Arbitrary Modulation of Average Dwell Time in Discrete-Time Markov Chains based on Tunneling Magnetoresistance Effect

Xihui Yuan, Jiajia Jian, Zheng Chai, Hengjia Wei, Xue Zhou, Yizhu Wen, Yadong Liu, Wentao Yan, Yingtong He, Jian Fu Zhang, Weidong Zhang and Tai Min

***Abstract*—Stochastic processes (SPs) are widely used in many real-world fields, especially AI algorithms and models. A discrete-time Markov chain (DTMC) is a fundamental SP where the probability of each event depends only on the state attained in the previous event. DTMC is extensively used in signal processing and information theory, but the hardware generation of DTMC remains hindered by the difficulty in arbitrarily modulating the averaged dwell times (ADTs), i.e. the average time that the DTMC stays at one state. In this paper, we propose a two-step procedure to modulate the ADTs of a DTMC generated from single magnetic tunnel junction (MTJ), without being limited by additional restrictions such as a fixed ratio between the ADTs, widening the SP applications scope of the MTJ-based DTMC. This method has been verified via mathematical derivation and electrical characterization. The generation throughput and power consumption can also be conveniently modulated. This procedure provides a new hardware solution for the generation of stochastic signal in semiconductor IC chips.**

***Index Terms*—Markov chain, magnetic tunnel junction, probabilistic switching, spintronic**

I. Introduction

Markov process is a stochastic process with a sequence of possible events, in which the probability of the next event depends only on the state attained in the present event, and is independent of the past events [1] [2]. It is different from the generic random bit stream generation where the random variables are totally independent. Markov processes with discrete time are called discrete-time Markov chains (DTMC), and are widely used in wireless communications [3] [4], Internet traffic modeling [5], information encryption [6] [7] and financial engineering [8], helping to provide stochasticity for the systems and meanwhile being mathematically trackable [9].

A DTMC is formed by segments of alternating consecutive appearing states. The lengths of the segments, often called dwell times (DT), are randomly but exponentially distributed [2]. The average dwell time (ADT) represents DTs' distribution and it is desirable that ADTs could be adjusted to any target sets, for practical applications such as the stabilization of stochastic systems in robust control theory [10] and fuzzy systems [11].

Most existing DTMCs are generated by software algorithms. However, for the Internet-of-Things (IoT) applications, it is desirable that DTMCs could be directly generated by hardware devices, for the performance, area, and power considerations. Recently, noise and variability of scaling/emerging devices [12] has been utilized for DTMC generation. Random telegraph noise induced by defect carrier trapping/de-trapping [13] [14] is a natural Markov chain, but its ADTs cannot be arbitrarily modulated for a determined defect. The variability in a resistive-switching memory (RRAM) has been used to generate DTMC, but with non-modulable transition matrix [12] [15], limited speed (~ms), and limited endurance (10^5).

Spintronic devices have been used to generate telegraphic switching signal [16] [17] similar to the random telegraph noise [13] [14] and also belongs to the Markov chain family [18], but it requires both a magnetic field and electrical pulses, the former of which is not easy to implement in on-chip circuitry. Our recent work proposed a DTMC generator with excellent endurance and uniformity [19], based on the probabilistic switching [20] of single magnetic tunnel junction (MTJ) [21] [22] and a specially designed waveform, with longer endurance, lower energy consumption and pure electrical operation.

In this work, we proposed a two-step procedure to control the ADTs of DTMC, using only one MTJ device and electrical operations. The arbitrary ADTs pairs of the two states can be electrically modulated. The correctness of this procedure has been mathematically derived and experimentally verified. High-quality DTMCs have been generated in a wide range. The method also shows promising potential in scaling the energy consumption and boosting the throughput, and provides a hardware solution for the practical hardware generation of DTMC in the IoT era.

Manuscript received xx. This work is supported by the National Key R&D Program of China (Grant No. 2022YFB4400200), the National Natural Science Foundation of China (No. 62104188 and 12327806), and in part by the major key project of Peng Cheng Laboratory under grant PCL2023AS1-2. The review of this letter was arranged by Editor B.G. Malm. (Xihui Yuan and Jiajia Jian contributed equally to this work.)

(Corresponding authors: Zheng Chai, Tai Min.)

Xihui Yuan, Jiajia Jian, Zheng Chai, Xue Zhou, Yadong Liu, Wentao Yan, Yingtong He and Tai Min are with the Center for Spintronic and Quantum Systems, State Key Laboratory for Mechanical Behavior of Materials, and School of Materials Science and Engineering, Xi'an Jiaotong University, Xi'an, China, and also with the Peng Cheng Laboratory, Shenzhen, Guangdong 518055, China, and also with the Pazhou Laboratory (Huangpu), Guangzhou, Guangdong 510555, China (e-mail: zheng.chai@xjtu.edu.cn, tai.min@mail.xjtu.edu.cn).

Hengjia Wei is with the Peng Cheng Laboratory, Shenzhen 518055, China. He is also with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China, and the Pazhou Laboratory (Huangpu), Guangzhou 510555, China.

Yizhu Wen is with the Spin-Quantum Co. LTD., Shaanxi, China.

Jian Fu Zhang and Weidong Zhang are with the School of Engineering, Liverpool John Moores University, L3 3AF Liverpool, U.K. (e-mail: w.zhang@ljmu.ac.uk).

This article has been accepted for publication in IEEE Electron Device Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LED.2024.3394711

II. Device and Experiment

The device is a bottom pinned perpendicular magnetization anisotropy (PMA) MTJ with a diameter of 78 nm. **Fig. 1a** demonstrates its magneto resistive switching between the low-resistance parallel (P) state, and the high-resistance anti-parallel (AP) state with the cross section scanning electron microscopy (SEM) image (**inset of Fig. 1a**). **Fig. 1b** shows the sigmoid-like dependence of switching probability on pulse amplitudes with fixed pulse width of 500 ns, which can be explained by the Néel-Brown equation [23]:

$$P(t) = 1 - exp(-t/\tau) \quad (1)$$

where *P(t)* is the thermal switching probability, and τ is the relaxation time:

$$\tau = f_0^{-1} exp(E_b/k_B T) \quad (2)$$

where f_0 is the attempt frequency, E_b is the energy barrier, and T is the temperature.

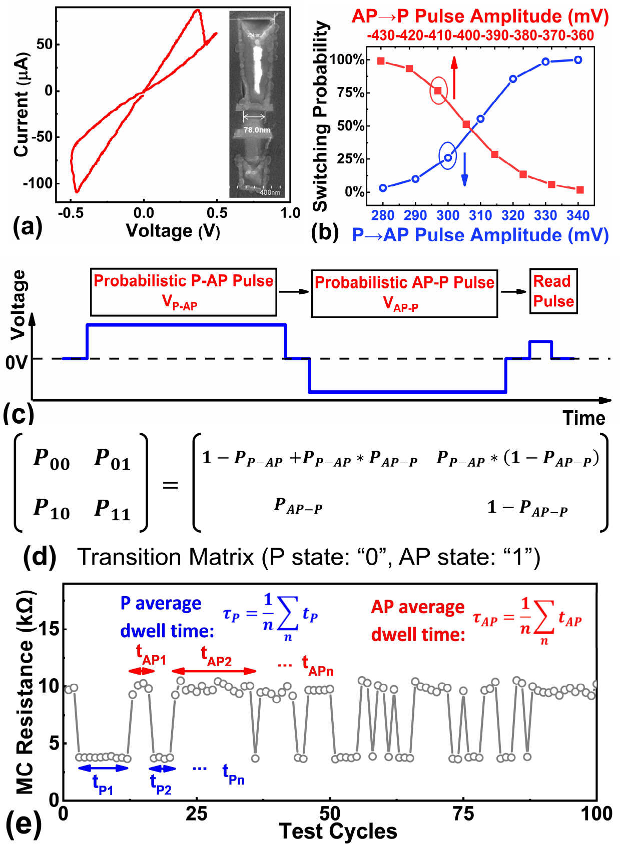

Fig.1. (a) The I-V curve of magneto-resistive switching. Inset: cross-section SEM image of the MTJ. (b) The sigmoid-like dependence of switching probability on pulse amplitudes with pulse width of 500 ns. (c) The three-pulse waveform. (d) The transition matrix. (e) Partial demonstration of the generated DTMC, with the dwell time (t_P and t_{AP}) and their average values (τ_P and τ_{AP}).

The DTMC is generated by consecutively applying the three-pulse waveform onto the MTJ (**Fig. 1c**). The details of DTMC generation setup and the transition matrix including transition probabilities P_{00}, P_{01}, P_{10}, P_{11} (**Fig. 1d**) can be found in [19]. Herein, the "time" in ADT refers to the number of consecutive identical bits, each of which is generated from one cycle. All electrical measurements were done using the pulse measurement units in a Keithley 4200 semiconductor analyzer.

III. Results and Discussions

In a DTMC, t_i denotes the dwell time, i.e. the time that the DTMC remains in one state i (in this work, i = 0 or 1, which corresponds to the state P or AP, respectively) until the next transition, and n_i denotes the number of segments in state i, while τ_i denotes the averaged value of t_i. **Fig. 1e** partially demonstrates the generated DTMC, together with the t_i and τ_i. Mathematically, the dependence of τ_i on the transition probabilities can be derived in the following:

In the DTMC, t_i is a random variable with probability distribution

$$\Pr[t_i = m] = P_{ii}^{m-1}(1 - P_{ii}). \quad (3)$$

Then its expectation can be calculated as follows:

$$E[t_i] = \sum_{m=1}^{\infty} m\, P_{ii}^{m-1}(1 - P_{ii}) = \frac{1 - P_{ii}}{P_{ii}} \sum_{m=1}^{\infty} m\, P_{ii}^{m} = \frac{1}{1 - P_{ii}}. \quad (4)$$

In a DTMC of infinite length, it follows from the law of large numbers that τ_i converges to $E[t_i]$ (which is equal to $\frac{1}{1-P_{ii}}$ when the DTMC length is infinite). According to the transition probabilities definition shown in **Fig. 1d**, the ADT of AP and P states, i.e. τ_{AP} and τ_P, converge to $\frac{1}{P_{AP-P}}$ and $\frac{1}{P_{P-AP}\cdot(1-P_{AP-P})}$, respectively.

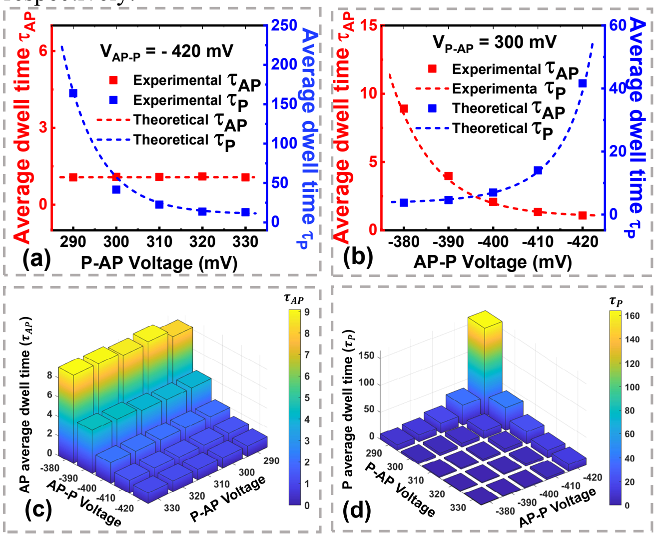

Fig.2. The dependence of τ_{AP} and τ_P on (a) V_{P-AP} and (b) V_{AP-P} is similar to the theoretical counterpart, with fixed V_{AP-P} and V_{P-AP}, respectively. The (c) τ_{AP} and (d) τ_P across a range of V_{P-AP} and V_{AP-P}.

Note that the P-AP (or AP-P) means that the device is intended to be switched from P (or AP) state to AP (or P) state with a specific probability. As show above, τ_{AP} is only determined by P_{AP-P}, while τ_P is determined by both P_{AP-P} and P_{P-AP}. Such dependency has been experimentally verified in **Fig. 2**: with a fixed V_{AP-P} of -420mV, when V_{P-AP} increases, τ_P gradually decreases, but τ_{AP} remains constant (**Fig. 2a**); if V_{P-AP} is fixed at 300mV and V_{AP-P} increases, τ_P increases but τ_{AP} decreases (**Fig. 2b**), both in agreement with the trend of theoretical values after calibration. Such dependency is further supported by the distribution of τ_P and τ_{AP} with V_{AP-P} and V_{P-AP}

covering a wider range, as visualized in the 3D plots in **Fig. 2c** and **2d** based on experimental data collected from a range of $V_{P\text{-}AP}$ and $V_{AP\text{-}P}$. ADTs cover a wide range as the switching probability varies from 0 to 1. The different dependences of τ_P and τ_{AP} on the pulse conditions provides the inspiration that a target set of τ_{AP} and τ_P combination can be reached following a two-step procedure (**Fig. 3**): (i) Since τ_{AP} is not affected by $P_{P\text{-}AP}$, for a target τ_{AP}, a corresponding $P_{AP\text{-}P}$ can be obtained, which is given by a suitable $V_{AP\text{-}P}$ from the sigmoid-like dependence in **Fig. 1b**. (ii) Next, for the target τ_P, since $P_{AP\text{-}P}$ is already determined, a suitable $P_{P\text{-}AP}$ can be calculated given by the $V_{P\text{-}AP}$.

In this way, the two-step procedure facilitates the electrical modulation of any arbitrary ADT pairs of DTMC generated from a single MTJ entirely under electrical operation. Note that this methodology also works, if the polarities of the two probabilistic pulses in **Fig. 1c** are reversed: τ_P can be realized first, then τ_{AP}.

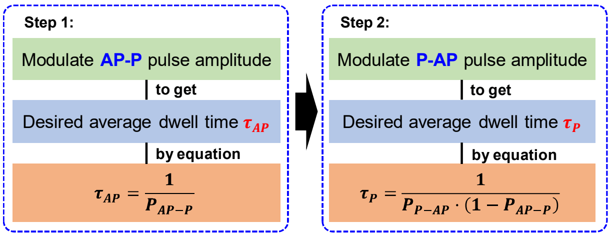

Fig.3. The two-step procedure to modulate ADT in a DTMC: (i) the target τ_{AP} is achieved by setting $V_{AP\text{-}P}$, regardless of $V_{P\text{-}AP}$. (ii) with fixed $V_{AP\text{-}P}$ and determined τ_{AP}, the target τ_P is achieved by setting $V_{P\text{-}AP}$.

Fig. 4 demonstrates the DTMCs with an example target set of ADTs (τ_{AP} = 2.5 and τ_P = 1.6), generated with four different pulse widths. As shown in **Fig. 4a**, the bits in a DTMC can be generated using probabilistic pulses with width scaling from 100 µs to 100 ns, limited by the instrument's minimum pulse width level and the measurement time [24], while maintaining the same set of ADTs stably. Note that the method is several orders faster than the existing MTJ-based superparamagnetic oscillation [25] [26] which are normally with millisecond time scale [27]. This indicates the potential energy efficiency scaling to lower than 10 pJ/bit, as the energy consumption scales linearly with pulse width in a log-log plot (**Fig. 4b**). Furthermore, the dwell time, t_P and t_{AP} of the generated DTMCs all follow the exponential distribution, as evidenced in the cumulative density function (CDF) in **Fig. 4c** and **4d**, supporting their Markovian property [1].

The randomness of the DTMC has been tested by a series of methods [19]: in **Fig. 5a**, the absence of any discernible pattern in the bitmap confirms the balance and randomness of the DTMC bits with pulse width of 100 ns. The Hamming distance (HD) and entropy of this DTMC are 0.49 (close to ideal value 0.5) and 0.96 (close to ideal value 1.0), respectively. The autocorrelation function (ACF) of the DTMC with various pulse width (**Fig. 5b**) shows the average 95% confidence boundary being only ~ 0.03, comparable to the established benchmarks [28].

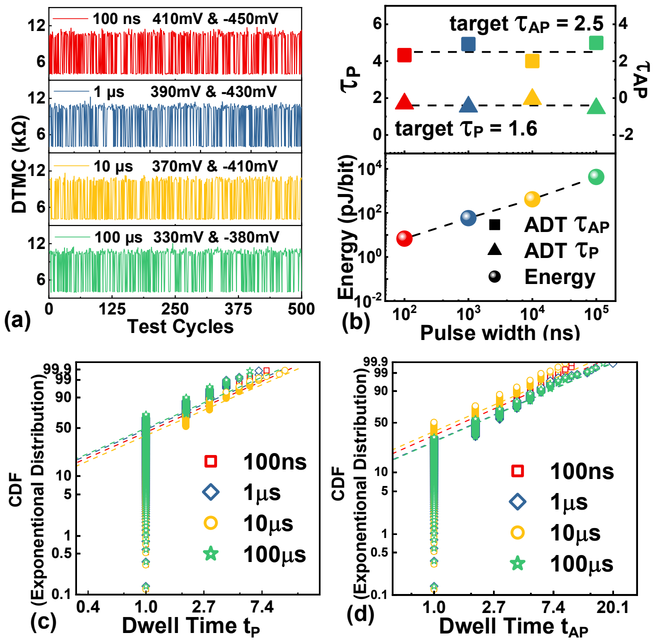

Fig.4. (a) The scalable generation speed with probabilistic pulse width reducing from 100 µs to 100 ns. (b) With pulse amplitude adjustment, the generated DTMCs show the same ADT, with energy per cycle scaled below 10 pJ. (c-d) The exponentially distributed CDF of t_P and t_{AP} in the DTMC, supporting their MC property.

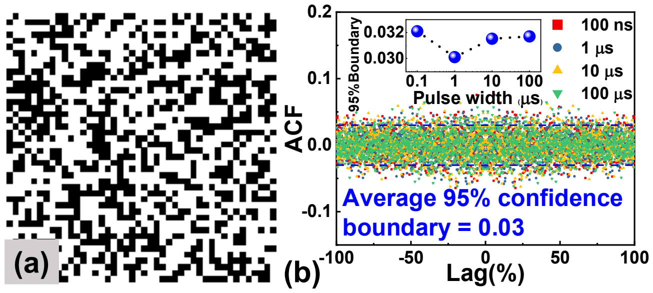

Fig.5. (a) The bitmap (ap: black; p: white) shows the superior randomness of the DTMC. (b) The ACFs of the obtained DTMC signals show average 95% confidence boundary as low as only ~ 0.03.

Recently, the concept of persistent dwell time (PDT) proves to be useful in investigating the Markov jump systems (MJSs) based on MCs. Compared with DT or ADT, PDT is more general because it can represent the intermittent occurrence of fast and slow switching existing in the practical and complex MJSs [11]. We will continue to work along this line and analyze the PDT in our DTMC which is a two-level switching signal generated from practical MTJ devices and could be affected by complex spintronic conditions.

IV. Conclusions

In this work, we proposed a two-step procedure to control the ADT of DTMC, using only one MTJ device with electrical operations. The arbitrary ADT pairs of the two states can be electrically modulated. The correctness of this procedure has been mathematically derived and experimentally verified. High-quality DTMCs have been generated with a wide range of ADTs. The proposed method also shows promising potential in scaling the energy consumption and boosting the throughput and provides a hardware solution for the practical hardware generation of DTMC in the IoT era.

REFERENCE

[1] W.-K. Ching, X. Huang, M. K. Ng and T.-K. Siu, Markov Chains: Models, Algorithms and Applications, Berlin: Springer Sci. Bus. Media, 2013, doi: 10.1007/0-387-29337-X.

[2] P. A Gagniuc, Markov Chains: From Theory to Implementation and Experimentation, Amsterdam: John Wiley & Sons, 2017, doi: 10.1002/9781119387596.

[3] Z. Hajjarian, M. Kavehrad, "Channel modeling in wireless optical communications using Markov chains," in *Canadian Conference on Electrical and Computer Engineering*, Niagara Falls, ON, Canada, 2008, pp. 002051-002056, doi: 10.1109/CCECE.2008.4564906.

[4] M. Ting, A. O. Hero, III, D. Rugar, C. -Y. Yip, and J. A. Fessler, "Near-Optimal Signal Detection for Finite-State Markov Signals With Application to Magnetic Resonance Force Microscopy," *IEEE Transactions on Signal Processing,* vol. 54, no. 6, pp. 2049-2062, 2006, doi: 10.1109/TSP.2006.874413.

[5] G. Aceto, G. Bovenzi, D. Ciuonzo, A. Montieri, V. Persico and A. Pescapé, "Characterization and Prediction of Mobile-App Traffic Using Markov Modeling," *IEEE Transactions on Network and Service Management,* vol. 18, no. 1, pp. 907-925, 2021, doi: 10.1109/TNSM.2021.3051381.

[6] V. R. Pamula, X. Sun, S. M. Kim, F. ur Rahman, B. Zhang and V. S. Sathe, "A 65-nm CMOS 3.2-to-86 Mb/s 2.58 pJ/bit Highly Digital True-Random-Number Generator With Integrated De-Correlation and Bias Correction," *IEEE Solid-State Circuits Letters,* vol. 1, no. 12, pp. 237-240, 2018, doi: 10.1109/LSSC.2019.2896777.

[7] V. R. Pamula, X. Sun, S. Kim, F. ur Rahman, B. Zhang and V. S. Sathe, "An All-Digital True-Random-Number Generator with Integrated De-correlation and Bias Correction at 3.2-to-86 MB/S, 2.58 PJ/Bit in 65-NM CMOS," in *IEEE Symposium on VLSI Circuits*, Honolulu, HI, USA, 2018, pp. 173-174, doi: 10.1109/VLSIC.2018.8502375.

[8] D.-M. Zhu and W.-K. Ching, "A New Estimation Method for Multivariate Markov Chain Model with Application in Demand Predictions," in *Third International Conference on Business Intelligence and Financial Engineering*, Hong Kong, China, 2010, pp. 126-130, doi: 10.1109/BIFE.2010.39.

[10] S. Battilotti, A. D. Santis, "Dwell-time controllers for stochastic systems with switching Markov chain," *Automatica,* vol. 41, no. 6, pp. 923-934, 2005, doi: 10.1016/j.automatica.2004.11.024.

[11] J. Wang, J. Xia, H. Shen, M. Xing and J. H. Park, "H_∞ Synchronization for Fuzzy Markov Jump Chaotic Systems With Piecewise-Constant Transition Probabilities Subject to PDT Switching Rule," IEEE Transaction on Fuzzy Systems, vol. 29, no. 10, pp. 3082-3092, 2021, doi: 10.1109/TFUZZ.2020.3012761.

[12] H. Tian, X.-F. Wang, M. A. Mohammad, G. -Y. Gou, F. Wu, Y. Yang and T.-L. Ren, "A hardware Markov chain algorithm realized in a single device for machine learning," *Nature Communications,* vol. 9, p. 4305, 2018, doi: 10.1038/s41467-018-06644-w.

[13] Z. Chai, J. Ma, W. Zhang, B. Govoreanu, E. Simoen, J. F. Zhang, Z. Ji, R. Gao, G. Groeseneken and M. Jurczak, "RTN-based defect tracking technique: Experimentally probing the spatial and energy profile of the critical filament region and its correlation with HfO2RRAM switching operation and failure mechanism," in *IEEE Symposium on VLSI Technology*, Honolulu, HI, USA, 2016, pp. 1-2, doi: 10.1109/VLSIT.2016.7573402.

[14] F. M. Puglisi, N. Zagni, L. Larcher and P. Pavan, "Random Telegraph Noise in Resistive Random Access Memories: Compact Modeling and Advanced Circuit Design," *IEEE Transactions on Electron Devices,* vol. 65, pp. 2964 - 2972, 2018, doi: 10.1109/TED.2018.2833208.

[15] G. Zoppo, A. Korkmaz, F. Marrone, S. Palermo, F. Corinto, R. S. Williams, "Analog Solutions of Discrete Markov Chains via Memristor Crossbars," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 68, no. 12, pp. 4910-4923, 2021, doi: 10.1109/TCSI.2021.3126477.

[16] B. R. Zink , Y. Lv, and J.-P. Wang, "Independent Control of Antiparallel- and Parallel-State Thermal Stability Factors in Magnetic Tunnel Junctions for Telegraphic Signals With Two Degrees of Tunability," *IEEE Transactions on Electron Devices,* vol. 12, no. 66, pp. 5353 - 5359, 2019, doi: 10.1109/TED.2019.2948218.

[17] B. R. Zink, Y. Lv and J.-P. Wang, "Telegraphic switching signals by magnet tunnel junctions for neural spiking signals with high information capacity," *Journal of Applied Physics,* vol. 15, no. 124, p. 152121, 2018, doi: 10.1063/1.5042444.

[18] S. Liu, "Characterizations and Stability Criteria for the 2-Mode Markovian Switched Nonlinear Systems," *IEEE Transactions on Automatic Control,* pp. 1-7, 2023, doi: 10.1109/TAC.2023.3310673 (Early Access).

[19] X. Yuan, J. Jian, Z. Chai, S. An, Y. Gao, X. Zhou, J. F. Zhang, W. Zhang and T. Min, "Markov Chain Signal Generation based on Single Magnetic Tunnel Junction," *IEEE Electron Device Letter*, vol. 44, no. 12, pp. 1963-1966, 2023, doi: 10.1109/LED.2023.3322194.

[20] H. Zhao, A. Lyle, Y. Zhang, P. K. Amiri, G. Rowlands, Z. Zeng, J. Katine, H. Jiang, K. Galatsis, K. L. Wang, I. N. Krivorotov and J.-P. Wang, "Low writing energy and sub nanosecond spin torque transfer switching of in-plane magnetic tunnel junction for spin torque transfer random access memory," *Journal of Applied Physics* , vol. 109, p. 07C720, 2011, doi: 10.1063/1.3556784.

[21] B. Dieny, I. L. Prejbeanu, K. Garello , P. Gambardella, P. Freitas, R. Lehndorff, W. Raberg, U. Ebels, S. O. Demokritov, J. Akerman, A. Deac, P. Pirro, C. Adelmann, A. Anane, A. V. Chumak, A. Hirohata, S. Mangin, S. O. Valenzuela, M. C. Onbaşlı, M. d'Aquino, G. Prenat, G. Finocchio, L. Lopez-Diaz, R. Chantrell, O. Chubykalo-Fesenko and P. Bortolotti, "Opportunities and challenges for spintronics in the microelectronics industry," *Nature Electronics,* vol. 3, p. 446–459, 2020, doi: 10.1038/s41928-020-0461-5.

[22] R. Beach, T. Min, C. Horng, Q. Chen, P. Sherman, S. Le, S. Young, K. Yang, H. Yu, X. Lu, W. Kula, T. Zhong, R. Xiao, A. Zhong, G. Liu, J. Kan, J. Yuan, J. Chen, R. Tong, J. Chien, T. Torng, D. Tang, P. Wang, M. Chen, S. Assefa, M. Qazi and J. DeBrosse, "A statistical study of magnetic tunnel junctions for high-density spin torque transfer-MRAM (STT-MRAM)," in *IEEE International Electron Devices Meeting*, San Francisco, CA, USA, 2008, pp. 1-4, doi: 10.1109/IEDM.2008.4796679.

[23] W. F. Brown, "Thermal Fluctuations of a Single-Domain Particle," *Phys. Rev.,* vol. 130, no. 5, pp. 1677-1686, 1963, doi: 10.1103/PhysRev.130.1677.

[24] C. Safranski, G. Hu, J. Z. Sun, P. Hashemi, S. L. Brown, L. Buzi, C. P. D'Emic, E. R. J. Edwards, E. Galligan, M. G. Gottwald, O. Gunawan, S. Karimeddiny, H. Jung, J. Kim, K. Latzko, P. L. Trouilloud, S. Zare and D. C. Worledge, "Reliable Sub-nanosecond MRAM with Double Spin-torque Magnetic Tunnel Junctions," in *IEEE Symposium on VLSI Technology and Circuits*, Honolulu, HI, USA, 2022, pp. 288-289, doi: 10.1109/VLSITechnologyandCir46769.2022.9830306.

[25] A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S Yuasa, J. Grollier and D. Querlioz, "Neural-like computing with populations of superparamagnetic basis functions," *Nature Communications*, vol. 9, pp. 1-11, 2018, doi: 10.1038/s41467-018-03963-w.

[26] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno and S. Datta, "Integer factorization using stochastic magnetic tunnel junctions," *Nature,* vol. 573, pp. 390-393, 2019, doi: 10.1038/s41586-019-1557-9.

[27] S. Heo, D. Kim, W. Choi, S. Ban, O. Kwon,and H. Hwang, "Experimental Demonstration of Probabilistic-Bit (p-bit) Utilizing Stochastic Oscillation of Threshold Switch Device," in *2023 IEEE Symposium on VLSI Technology and Circuits*, Kyoto, Japan, 2023, pp. 1-2, doi: 10.23919/VLSITechnologyandCir57934.2023.10185217.

[28] T. Gong, Q. Hu, D. Dong, H. Jiang, J. Yang, X. Xu, X. Chen, Q. Luo, Q. Liu, S. S. Chung, H. Lyu and M. Liu, "A 128kb Stochastic Computing Chip based on Flicker Noise with High Noise Density and Nearly Zero Autocorrelation on 28-nm CMOS Platform," in IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2021, doi: 10.1109/IEDM19574.2021.9720588.