model of stochastic switching in spin-transfer torque magnetic memory. IEEE Trans Electron Devices 60(10):4176–4182
21. Kan JJ et al (2017) A study on practically unlimited endurance of STT-MRAM. IEEE Trans Electron Devices 64(9):3639–3646
22. Carboni R et al (2018) Modeling of breakdown-limited endurance in spin-transfer torque magnetic memory under pulsed cycling regime. IEEE Trans Electron Devices 65(6):2470–2478
23. Yoshida C et al (2009) A study of dielectric breakdown mechanism in CoFeB/MgO/CoFeB magnetic tunnel junction. In: IEEE international reliability physics symposium (IRPS)
24. Xu N et al (2015) Physics-based compact modeling framework for state-of-the-art and emerging STT-MRAM technology. In: IEEE international electron devices meeting (IEDM), pp 28.5.1–28.5.4
25. Ielmini D, Milo V (2017) Physics-based modeling approaches of resistive switching devices for memory and in-memory computing applications. J Comput Electron 16
26. Heindl R et al (2011) Validity of the thermal activation model for spin-transfer torque switching in magnetic tunnel junctions. J Appl Phys 109(7)
27. Carboni R, Vernocchi E et al (2019) A compact model of stochastic switching in STT magnetic RAM for memory and computing. In: ACM/IEEE international symposium on nanoscale architectures (NANOARCH), pp 5.2.1–5.2.6,
28. Lv Y, Wang J-P (2017) A single magnetic-tunnel-junction stochastic computing unit. In: IEEE international electron devices meeting (IEDM),
29. Grollier J et al (2016) Spintronic nanodevices for bioinspired computing. Proc IEEE 104(10):2024–2039
30. Milo V et al (2018) A 4-transistors/1-resistor hybrid synapse based on resistive switching memory (RRAM) capable of spike-rate-dependent plasticity (SRDP). IEEE Trans Very Large Scale Integr (VLSI) Syst 26(12):2806–2815

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.