


Nanosecond True-Random-Number Generation with Superparamagnetic Tunnel Junctions: Identification of Joule Heating and Spin-Transfer-Torque Effects

Leo Schnitzspan^{1,2}, Mathias Kläui^{1,2,*}, and Gerhard Jakob^{1,2}

¹*Institute of Physics, Johannes Gutenberg-Universität Mainz, Mainz 55122, Germany*

²*Max Planck Graduate Center Mainz, Mainz 55122, Germany*

 (Received 13 January 2023; revised 24 May 2023; accepted 28 June 2023; published 1 August 2023)

This work investigates nanosecond superparamagnetic switching in 50-nm-diameter in-plane magnetized magnetic tunnel junctions (MTJs). Due to the small in-plane uniaxial anisotropy, dwell times below 10 ns and autocorrelation times down to 5 ns are measured for circular superparamagnetic tunnel junctions (SMTJs). SMTJs exhibit probabilistic switching of the magnetic free layer, which can be used for the generation of true random numbers. The quality of the random bit streams generated by our SMTJ is evaluated with a statistical test suite (NIST STS, sp 800-22) and shows true randomness after three exclusive OR (XOR) operations of four random SMTJ bit streams. A low-footprint CMOS circuit is proposed for fast and energy-efficient random-number generation. We demonstrate that the probability of a 1 or 0 can be tuned by spin-transfer torque (STT), while the average bit-generation rate is mainly affected by the current density via Joule heating. Although both effects are always present in MTJs, most often Joule heating is neglected. However, with a resistance-area (RA) product of $15\ \Omega\ \mu\text{m}^2$ and current densities of the order of $1\ \text{MA}/\text{cm}^2$, an increasing temperature at the tunneling site results in a significant increase in the switching rate. As Joule heating and STT scale differently with current density, the device design can be optimized based on our findings.

DOI: [10.1103/PhysRevApplied.20.024002](https://doi.org/10.1103/PhysRevApplied.20.024002)

I. INTRODUCTION

A superparamagnetic tunnel junction (SMTJ), acting as a random stochastic noise source, is the counterpart to nonvolatile random access memory (MRAM) [1], where high temporal stability (switching energy barrier $E_b > 40 k_B T$) of the state is required to store the information of a bit. The volatile behavior of the superparamagnetic junction originates from the ambient thermal energy acting on the magnetization of the free layer. This thermal energy is high enough to overcome the low energy barrier of a few $k_B T$, resulting in a superparamagnetic state with fluctuation times of milliseconds down to nanoseconds at room temperature. The time spent in each state, called the dwell time, can be controlled by an applied external magnetic field, by spin-transfer-torque, or by spin-orbit torque [2]. Due to the inherent probabilistic nature of SMTJs and their state controllability and energy efficiency, SMTJs have been proposed for various computational concepts [3], such as invertible logic [4], Boltzmann machines [5], reservoir computing [6], spiking neural networks [7,8], or stochastic computing [9]. These concepts can provide advantages over pure digital complementary metal-oxide-semiconductor (CMOS) based computational

logic, since the fundamental building block is a stochastic bit. Despite the tremendous development of conventional CMOS-based deterministic computers over recent decades, there are still classes of problems that cannot be addressed efficiently, due to the deterministic nature of von Neumann computers. Many numerical-computation techniques are based on the Markov chain Monte Carlo (MCMC) methods [10] and require many random numbers. Since SMTJs are inherently probabilistic, as opposed to deterministic CMOS circuits, they can provide random signals with low power and a low areal footprint, which can be transformed to random bit streams using a few CMOS transistors. MTJs are already CMOS compatible on scale for different applications such as MRAM [1]. Through the combination of CMOS logic and a set of probabilistic SMTJs (called p-bits [4]), it has been shown that computationally hard problems such as the “traveling-salesman problem” can be solved efficiently [11]. The performance of such a “p-computer” will then be a consequence of the p-bit hardware density and the average fluctuation rate, described by the Néel-Arrhenius law. Fast fluctuations and dwell times in the nanosecond range are desired. For in-plane easy-axis MTJs (ip MTJs), dwell times of the order of milliseconds [12], microseconds [13–15], and down to nanoseconds [16,17] have been reported. Compared to out-of-plane MTJs (oop MTJs), ip

*mathias.klaui@klaui-lab.de

MTJs often exhibit shorter dwell times, due to a different contribution of in-plane and out-of-plane anisotropy energies [18]. Out-of-plane MTJs often comprise larger time scales of milliseconds [19,20] to microseconds [21–23] and for this reason, we have decided to design our MTJ stack with in-plane ferromagnetic layers. We demonstrate that two phenomena occur when a current is applied to an MTJ nanopillar: spin-transfer torque (STT) and Joule heating. STT refers to the transfer of angular momentum from the flowing electrons to the magnetic moments in the ferromagnetic layer. Joule heating, on the other hand, refers to the generation of heat induced by the flow of charge current through the device. Both effects are always present in an MTJ and have to be considered to understand the dependence of the fluctuation times and the current density. So far, Joule heating and STT have been studied together on superparamagnetic nanoislands with spin-polarized scanning tunneling microscopy [24] but Joule heating has mostly been ignored or neglected for SMTJs, where only STT effects have been considered.

In this work, we separate and extract the contributions of STT and Joule heating for different current densities to understand the mechanism of the superparamagnetic switching dependence. We show that STT affects the state energy linearly with applied current, while Joule heating has a quadratic dependence on the current. Due to the low energy barrier between the two states, Joule heating results in an overall faster switching, while STT tunes the state probability differently. In the case of a random-number generator, it is important to control the state probability to a desired value. An implementation idea based on logic exclusive OR (XOR) gates is proposed in order to generate a stream of true random bits. The quality of the randomness is quantified by the Statistical Test Suite from NIST [25]. We show that true randomness, with a bit-generation rate of 200 Mbit/s, is achieved by using four SMTJs and a combination of three XOR gates. It is also demonstrated that for the generation of multiple independent true random bit streams, two XOR gates and two SMTJs are sufficient.

This circuit design saves energy and space and therefore has the potential for applications where many random bits are required.

II. METHODS

A. MTJ sample preparation

TMR stacks were deposited at room temperature on oxidized Si substrates using rf- and dc-magnetron sputtering (Singulus Rotaris) with the composition Ta(10 nm)/Ru(10 nm)/Ta(10 nm)/(Pt-Mn)(20 nm)/(Co-Fe)(2.2 nm)/Ru(0.8 nm)/(Co-Fe-B)(2.4 nm)/MgO(1.1 nm)/(Co-Fe-B)(3.0 nm)/Ta(10 nm)/Ru(20 nm) and is based on an optimized stack, developed earlier [26]. The stack is illustrated in Fig. 1(a), where the function of each layer is specified. The tunnel magnetoresistance (TMR) ratio of a continuous film of our stack is found to be approximately 150% and for patterned nanopillars it is in the range of 100–150%.

The state of the MTJ (parallel or antiparallel) is determined by the magnetization orientation of the free layer with respect to the reference layer. The Co-Fe-B reference layer together with a Co-Fe pinned layer forms a synthetic antiferromagnet (SAF), in order to compensate stray fields at the free-layer site. In addition, the pinned Co-Fe layer is exchange biased [27] by a Pt-Mn antiferromagnet (AFM). Even though a SAF is integrated into the stack, the stray field at the free-layer position will never be compensated exactly to zero [28]. For this reason, an in-plane easy-axis field of a few millitesla is typically applied in order to compensate for this “offset” and to set the SMTJ into equal state probability. This offset is approximately 1.1 mT in Fig. 1(b) but can vary from device to device due to variations in the sample fabrication process. In general, in the case of a stochastic MTJ, an in-plane offset field is selected to achieve an equal occupancy of both the parallel (P) and antiparallel (AP) states, aiming for an approximately 50% state probability. As shown in Fig. 1(b), the state probability can be tuned by the applied in-plane easy-axis field. A

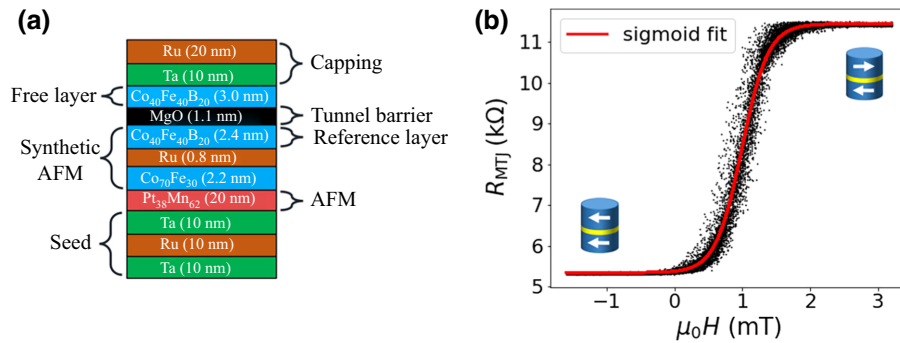


FIG. 1. (a) A TMR stack with an RA product of $15\Omega \mu\text{m}^2$ and a TMR ratio of 150%, used for SMTJs. (b) The MTJ resistance versus the external applied easy-axis field of a SMTJ, plotted for 20 repetitions. The red line indicates the sigmoidal relation of the resistance and the magnetic field. Negative fields stabilize the parallel state, while positive fields stabilize the antiparallel state.

change in the field strength results in a shift of the energy levels, thus stabilizing either the P or the AP state. The relative increase in resistance from the P to the AP state is defined as the TMR ratio: $\text{TMR} = R_{\text{AP}}/R_{\text{P}} - 1$. To obtain a high TMR ratio, annealing was carried out in a 300-mT in-plane field at 300 °C for 1 h. The current in-plane tunneling (CIPT) measurements [29] of our unpatterned stack exhibit a resistance-area (RA) product of $15 \pm 2 \text{ } \Omega \text{ } \mu\text{m}^2$ and a TMR ratio of $150 \pm 3\%$. The MTJ nanopillars are patterned using 30-kV electron-beam lithography, to define a circular etch mask. Argon-ion etching under 55° and 20° is used to structure the MTJ pillar. As a passivation layer, SiN_x was chosen for its hardness and good insulating properties. It was sputtered under a 50:50 mixed argon-nitrogen atmosphere at 5 Pa. Above the MTJ nanopillar, a 60-nm-thick gold pad was sputtered as a top-electrode lead in order to be able to bond it to a measurement sample holder. By using the top electrode and the conducting seed layer as a bottom electrode, a voltage was applied across the junction and the switching behavior was studied.

B. Setup and measurement

Electrical measurements were carried out at room temperature with an MTJ in series with a shunt resistor R_s . The amplified signal of the SMTJ fluctuation was measured with a 3.5-GHz-bandwidth digital oscilloscope (Tektronix DPO7543, 40 Gs/s). A 50- Ω low-noise amplifier (ZFL-1000LN+, mini-circuits) with a gain of 20 dB was used to amplify the signal. To ensure that the time-series measurement has a high time resolution and is not limited by the RC time constant of the setup, the parasitic capacitance and resistances of the setup have to be small. The 50- Ω input impedance of the amplifier results in a small signal voltage drop of the order of 1 mV at the amplifier. Since the MTJ resistor (5–10 k Ω) is in parallel with the 50- Ω input impedance of the amplifier, the equivalent resistance is approximately also 50 Ω (neglecting the impedance of C_i for high frequencies). Forming a voltage divider with R_s leads to a small ac voltage under stochastic fluctuation of R_{MTJ} [see Eq. (VI) in Appendix A 1]. In addition, a low

resistance lowers the RC time constant of the setup, thus leading to a high time resolution of approximately 1 ns ($R \approx 50 \text{ } \Omega$ and $C \approx 20 \text{ pF}$). Therefore, the time series of the fluctuating voltage signal are typically recorded with a sampling rate of 1 GS/s for a few milliseconds. The measurement setup is illustrated in Fig. 2(a) and the recorded nanosecond time series is plotted in Fig. 2(b).

III. DWELL TIMES AND AUTOCORRELATION TIMES

A high random bit-generation rate is primarily limited by the average fluctuation rate or dwell time. The dwell time in the macrospin approximation for $E_b/k_B T \geq 1$ follows the Néel-Arrhenius law [30]: $\tau_{\text{P,AP}} = \tau_0 \exp(E_b/k_B T)$, where E_b is the energy barrier between the states, T is the temperature, and τ_0 is the attempt time. The assumption of the Néel-Arrhenius model is switching between two distinct energy minima, i.e., two distinct resistance states. For very low energy barriers with respect to the thermal energy, a Fokker-Planck–based uniaxial low barrier solution as formulated by Brown will be needed [31]. As we can be close to that regime, our extracted barriers should be considered as the effective barriers for a Néel-Arrhenius–model description. In order to generate true random bits, the prerequisite is the independence of each generated bit with respect to the previously generated bit. An important metric for this is the autocorrelation time τ_{ac} of the SMTJ signal, which can be interpreted as the time scale for randomness and describes for which time interval the signal is still correlated to its past. As shown theoretically [32], MTJs with in-plane magnetic anisotropy exhibit shorter autocorrelation times than MTJs with out-of-plane anisotropy due to the difference in precessionlike fluctuations, attributed to the large demagnetization field for ip MTJs. In the macrospin approximation with a single normalized magnetization vector $m = M/M_s$ fluctuating randomly on a Bloch sphere, the anisotropy energies will restrict the fluctuation to a specific region in the sphere. The demagnetization energy constrains the fluctuation to the equator of the Bloch sphere and the uniaxial anisotropy

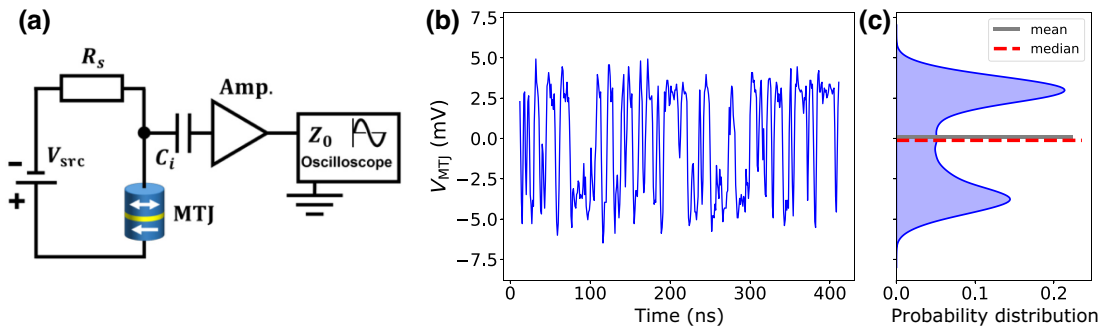


FIG. 2. (a) The circuit diagram of the measurement setup. (b) An example of the nanosecond stochastic switching of the amplified SMTJ signal. (c) The probability distribution (kernel-density estimate) of the SMTJ signal for 10^5 data points.

to regions near the easy-axis directions. In order to measure these superparamagnetic fluctuations, a circular nanopillar MTJ with 50 nm in diameter was patterned by electron-beam lithography and structured with argon-ion etching. Next, a small in-plane offset field of a few millitesla was applied to tune the stochastic MTJ to a 50:50 state probability. With the measurement setup shown in Fig. 2(a), stochastic fluctuations have been acquired, where the dc component of the signal is filtered out by the capacitance in the circuit. The time between two consecutive switching events is defined as the dwell time and is measured for the parallel as well as the antiparallel state. The dwell times can be determined in two ways. First, a time series measurement of 0.5 s (sampling rate of 1 GS/s) is binarized, by using the median as a reference threshold, and then the series is divided into an array of AP (= 1) and P (= 0) dwell-time intervals. The mean of all AP and P time intervals then corresponds to the average dwell time for the AP and P states. The average dwell time τ is calculated by $\sqrt{\tau_P \tau_{AP}}$, as suggested in the literature [17]. For the average dwell time, we find $\tau = 7.3 \pm 0.2$ ns. Second, the switching mechanism can be described by a Poisson process, where the probability for a switching event follows the Poisson distribution, similarly to what happens in a radioactive decay, where the decay probability is described by a Poisson process. By fitting the number of switching events for the AP and P states with $N = N_0 \exp(-t/\tau_{(p,ap)})$, the average dwell times τ_P , τ_{AP} can be extracted in a more robust manner, as shown in Fig. 3. All of the calculated (average) dwell times in this work are determined by the second method, through fitting a Poisson distribution to the dwell times. Figure 3 indicates that the distribution of dwell times is in agreement with the theory, as shown by a linear decrease in the logarithmic plot. The average dwell time is found to be $\tau = 6.7 \pm 0.1$ ns, which is orders of magnitude shorter than most measured dwell times of oop-SMTJs or ip-SMTJs. We further determine the autocorrelation time of the binarized signal, which is the limiting factor for the generation of random numbers. The autocorrelation function (ACF) is calculated by the following equation:

$$\text{ACF}(t) = \frac{\sum_{i=1}^{N-t} [s_i - \bar{s}_i][s_{i+t} - \bar{s}_i]}{\sum_{i=1}^N (s_i - \bar{s}_i)^2}, \quad (1)$$

where s is the binarized signal, \bar{s}_i the mean of the signal, and t is the time lag of the time series of length N . The ACF is plotted in Fig. 4(a) and shows an exponential decrease. After the autocorrelation time τ_{ac} , the signal is uncorrelated with any signal of its past. It is defined as the maximum time by which the integral of the ACF reaches 99%. This results in an autocorrelation time of $\tau_{ac} = 5.1 \pm 0.3$ ns. Next, we consider the rise time of the signal, which was determined by averaging 100 rising edges in a time series (sampling rate 2.5 GHz). Typically, the definition of the

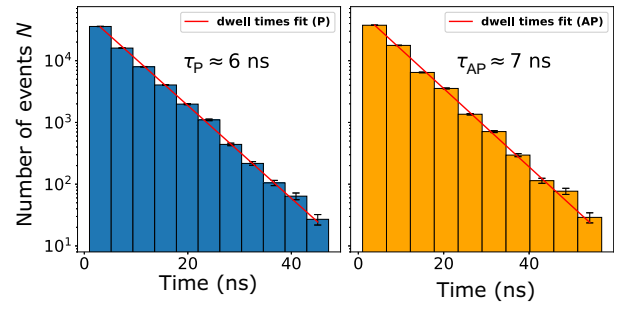


FIG. 3. A histogram of the dwell times for a 1-ms time series for the P and AP states. The bins can be interpreted as a waiting-time interval in a Poisson process, for which exactly one switching event will occur in this interval. From a fit (red line), the average dwell times τ_P and τ_{AP} are determined, according to a Poisson process.

rise time τ_{rise} is a 10%–90% change of amplitude on the rising edge of a signal. However, this stochastic telegraph noise has a noise contribution, coming from the setup, as well as a small stochastic variation in the signal for both the P and AP states due to the thermal excitation of the magnetization vector. For this reason, we define the rise time as the time between the midrange point minus one standard deviation to the midrange point plus one standard deviation: $\tau_{rise} = t_{mid+\sigma} - t_{mid-\sigma}$. The midrange point is defined as $1/2(\max(s) + \min(s))$. Of 100 rising edges, the average rise time is determined to be $\tau_{rise} = 1.5 \pm 0.6$ ns, as shown in Fig. 4(b). The RC time constant has to be shorter than the rise time and is approximately of the order of 1 ns for $R \approx 50\Omega$ and $C \approx 20$ pF. The rising edge (or falling edge) does not indicate an exponential increase (decrease), suggesting that the time resolution of our setup is sufficiently high to acquire the actual signal and that it is not affected by parasitic capacitances in the circuit, which would result in an exponential charge (discharge) characteristics of the signal. For this reason, we only show the shortest measured rise times.

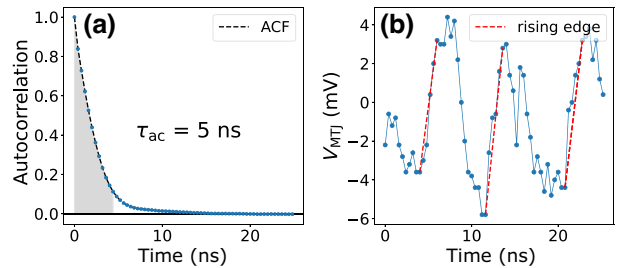


FIG. 4. (a) The autocorrelation function (ACF), plotted for an SMTJ time-series measurement. τ_{ac} is defined as the 99% integral under the ACF curve. (b) A detailed view of the rising and falling edges of the time series. The three red dashed lines show three rising edges, which are approximately linear and of the order of 2 ns.

IV. TRUE-RANDOM-NUMBER GENERATION

Due to the intrinsic stochastic nature of SMTJs, the potential application as an energy-efficient true-random-number generator is of interest and has been proposed and studied by various groups [34–38]. To evaluate the quality of randomness of our stochastic MTJ signal and to clarify the use for potential cryptographic applications as a true-random-number generator, the National Institute of Standards and Technology Statistical Test Suite (NIST STS) [25] was applied to the binarized MTJ signal. The amplified analog output signal was acquired by an oscilloscope and was converted into a Bernoulli sequence of 0s and 1s by comparing the output signal with the median of the acquired time series. This ensures an equal probability of 0 and 1 of this Bernoulli sequence and is a prerequisite for a random bit stream. The random bit stream can then be interpreted as the result of flipping an unbiased coin (head 0, tail 1) with a probability of 50% yielding to a head or a tail, where for true randomness the flips are independent of each other, such that any past result of a coin flip does not affect any future results. The NIST STS provides 15 different tests (188 tests including all subtests) and requires a minimum bit-stream length of the order of 10^6 . There is a considerable amount of work that has been published recently regarding the quality of random-number generation by SMTJs [14,16,22,34]. It has been shown that cryptographic quality randomness, by passing all NIST STS tests, is not reached by a single SMTJ bit

stream but, rather, by combining multiple bit streams with XOR gates (also called “XOR whitening”). This is a common method to improve the quality of randomness and is easy to implement by using standard CMOS-based XOR gates. To further improve the quality of the randomness, nested XOR operations can be applied to multiple bit streams. In an XOR² operation, two rounds of XOR operations are applied to four input bit streams [see blue rectangle in Fig. 5(e)], whereas in an XOR³ operation three rounds of XOR operations are applied to eight input bit streams. With this method, the resulting bit stream of an XOR² operation passes all NIST STS tests, thus providing true randomness, which can be used for cryptographic applications. With an average autocorrelation time of 5 ns and four SMTJs as an input for an XOR² gate, true random bit generation of 200 Mbit/s can be achieved in a very energy-efficient way. Without the consideration of any additional CMOS circuitry, the energy dissipation of a single SMTJ can be of the order of nJ/bit or fJ/bit [34], depending on the applied voltage and resistance of the tunnel-junction device. For the evaluation, 5×10^8 data points were recorded, transformed into a binarized sequence, and divided into eight bit streams of equal size, in order to be able to feed the inputs to the XOR³ gate. Other works achieve cryptographic randomness by passing all NIST STS tests for a bit stream after a XOR² operation [12] or XOR³ operation [34]. Here, we also observe true randomness after an XOR² operation for a sampling time of 5 ns, as shown in Table I. This

TABLE I. The results of the p values for NIST STS tests after applying different XOR operations to the binarized SMTJ telegraph signal. A randomness test is passed and highlighted in green for p values above the significance level of 0.01. For multiple subtests, the particular p values were combined and tested according to Fisher’s method [33]. The sampling times of 3, 5, and 10 ns are chosen to be close to τ_{ac} .

Type Sampling time	Raw			XOR ¹			XOR ²		
	3 ns	5 ns	10 ns	3 ns	5 ns	10 ns	3 ns	5 ns	10 ns
(1) Frequency	0	0	0	0	0.141	0	0.373	0.739	0.904
(2) Block frequency ($m = 128$)	0	0	0	0	0	0	0	0.703	0.236
(3) Runs	0	0	0	0	0	0	0	0.069	0.817
(4) Longest run	0	0	0	0	0.019	0.001	0.001	0.990	0.185
(5) Binary matrix rank	0.335	0.028	0.897	0.507	0.251	0.072	0.517	0.692	0.632
(6) Discrete Fourier transform (spectral)	0	0.871	0	0	1	0.026	0.765	0.448	0.508
(7) Nonoverlapping template matching	0	0	0	0	0	0	0	0.039	0.313
(8) Overlapping template matching	0	0	0	0	0	0	0	0.623	0.944
(9) Maurer’s “universal statistical”	0	0	0	0	0.074	0	0.182	0.472	0.761
(10) Linear complexity ($M = 500$)	0.467	0.137	0.737	0.967	0.929	0.986	0.174	0.085	0.673
(11) Serial ($m = 16$)	0	0	0	0	0	0	0	0.687	0.900
(12) Approximate entropy ($m = 10$)	0	0	0	0	0	0	0	0.118	0.151
(13) Cumulative sum	0	0	0	0	0	0	0.162	0.172	0.010
(14) Random excursion	0	0.027	0.068	0.068	0.073	0.042	0.073	0.074	0.074
(15) Random-excursion variant	0	0	0	0	0.015	0	0.014	0.015	0.028

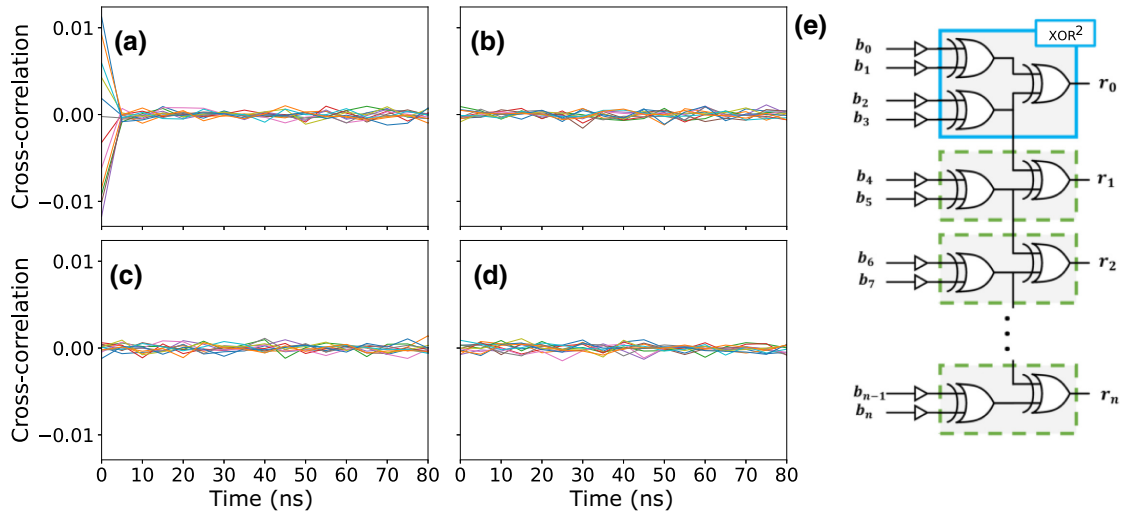


FIG. 5. (a)–(d) Cross-correlations of $r_0 = \text{XOR}^2(b_0, b_1, b_2, b_3)$ with (a) $\text{XOR}^2(b_0, b_1, b_2, b_*)$, (b) $\text{XOR}^2(b_0, b_1, b_*, b_*)$, (c) $\text{XOR}^2(b_0, b_*, b_*, b_*)$, and (d) $\text{XOR}^2(b_*, b_*, b_*, b_*)$ for 20 different raw SMTJ bit streams b_* . (e) Raw bit streams from SMTJs b_i are combined with XOR gates to generate random bit streams r_i . This circuit reduces the required number of SMTJs and XOR gates for generating multiple true random numbers.

sampling time matches the autocorrelation time measured for this time series. Larger sampling times provide the same quality of randomness, whereas for sampling times below 5 ns, each state would be sampled too frequently, thus resulting in nontrue randomness with artificial and nonrandom chunks of only 0 s and only 1 s.

Randomness is a probabilistic property; therefore, a random sequence can be evaluated in terms of the probabilities of the occurrence of a specific pattern in the sequence by means of hypothesis tests. The resulting p values give the probabilities of obtaining the observed pattern assuming that the underlying sequence is random. A common level of significance in cryptography for statistical hypothesis tests is $\alpha = 0.01$, which was also chosen in this work. For each of the 15 different NIST STS tests, the p value was calculated after applying different XOR operations to the binarized SMTJ telegraph signal. All results are summarized in Table I. The raw time-series data in our case never pass all of the NIST STS tests, independent of the chosen sampling rate. If a p value is greater than 0.01 (significance level) for a particular test, then the bit stream is characterized as random and passes this test, whereas for a p value below 0.01, the null hypothesis (of true randomness) is rejected and fails this test.

XOR operations can easily be implemented in CMOS and in our case, a simple random-number generator would require three XOR gates and four SMTJs. However, for further random-number generators, less hardware is necessary, since it is possible to tap the adjacent source randomness and XOR it with a new bit stream coming from the next two SMTJs. This is depicted schematically in Fig. 5(e), where raw bit streams from SMTJs b_i are combined with

XOR gates to generate random bit streams r_i . If r_0 , which is generated by $b_0 - b_3$, is truly random, then $r_1 - r_n$ will provide true random numbers too, due to the same XOR^2 operation. Nonetheless, an additional true random number only requires two more SMTJs and two more XOR gates (for r_1 , b_4 and b_5 are added). A further condition requires the true random numbers $r_0 - r_n$ to be uncorrelated. Therefore, we determined the cross-correlation of any new random number r_* with r_0 , as shown in Figs. 5(a)–5(d). Cross-correlations are calculated for 20 different bit streams ($b_0 - b_{19}$) of length 5×10^6 . Figure 5(a) points out a significant cross-correlation at 0 time lag for r_0 and $r_* = \text{XOR}^2(b_0, b_1, b_2, b_*)$, where b_* is the added new SMTJ bit stream. When two or more new SMTJs are combined, there is no further cross-correlation observed, meaning that the output will be uncorrelated with r_0 . The sequence of SMTJ bit streams to the XOR^2 circuit does not affect the outcome, since the XOR operation is permutation invariant. As long as two new bit streams (b_* and b_*) are combined with two already used bit streams (b_2 and b_3), it will result in a true random and uncorrelated number. It is also confirmed by the comparison with the cross-correlation in Fig. 5(d) with $r_* = \text{XOR}^2(b_*, b_*, b_*, b_*)$, where no qualitative difference is recognizable. Here, all SMTJs are exchanged; thus r_* has to be uncorrelated with r_0 . In addition, we wish to emphasize that the permutation-invariance property of the XOR operation renders possible cross-correlations of the form (r_a, r_b) , with $a \neq b$, to become uncorrelated. Our proposed circuit also offers a significant advantage in its ability to utilize SMTJs that are not precisely set to a balanced 50% equilibrium state. Even if there is a biased probability or deviation from the equilibrium state

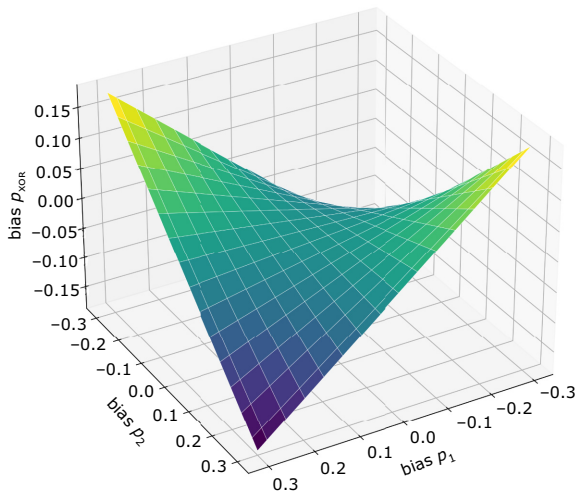


FIG. 6. The stochastic XOR function for two input bit streams with probabilities p_1 and p_2 , which are biased. The bias is the deviation from the “balanced” 0.5, or 50% equilibrium, state.

of SMTJs, a true-random-output bit stream can still be generated as long as the bias (deviation from 50%) is below roughly 10%. This means that the probabilities for each SMTJ can be in the range of 40–60%, which can accommodate fabrication inhomogeneities in the devices and simplify the fabrication process. The balancing effect of the XOR logic is responsible for this property. In stochastic computing, an XOR function performs the following computation on two bit streams with probabilities p_1 and p_2 [39]: $p_{\text{XOR}} = -2p_1p_2 + p_1 + p_2$. When the input probabilities are biased, the XOR logic tends to balance the output toward 50%. This balancing property is demonstrated in Fig. 6 for two biased input probabilities. In the case of our XOR² operation logic, this effect is even greater, allowing for large biases of the SMTJs. Additionally, it should be noted that if biases are antisymmetric, they can even be greater than 10%, since they partially cancel each other out. This feature further demonstrates the versatility and robustness of our proposed circuit. Next, we evaluate how an applied MTJ current controls the bit-stream generation rate and modifies the state probability. Both features are important for the performance of a true-random-number generator.

V. JOULE HEATING AND STT EFFECTS

The application of a current to an SMTJ can affect its switching speed, state probability, and dwell times not only by STT but also by Joule heating at the site of the free layer at the tunneling barrier. Joule heating has most often been studied in nonvolatile MTJs, where a reduction of the critical switching current is desired in order to perform fast and energy-efficient free-layer switching [40,41]. The current through an MTJ for an applied voltage is mainly determined by the RA product and the areal size of the

junction. The RA product, which is strongly dependent on the thickness of the tunneling barrier, defines the current density of any nanopillar size because of the relation $J = V/RA$. For an RA product of $15\Omega \mu\text{m}^2$ and a few hundred millivolts, the typical current density is of the order of 1 MA/cm^2 , where Joule heating can be significant. A simulation for a 50-nm-diameter MTJ suggests that the barrier temperature could increase by more than 100 K after applying 8 MA/cm^2 for a few hundred nanoseconds [41] and according to the Néel-Arrhenius law this will significantly decrease the dwell time of an SMTJ. Here, we have measured dwell times for different current densities and determined the contributions of STT and Joule heating.

Figures 7(e) and 7(f) illustrate the effect of STT and Joule heating. Initially, without STT, both states (P and AP) are tuned to the same energy level, where the free-layer macrospin is considered as a quasiparticle in a one-dimensional potential well. The ambient thermal energy (approximately 25 meV) raises the energy of the quasiparticle, such that the likelihood of overcoming the energy barrier is enhanced. By applying a current to the MTJ, the energy landscape is modified as illustrated in Fig. 7(f), where the AP-state energy level is decreased and the P-state energy level is enhanced by STT in our circuit. In the quasiparticle picture, this results in an energy shift ΔE_{STT} in the energy landscape and is defined here as a positive energy shift at the P state [see Fig. 7(f)]. Higher current densities result in a higher energy level for the P state yielding shorter dwell times, as can be seen for τ_P in Fig. 7(a). The energy shift ΔE_{STT} is positive for lower current densities, which means that the AP energy state is decreased, whereas the P energy state is increased. Furthermore, Joule heating, originating from electron-phonon interaction, raises the energy level of the quasiparticle in the valleys, symmetrically leading to overall faster switching. The temperature increase with respect to room temperature, ΔT_{JH} , reveals a quadratic relation to J [see Fig. 7(c)], as expected for the heating power, which scales with J^2 . Figure 7(b) reveals a similar trend of the average fluctuation rate, which is caused by the Joule-heating effect. This would correspond to a higher bit-generation rate for a random-number generator. The maximum temperature increase at 5.6 MA/cm^2 reaches $36 \pm 5 \text{ K}$ and corresponds to an applied voltage at the tunnel junction of approximately 0.6 V. In order to validate this temperature increase caused by Joule heating, a temperature-dependent dwell-time measurement would provide valuable insights and would allow for a detailed comparison. However, this measurement is part of a future project currently in preparation. The time series for the lowest and highest applied current densities are plotted in Fig. 8. Faster fluctuations with larger amplitude can be observed at high current densities, whereas for low current densities the fluctuation rate and amplitude are lower. By this measurement, we demonstrate that Joule heating has a significant effect on the dwell

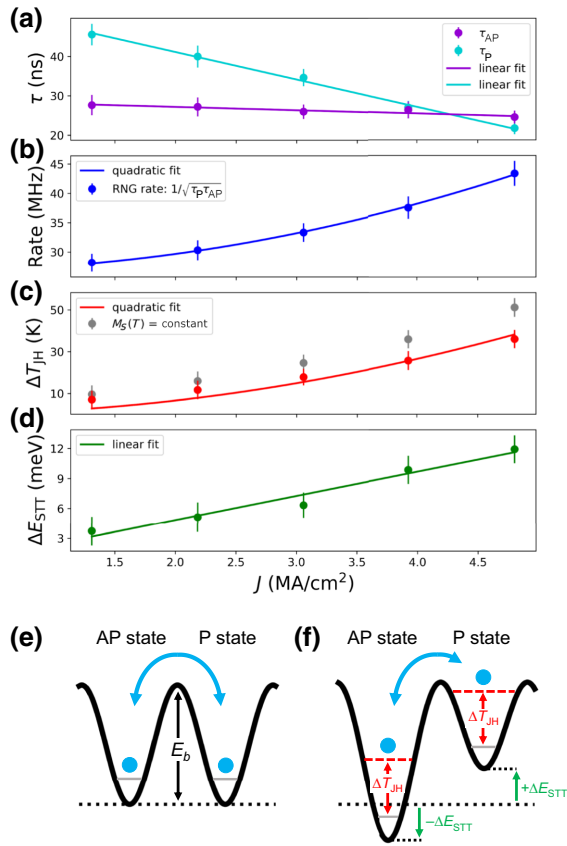


FIG. 7. (a) SMTJ dwell times for current densities ranging from 1.5 to 5.6 MA/cm². Linear fits are applied to indicate the trends. (b) The average fluctuation rate (blue), fitted with a quadratic function. (c) The temperature increase due to Joule heating (red). The nonlinear increase is fitted with a quadratic function. The gray data points indicate the calculated temperature increase without the consideration of temperature-dependent M_s . (d) The energy shift due to the STT (green). Here, a positive energy shift destabilizes the P state. (e) A schematic of the free-layer magnetization orientation described as a quasiparticle in a one-dimensional symmetrical potential well, with two metastable states separated by an energy barrier. (f) The effect of STT antisymmetrically shifts the energy level of the P and AP states and the effect of Joule heating raises the energy level of the quasiparticle symmetrically for both states.

times of an SMTJ and cannot be neglected. For a potential application as a random-number generator or as a p-bit for a probabilistic computer, the random-number-generation rate can be tuned by the current density. When considering an SMTJ as a spiking neuron for neuromorphic computing [42], the current through the junction will modify the spike rate, thus enabling, e.g., rate coding in a neuromorphic network.

A. Calculation of ΔT_{JH} and ΔE_{STT}

The energy shift of the states ΔE_{STT} affects the dwell times according to

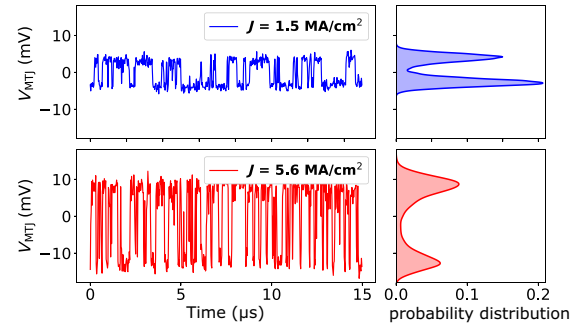


FIG. 8. Time-series measurements of the amplified fluctuating MTJ signal at low (1.5 MA/cm²) and high (5.6 MA/cm²) current density together with the respective probability distribution of each signal.

$$\tau_{P,AP} = \tau_0 \exp\left(\frac{E_b \mp \Delta E_{STT}}{k_B T}\right), \quad (2)$$

where $\tau_{P,AP}$ are the dwell times, T is the temperature, τ_0 is the attempt time, and E_b is the energy barrier between the metastable states. The energy barrier E_b for given dwell times can be determined by

$$E_b = 1/2 k_B T \log(\tau_P \tau_{AP} / \tau_0^2), \quad (3)$$

since the product of τ_P and τ_{AP} is independent of the energy-barrier modification by STT. For this reason, the product of the dwell times for an enhanced current density is also independent of STT:

$$\tau_P^* \tau_{AP}^* = \tau_0^2 \exp\left(\frac{2E_b^*}{k_B T^*}\right). \quad (4)$$

Here, T^* and E_b^* are considered as the altered (in our case, increased) temperature and energy barrier with dwell times τ_P^* and τ_{AP}^* . The increased temperature can be described as $T^* = T + \Delta T_{JH}$ with $T = 293$ K. The energy barrier at elevated temperature, $E_b^* = E_b(T^*) = E_b(M_s(T^*))$, is a function of the temperature-dependent saturation magnetization M_s of the free layer made of Co₄₀Fe₄₀B₂₀, which approximately follows Bloch's $T^{3/2}$ law: $M_s(T) = M_s(0)(1 - (T/T_c)^{3/2})$. For the Curie temperature T_c of Co₄₀Fe₄₀B₂₀, we have chosen 895 K [43]. The in-plane crystalline anisotropy is considered to be constant since a temperature increase of a few kelvin is unlikely to affect the polycrystalline phase of the free layer. The resulting temperature-dependent energy barrier E_b^* is then described as

$$E_b^* = E_b \frac{1 - (T^*/T_c)^{3/2}}{1 - (T/T_c)^{3/2}}. \quad (5)$$

The detailed equation between the dwell times and the temperature for the determination of ΔT_{JH} can be found in

Appendix A 2 [Eq. (A1)]. The asymmetric contribution to the energy barrier, due to STT, is calculated by $\Delta E_{\text{STT}} = 1/2k_B(T + \Delta T_{\text{JH}}) \log(\tau_{\text{AP}}/\tau_{\text{P}})$. However, one has to take into account that the initial P and AP states at low current densities are not in equilibrium, because $\tau_{\text{P}} \neq \tau_{\text{AP}}$ [see Fig. 7(a)]. Shorter AP dwell times arise due to an uncompensated stray field from the TMR stack itself and/or from an external applied in-plane field. With linear fits to the dwell times, we can estimate the theoretical dwell-time values at zero applied current, $\tau_{\text{P}}(J = 0)$ and $\tau_{\text{AP}}(J = 0)$, which are not measurable. However, at $J = 0$ there is neither a temperature increase ($\Delta T_{\text{JH}} = 0$) nor an STT effect ($\Delta E_{\text{STT}} = 0$) and the energy barrier at 293 K can be determined by Eq. (3) to be 95.3 ± 1.1 meV. Since the estimated dwell times at $J = 0$ are not equal and the energy landscape is asymmetric at this point, this shows an initial energy shift of the states due to the Zeeman energy, which is here approximately 11 meV. From Eq. (4), the temperature increase can be calculated numerically [see also Eq. (A1) in Appendix A 2].

VI. CONCLUSIONS

In this work, nanosecond superparamagnetic switching in in-plane MTJs with average dwell times of 6.7 ± 0.1 ns at room temperature is demonstrated and the potential application as a true random bit-stream generator is evaluated by the NIST STS. By employing an XOR² operation on four independent SMTJ bit streams, we demonstrate the successful generation of true random bits with a significantly high throughput rate of 200 Mbit/s. The scalability of our true-random-number generator is demonstrated by a circuit design allowing for multiple uncorrelated true random numbers with fewer stochastic MTJs and a smaller CMOS-area footprint. Further analysis reveals that even SMTJs with biased state probabilities can yield unbiased bit streams due to the balancing effect induced by the XOR operation, effectively passing all NIST tests. The influence of an applied current on the bias or state probability and switching rate of an SMTJ is investigated in detail by the analysis of dwell-time measurements. In particular, we show that the dwell times are linearly dependent on the applied current density and that Joule heating and STT both have a significant effect on the dwell times as well as the dwell-time ratio $\tau_{\text{P}}/\tau_{\text{AP}}$. STT shifts the energy levels of the MTJ states antisymmetrically and has a linear relation to the current density. Joule heating raises both energy levels equally and has a quadratic dependence on the current density. At the free layer, it can reach up to 329 ± 5 K for 5.6 MA/cm² at ambient room temperature. Since the dwell time is strongly dependent on the temperature, the shortest dwell times are observed for the highest current densities. This effect is not negligible and therefore should always

be considered in superparamagnetic MTJs with a low RA product.

The data that support the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGMENTS

This work was supported by the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC) and used infrastructure provided by For-Lab MagSens. We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 268565370 (SFB TRR173 Projects A01 and B02) as well as TopDyn and the Zeiss Foundation through the Center for Emergent Algorithmic Intelligence and the Horizon Europe Project No. 101070290 (NIM-FEIA). We would also like to thank T. Reimer for technical support during the development of the samples.

APPENDIX: FURTHER INFORMATION

1. SMTJ voltage signal

For the calculation of the voltage drop at the amplifier input, the circuit can be considered as a voltage divider circuit with shunt resistance R_s , MTJ resistance R_{MTJ} , in either low (parallel) R_{P} or high (antiparallel) resistance state, and input impedance Z_i . The measured voltage drop V can then be determined by:

$$V = V_{\text{src}} \frac{R_{\text{MTJ}} \parallel Z_i}{R_s + R_{\text{MTJ}} \parallel Z_i} = \frac{\frac{R_{\text{MTJ}} Z_i}{R_{\text{MTJ}} + Z_i}}{R_s + \frac{R_{\text{MTJ}} Z_i}{R_{\text{MTJ}} + Z_i}},$$

where V_{src} is the applied dc source voltage. The peak-to-peak voltage signal V_{pp} , which results from the MTJ resistance change, can be calculated by:

$$\begin{aligned} V_{\text{pp}} &= V_{\text{AP}} - V_{\text{P}}, \\ &= V_{\text{src}} \left[\frac{\frac{R_{\text{AP}} Z_i}{R_{\text{AP}} + Z_i}}{R_s + \frac{R_{\text{AP}} Z_i}{R_{\text{AP}} + Z_i}} - \frac{\frac{R_{\text{P}} Z_i}{R_{\text{P}} + Z_i}}{R_s + \frac{R_{\text{P}} Z_i}{R_{\text{P}} + Z_i}} \right]. \end{aligned}$$

For an MTJ resistance of 5–10 k Ω , $R_s = 50$ Ω and $V_{\text{src}} = 500$ mV, the fluctuating MTJ voltage signal (V_{pp}) is in the order of a millivolt.

2. Derivation of the effect of Joule heating

The product of dwell times at temperature T is given by $\tau_{\text{P}} \tau_{\text{AP}} = \tau_0^2 \exp(2E_b/k_B T)$. For an elevated temperature $T^* = T + \Delta T_{\text{JH}}$, the energy barrier is modified according to Bloch's $T^{3/2}$ law [see Eq. (5)]. The following equations

represent the dependence of the Joule heating ΔT_{JH} on the dwell times and can be solved numerically for ΔT_{JH} :

$$\tau_P^* \tau_{AP}^* = \tau_0^2 \exp\left(\frac{2E_b^*}{k_B T^*}\right), \log\left(\frac{\tau_P^* \tau_{AP}^*}{\tau_0^2}\right) = \frac{2E_b}{k_B T^*} \times \frac{1 - (T^*/T_c)^{3/2}}{1 - (T/T_c)^{3/2}} = \frac{T \log(\tau_P \tau_{AP} / \tau_0^2) (1 - ((T + \Delta T_{JH})/T_c)^{3/2})}{(T + \Delta T_{JH})(1 - (T/T_c)^{3/2})},$$

$$\frac{\log(\tau_P^* \tau_{AP}^* / \tau_0^2)}{\log(\tau_P \tau_{AP} / \tau_0^2)} = \frac{T(1 - ((T + \Delta T_{JH})/T_c)^{3/2})}{(T + \Delta T_{JH})(1 - (T/T_c)^{3/2})}. \quad (\text{A1})$$

Here, τ_P^* and τ_{AP}^* refer to the dwell times at an increased current density, while τ_P and τ_{AP} consider (reference) dwell times at a current density close to 0, where no Joule-heating effect is present.

3. Fisher's method

Fisher's method [33] is often used in meta-analysis to combine the results, namely, the p values, of independent hypothesis tests (each having the same hypothesis) to give a new null hypothesis suggesting that all null hypotheses are true. For a rejection H_0 , at least one of the null hypotheses is rejected. Since the p value of a hypothesis test follows a uniform distribution on the interval $[0,1]$, it can be shown that $-\sum_{i=1}^n \log(p_i)$ follows a chi-squared distribution χ_{2n}^2 with $2n$ degrees of freedom. In this work, the p values of several subtests are combined and evaluated according to Fisher's method and the resulting p value is noted in Table I. This holds for tests 11, 13, 14, and 15.

-
- [1] D. Apalkov, B. Dieny, and J. M. Slaughter, Magnetoresistive random access memory, *Proc. IEEE* **104**, 1796 (2016).
- [2] V. Ostwal and J. Appenzeller, Spin-orbit torque-controlled magnetic tunnel junction with low thermal stability for tunable random number generation, *IEEE Magn. Lett.* **10**, 1 (2019).
- [3] J. Grollier, D. Querlioz, K. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Neuromorphic spintronics, *Nat. Electron.* **3**, 360 (2020).
- [4] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-Bits for Invertible Logic, *Phys. Rev. X* **7**, 031014 (2017).
- [5] J. Kaiser, W. A. Borders, K. Y. Camsari, S. Fukami, H. Ohno, and S. Datta, Hardware-Aware *In Situ* Learning Based on Stochastic Magnetic Tunnel Junctions, *Phys. Rev. Appl.* **17**, 014016 (2022).
- [6] S. Ganguly, K. Y. Camsari, and A. W. Ghosh, Reservoir computing using stochastic p-bits (2017), *ArXiv:1709.10211*.
- [7] A. Sengupta, P. Panda, P. Wijesinghe, Y. Kim, and K. Roy, Magnetic tunnel junction mimics stochastic cortical spiking neurons, *Sci. Rep.* **6**, 1 (2016).
- [8] A. Sengupta, M. Parsa, B. Han, and K. Roy, Probabilistic deep spiking neural systems enabled by magnetic tunnel junction, *IEEE Trans. Electron Devices* **63**, 2963 (2016).
- [9] M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, Energy-Efficient Stochastic Computing with Superparamagnetic Tunnel Junctions, *Phys. Rev. Appl.* **13**, 034016 (2020).
- [10] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (CRC Press, Chapman & Hall/CRC, London, 1995).
- [11] K. Y. Camsari, B. M. Sutton, and S. Datta, p-bits for probabilistic spin logic, *Appl. Phys. Rev.* **6**, 011305 (2019).
- [12] T. Kim, H. Park, K.-H. Han, Y.-J. Nah, H. C. Koo, B.-C. Min, S. Hong, and O. Lee, Demonstration of in-plane magnetized stochastic magnetic tunnel junction for binary stochastic neuron, *AIP Adv.* **12**, 075104 (2022).
- [13] B. R. Zink, Y. Lv, and J.-P. Wang, Telegraphic switching signals by magnet tunnel junctions for neural spiking signals with high information capacity, *J. Appl. Phys.* **124**, 152121 (2018).
- [14] M. Bapna and S. A. Majetich, Current control of time-averaged magnetization in superparamagnetic tunnel junctions, *Appl. Phys. Lett.* **111**, 243107 (2017).
- [15] B. Parks, A. Abdelgawad, T. Wong, R. F. Evans, and S. A. Majetich, Magnetoresistance Dynamics in Superparamagnetic Co-Fe- B Nanodots, *Phys. Rev. Appl.* **13**, 014063 (2020).
- [16] C. Safranski, J. Kaiser, P. Trouilloud, P. Hashemi, G. Hu, and J. Z. Sun, Demonstration of nanosecond operation in stochastic magnetic tunnel junctions, *Nano Lett.* **21**, 2040 (2021).
- [17] K. Hayakawa, S. Kanai, T. Funatsu, J. Igarashi, B. Jinnai, W. Borders, H. Ohno, and S. Fukami, Nanosecond Random Telegraph Noise in In-Plane Magnetic Tunnel Junctions, *Phys. Rev. Lett.* **126**, 117202 (2021).
- [18] S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, Theory of relaxation time of stochastic nanomagnets, *Phys. Rev. B* **103**, 094423 (2021).
- [19] G. Reiss, J. Ludwig, and K. Rott, Superparamagnetic dwell times and tuning of switching rates in perpendicular CoFeB/MgO/CoFeB tunnel junctions (2019), *ArXiv:1908.02139*.
- [20] M. Bapna, S. K. Piotrowski, S. D. Oberdick, M. Li, C.-L. Chien, and S. A. Majetich, Magnetostatic effects on switching in small magnetic tunnel junctions, *Appl. Phys. Lett.* **108**, 022406 (2016).

- [21] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, Integer factorization using stochastic magnetic tunnel junctions, *Nature* **573**, 390 (2019).
- [22] B. Parks, M. Bapna, J. Igbokwe, H. Almasi, W. Wang, and S. A. Majetich, Superparamagnetic perpendicular magnetic tunnel junctions for true random number generators, *AIP Adv.* **8**, 055903 (2018).
- [23] K. Kobayashi, W. A. Borders, S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, Sigmoidal curves of stochastic magnetic tunnel junctions with perpendicular easy axis, *Appl. Phys. Lett.* **119**, 132406 (2021).
- [24] S. Krause, G. Herzog, A. Schlenhoff, A. Sonntag, and R. Wiesendanger, Joule Heating and Spin-Transfer Torque Investigated on the Atomic Scale Using a Spin-Polarized Scanning Tunneling Microscope, *Phys. Rev. Lett.* **107**, 186601 (2011).
- [25] L. E. Bassham III, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, S. D. Leigh, M. Levenson, M. Vangel, and D. L. Banks, *et al.*, *Sp 800-22 rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications* (National Institute of Standards and Technology, Gaithersburg, MD, 2010).
- [26] L. Schnitzspan, J. Cramer, J. Kubik, M. Tarequzaman, G. Jakob, and M. Kläui, Impact of annealing temperature on tunneling magnetoresistance multilayer stacks, *IEEE Magn. Lett.* **11**, 1 (2020).
- [27] R. Stamps, Mechanisms for exchange bias, *J. Phys. D: Appl. Phys.* **33**, R247 (2000).
- [28] S. Jenkins, A. Meo, L. E. Elliott, S. K. Piotrowski, M. Bapna, R. W. Chantrell, S. A. Majetich, and R. F. Evans, Magnetic stray fields in nanoscale magnetic tunnel junctions, *J. Phys. D: Appl. Phys.* **53**, 044001 (2019).
- [29] D. Worledge and P. Trouilloud, Magnetoresistance measurement of unpatterned magnetic tunnel junction wafers by current-in-plane tunneling, *Appl. Phys. Lett.* **83**, 84 (2003).
- [30] P. Hänggi, P. Talkner, and M. Borkovec, Reaction-rate theory: Fifty years after Kramers, *Rev. Mod. Phys.* **62**, 251 (1990).
- [31] W. F. Brown Jr, Thermal fluctuations of a single-domain particle, *Phys. Rev.* **130**, 1677 (1963).
- [32] O. Hassan, R. Faria, K. Y. Camsari, J. Z. Sun, and S. Datta, Low-barrier magnet design for efficient hardware binary stochastic neurons, *IEEE Magn. Lett.* **10**, 1 (2019).
- [33] F. Mosteller and R. A. Fisher, Questions and answers, *Am. Stat.* **2**, 30 (1948).
- [34] D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. S. Friedman, A. F. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, and S. Yuasa, *et al.*, Low-Energy Truly Random Number Generation with Superparamagnetic Tunnel Junctions for Unconventional Computing, *Phys. Rev. Appl.* **8**, 054045 (2017).
- [35] A. Fukushima, T. Seki, K. Yakushiji, H. Kubota, H. Imaura, S. Yuasa, and K. Ando, Spin dice: A scalable truly random number generator based on spintronics, *Appl. Phys. Express* **7**, 083001 (2014).
- [36] X. Chen, J. Zhang, and J. Xiao, *et al.*, Magnetic-Tunnel-Junction-Based True Random-Number Generator with Enhanced Generation Rate, *Phys. Rev. Appl.* **18**, L021002 (2022).
- [37] N. Rangarajan, A. Parthasarathy, and S. Rakheja, A spin-based true random number generator exploiting the stochastic precessional switching of nanomagnets, *J. Appl. Phys.* **121**, 223905 (2017).
- [38] Z. Fu, Y. Tang, X. Zhao, K. Lu, Y. Dong, A. Shukla, Z. Zhu, and Y. Yang, An overview of spintronic true random number generator, *Front. Phys.* **9** (2021).
- [39] A. Alaghi and J. P. Hayes, in *Proceedings of the 25th edition on Great Lakes Symposium on VLSI* (2015), p. 331.
- [40] D. Lee and S. H. Lim, Increase of temperature due to joule heating during current-induced magnetization switching of an MgO-based magnetic tunnel junction, *Appl. Phys. Lett.* **92**, 233502 (2008).
- [41] J. J. Kan, C. Park, C. Ching, J. Ahn, Y. Xie, M. Pakala, and S. H. Kang, A study on practically unlimited endurance of STT-MRAM, *IEEE Trans. Electron Devices* **64**, 3639 (2017).
- [42] A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S. Yuasa, J. Grollier, and D. Querlioz, Neural-like computing with populations of superparamagnetic basis functions, *Nat. Commun.* **9**, 1 (2018).
- [43] K.-M. Lee, J. W. Choi, J. Sok, and B.-C. Min, Temperature dependence of the interfacial magnetic anisotropy in W/CoFeB/MgO, *AIP Adv.* **7**, 065107 (2017).