Tunable Spintronic Devices with Different Switching Mechanisms for Probabilistic and Stochastic Computing

Brandon R. Zink,[1] Yang Lv,[2] Deyuan Lyu,[2] Brahmdutta Dixit,[2] Qi Jia,[2] Yifei Yang,[2] Yu-Chia Chen,[2] Sreevatsan Sangaprasad,[2] and Jian-Ping Wang,[2*]

[1]Department of Biochemistry, University of Maryland College Park, College Park, MD 20742, USA
[2]Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN 55455, USA

ABSTRACT. Probabilistic spin logic (PSL) has recently been proposed as a novel computing paradigm that leverages random thermal fluctuations of interacting bodies in a system rather than deterministic switching of binary bits. A PSL circuit is an interconnected network of thermally unstable units called probabilistic bits (p-bits), whose output randomly fluctuates between bits 0 and 1. While the fluctuations generated by p-bits are thermally driven, and therefore, inherently stochastic, the output probability is tunable with an external source. Therefore, information is encoded through probabilities of various configuration of states in the network. Recent studies have shown that these systems can efficiently solve various types of combinatorial optimization problems and Bayesian inference problems that modern computers are unfit for. Previous experimental studies have demonstrated that a single magnetic tunnel junctions (MTJ) designed to be thermally unstable can operate tunable random number generator making it an ideal hardware solution for p-bits. Most proposals for designing an MTJ to operate as a p-bit involve patterning the MTJ as a circular nano-pillar to make the device thermally unstable and then use spin transfer torque (STT) as a tuning mechanism. However, the practical realization of such devices is very challenging since the fluctuation rate of these devices are very sensitive to any device variations or defects caused during fabrication. Despite this challenge, MTJs are still the most promising hardware solution for p-bits because MTJs are very unique in that they can be tuned by multiple other mechanisms such spin orbit torque, magneto-electric coupling, and voltage-controlled exchange coupling. Furthermore, multiple forces can be used simultaneously to drive stochastic switching signals in MTJs. This means there are a large number of methods to tune, or termed as bias, MTJs that can be implemented in p-bit circuits that can alleviate the current challenges of conventional STT driven p-bits. This article serves as a review of all of the different methods that have been proposed to drive random fluctuations in MTJs to operate as a probabilistic bit. Not only will we review the single-biasing mechanisms, but we will also review all the proposed dual-biasing methods, where two independent mechanisms are employed simultaneously. These dual-biasing methods have been shown to have certain advantages such as alleviating the negative effects of device variations and some biasing combinations have a unique capability called 'two-degrees of tunability', which increases the information capacity in the signals generated.

I. INTRODUCTION.

In recent years, transistor size scaling is reaching a physical limit [1] and the pace of scaling will fall behind Moore's law. Furthermore, energy consumption per transistor has not scaled well with transistor size, meaning that total energy consumption continues to increase rapidly as computing demand increases [2]. To further compound the problem of massive global energy consumption, increased computing demand has coincided with major developments in machine learning and AI [3–5]. Modern computing methods that rely on binary numbers to encode information are ill-suited to perform various types of AI problems, particularly hard optimization problems. Since further transistor scaling is not a sustainable, long-term solution for solving future AI demands, many researchers have been investigating alternative methods of processing and encoding information to perform AI tasks.

One such alternative computing approach that has received attention in recent years is probabilistic computing [6–9], where random noise of interacting bodies in a system is leveraged to solve certain types of NP-hard problems [8,9]. Due to the inherent stochasticity of the system, a probabilistic computer can accelerate a broad class of algorithms such as Monte Carlo [10], Bayesian inference [11–13] and Boltzmann machine learning [13,14]. Also, these systems are ideal for solving optimization problems due to their natural mapping to Ising models. Combinatorial optimization can be solved by formulating the constraints of the problem, which are used to create an energy landscape of the physical

*Contact authors: bzink@umd.edu, jpwang@umn.edu

system [15–17]. The system has the highest probability of residing in the minimum of the energy landscape, and the configuration of states that causes energy minimization corresponds to the solution of the optimization problem.

Note that conventional deterministic approaches can also be used to solve optimization problems. For example, Yamoka et al designed a CMOS prototype of a natural computer for max-cut problems [18] and Moy et al proposed a ring oscillator based Ising machine [19]. Both studies showed significant reduction in power consumption when compared to traditional von-Neumann designs, however, the configuration of spins in any deterministic system tends to reside in local minimums, leading to inaccurate results. This can be avoided through a process called CMOS annealing, but this requires additional peripheral circuitry and computation overhead. The benefit of probabilistic approaches for combinatorial optimization is that the injection of random noise occurs naturally.

The elementary building block of a probabilistic computer is a probabilistic bit, or a p-bit [10,20]. Hardware implementation of a p-bit requires a stochastic element, an input terminal to bias the stochastic element, and an output terminal to read the element. The behavior of the stochastic element of the p-bit is the same as a binary stochastic neuron [21], where the state fluctuates randomly between 0 and 1. While the fluctuations of this stochastic element are random, the time-averaged value of these fluctuations need to be tunable with the input. In a circuit of interconnected p-bits, the input of each p-bit is proportional to the weighted sums of the outputs of the other p-bits plus an additional biasing term. The weights and biases of the p-bits are determined based on the constraints of the problem.

One such device that possesses all of the desired characteristics of a p-bit is the magnetic tunnel junction (MTJ), which is a two-terminal binary device whose information is encoded in the relative magnetization orientation of two ferromagnetic layers separated by a tunneling barrier. The process of switching magnetization states in MTJs is inherently stochastic, therefore, a single MTJ can be used as a true random number generator [22]. Previous studies have demonstrated that MTJ based random number generators are 2-4 times more energy-efficient than CMOS based random number generators [23]. The CMOS compatibility of MTJs arises from the fact that the magnetization state can be read through a small read current applied across the tunnel barrier since the resistance of the MTJ changes with magnetization orientation due to tunneling magnetoresistance (TMR) [24]. Furthermore, the magnetization state can be manipulated by the use of an applied current via spin transfer torque (STT) [25,26] or spin orbit torque (SOT) [27–30]. In recent years, additional switching methods involving an electric field have also been explored to reduce the write energy of MTJs. These include voltage controlled magnetic anisotropy (VCMA) [31–34] and voltage-controlled exchange coupling (VCEC) [35].

Typically, MTJs are designed for non-volatile memory units in magnetoresistive random-access memory (MRAM), where the state retention can exceed 10 years [36]. However, through simple modifications in the MTJ stack and pillar design, the MTJ will become unstable and thermal noise will drive random fluctuations between the two magnetic states, thus providing an entropy source for probabilistic bits [37–40]. Furthermore, biasing mechanisms such as STT [41], SOT [42], VCMA [43], and VCEC [44] can be used to tune these fluctuations to favor one resistance state over the other, thus providing the non-linear transfer characteristics. In this paper, we present an overview of the experimental and theoretical work on MTJ based probabilistic bits driven by various switching mechanisms. We will compare the performance of p-bits driven by STT, SOT, VCMA, and VCEC in terms of speed, energy efficiency, and compatibility with modern fabrication methods.

Another key feature of MTJs is that their magnetization states can be manipulated by multiple forces simultaneously. For example, the current paths for STT switching and SOT switching are independent of one other, therefore, the two switching mechanisms can be applied in conjunction with each other. Other driving forces that can be employed simultaneously with STT are an external field and magneto-electric (ME) coupling. Using two or more biases in MTJs to generate random switching signals in MTJs has several key advantages. First is that the overall switching energy and switching speed can both be reduced under certain biasing conditions such as STT and SOT [45–48] and others, such as VCEC and SOT, can be combined to reduce the switching current density by several orders of magnitude [49]. A second advantage is that the use of simultaneous biases can alleviate one of the major challenges of probabilistic circuits that use thermally unstable MTJs, which is that their switching speeds and transfer characteristics are highly susceptible to device defects and variations that arise from the fabrication processes [50–52]. A third advantage of using two driving forces is that it provides two degrees of tunability, which means that the high and low-state dwell times can be tuned separately thus doubling the information capacity in the signals generated. Because of these advantages, this paper will also review several different ways in which different mechanisms can be used in

conjunction to optimize p-bit performance such as STT + magnetic field, STT + ME coupling, STT + SOT, and VCEC + SOT.

The organization of our review goes as follows. In section 2, we will provide the necessary background on circuits and applications of probabilistic bits, MTJ basics, and the physics of the various switching mechanisms. An overview of recent research for probabilistic bits driven by STT, SOT, VCMA, and VCEC switching will be presented in section 3. Section 4 will review research where p-bits are driven by a combination of two forces, namely, STT + field, STT + ME coupling, STT + SOT, and VCEC + SOT. A comparison in performance of all the different driving mechanisms in p-bits will also be presented in section 4. A roadmap of probabilistic computers and future MTJ technologies will be discussed in section 5. Finally, the paper will be concluded in section 6.

II. Background

A. Magnetic Tunnel Junction

Experimental work on magnetic tunnel junctions (MTJs) can be traced back to 1988, when the giant magnetoresistance (GMR) effect was discovered independently by groups led by Fert [53] and Grünberg [54], both were awarded the Nobel Prize in Physics in 2007. The GMR phenomenon revealed that electrons with majority-spin (aligned with the local magnetization) experience significantly lower scattering rates compared to minority-spin electrons, leading to measurable changes in electrical resistance. This discovery sparked the field of spintronics, wherein the electron's spin degree of freedom is exploited alongside its charge to enable new device functionalities.

A classic demonstration of GMR is the ferromagnetic (FM)/non-magnetic (NM)/FM trilayer, often termed a spin-valve structure. In this configuration, a current passing through the first FM layer emerges as spin-polarized, carrying more majority-spin electrons. As these electrons traverse the NM spacer and enter the second FM layer, they encounter spin-dependent scattering processes that differ depending on whether the second layer's magnetization is parallel or antiparallel to the first. In the parallel alignment, most majority-spin electrons pass through with relatively low scattering, yielding a lower device resistance. Conversely, in the antiparallel alignment, the mismatch in the electrons' spin polarization and local magnetization causes more scattering, thereby increasing the resistance.

While the GMR-based spin-valves led to substantial technological developments, particularly in magnetic field sensing and read heads for hard disk drives, the magnitude of GMR at room temperature usually remains only a few percent [55]. One reason for this limit is the series resistance, which dilutes the spin-dependent signal. TMR was first observed in an experiment by M. Jullier in 1975 [56], which reported change in conductance in Fe/Ge/Fe films with an applied voltage, which was attributed to the spin dependent tunneling effect. In 1995, Moodera *et al* [57] and Miyazaki *et al* [58] replaced the metallic spacer (e.g., Cu or Cr) with an insulating tunnel barrier (e.g., amorphous AlO_X), thereby creating an MTJ structure in which electrons can tunnel from one FM electrode to another. The insulating tunnel barrier increases the areal resistance and enables room-temperature tunneling magnetoresistance (TMR) of a few tens of percent [57,58].

A major breakthrough came in 2001, when theoretical works by Butler *et al.* [59] and Mathon and Umerski [60] predicted that a (001)-oriented MgO barrier, coupled with Fe(001) electrodes, could yield TMR values exceeding 1000%. The key lies in the coherent tunneling of Bloch states with Δ_1 symmetry in Fe, which experience minimal decay through the MgO layer for majority-spin electrons. In 2004, Yuasa *et al.* [61] and Parkin *et al.* [62] experimentally demonstrated TMR ratios of about 200% at room temperature using Fe/MgO/Fe-based MTJs. Today, the combination of a crystalline MgO tunnel barrier and Fe-based electrodes represents the mainstream MTJ architecture. Ever-higher TMR values have continued to be reported. In 2023, Scheike *et al.* achieved a TMR of 631% [63], which, to our best knowledge, is the highest TMR at room temperature.

B. Switching Mechanisms

Despite continuously improving TMR and thus enhanced *read* performance, the *write* mechanism of MTJs remains a key research focus. Conventionally, switching an MTJ between parallel and antiparallel magnetization states relies on an external magnetic field. In this scheme, one FM electrode is engineered to have sufficiently high coercivity so it can maintain its original magnetization direction and act as a *reference*, while the other FM electrode has lower coercivity and can be *freely* manipulated by the external field. As a result, the switching between low and high resistance states of an MTJ is realized by field-driven magnetic reversal of its free layer. However, this field-driven switching approach is not readily scalable because: 1) Generating the necessary magnetic field typically demands large currents, resulting in significant power consumption; 2) The required field strength is inversely proportional to the area of the free layer, complicating device miniaturization; and 3) The switching field of one

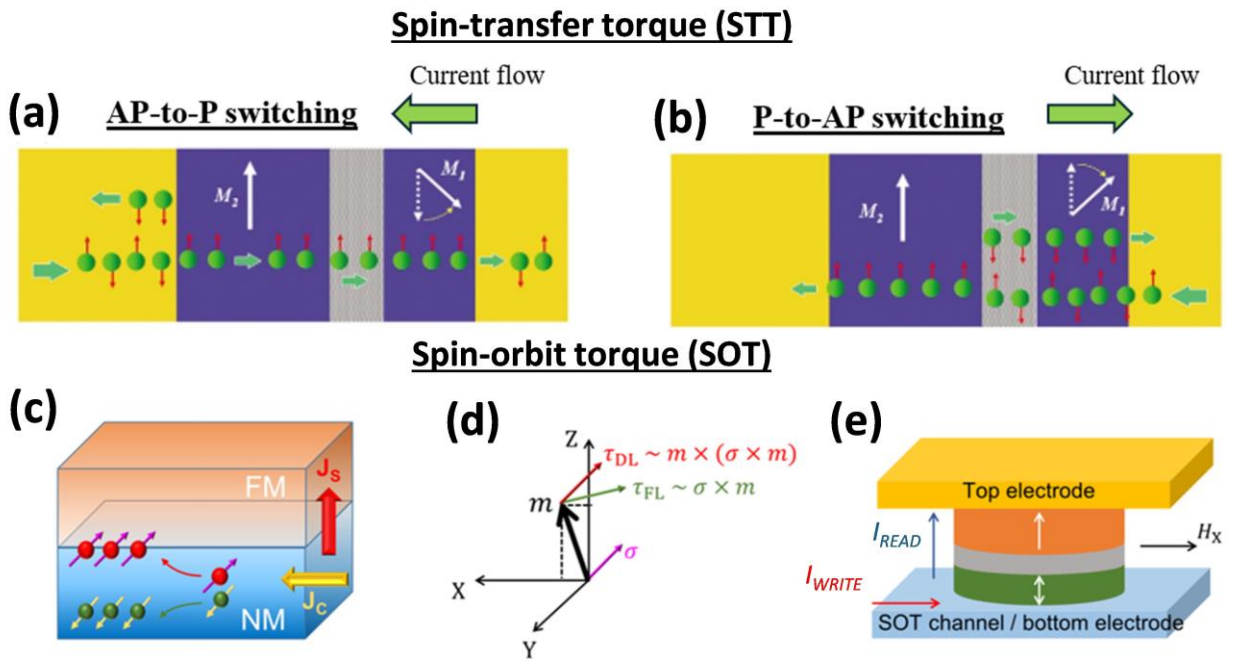

Fig. 1. Illustration of **(a-b)** spin-transfer torque (STT) and **(c-e)** spin-orbit torque (SOT) switching mechanisms. **(a-b)** Spin-dependent tunneling occurs between the two ferromagnetic layers, M_1 and M_2, which generates a net torque on magnetization M_1. The direction of the torque depends on the direction of current flow (Reprinted from [24] with permission from Elsevier). **(c)** Schematic of spin Hall effect, where a charge current (J_C) is converted to a spin current (J_S). **(d)** Damping-like (DL) and field-like (FL) spin-orbit torques exerted on the magnetization. **(e)** Structure of a SOT-MTJ, where read and write currents are decoupled.

MTJ can interfere with its neighboring junctions, especially when devices are densely packed. Due to these limitations, field-driven MTJs can only be scaled down to about 90 nm [64].

To overcome this size limit and improve MTJs' *write* performance, researchers have pursued various field-free switching mechanisms. The most technologically significant approach leverages Spin-Transfer Torque (STT), a concept independently predicted by Slonczewski [25] and Berger [26] in 1996. They showed that a spin-polarized current can transfer its spin angular momentum to the local magnetization of the material it traverses, exerting an effective torque, as illustrated in Fig. 1(a-b). If this torque is strong enough, it can reorient the magnetization. In an MTJ, when electrons flow from the reference layer to the free layer, they first become spin-polarized in the reference layer and then transfer their spin angular momentum to the free layer, driving its magnetization to align parallel to that of the reference layer. Conversely, when electrons flow in the opposite direction, they become spin-polarized by the free layer. As majority-spin electrons experience less scattering in the reference layer, so the electrons reflected back generally carry the minority spin, pushing the free layer's magnetization to align antiparallel to that of the reference layer. This bidirectional, field-free switching mechanism underpins STT-MTJ devices, which have already seen commercial adoption [65].

While STT holds great potential for spintronics applications, it faces inherent limitations, including low reliability and high power consumption. A distinct type of spin torque, known as spin-orbit torque (SOT), has been discovered in materials with strong spin-orbit coupling [27,28,30,66–70], offering improved reliability and device performance [71–73]. SOT is mainly generated by a mechanism called spin Hall effect (SHE) [74,75]. As illustrated in Fig. 1(c), when a charge current (J_C) is applied in a non-magnetic material (NM) with large spin-orbit coupling, electrons with different spin polarizations will be deflected in opposite directions. Therefore, a spin current (J_S) with spin-polarized electrons is generated and flows into the adjacent ferromagnetic material (FM). In addition to SHE, Rashba-Edelstein effect is another mechanism that can generate SOT at interfaces [76–78] or through the topological surface states [79–81]. In both cases, the spin accumulation exerts a torque on the FM's magnetization, enabling its switching, as depicted in Fig. 1(d). By integrating the NM/FM bilayer into a magnetic tunnel junction (MTJ), a SOT-MTJ structure can be realized, as shown in Fig. 1(e). Compared to traditional STT-MTJs, SOT-

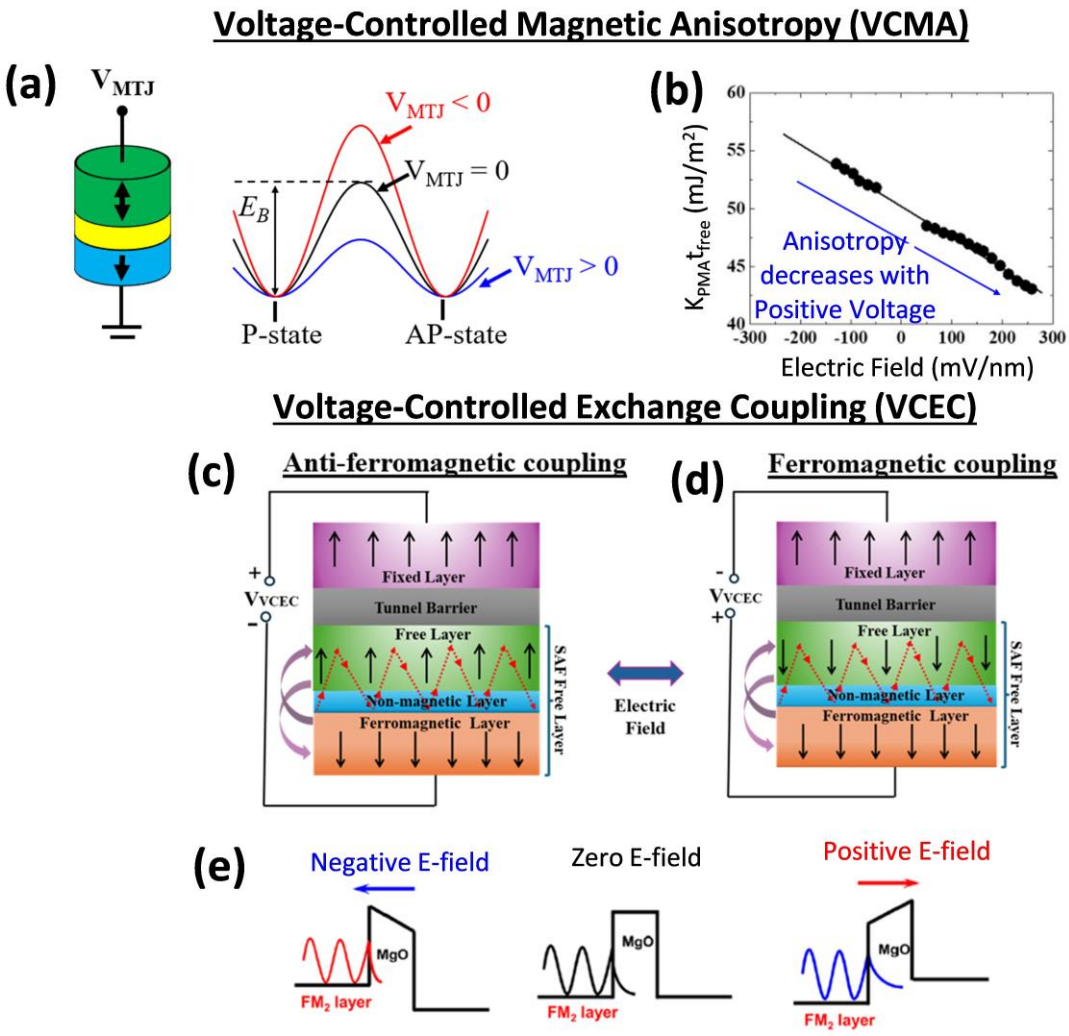

Fig. 2. (a) Influence of voltage (V_{MTJ}) on energy barrier (E_B) showing a decrease in anisotropy at positive voltages, which can allow for thermal fluctuations. **(b)** Plot showing linear dependence of perpendicular anisotropy (K_{PMA}) with electric field induced by voltage (reproduced from ref. [31]). **(c,d)** The interlayer exchange coupling (IEC) between the bottom ferromagnetic layer (FM1) and the free layer (FM2) can be tuned by altering the exchange interaction from ferromagnetic (FM) to antiferromagnetic (AFM) through an external electric field. This switching subsequently enables the free layer (FM2) and the upper fixed layer (FM3) to adopt either high- or low-resistance configurations. **(e)** VCEC emerges from changes in reflectivity at the interface between the nonmagnetic layer (NM) and the second ferromagnetic layer (FM2), influenced by interactions at the FM2 and tunnel barrier interface. When an electric field is applied, electron wave functions extend into the tunnel barrier, modifying both the reflection phases and penetration depths of spin-up and spin-down electrons. Such alterations result in a shift between ferromagnetic (FM) and antiferromagnetic (AFM) exchange coupling.

MTJs offer higher power efficiency, faster switching speeds, and improved reliability [72,73].

One of the key challenges for both STT and SOT switching is the high current densities required, which limits the energy efficiency of both STT- and SOT-MRAM. In the pursuit of ultra-low energy consumption for MTJ devices, voltage-controlled magnetic anisotropy (VCMA) effect plays a crucial role in further reducing both the write and read energy [31–34]. VCMA switching is realized by applying a voltage pulse across the MTJ and the electric field modifies the interfacial electron distribution, leading to either a reduction or an increase in magnetic anisotropy depending on the direction of the applied electric field, as illustrated in Fig. 2(a-b). When the VCMA effect is combined with thermal fluctuations, the reduced energy barrier increases the probability of stochastic magnetization

switching, making the switching process inherently probabilistic [82,83]. However, when VCMA is integrated with STT or SOT, it facilitates deterministic magnetization switching by reducing the critical current required for the switching process [84–86]. This synergistic effect enables more efficient and reliable magnetization control, making VCMA a promising candidate for next-generation energy-efficient spintronic devices.

VCMA has emerged as the most widely studied voltage-controlled write mechanisms of MTJs, however, achieving nanosecond-scale precessional switching typically requires the presence of an in-plane magnetic field or an SOT current to assist the process. This limitation highlights the need for a voltage-driven alternative to spin-transfer torque (STT) that can achieve fast and deterministic switching without the reliance on an external magnetic field. Voltage-Controlled Exchange Coupling (VCEC) has demonstrated its potential by enabling bipolar magnetization switching in magnetic tunnel junctions (MTJs), making it a promising candidate for energy-efficient spintronic devices. VCEC can occur in an MTJ where the free layer consists of a ferromagnetic /non-magnetic /ferromagnetic (FM /NM/ FM) structure, where the two FM layers are expected to couple via the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction through the NM spacer layer, provided the NM layer is sufficiently thin. This coupling oscillates between P and AP alignments depending on the spacer layer thickness. The concept was first proposed by Bruno, who suggested that the RKKY interaction could be manipulated through an applied voltage [87], which modulates the spin reflectivity at the interfaces. This voltage-induced modification of the RKKY interaction was later demonstrated by Zhang et al to drive MTJ switching [35]. Using conductive atomic force microscopy (C-AFM), bipolar switching behavior was observed in the MTJ, providing direct evidence of voltage-driven exchange coupling. Further studies confirmed this effect by analyzing shifted minor loops under different voltages in patterned devices [49] and observing the generation of stochastic signals [44], reinforcing its feasibility for practical applications in next-generation spintronic memory and logic technologies.

C. Overview of Probabilistic Computing and Need for New Devices

P-bits serve as the fundamental computational units of probabilistic computing, analogous to binary bits in deterministic digital systems, yet uniquely embodying characteristics that bridge classical and quantum paradigms. Functioning as stochastic entities, P-bits dynamically fluctuate between 0 and 1, with their probability distribution governed by the switching mechanisms of Magnetic Tunnel Junctions (MTJs). This inherent stochasticity enables them to act as classical counterparts to quantum bits (qubits), effectively emulating superposition states within a classical framework and thus facilitating applications across both classical and quantum-inspired computing domains. While P-bits can be realized through various tunable stochastic components [10,88], Low-Barrier Nanomagnets (LBMs) and stable nanomagnets with adjustable energy barriers emerge as the most energy-efficient and practical candidates for implementation [89]. Specifically, MTJs engineered with tailored free layers and optimized energy barriers enable highly efficient P-bit realization, leveraging the same switching physics such as Spin-Transfer Torque (STT), Spin-Orbit Torque (SOT), Voltage-Controlled Magnetic Anisotropy (VCMA), and Voltage-Controlled Exchange Coupling (VCEC)—that underpin Magnetic Random Access Memory (MRAM) technology. Additionally, alternative control mechanisms, including strain-induced and field-driven tuning, further expand the configurability of P-bits. This study provides a comprehensive exploration of P-bit behavior from a magnetics perspective, wherein their probabilistic dynamics are described by following mathematical equations.

$$m_i = sign[\tanh(\beta I_i) - r_{[-1,+1]}] \quad (1)$$

$$I_i = \sum_j W_{ij} m_j + h_i \quad (2)$$

Here, m_i, the output of p-bit is the time averaged magnetization of the nanomagnet, defined as a bipolar variable (m belongs to {-1, +1}) and r is a uniform random number drawn from the interval [-1,+1]. W in Eqn 2 is the coupling strength matrix between p-bits, β in Eqn. 1 is the inverse temperature used to control noise like thermal noises or equivalent in the systems and {h} is the bias vector [6]. This very intrinsic sigmoidal response of nanomagnets, stemming from its tunable inherent stochasticity, enables its native mapping to a broad category of Monte Carlo algorithms [6]. Furthermore, the behavior of p-bit (Eqn. 1) serendipitously aligns with that of binary stochastic neurons (BSN), the building blocks of stochastic neural networks, whose subsets are Boltzmann machines (BM), its bipartite variant Reduced Boltzmann Machines (RBM) and Bayseian/Belief networks [20,90]. Significant advancements have been achieved by interconnecting p-bits as Binary Stochastic Neurons (BSNs) in a neural network architecture, forming Boltzmann machines,

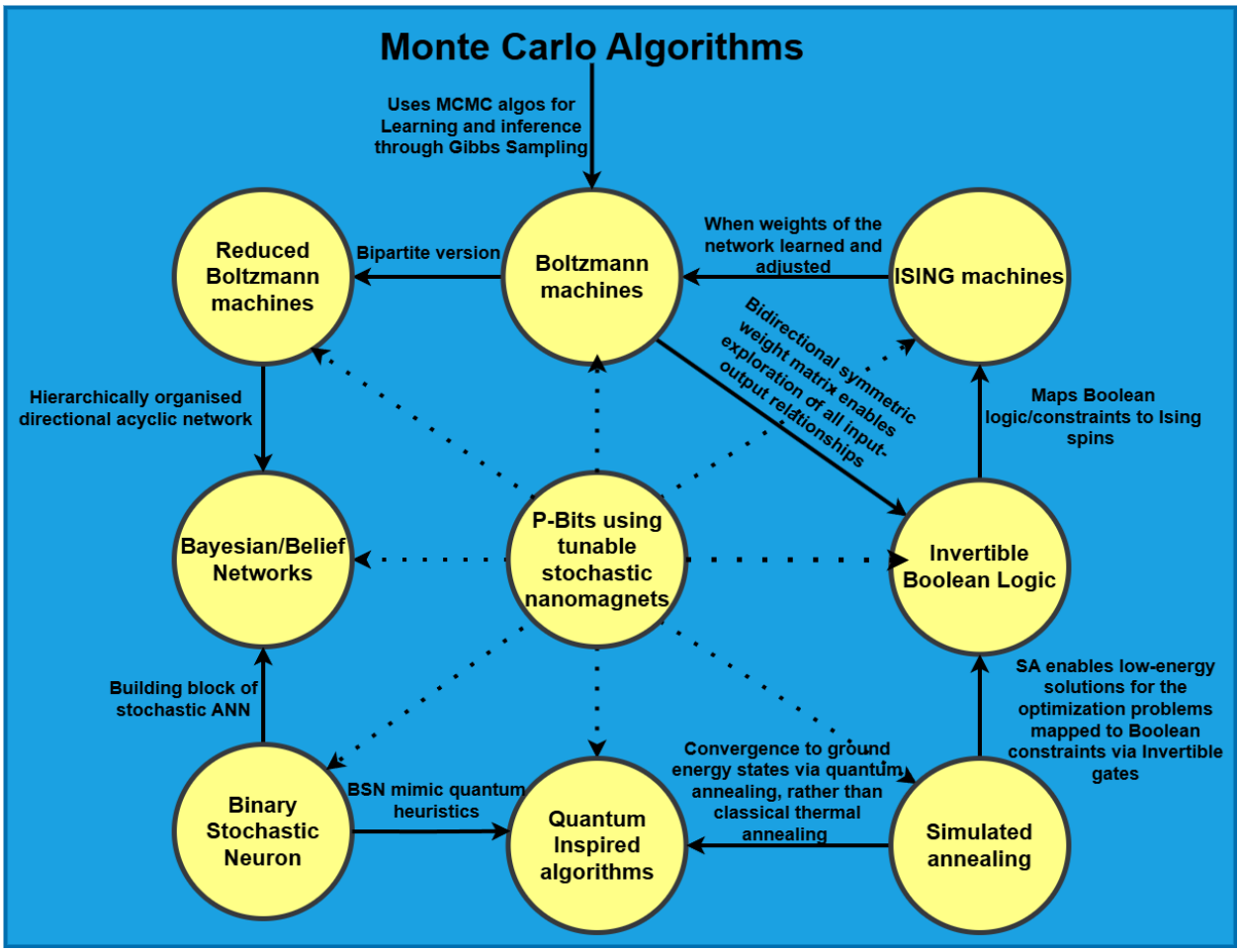

Fig. 3. A high-level schematic illustrating the interconnections among key probabilistic computing applications using Monte-Carlo algorithms. Central to all applications are the P-bits realised using Spintronic devices, which enable their energy-efficient hardware-level implementation. The figure highlights how Boltzmann machines, Bayesian/Belief networks, Ising machines, invertible Boolean logic, simulated annealing, quantum-inspired algorithms, and other paradigms interrelate and can be viewed as special cases or extensions of one another.

where the weight matrix (Eqn. 1) is symmetrized ($W_{ij} = W_{ji}$) and systematically adjusted to align with specific training objectives and functional requirements [13]. In particular, the iterative updating of each p-bit in a Boltzmann Machine through well-defined probabilistic techniques such as Gibbs Sampling ensures that the system attains a steady-state distribution. This can be further extended to implement Bayesian or belief networks, which are structured as directed acyclic graphs (DAGs) ($W_{ij} \neq W_{ji}$), where the edges encode causal conditional probabilities between parent and child nodes, governed by Bayes' theorem [11]. Such a belief network is realized by hierarchically organizing Boltzmann networks, incorporating additional layers, as demonstrated in [91], to construct Bayesian Neural Networks. Additionally, Probabilistic Spin Logic (PSL), unlike conventional deterministic logic, inherently introduces invertibility and Boolean satisfiability by allowing input states to fluctuate when the output is clamped. This characteristic is particularly pronounced in Boltzmann-based implementations, where the symmetric weight matrix enables bidirectional state updates, following the Boltzmann distribution. The system, in its natural course of energy minimization, explores all possible input-output relationships, and upon reaching the ground state, the correct input states are inherently deduced for any clamped output, thereby endowing Boolean gates with polymorphic computational capabilities [92].

The algorithms that benefit most from p-bit hardware are those dominated by repeated stochastic sampling or combinatorial search—namely Monte-

Carlo inference in Bayesian networks, Ising-model optimization (e.g., MAX-CUT, knapsack), and quantum-Monte-Carlo emulation. In software these tasks scale exponentially because each new variable requires fresh random numbers and costly convergence checks; the same kernels mapped to arrays of MTJ-based p-bits run in effectively linear time, because every device delivers an independent, nanosecond-scale, true-random sample while the on-chip "kernel" accumulates and feeds back results in one clock cycle. Kaiser and Datta report $>10^2$-$10^3\times$ speed-ups versus multithreaded CPU or GPU baselines at a fraction of the energy (few-fJ per update) and demonstrate that solution quality remains within 1 % of optimum for problem sizes exceeding 10^4 variables [10]. By collapsing the Monte-Carlo loop into physics, p-bit hardware also eliminates latency overheads and DRAM traffic, yielding further order-of-magnitude improvements in energy-delay product that scale nearly linearly with the number of parallel p-bits, thereby providing the strongest computational case for the MTJ-based approaches surveyed in this review.

Many real-world computational challenges inherently exhibit multi-body interactions, mirroring complex physical systems. By mapping these problems onto their fundamental physical counterparts, we can exploit nature's intrinsic optimization mechanisms for enhanced computational efficiency. In this context, the principles of Spintronics, rooted in the interactions between electron spins, offer a compelling framework. The behavior of p-bits, stochastic probabilistic entities, described by Eqn. 1 encapsulates the mathematical formalism of the Ising model, a foundational paradigm in equilibrium statistical physics. The Ising model inherently adheres to the Boltzmann distribution, where spin interactions minimize system energy under thermal fluctuations, leading to a greater likelihood of lower-energy states while suppressing higher-energy configurations [13]. This intrinsic correspondence between Ising models and Boltzmann machines facilitates their mutual interchangeability, enabling one to be construed as a specialized instance of the other. For instance, Boltzmann machines can be interpreted as Ising machines, wherein the interaction weights are dynamically learned and adjusted during training [13]. This flexibility to reinterpret one model within the formalism of another significantly streamlines the process of mapping complex optimization problems to probabilistic paradigms such as Ising models. For instance, in [6], key combinatorial optimization problems including maximum satisfiability, number partitioning, and knapsack problems are mathematically condensed and mapped topologically to invertible Boolean circuits, then these invertible circuits are encoded to Ising energy states. Therefore, this facilitates translation of p-bit nodes in neural networks to Ising models through invertible Boolean logic using techniques like Hopfield networks to represent any Boolean truth table as weight matrix in Ising models. Building upon this understanding of interconnection of multiple p-bit applications, we can extend these interpretations to quantum algorithms such as simulated annealing, Adiabatic Quantum Computing (AQC), where a network of qubits with energy E corresponding to multi-body interactions is captured using a network of p bits, whose electrical connections mimic multi-body interactions. Such a Boltzmann network through correlation of stochastic nanomagnets achieves analytical synaptic weight update rather than learning similar to AQC [93]. Similarly combinatorial problems such as max-cut can be mapped to simulated annealing models, which utilize the minimization of energy or cost function in perspective to NP hard optimization problems much like Ising model working on bringing the system to ground state, corresponding to optimal solution for the optimization problem [16].

While traditional applications of probabilistic or stochastic computing primarily manifest through Monte Carlo algorithms or as derivatives of random number generators, lesser-explored potentials of Low Barrier Magnets remain underreported. This work seeks to illuminate these unconventional capabilities of LBM-based p-bits, fostering interest in their novel applications and expanding their scope. Notably, the intrinsic thermally driven randomness of LBMs is leveraged to develop on-chip thermal sensors, wherein temperature is inferred from the switching probability of Magnetic Tunnel Junctions (MTJs) [94] and through the analysis of the temperature dependence of coercivity. The sensor functions by recording switching probabilities over multiple "write-read-reset" cycles, extracting temperature information from the statistical distribution of switching events. In the superparamagnetic regime, where the energy barrier approximates k_BT, the magnetization exhibits rapid telegraphic switching, enabling real-time temperature sensing with enhanced conversion rates and lower power consumption compared to conventional CMOS-based sensors. A similar methodology can be employed to monitor process variations in Magnetic Tunnel Junctions (MTJs) and assess the health of transistors integrated within the Back-End-of-Line (BEOL), where fabrication-induced variations manifest as fluctuations in voltage and energy barriers.

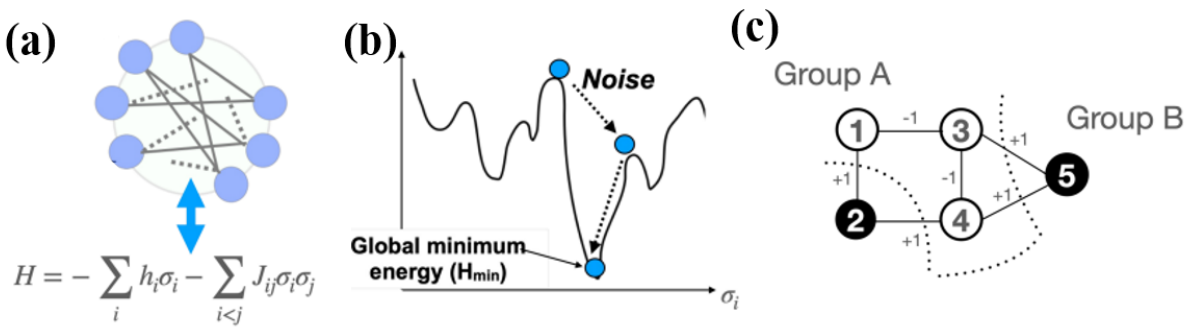

Fig. 4(a) p-bit architecture for representing an Ising model of a combinatorial optimization problem. (b) Mapping the Ising model to a simulated annealing process, where p-bit states are updated under a gradually increased pseudo-inverse temperature to reach the global minimum energy. (c) Five-node MAX-CUT example, which can be mapped to the Ising formulation by assigning binary spin states to nodes. Reproduced from ref. [16], https://creativecommons.org/licenses/by/4.0/.

The MTJ-based variation monitoring approach detailed in [95] involves subjecting MTJs to stress voltages or currents, effectively reducing retention time and enhancing thermal activation probability. This phenomenon is inherently present in superparamagnetic p-bits, where a diminished energy barrier leads to intrinsically low retention times and thermally driven switching. The switching rate of stressed MTJs is subsequently measured and compared against a predefined threshold to dynamically infer process and temperature variations. By engineering the MTJ such that its energy barrier approximates $40k_BT$, a typical threshold for stable nanomagnets, the Voltage-Controlled Magnetic Anisotropy (VCMA) voltage applied to induce superparamagnetic behavior can be repurposed as a sensitivity mechanism for detecting voltage variations. In an ideal scenario with no voltage variation, the applied VCMA voltage transitions a stable magnet into a p-bit, as evidenced by fluctuating MTJ states. However, if say +3% variation occurs, the switching rate increases due to a further reduction in the energy barrier, while a -3% variation results in the MTJ maintaining its stable magnetic state, as reflected in its data retention characteristics. This methodology enables precise detection of voltage variations stemming from process inconsistencies, facilitating the design of compensatory circuits to mitigate their impact. This synergy between physics-inspired computing and stochastic magnetization dynamics paves the way for organic execution of novel algorithms/applications using low power LBM MTJ P-bits, instead of the regular software implementation with conventional CMOS hardware suffering from Von-Neumann bottleneck.

In this subsection we aim to present an insightful albeit non-exhaustive review highlighting the innate interconnection of various probabilistic algorithms and computation models as summarized in Fig. 3. Here in Fig. 3, the edges denote specific conditions or transformations that link one model to another. The arrow directions originate at the source nodes, indicating that the corresponding computing paradigm can be reinterpreted or transformed into the model represented at the destination node. Although the arrows are drawn in one direction for clarity, these relationships are inherently bidirectional, allowing for the reverse interpretation as well. Furthermore, this is not an exhaustive encapsulation of all possible mappings amongst various probabilistic computing paradigms, there can be other edges between nodes, for instance Fig. 4 illustrates how Ising models utilize Simulated annealing and how Invertible boolean gates can be mapped to Ising model.

Additionally, we elucidate how mapping of applications as a combination of one or more of these models can effectively tackle a broad range of challenges spanning from optimization to fabrication limitation mitigation. Subsequently we can accredit the LBM MTJ hardware fostering tunable stochasticity of p-bits, the central tenet to above mentioned applications. Careful energy barrier engineering defined by the following equation, has been done in [50,89,96] to achieve LBM MTJ based p-bits.

$$E_B = \frac{H_K M_S Vol}{2} \tag{3}$$

Here, H_k is the anisotropy field, M_s is the saturation magnetization and Vol is the volume. Drawing parallels to the ongoing pursuit of enhanced gate

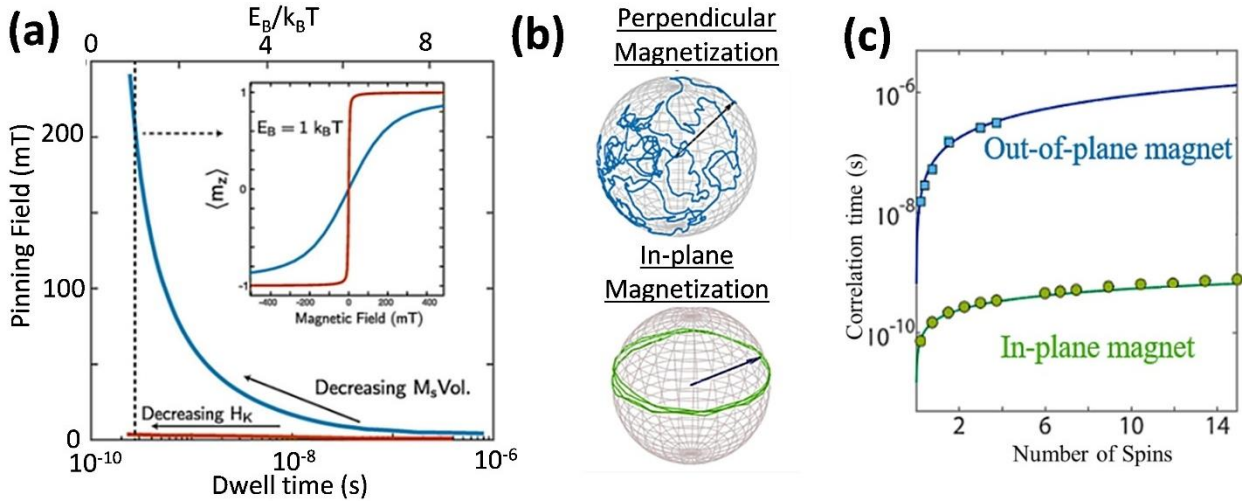

Fig. 5. Design considerations for single-biased p-bits. **(a)** Pinning field vs thermal stability plot (reproduced from ref. [97]) showing that reducing M_SVol as a method of minimizing thermal stability also increases the pinning field whereas reducing H_K has no influence on the pinning field illustrating that reducing H_K is a more energy efficient solution. **(b)** Bloch spheres illustrating the magnetization switching dynamics in low-barrier nanomagnets with perpendicular vs in-plane anisotropy (adapted from ref. [90]). The magnetization in the nanomagnet with perpendicular anisotropy is allowed to fluctuate anywhere inside the Bloch sphere whereas the magnetization direction of the in-plane nanomagnet is constrained to the x-y plane. This constraint in the in-plane nanomagnet leads to a much fast correlation time when compared to the perpendicular nanomagnet, as illustrated in **(c)** (reproduced from ref. [90]).

control in conventional transistor scaling, this work emphasizes tuning/biasing mechanisms and the underlying switching physics of MTJs, which control the probability distribution of random numbers generated by P-bits in the subsequent sections.

III. SINGLE-BIASING MECHANISMS FOR P-BITS

In this section, we will discuss four different methods for generating telegraphic switching signals in MTJs for probabilistic bits that use a single bias scheme. These schemes tend to be the most simple configurations of p-bits since their output signals rely on a single switching mechanism at the input rather than a combination of mechanisms.

The two most common switching methods for p-bits are spin transfer torque (STT) and spin orbit torque (SOT). In both cases, the MTJs should be designed so that their switching energy (E_B) is less than k_BT, so that thermal fluctuations drive random magnetic switching. Since $E_B = H_KM_SVol/2$, reducing E_B to values less than k_BT can be achieved by reducing H_K or M_S. For MTJs with in-plane magnetic anisotropy (IMA), H_K is primarily dominated by shape anisotropy and therefore can be reduced to ≈ 0 by patterning the MTJs as circular nano-pillars rather than elliptical nano-pillars. M_S can be reduced by using MTJs with a permalloy (Py) free layer, followed by post fabrication heat treatment [97], rather than a CoFeB free layer, which is conventionally used in MTJs. Work done by P. Debashis et al showed that reducing H_K is a much more effective and scalable method for reducing E_B than reducing M_S, as shown in Fig. 5(a) [97]. This is because the field required to pin the magnetization increases as M_SVol decreases, thus increasing the energy consumption for tuning the p-bit. However, when H_K is reduced, the pinning field remains constant, thus making reduction of H_K the favorable method of reducing E_B in IMA MTJs.

MTJs with perpendicular magnetic anisotropy (PMA) can also be used as p-bits since they also can be designed to be thermally unstable. In PMA MTJs, H_K is dominated by interfacial anisotropy, which is modified through free layer thickness. By designing a PMA MTJ so that the free layer thickness is on the threshold of in-plane or perpendicular anisotropy, the effective H_K becomes ≈ 0. While PMA MTJs are certainly useful for modern STT-MRAM [98] due to their lower write current and superior scalability, a study by O. Hassan et al showed that PMA MTJs have longer correlation times than IMA MTJs [90]. This is because the magnetization during thermal fluctuations in PMA MTJs can be oriented anywhere in the Bloch sphere whereas the magnetization in IMA MTJs is

strictly limited to the x-y plane due to a high demagnetization field along the z-direction, as illustrated in Fig. 5(b). Note that full magnetization switching is needed for true randomness so that the p-bit can lose memory of previous states. This restriction of the magnetization orientation in IMA MTJs results in correlations times that are two orders of magnitude smaller than PMA MTJs [90], as illustrated in Fig. 5(c).

Since the work by O. Hassan et al [90] and P. Debashis et al [97] indicates that IMA MTJs patterned as circular nano-pillars are the most effective STT and SOT based p-bits, they are the main focus of most theoretical work on p-bit hardware and circuits. However, the switching rate and transfer characteristics of these devices are highly susceptible to device imperfections [50,52]. This makes medium-scale to large-scale experimental demonstrations of p-bit circuits using circular IMA MTJs very unrealistic. To overcome this challenge, many experimental studies have to compromise and use other techniques of generating stochastic switching signals such as modifying free layer thickness in PMA MTJs or hard-axis initialization of thermally stable MTJs using a fixed external magnetic field.

In addition to STT and SOT p-bits, several recent studies have proposed p-bits where the output signals are manipulated by an electric field. One E-field based mechanism is voltage-controlled magnetic anisotropy (VCMA), where E_B is momentarily reduced by an E-field, thus generating telegraphic switching signals through the duration of the E-field. The other E-field switching mechanism is voltage-controlled exchange coupling (VCEC) where the E-field controls the interlayer exchange coupling in a synthetic antiferromagnetic (SAF) free layer. There has been significantly less research done on VCMA and VCEC based p-bits as their STT and SOT counterparts. Yet, p-bits controlled by an E-field have the potential for significant reduction in write energy, and therefore, the prospects and challenges of these devices will be discussed in this section.

A. Spin-Transfer Torque p-Bits

STT-MTJs have been widely investigated for nonvolatile memory applications, and they have recently garnered increasing interest for p-bit implementations. In early 2017, Bapna and Majetich successfully fabricated STT-driven stochastic MTJs (s-MTJs) that generate true randomness, as confirmed by their successful performance on the NIST Statistical Test Suite (STS) [41]. Their results, shown in Fig. 6(a), demonstrate that the time-averaged magnetization in the telegraph signal can be tuned by adjusting the bias. This is understandable, since the STT acting on the free layer can align its magnetization in a particular direction. However, many p-bit design proposal involve an sMTJ connected in series with a transistor and the output probability is tuned with the gate voltage of the transistor [40]. Therefore, such bias-dependence is not always desirable because the stochastic output is influenced by the STT itself. To mitigate this issue, Camsari *et al.* proposed an s-MTJ structure without a reference layer, instead featuring two free layers [99]. As shown in Fig. 6(b), the double-free-layer (DFL) s-MTJ is predicted to exhibit nearly bias-independent behavior over a wide range of voltages. This prediction was later experimentally verified by Ota *et al.* [100] In Fig. 6(c), their DFL sMTJ—where both free layers have the same effective thickness—shows a nearly constant ratio of the relaxation times for the parallel (P) and antiparallel (AP) states under different bias values. Although these advances effectively mitigate the influence of STT and make STT-driven sMTJs more promising, the designs typically require a constant external field to compensate for stray fields from the FM layers. To eliminate external-field requirements and improve robustness against magnetic field perturbations, Selcuk *et al.* proposed a novel DFL s-MTJ structure in which both free layers are replaced by synthetic antiferromagnets (SAFs) [101]. As illustrated in Fig. 6(d), their model demonstrates uniform randomness in the magnetization angle θ between the two free layers.

Beyond the pinning effect of STT, another critical concern for STT-driven s-MTJs is their operation speed. In 2019, Borders *et al.* experimentally demonstrated the factorization of integers up to 945 using STT-driven s-MTJs, an achievement that significantly encouraged further research in the field [93]. However, as illustrated in Fig. 6(e), their perpendicular magnetic anisotropy (PMA) sMTJs operated on the millisecond timescale, which is unfavorable for practical applications. Similar millisecond-scale random telegraph noise in PMA s-MTJs was also reported by Kobayashi *et al.* in 2021 [102]. To ensure true randomness in the sMTJ output, the sampling rate must be limited so that the magnetization has enough time to lose its “memory” of past states. Theoretically, Hassan *et al.* showed in 2019 that the correlation time of a nanomagnet with in-plane magnetic anisotropy (IMA) can be two orders

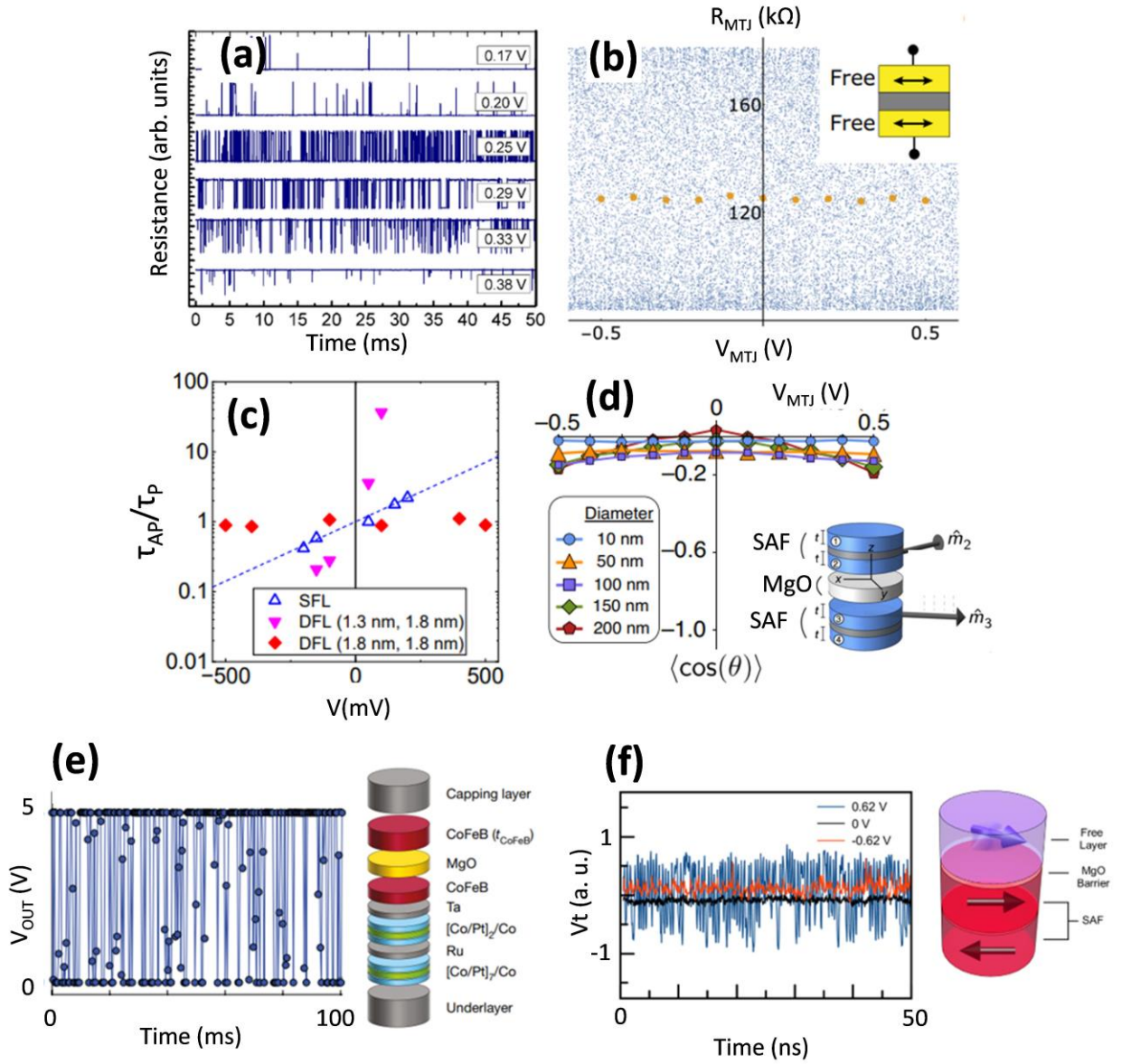

Fig. 6. (a) R vs time traces measured at different bias values (reprinted from [41] with the permission from AIP publishing) **(b)** R vs V characteristics for a DFL s-MTJ. Inset: Structure of DFL s-MTJ. (Reproduced from ref. [99] **(c)** Bias dependence of τ_{AP}/τ_P of a single-free-layer (SFL) s-MTJ, a DFL sMTJ with (t^*_{bottom}, t^*_{top}) = (1.8 nm, 1.3 nm), and a DFL sMTJ with that of (1.8 nm, 1.8 nm). Here, t* denotes effective free layer thickness, excluding magnetic dead layer (reprinted from [100], with the permission from AIP publishing). **(d)** Time-averaged cos(θ) as a function of bias voltage for double-SAF-free-layer sMTJs. Inset: Structure of double-SAF-free-layer s-MTJ. (Reproduced from ref. [101]). **(e)** Left: Time snapshots of output voltage V_{OUT} with input V_{IN} = 1.950 V. Right: Structure of the PMA sMTJ tested (reprinted from [93], with the permission from Springer Nature). **(f)** Left: Time traces of the IMA sMTJ signal under different bias. Right: Schematic of the IMA sMTJ tested (reprinted with permission from [103]. Copyright 2025 American Chemical Society).

of magnitude shorter than that of its PMA counterpart [90]. This shorter timescale arises because the large demagnetization field in IMA magnets confines their magnetization to the sample plane, which leads to a much smaller for IMA nanomagnets (recall Figs. 5(b-c)). In 2021, Safranski *et al.* and Hayakawa *et al.* experimentally demonstrated nanosecond operation speeds in STT-driven IMA sMTJs [103,104]. In 2023, Schnitzspan *et al.* used the NIST STS to verify the true randomness of high-speed

stochastic bit streams generated by STT-driven IMA sMTJs [105]. Their study also confirmed the previously noted tunability of the telegraph signal via STT and highlighted the role of Joule heating—induced by the large STT current—in increasing the device temperature and, consequently, the switching rate. Also in 2023, Sun et al. fabricated STT-driven IMA sMTJs in which both the free layer and the reference layer consist of SAFs [106]. The high-speed stochastic bit streams produced by their devices also passed the NIST STS, further underscoring the viability of STT-driven IMA sMTJs for practical random number generation and p-bit applications.

In addition to the single-device performance of STT-driven sMTJs, the collective behavior of coupled devices is also crucial for hardware implementations of p-circuits. As early as 2016, Mizrahi *et al.* discussed the synchronization phenomenon in two serially connected sMTJs, showing through simulations that the devices could be phase-locked via STT [107]. In 2023, Schnitzspan *et al.* experimentally investigated the effective coupling of two sMTJs connected in series, demonstrating cross-correlation in the system [108]. Meanwhile, parallel coupling has also garnered attention. In 2021, Talatchian *et al.* mutually coupled two sMTJs in a parallel circuit, observing significant cross-correlation [109]. In 2024, Gibeault *et al.* further reported an experimental demonstration of programmable electrical coupling between sMTJs [110]. Across these studies, STT plays a key role, linking the change in charge current to the switching of device resistance, which in turn modulates the current flow and STT itself. As STT-driven sMTJs have attracted growing research interest, numerous related studies, for instance, the temperature dependence [111], underlying physics [112], and compatibility with CMOS techniques [113] and 2D electronics [114] of STT-driven sMTJs, have been conducted.

B. Spin-Orbit-Torque p-Bits

Another approach to implementing p-bits leverages the tunability of magnetization switching by spin-orbit torque (SOT), which offers robustness and high power efficiency [115]. As illustrated in Fig. 7(a), there are three types of p-bit implementations using SOT-MTJs, each featuring a different MTJ configuration [116]. The type I configuration utilizes MTJs with a low barrier that is comparable to k_BT. In this configuration, thermal noise induces spontaneous magnetization fluctuations between two states [42,90,117–120]. One example is shown in Fig. 7(b), where a thermally unstable CoFeB with PMA was fabricated on the SOT material Ta, exhibiting stochastic magnetization switching due to thermal noise [119]. The magnetization states were measured for certain time interval (see Fig. 7(b), right panel) and the average values were calculated from the telegraphic output. The left panel of Fig. 7(b) demonstrates how SOT from Ta tunes the average magnetization without an external field, which was attributed to the tilted anisotropy of the magnet [119].

Although low-barrier MTJs operate via thermal noise and consume minimal power, they lack the ability to retain information, making them unsuitable for applications such as Bayesian neural networks [116]. To address this limitation, high-barrier SOT-MTJs should be used instead. Type II implementation of p-bits are realized using high-barrier magnets with PMA, as depicted in the middle panel of Fig. 7(a). Since conventional SOT materials generate spin currents with in-plane spin polarizations [121–123], the magnetization can be initialized to the in-plane direction by the spin torque from SOT materials. Once the SOT is removed, the magnetization relaxes into either the spin-up or spin-down state with equal probability. An external magnetic field can be used to adjust this probability [97,124]. An implementation example is shown in Fig. 7(c)-(e), where a Ta/CoFeB stack with PMA was fabricated [124]. During device testing, an SOT current was first applied to initialize the magnetization into the hard-axis (in-plane). Upon stopping the current, the magnetization relaxed into either the up or down state with a 50% probability in the absence of an external field. By applying a magnetic field, either through a metal ring (Fig. 7(c)) or directly (Fig. 7(d)), the average magnetization was tuned, demonstrating controlled probabilistic outputs, as shown in Fig. 7(e) [124].

Since conventional SOT materials generate only in-plane spin polarization, it is ineffective in directly tuning the switching probability of a perpendicular magnet. To overcome this limitation, Type III implementations use in-plane magnets for the MTJ, enabling direct SOT-controllable switching (see right panel of Fig. 7(a)) [116,125–128]. A representative work is shown in Figs. 7(f-g). A high-barrier SOT-MTJ stack consisting of W/CoFeB/MgO/CoFeB was fabricated, ensuring data storage and non-volatility [116]. The switching probability of the free layer was modulated by varying the SOT bias voltage, as shown in Fig. 7(f). Furthermore, such SOT-MTJs

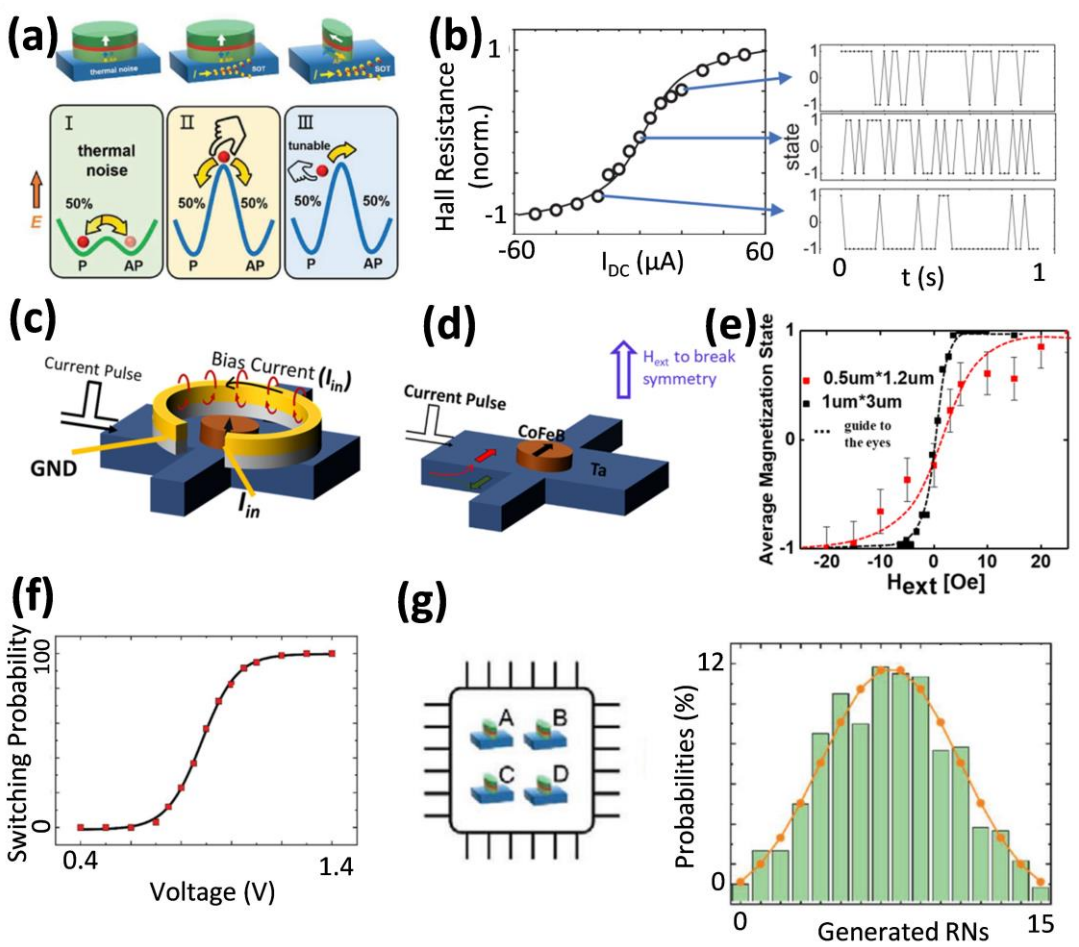

Fig. 7. (a) Three types of p-bits implementations using SOT-MTJ with different MTJ configurations. Type I: Low-barrier MTJ sensitive to thermal noise. Type II: High-barrier MTJ with PMA. Type III: High-barrier MTJ with in-plane magnetic anisotropy. Reproduced from ref. [116], https://creativecommons.org/licenses/by/4.0/. **(b)** Example of Type I implementation: Tunability of the anomalous Hall effect (AHE) resistance in a thermally unstable Ta/CoFeB (PMA) system by the input current in the SOT channel (Ta). The left panel shows the averaged values of the telegraphic output displayed in the right panel. Reproduced from ref. [119]. **(c)-(e)** Example of Type II implementation: A Ta/CoFeB (PMA) stack where the magnetization is controlled by a magnetic field generated by a metal ring **(c)** or externally applied **(d)** after initialization via SOT. The field-dependent average magnetization is shown in **(e)**. Reproduced from ref. [124], https://creativecommons.org/licenses/by/4.0/. **(f),(g):** Example of Type III implementation: A high-barrier W/CoFeB/MgO/CoFeB SOT-MTJ. **(f)** Demonstrates the sigmoid-shaped switching probability as a function of the SOT bias voltage. **(g)** Illustrates a TRNG implementation using four SOT-MTJs in a Bayesian network, producing a Gaussian-distributed output. Reproduced from ref. [116], https://creativecommons.org/licenses/by/4.0/.

can be utilized to create true random number generators (TRNGs) with configurable probability distribution functions (PDFs), including uniform, Gaussian, and exponential distributions.

These distributions can be achieved using 4 SOT MTJs that switch stochastically to represent a 4-bit random number, as illustrated in the left panel of Fig. 7(g) [116]. The probability of each MTJ outputting either 0 or 1 is controlled by the SOT bias voltage according to a predefined PDF. Each MTJ is updated sequentially. The probability of the first MTJ follows the PDF shown in Fig. 7(f) whereas the probability of the following MTJs are dependent on the binary states of the previous MTJs. The Bayesian network – based approach in [116] can follow any discrete distribution depending on the conditional probability functions of the SOT MTJs. A demonstration of a Gaussian-

distributed TRNG output is shown in the right panel of Fig. 7(g).

C. Voltage-Controlled Magnetic Anisotropy p-Bits

The structure of a VCMA-based p-bit is similar to an STT p-bit in that it consists of a 2-terminal MTJ and the signals are being driven by applying a voltage across the tunnel junction. However, the physical mechanism that is driving magnetic switching is different. In STT switching, thermal fluctuations are causing random magnetic switching and a small voltage applied to the MTJ is changing the time-averaged output. However, for VCMA switching, a large voltage is applied to the MTJ to momentarily reduce its switching energy barrier to allow for random fluctuations to occur. If the voltage is less than the critical switching voltage, then thermal fluctuations will cause random switching, however, if the voltage is greater than the critical switching voltage, then the MTJ will operate in the precessional switching regime, as illustrated in Fig. 8(a) One advantage that VCMA has over STT switching is that VCMA switching can occur in MTJs with MgO tunnel barrier thicknesses of 1.5 nm or thicker. This means that VCMA MTJs are significantly less influenced by variations in MgO thickness and more importantly, less conductive. This absence of current flow in VCMA MTJs allows for potential of write energies less than 10 fJ/bit [129]. Since a larger voltage is required for VCMA switching than for STT switching, a synchronous read-out method is preferred where a series of write and read pulses are applied to the MTJ. Now, the speed is defined as the period for each write and read cycle rather than mean dwell times.

While the voltage pulse is applied, the magnetization precesses around the easy anisotropy axis, then randomly settles into one of the two magnetization states once the voltage pulse is turned off, as illustrated in Fig. 8(a). The precession behavior of the magnetization during the voltage pulse is shown in Fig. 8(b), which shows the probability of the magnetization settling in the opposite state of the original (pre-pulse) state with respect to time. This figure shows for pulse widths less than 10 ns, the switching probability oscillates between 0 and 100 %. The amplitude of the oscillations decrease with each cycle, then settles to a 50 % switching probability after around 10 – 15 ns. Note that in order to achieve this behavior, an additional external bias (such as an in-plane magnetic field or an SOT current) needs to be applied to define the easy-axis direction during the application of the VCMA voltage pulse. There are two approaches to using VCMA as an entropy source in p-bits, which we will define as short pulse and long pulse methods for our analysis. In the short pulse method, ultra-fast (~100 ps) pulses are applied to try and capture the magnetization dynamics during a single precession period. In the long pulse method, voltage pulses that over 10 ns are applied to capture enough precession cycles to ensure 50% switching. As we will discuss in this section, both of these methods have their trade-offs. The advantages of the short pulse method are that it allows for GHz operation speed of the write + read cycle and the switching probability can be tuned with the pulse width. For example, a 50% switching probability is achieved by setting the pulse width to half of the precession period. However, the output probability is very noise sensitive since a 100 ps increase or decrease in the pulse width can change the switching probability significantly, as illustrated in Fig. 8(b). Furthermore, the precession period is influenced by any variations in the voltage amplitude or device properties, which can also cause unpredictable shifts in the switching probability. On the other hand, the long pulse method has much longer write + read cycle periods, which leads to higher energy consumption and lower computation speed, and has a fixed output probability of 50 %, meaning that external circuitry is required to produce the sigmoidal transfer curve desired by p-bits. However, this method is very robust against device variations and insensitive to noise. This method was used to provide an experimental demonstration of integer factorization [130].

Currently, there are no studies that investigate the performance of VCMA driven p-bits using the short pulse method. However, several studies have proposed some version of the short pulse method to generate random numbers in restricted Boltzmann machines [131,132] and Bayesian networks [133]. The results from these studies can be used to estimate certain performance metrics in VCMA driven p-bits. In each of these studies, a 0.5 ns pulse is used for random bit generation. Each of these studies showed promising results in terms of high accuracy (>96%) and high throughput (166 MHz) [132]. The energy consumed/bit depended on the biasing method of setting the easy anisotropy axis. In refs. [132,133], this easy axis direction was set using voltage-induced strain, in ref. [131], an SOT current was used, and in ref. [134], an external field was applied in the in-plane direction. Using the SOT current reported the lowest energy consumed per bit at 1 fJ/bit [131] (which was

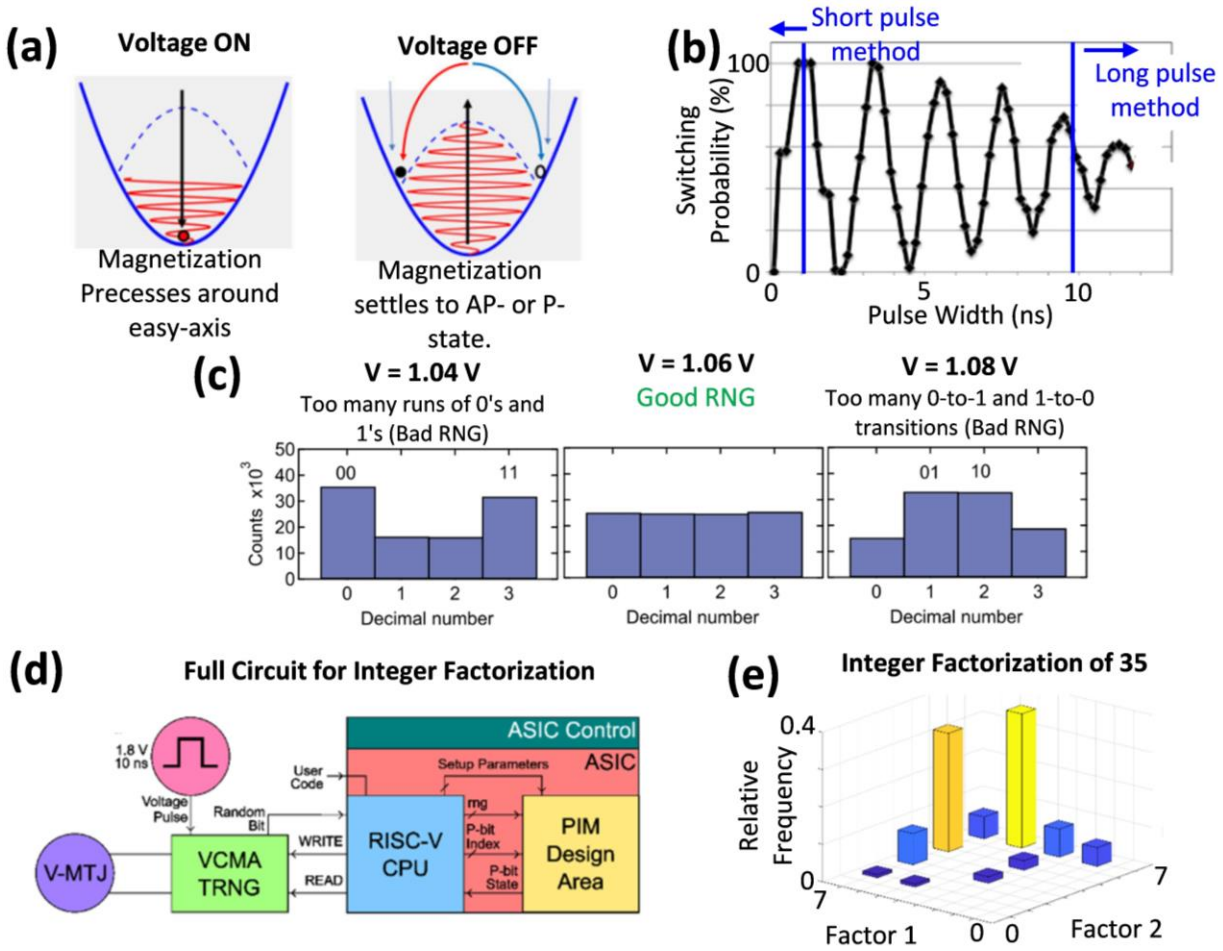

Fig. 8. (a) Illustration of the method of generating random numbers using VCMA MTJs operating in the precessional regime (reproduced from ref. [135], https://creativecommons.org/licenses/by/4.0/). When the voltage is applied, the magnetization precesses around the in-plane easy axis and when the voltage is turned off, the magnetization randomly settles into one of the two magnetization states. **(b)** Dependence of switching probability with pulse duration when operating in the precessional switching regime, showing that the switching probability settles to 50% for durations above 10 ns (reproduced from ref. [82], https://creativecommons.org/licenses/by/4.0/). **(c)** Plots showing the frequency of occurrence for 0-to-0, 0-to-1, 1-to-0, and 1-to-1 transitions at various pulse amplitudes demonstrating that there is degradation in randomness quality when tuning the switching probability with pulse amplitude (reproduced from ref. [134], https://creativecommons.org/licenses/by/4.0/). **(d)** Block diagram of the experimental demonstration of integer factorization via VCMA driven p-bits (long pulse method) presented in [130] illustrating the external circuitry required to implement a probabilistic circuit with long pulse method. **(e)** Experimental results of integer factorization of the number 35 (reproduced from ref. [130], https://creativecommons.org/licenses/by/4.0/).

1000X lower than the STT method of random bit generation), whereas voltage-induced strain reported >100 fJ/bit [132]. The energy consumption using an external field is unknown as it depends on the method of generating an external field. In each of these studies, the switching probability can be tuned with both the pulse width as well as the pulse amplitude. However, as previously stated, the switching probability in this regime is very sensitive to noise and variations between devices. S Nasrin et al claimed that the effects of these variations can be mitigated through online training [132], however, this may not apply to certain types of probabilistic circuits with preset weights determined by the J-matrix. Furthermore, A. Fukushima et al [134] showed that tuning the switching probability with the VCMA voltage amplitude introduces another problem in that the quality of randomness decreases when the voltages are

too high and too low, as illustrated in Fig. 8(c). This plot shows that at low VCMA voltages, the bit streams have too many runs of 0's and 1's and do not pass the NIST statistical test suite (STS) and at high VCMA voltages, the bit streams oscillate between 0 and 1 too frequently and still do not pass the NIST STS. Their results showed that even at the intermediate VCMA voltages, an XOR operation is still required to avoid unwanted correlations and provide high enough quality of randomness to pass the NIST STS. While the short pulse method provides an exciting possibility for VCMA driven p-bits with high computation speed and low energy consumption, there are several challenges that need to be overcome for it to be a feasible solution for hardware implementation of probabilistic circuits.

Several studies have also proposed designs of random number generators based on VCMA driven MTJs using the long pulse method. Typically, these studies use a pulse width ≥ 10 ns to ensure 50 % switching probability. As expected, this method is much slower in terms of computational throughput than the short pulse method, although Gbps throughputs can be achieved using a multibit RNG [82]. However, the multibit RNG is not applicable in p-bit circuits. Additionally, there may be a concern with high energy consumption with this method. Energy calculations by S. Liu et al [135] showed that RNGs driven by VCMA (long pulse method) consume nearly 20 pJ/bit, which is 10X higher than for STT based RNGs and 1000X higher than for SOT based RNGs. Note that 20 pJ/bit is significantly higher than the 311 fJ/bit reported by H. Lee et al [82]. The reduction in energy/bit in the calculations by H. Lee et al can be attributed to shorter write pulses used and more optimal stack for VCMA switching (ie. larger VCMA coefficient). However, the calculations by S. Liu et al do provide some concern regarding the energy efficiency of VCMA RNGs (and p-bits) relative to STT and SOT RNGs.

As previously stated, the main advantage to the long pulse method in VCMA driven p-bits is its robustness to noise and device variation. This can be attributed to the fact that the long pulse is enough time for the precession to settle to a predictable 50 % switching probability, therefore, any variations in precession rate or response to voltages becomes irrelevant at these time scales. This robustness makes VCMA p-bits driven at 10 ns pulses the most viable biasing option for short-term development of p-bit circuits. Two recent papers [43,130] have provided experimental demonstration of factorization of a 6-bit integer using a circuit of 1143 p-bits, as shown in Fig. 8(d-e). This demonstration achieved an energy consumption of 0.43 pJ/bit at write pulse widths of 10 ns. It should be noted that their circuit was an MTJ/CMOS hybrid design, where the all of the computation and p-bit functionality was performed on an application-specific integrated circuit (ASIC) and the VCMA MTJs simply provided an entropy source that randomly selected p-bits and whether they should be updated each cycle. Furthermore, a fixed external field was used to set the easy-axis direction, however, this was done using permanent magnets, therefore, the external field did not add to the total energy consumption. The key short-coming to the set-up in ref. [130] was that the VCMA MTJs were not integrated onto the ASIC, which caused significant computation delays and limited the overall throughput. Even if the VCMA MTJs were fabricated onto the ASIC, the throughput would still be maximized to 100 Mbps. Despite these short-comings, the work done by Y. Shao et al [130] and C. Duffee et al [43] are very impressive achievements and both are a large steps towards prototyping of probabilistic circuits.

D. Voltage-Controlled Exchange Coupling p-Bits

Each of the biasing mechanisms discussed at this point have certain short-comings. STT and SOT effects rely on generating spin current from current flow, which inherently leads to energy dissipation through Joule heating. To achieve more energy-efficient magnetization control in MTJs, voltage-controlled approaches are highly desirable. The widely studied VCMA effect modulates the free layer's energy barrier but does not directly influence the output probability. In contrast, the voltage-controlled exchange coupling (VCEC) effect enables bipolar control, similar to STT, where the MTJ preferentially switches to either the AP or P state depending on the applied voltage polarity. Leveraging this characteristic, the VCEC effect offers a promising alternative to STT for enabling energy-efficient MTJ p-bit control. Q. Jia et al [44], fabricated MTJs with synthetic anti-ferromagnetic (SAF) free layers with perpendicular magnetic anisotropy, which are needed to facilitate VCEC switching, to study the potential for VCEC driven p-bits. At room temperature, thermal agitation induces spontaneous switching of the free layer between the AP (antiparallel) and P (parallel) states, as illustrated in Figs. 9(a-b), allowing the MTJ to function as a p-bit. The time-averaged resistance state (measured as the AP-rate in Fig. 9(c)) can be effectively controlled through an applied voltage via the VCEC effect, as shown in Fig. 9(c), achieving low-

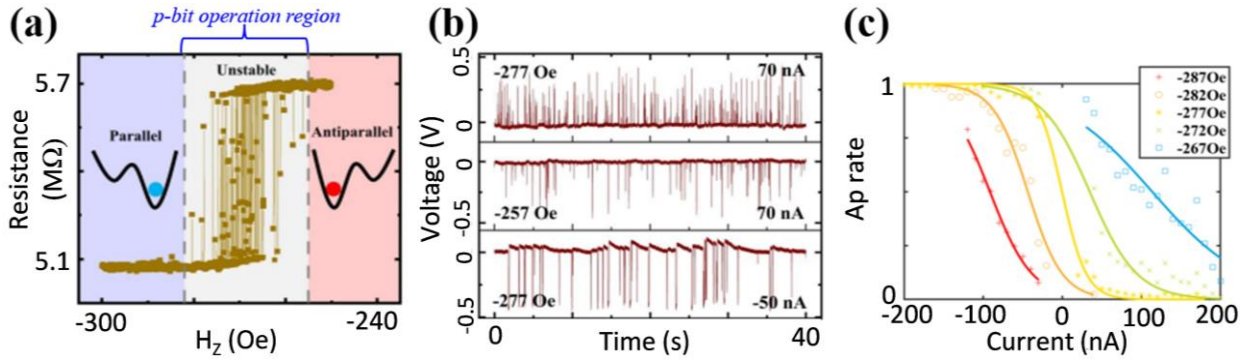

Fig. 9. (a) Resistance hysteresis plot of an MTJ with a SAF free layer as a function of the external magnetic field (H_Z) showing a thermally unstable region for potential p-bit operation. **(b)** time-domain plots showing thermal fluctuations between the AP- and P-state resistances under various fixed H_Z and applied currents. **(c)** Impact of current on the time-averaged resistance (AP rate) at various H_Z. Reprinted with permission from [44] Copyright 2025 American Chemical Society.

energy operation with a significantly reduced switching current compared to STT. Note that the current required to operate the VCEC sMTJ in Fig. 9(c) is less than 100 nA whereas the current required to operate the STT driven sMTJ in [41] was over 100 μA. VCEC facilitates AP-to-P switching at an exceptionally low current density, highlighting its potential for future low-power p-bit applications. In scenarios such as neuromorphic computing and probabilistic computing networks, where a large number of MTJs operate simultaneously, the significantly reduced power consumption per MTJ translates into substantial energy savings, making VCEC a highly promising approach for next-generation spintronic computing architectures.

It should be noted that at this time, ref. [44] is the first and only study that directly studies the VCEC mechanism for probabilistic bits. While this study clearly shows that VCEC is a viable biasing mechanism for p-bits, the lack of additional research means that there is still a lot of uncertainty regarding the overall performance of VCEC driven p-bits. Furthermore, the mechanism that causes a VCEC switchable MTJ to be thermally unstable is still unclear since thermal stability can be modified through free layer thickness as well as the non-magnetic spacer thickness. This leads to uncertainties in key performance metrics in p-bit devices such as maximum achievable switching speed and susceptibility to device variations. It should be noted that ref [136] demonstrated sub-nanosecond VCEC switching. However, more research is needed on p-bits biased with VCEC to determine if this translates to ultrafast switching in p-bits. However, the data that is currently available shows that VCEC driven p-bits can operate at currents less than 100 nA and on MTJs with MgO thicknesses of 2 nm [44], thus relaxing the requirement for ultra-thin (<1 nm) MgO thickness that is required for STT switching. These properties warrant further investigation into VCEC p-bits as they have potential for p-bits with energy consumptions below 1 fJ/bit.

IV. Dual-Biasing Mechanisms for p-Bits

Each of the tuning mechanisms presented in section 3 have their advantages, however, each also have their challenges. The key challenge for all of the single-biased p-bits is their susceptibility to defects and device variations in the MTJs themselves caused during the fabrication process. These defects cause drastic variations in switching rate and input response between MTJs. Previous studies have shown that the switching rate of an sMTJ influences the overall resistance distribution of the signals generated, as shown in Fig. 10. SMTJs with slower switching rates tend to follow a bimodal resistance distribution. As the switching rate increases, this resistance distribution becomes more parabolic, as shown in Fig. 10(b), and eventually becomes more continuous. Probabilistic circuits that contain devices with both distributions exhibit unwanted plateaus in the probabilistic output, which lead to errors [51], which will be discussed in further detail in section 5(b). Furthermore, the performance of Boolean satisfiability problems have been shown to improve drastically when using sMTJs with parabolic or continuous resistance distributions when compared to sMTJs with bimodal distributions [39], thus demonstrating the importance of these resistance distributions in the signals generated in each p-bit. To tackle unwanted plateaus, recent work by Lv, Yang et.al shows that introducing a moderate VCMA coefficient (~200 fJ V^{-1} m^{-1}) sharpens the switching probability transfer curve of an MTJ, turning the broad, plateau-prone profile into a clean sigmoid. This single knob cut logic-error rates

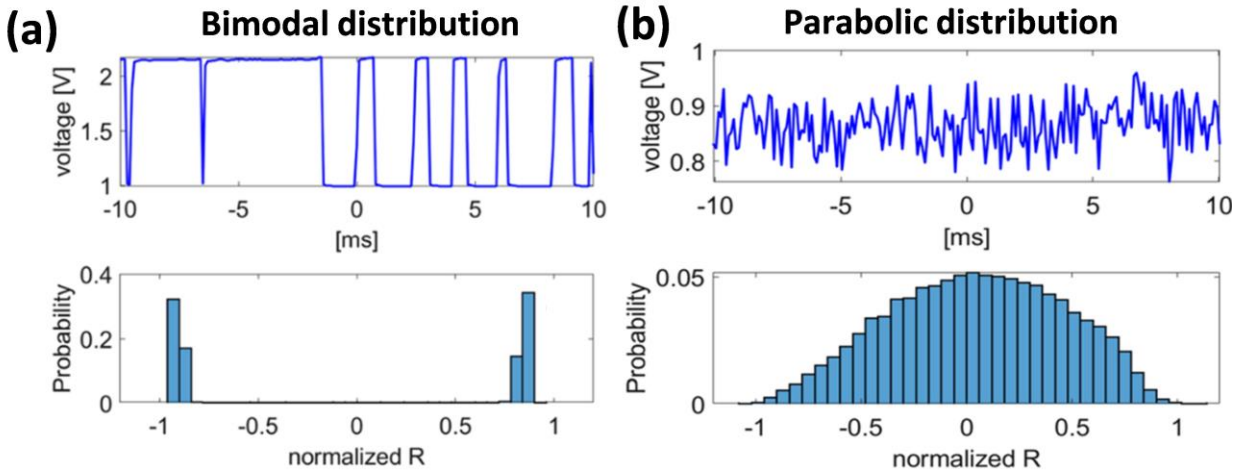

Fig. 10. Illustration of changes in resistance distribution (represented as a normalized signal) as a function of switching rate. **(a)** Illustration of an sMTJ with ≈1 ms dwell times, resulting in a resistance histogram showing bimodal distribution. **(b)** sMTJ with faster (≈10-100 μs) dwell times, resulting in a resistance histogram showing a parabolic distribution. These plots illustrate the potential negative effects of variations in switching rate in p-bit circuits. Plots were reproduced from ref. [39].

by ~61 %, without altering device geometry. Applying the same VCMA overlay to sMTJs used as p-bits will help to compress the bimodal resistance spread into a tighter distribution and substantially reduce plateau-driven probabilistic errors [137].

Other ways to solve these shortcomings is by employing two tuning mechanisms on a single p-bit, which we will refer to as dual-biasing. One benefit is that certain combinations of tuning mechanisms can actually improve the overall energy efficiency and endurance of the p-bit. Another benefit is that dual-biased p-bits have the potential to mitigate the negative effects of device-to-device variations. One way these effects are mitigated is that employing two biases can reduce the effective thermal stability in MTJs, thus generating random telegraphic switching signals in thermally stable MTJs. This relaxes the criteria for the MTJ to have a near 0 k_BT energy barrier, therefore, MTJs with a broader range of thermal stability factors can be employed in p-bit circuits. Furthermore, some dual-biasing mechanisms can modify the average switching rate of the stochastic MTJs by nearly three orders of magnitude. By utilizing a secondary bias in p-bit circuits, the average switching rate for all p-bits can be reduced drastically, thus solving any potential issues of different resistance distributions between p-bits.

Dual-biased p-bits also have additional potential benefits. For example, having an additional biasing term also allows for more flexibility in how an optimization problem is mapped onto the p-bit circuit. For example, one bias term can represent the sum of the weighted outputs while the other behaves as the {h} bias. Lastly, some mechanisms for dual-biased p-bits have a unique capability called two degrees of tunability, where the two biasing mechanisms can control the AP-state dwell times and P-state dwell times separately. The implications of this novel capability means that these dual-biased p-bits have two times more information capacity than single-biased p-bits.

In this section, we will discuss the advantages of various dual-biasing strategies. It should be noted that we will only be discussing strategies where the two tuning mechanisms can be controlled independently. For example, STT and SOT have separate current paths that do not rely on each other. Therefore, STT and SOT currents can be controlled separately to optimize p-bit performance and is a valid dual-biasing method. On the other hand, any combination of STT, VCMA, or VCEC will not be considered as a valid dual-biasing strategy. Even though the physical mechanisms that cause magnetization switching are different between STT, VCMA, and VCEC, there is no way to isolate these terms from one another since they all share the same signal path. The dual-biasing strategies that will be discussed in this section are STT + external magnetic field, STT + magneto-electric coupling, STT + SOT, and VCEC + SOT.

A. STT + External Field Driven p-Bits

In this section, we will investigate the performance of p-bits driven by the combination of STT and an

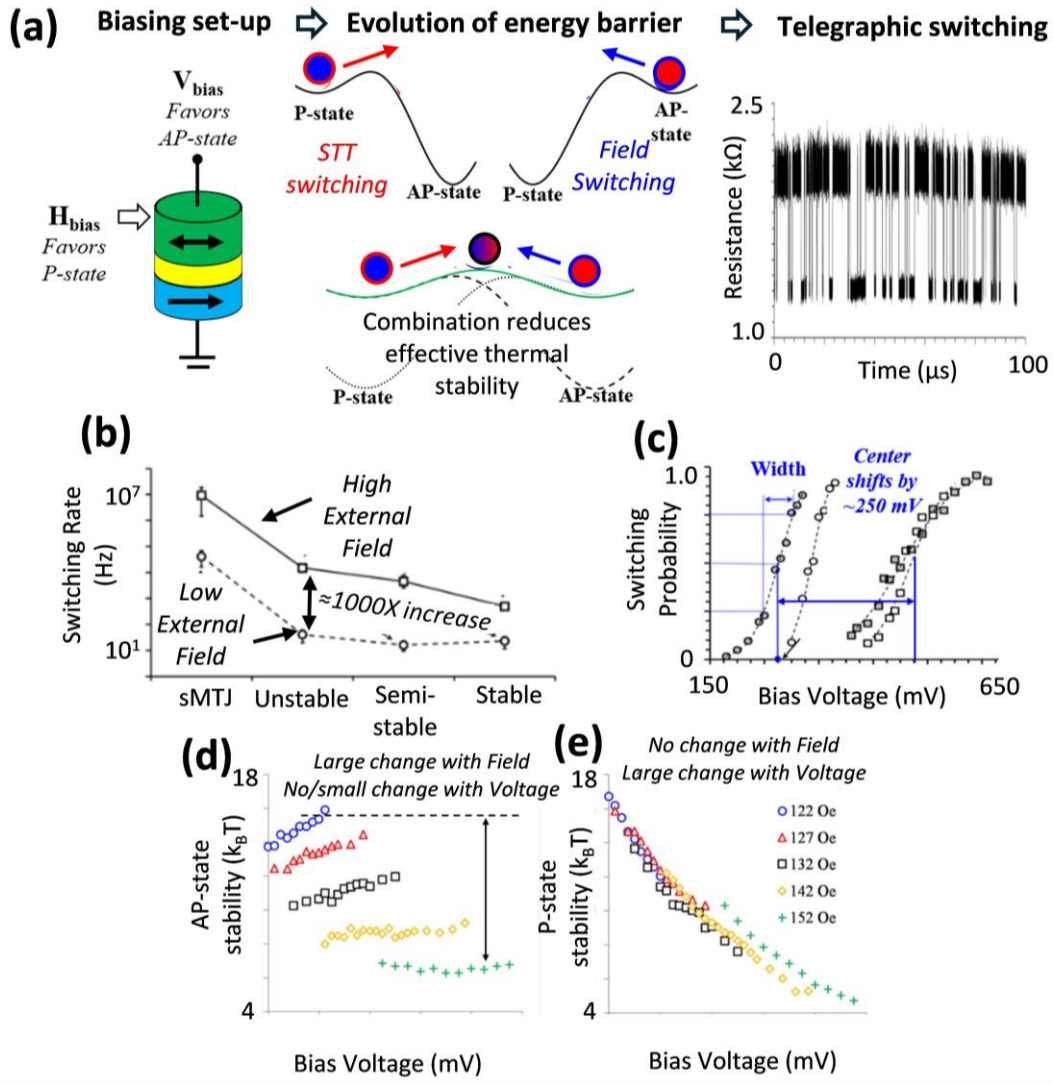

Fig. 11. (a) Illustration of the orientation of the voltage bias and field bias for STT + field dual-biasing. The combination of these two biasing terms causes a time-evolution of the energy barrier, which reduces the effective anisotropy energy and results in random telegraphic switching (time-domain plot reproduced from ref. [141]). **(b-c)** Plots showing the dependence on peak frequency of signals generated and center and width of transfer characteristics with respect to H_{Bias} (reproduced from ref. [141], reprinted with permission from IEEE). These plots illustrated that the average switching rate can be increased by 3 orders of magnitude and the center of the transfer plot can be shifted by 250 mV by increasing H_{Bias}. **(d-e)** Plots showing the dependence of the AP-state and P-state dwell times with respect to **(d)** V_{Bias} and **(e)** H_{Bias} (reproduced from ref. [140], reprinted with permission from IEEE). These plots illustrate the unique capability that is achieved through STT + field dual-biasing which is two-degrees of tunability, where the P-state dwell times are significantly influenced by V_{Bias} but not H_{Bias} whereas the AP-state dwell times are influenced by H_{Bias} but not V_{Bias}. Having this separation in the control between AP- and P-state dwell times means that the signals generated via STT + field dual-biasing have double the information capacity as those generated by single biasing methods.

external field. It should be noted that an external field is used in some of the studies mentioned in section 3(a) [41,104]. However, in these experiments, the external field is fixed and is simply used to compensate for the stray field induced by the fixed layer of the MTJ. In the experiments discussed in this section, the field serves as an additional input term or bias term and needs to be tunable. There are two key disadvantages to this method that should be mentioned. One is that the cell size in a circuit with a tunable magnetic field at each bit will have a lower limit of around 90 nm so that the applied field does not

influence neighboring cells [64]. The second is that inducing an external magnetic field will drastically increase the energy consumption at each cell during write operation. Note that the energy consumption from an external magnetic field is only a factor when the field is tunable but not when it is fixed. This is because a fixed magnetic field can be induced by including a hard magnetic layer in the MTJ stack and therefore, does not consume any energy [138]. Despite these challenges, p-bits driven by STT and an external field have three unique advantages over single biased p-bits, each of which will be discussed in this section.

As previously mentioned, one of the biggest challenges in the implementation of p-bits driven only by STT and SOT in large-scale circuits is their susceptibility to device variations. There are two ways in which these variations negatively impact the performance of p-bit circuits. The first is that random defects in p-bits cause them to fluctuate at different speeds which results in different resistance distribution (recall Fig. 10). . The second way in which variations negatively impact probabilistic circuits is that they can change the position and width of the transfer curves among each p-bit. These variations in the voltage-dependence of the average output means that the circuit as a whole is unpredictable. A straight-forward solution to both of these effects is to introduce an additional tuning mechanism, an external field in this case, to compensate for these effects. Previous studies have shown that the average switching rate of sMTJs can be increased over 2-3 orders by proper tuning of the magnetic field [139–141] and also the position of the transfer curve can be adjusted by 250 mV [141], as shown in Fig. 11(b) and 11(c), respectively. This flexibility of the output response with voltage means that dual-biased p-bits can be tuned to compensate for device variations with on-board corrections.

Another benefit of STT and external field biasing is that random telegraphic switching signals can be generated in MTJs that are intrinsically thermally stable. This means that p-bits driven by STT + field are compatible with state-of-the-art MRAM, and thus, less susceptible to device variations. To induce random telegraphic switching signals in thermally stable MTJs, a constant voltage needs to be applied to favor AP-state switching and the external field is applied to favor P-state switching. This specific orientation of these biases causes an evolution of the energy barrier, as illustrated in Fig. 11(a). When the MTJ is in the P-state, the current through the MTJ is larger than is the AP-state due to the change in resistance. Therefore, STT will cause P-to-AP switching. When the MTJ is in the AP-state, the energy barrier will shift to favor P-state orientation due to the decrease in STT. Therefore, the field will cause AP-to-P switching. This constant toggling between P and AP switching results in a reduction in the effective energy barrier, thus inducing random telegraphic switching signals.

An important point to make is that the two benefits mentioned thus far, correct for device variations and compatibility with state-of-the-art MRAM, are not unique to STT + field dual-biasing. These can be accomplished via other dual-biasing methods, which will be discussed in later sections. However, there is a third benefit to STT + field dual-biasing called two-degrees of tunability, which at this point, has not been experimentally demonstrated by any other dual-biasing method. Two degrees of tunability is observed when comparing the dependence of the P- and AP-state dwell times with the bias voltage and the external field, as shown in Fig. 11(d-e). These plots show that P-state dwell times can be reduced by over 3 orders of magnitude with respect to the bias voltage, whereas the AP-state dwell increase, but by less than one order of magnitude. The opposite trend is observed when examining the dependence on external field. Figure 11(d) show that the AP-state dwell times can decrease by around 3 orders of magnitude with respect to field whereas the P-state dwell times are not influenced significantly. The two-degrees of tunability capability has been observed when applying the STT + field dual-biasing method on both thermally stable MTJs and superparamagnetic MTJs [141].

This two degrees of tunability capability means that the signals generated by the STT + field dual-biasing method have double information capacity compared to single-biased p-bits. This is because encoding schemes for random telegraphic switching signals are typically based on either the average output or the number of pulses generated. However, this separation of control in AP- and P-state dwell times with the external field and voltage means that the average output and the number of pulses can also be controlled separately, as illustrated in Fig. 11(d-e). Furthermore, independent control over AP- and P-state dwell times offers the capability for stochastic signals to be encoded temporally [142]. Therefore, a dual-biased p-bit benefits from the high information capacity of temporal encoding along with the noise resilience of rate encoding on stochastic signals.

The flexibility in tunability that STT + Field dual-biasing offers has been a great asset in recent small-

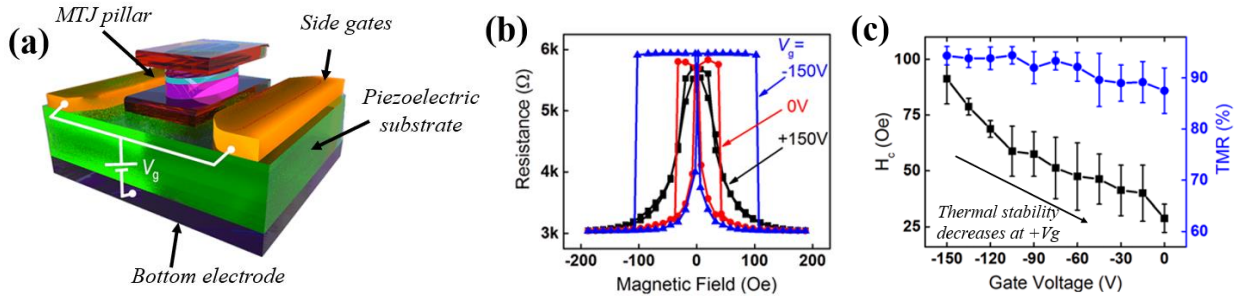

Fig. 12. (a) Schematic of an MTJ fabricated on a piezoelectric substrate. Strain is induced by applying a gate voltage at the side gates. **(b)** field switching hysteresis measurements obtained at various gate voltages demonstrating the influence on coercivity of the MTJ. **(c)** Plot of coercivity vs gate voltage, illustrating that voltage-induced strain is a valid method to modify the thermal stability of an MTJ. Reprinted from [145] with the permission of AIP publishing.

scale experimental demonstrations. In 2019, the first experimental demonstration of the invertible AND gate using probabilistic spin logic was carried out and made possible by 3 MTJs driven by STT + Field dual-biasing [143]. In 2022, STT + Field dual-biasing was applied to 2 MTJs connected in series, which created a bipolar random number generator [144]. In this method, the two-degrees of tunability feature of STT + field dual-biasing made it possible to fix the output probability at 50% over a large range of bias voltages. However, as mentioned earlier in this section, the large energy consumption and lower limit on cell size makes large-scale implementations of STT + field biased p-bits impractical. Therefore, for the remainder of this section, we will investigate other dual-biasing methods and determine if these other schemes are more feasible for large-scale implementation in terms of energy consumption while still achieving two degrees of tunability.

B. STT + Magneto-electric Coupling Driven p-Bits

As mentioned in the previous section, STT + field is not a viable dual-biasing method for p-bit circuits since a tunable magnetic field applied at each cell requires large amounts of energy and sets a lower limit on cell size. An alternative solution will likely involve using a second bias term that induces an effective local magnetic field on the MTJ rather than an external magnetic field. One way to accomplish this is through magneto-electric (ME) coupling.. One way to generate ME coupling on an MTJ is by fabricating an MTJ on a piezoelectric film [145,146], and applying an electric field parallel to the poling direction of the piezo material, as illustrated in Fig. 12(a). The strain induced by this electric field will generate a transverse magnetic field in the MTJ via the magneto-electric effect, thus modifying its magnetic properties. This effective field will either increase or decrease the shape anisotropy of the MTJ, depending on the polarity of the electric field, as shown in Fig. 12(b-c). This strategy has been proposed as a solution to reduce the write energy in STT-MRAM [145]. However, this modification of shape anisotropy means that voltage-induced strain may also be used as a biasing mechanism for p-bits, since many p-bit designs are dependent on low-E_B MTJs. In this section, we will review previous studies investigating ME coupling drivenp-bits and determine if the combination of STT+ME coupling is a viable dual-biasing method.

Reference [147] was one of the first studies that indicated that ME coupling can be utilized as a biasing mechanism in p-bits. In this study, K. Y. Camsari described an equivalent circuit model for magneto-electric random-access memory [148,149], as illustrated in Fig. 13(a), and suggested a mode of operation where states are represented by changes in the easy-axis orientation of the ferromagnetic layer rather than net magnetization. Even though this study was not primarily focused on probabilistic computing, it did show that random signals can be generated in this circuit when the ferromagnetic layer is patterned as a circular nanopillar (no shape anisotropy), which can be tuned with the input voltage, as shown in Fig. 13(b). It should be noted that the configuration described in ref. [147] is not a dual-biasing scheme since the STT and ME biases cannot be separated. However, it does illustrate how the magneto-electric effect does produce the desired output response for p-bit functionality.

A configuration that would allow for STT + ME coupling dual-biasing is illustrated in Figs. 13(c) and 12(d). This schematic allows for an STT signal path (applied across the MTJ) that is separate from voltage-induced strain (which is applied at the electrodes on the piezo film). One way that this configuration can be utilized in probabilistic circuits is through the method

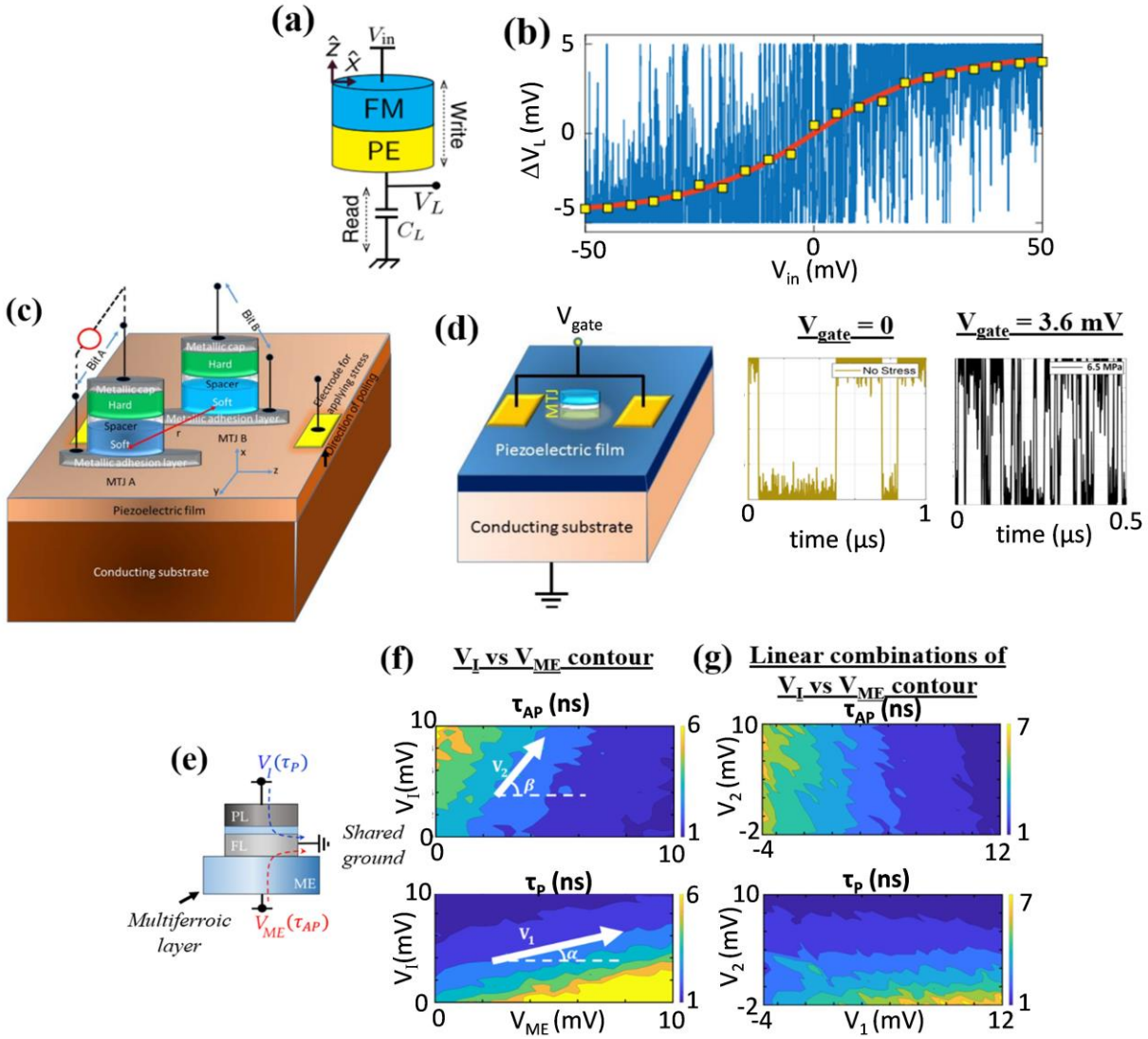

Fig. 13. (a) Equivalent circuit for magneto-electric random-access memory and **(b)** tunable random signals generated when the piezoelectric/ferromagnetic structure is fabricated as a circular nano-pillar with no shape anisotropy (reproduced from ref. [147]). (c-d) Schematic of sMTJs fabricated on a piezoelectric film where **(c)** two sMTJs are fabricated on the substrate and voltage induced strain is used to modify the dipole coupling between the two MTJs (reproduced from ref. [150], https://creativecommons.org/licenses/by/4.0/) and **(d)** a single sMTJ fabricated on the substrate and the gate voltage is used to tune the average switching rate, as shown in the time-domain plots (reproduced from ref. [151], https://creativecommons.org/licenses/by/4.0/). **(e)** Schematic of an MTJ fabricated on a multiferroic material ($BiFeO_3$ in this case) where V_I controls the STT switching of the MTJ and V_{ME} induces an exchange bias field at the multiferroic/MTJ interface (reproduced from ref. [151]). **(f)** Contour plots of the AP- and P-state dwell times (top and bottom, respectively) showing some separation in their tunability. **(g)** Contour plots from **(f)** but modified so that V_1 and V_2 are linear combinations of V_I and V_{ME} showing true separation between the tunability in the AP- and P-state dwell times, thus demonstrating the potential for two-degrees of tunability. Images **(e-g)** were reproduced from ref. [142].

proposed in ref. [150] where two MTJs are fabricated on a single piezo-electric substrate. In the set-up shown in Fig. 12(a), MTJ A provides the input signal and MTJ B provides the output signal. Random telegraphic switching signals are generated in MTJ A via STT, which generates signals in MTJ B through dipole coupling. The strain generated via a gate voltage is used to control the coupling strength between MTJ A and B. In this method, the strain acts as a tunable weight term between two adjacent p-bits.

A more straight-forward approach to strain-based p-bits is to fabricate a single MTJ on a piezoelectric film, as illustrated in Fig. 13(d). The study in ref. [151] shows that using voltage-induced strain in this configuration can modify E_B. As with STT + field dual-biasing, ref. [151] shows that strain + STT dual-basing can serve as a method of solving device-to-device variations in p-bit circuits through control of the switching rates and control of the width of the input-output characteristics. Most importantly, the application of strain in the configuration shown in Fig. 13(d) allows the sMTJ to transition between operating as a binary stochastic neuron to an analog stochastic neuron, meaning that it can be used to fix issues in differences in resistance distribution (recall Fig. 10). Alternatively, the strain voltage can be used as a temperature parameter for simulated annealing for training in optimization problems [151]. The theoretical studies in refs. [151] and [150] show that STT+ME coupling is potentially a more viable dual-biasing method for large-scale p-bit circuits since the additional energy required to induce strain is less than 260 aJ.

A numeric study by K. Yang et al shows that two-degrees of tunability can potentiallybe achieved with STT + ME coupling dual-biasing [142]. This is done by using a similar version of the gating scheme in Fig. 13(c-d), except modified so that the STT voltage and magneto-electric voltage have a shared ground, which is illustrated in Fig. 13(e). Furthermore, there is no strain applied to the MTJ, but rather the MTJ is fabricated on a multiferroic substrate, which induces a localized transverse magnetic field in the presence of an electric field across the substrate. The contour maps of $\boldsymbol{\tau}_{AP}$ and $\boldsymbol{\tau}_{P}$ with STT voltage and magnetoelectric voltage (V_I and V_{ME}, respectively) are shown in Figs. 13(f). These plots illustrate the two-degrees of tunability capability of STT + magnetoelectric coupling dual-biasing under this configuration. It should be noted that changes in $\boldsymbol{\tau}_{AP}$ and $\boldsymbol{\tau}_{P}$ are not completely independent of one another, however, K. Yang et al suggest a simple solution where V_I and V_{ME} are linear combinations of V_1 and V_2, implemented through voltage divider circuits. The resulting contour maps for V_1 and V_2 are shown in Fig. 13(g), which show completely independent control over $\boldsymbol{\tau}_{AP}$ and $\boldsymbol{\tau}_{P}$ with V_1 and V_2.

While the experimental work on STT + ME coupling dual-biased p-bits is lacking, the theoretical work is very promising. Combining the STT and the magneto-electric effects provides an energy-efficient method of correcting for device variations in switching rate and transfer characteristics. Additionally, this method offers other unique capabilities beneficial for p-bits circuits. One of these capabilities is that strain can be used as a method for tuning coupling strength (element in the weight matrix) between p-bits. Another is that two-degrees of tunability can be achieved, thus increasing the information capacity in the signals generated. This capability was also achieved in STT + Field dual-biasing, but STT + ME coupling is much more energy efficient. However, the experimental work on inducing ME coupling via strain show that voltages over 100 V are required to generate enough strain for magnetic switching [152–154]. Even if ≈100 aJ switching is possible, the large write voltages still make strain based MTJs incompatible with CMOS technology. Furthermore, the theoretical studies that predict energies as low as 100 aJ typically assume that the voltage required to generate strain will scale with piezo thickness, however, this prediction does not account for the fact that the piezo constant decreases with piezo thickness for piezo substrates with thicknesses below 1 μm [155]. There are potential strategies to lower the magneto-electric coupling voltage such as using multiferroic materials [156] or inducing strain via a surface acoustic wave (SAW) [157,158] but more experimental work is needed to confirm these as valid strategies. For the remainder of this section, we will investigate dual-biasing methods that involve an SOT current as a secondary biasing term, which is more compatible with CMOS.

C. STT + SOT Driven p-Bits

While the combination of ME coupling and STT offers an energy-efficient method for dual-biasing in p-bits, current experimental work indicates that it is not CMOS compatible due to large ME coupling voltages. Another dual-biasing approach is to combine STT and SOT switching as this can be accomplished using the same hardware as an SOT-MRAM cell. A key benefit of combining SOT and STT mechanisms is that it allows for both binarization and tunability of the p-bit, thus enabling implementation of several Ising Machine paradigms [159]. A. Grimaldi et. al. proposed a design where STT + SOT driven p-bits were used to solve several max-cut problems [159]. At this point, there are very few studies that investigate STT + SOT biased p-bits. However, there are a multitude of studies in recent years that show that combining STT and SOT write mechanisms in MRAM has major benefits. One benefit is that field-free SOT switching in perpendicular MTJs can be

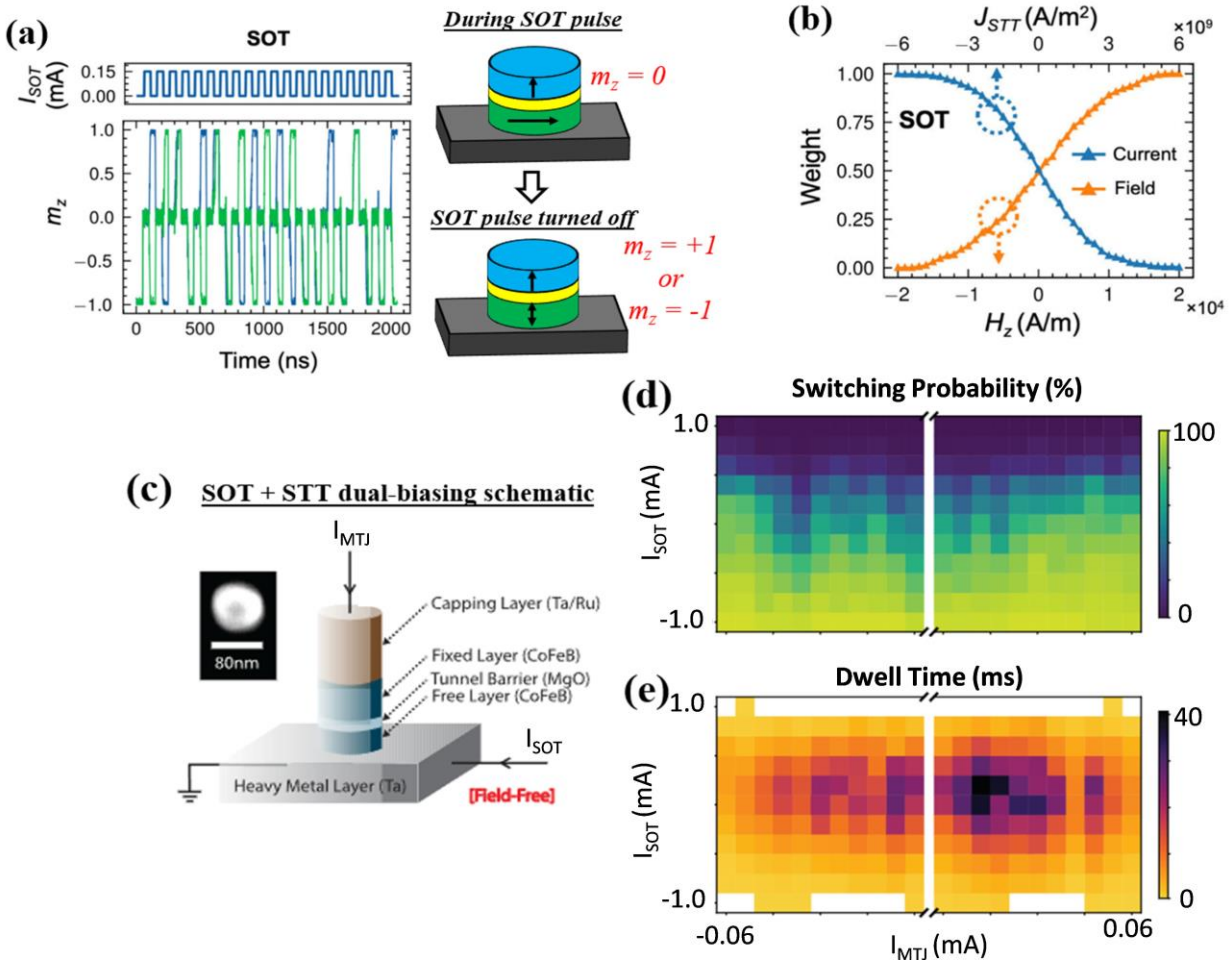

Fig. 14. (a) Illustration of I_{SOT} pulse timing with respect to magnetization dynamics in the $\hat{z}$ direction (m_z). **(b)** Switching probability (represented as 'weight') for the SOT + STT method with respect to STT current density (J_{STT}) at a constant SOT current of 0.15 mA. Plots **(a-b)** were reproduced from ref. [135] (https://creativecommons.org/licenses/by/4.0/). **(c)** Schematic of sMTJ fabricated on a Ta SOT channel and contour plots of **(d)** AP-state probability and **(e)** mean dwell time vs SOT current (I_{SOT}) and STT current (I_{MTJ}) showing dual current controllability. Images **(c-e)** were reprinted with permission from [165] Copyright 2025 American Chemical Society.

accomplished with STT assistance [160]. Other methods of achieving field-free switching in p-MTJs include using to antiferromagnetic/ferromagnetic interface to induce an exchange bias field [161], introducing lateral asymmetry in the SOT channel [162,163], and fabricating bi-layered SOT channels with competing spins [164]. However, these methods all require large write current densities whereas STT assisted SOT switching is the only method that actually reduces the overall write energy [45–47]. Another benefit is that sub-nanosecond write speed can be accomplished. In a recent study, E. Grimaldi et al demonstrated a write speed of 210 ps using STT-assisted SOT switching [48].

Due to the limited research available, it is not clear if the benefits of lower write energy and faster switching speeds observed in STT-assisted SOT-MRAM will extend to STT + SOT biased p-bits. However, there are two studies that we will investigate in this section that will provide some helpful insight into the potential for STT + SOT dual-biased p-bits. The first is by S. Liu et. al. [135], which uses an analytical model to compare STT, STT + SOT, and VCMA switching mechanisms for random bit generation. In this study, the stochastic signals are not represented in random asynchronous telegraphic signals generated from thermally unstable MTJs, but rather through a series of synchronous perturb and read pulses applied to thermally stable, perpendicular

MTJs. Methods for generating random bits via SOT + STT switching are illustrated in Fig. 14(a). For STT switching, the MTJ needs to be initialized to the P-state each cycle, therefore, a reset pulse is applied before each perturb pulse. For STT + SOT switching, the SOT pulse is applied to set the magnetization orientation of the MTJ along the in-plane direction, then the magnetization will jump to one of the two out-of-plane orientations once the SOT current is removed. In the absence of an STT-current, the probability of being in the AP-state is 50%, however, an STT current is applied in conjunction with the SOT current to tune the probability between 0 to 100%, as illustrated in Fig. 14(b). In this method, the SOT current generates entropy in the system and the STT current sets the weight. The results in ref. [135] show that the STT + SOT method consumes 0.1 pJ per bit whereas the STT method consumes 1 pJ per bit. However, this comparison may not be valid for this review since we are interested in thermally unstable p-bits that fluctuate asynchronously, which do not require a reset operation for STT switching. For a more valid comparison, we will evaluate the energy consumption of STT switching without the reset operation. Based on the pulse amplitude pulse widths provided in [135], the reset operation in the STT method accounted for approximately 62% of the total energy consumption, meaning that the energy consumption for the SOT + STT switching method was still over 3X lower. This indicates that SOT + STT driven p-bits are potentially more energy efficient than those driven by STT or SOT alone.

While the results in ref. [135] shows potential for energy reduction in SOT + STT driven p-bits, it does not provide any insight into how the STT and SOT biases influence dwell times. More importantly, is STT + SOT dual-biasing capable of two degrees of tunability? The second study that will be reviewed in this section is by D. Koh et al [165], which is an experimental study where a circular, in-plane MTJ has fabricated on an SOT channel in the set-up shown in Fig. 14(c). This is the first experimental study of p-bits that investigates the interplay between STT and SOT biases. The surface plots in Figs. 14(d) and 1(e) show the influence of the SOT and STT currents on the AP-state probability and average dwell time, respectively. These plots show that the AP-state probability was mostly controlled by the SOT current whereas the average dwell time was controlled by both the SOT and STT currents. In this configuration, the SOT current is used to tune the output probability and the STT current could tune the fluctuation rate. In their experiment, the MTJ tested has a high resistance-area product, which explains the weak dependence of the output probability on the STT current. It should be noted that the changes in overall switching rate is much smaller with SOT+STT dual-biasing than the changes observed for STT+field or STT+ME coupling. One key difference is that a voltage source rather than a current source was used to provide a bias for the STT term in STT+field and STT+ME coupling. Using a constant voltage source for the STT bias means that the strength of STT was state dependent (see ref. [140] for more details). Since ref. [165] used a current source for the STT term, it is unclear if two degrees of tunability is achieved with SOT + STT dual-biasing. However, the results in ref. [165] does successfully demonstrate that dual-control of the SOT and STT biases can be used to address the issue of device variation in sMTJs.

D. VCEC + SOT Driven p-Bits

Similar to STT, the VCEC effect can be combined with SOT to achieve dual-bias control of a p-bit. This was demonstrated in a VCEC-based sMTJ [44], where the underlying conductive layer consists of SOT materials with a large spin Hall angle: Ta (5.0 nm) / Pd (2.0 nm), as shown in Fig. 15(a). Since Ta and Pd are expected to have opposite spin Hall angles, their combination is designed to enable field-free switching of a magnetic layer with perpendicular magnetic anisotropy (PMA). The bilayered spin Hall channel induces an effective magnetic field along the easy-axis of the MTJ's SAF free layer, directly influencing its magnetization, further enhancing the controllability of p-bit operation.

A deeper analysis into the interplay between the VCEC and SOT effects show that the effective magnetic field induced by SOT is primarily dominated by thermal effects, as shown in Fig. 15(b) and also confirmed in ref. [49]. The resulting temperature increase causes a shift in the exchange coupling field, which is independent of the current direction and proportional to the square of the current. Figure 15(b) also shows that the SOT current induces effective field along the z-axis due to the presence of the bilayered spin Hall channel [49], which changes sign with the reversal of the current direction, however, this effective field has much less influence on the thermal stability than Joule heating. Notably, SOT and VCEC can independently manipulate the time-averaged output (or AP rate) without interference, as shown in Fig. 15(c). Since VCEC itself does not produce a

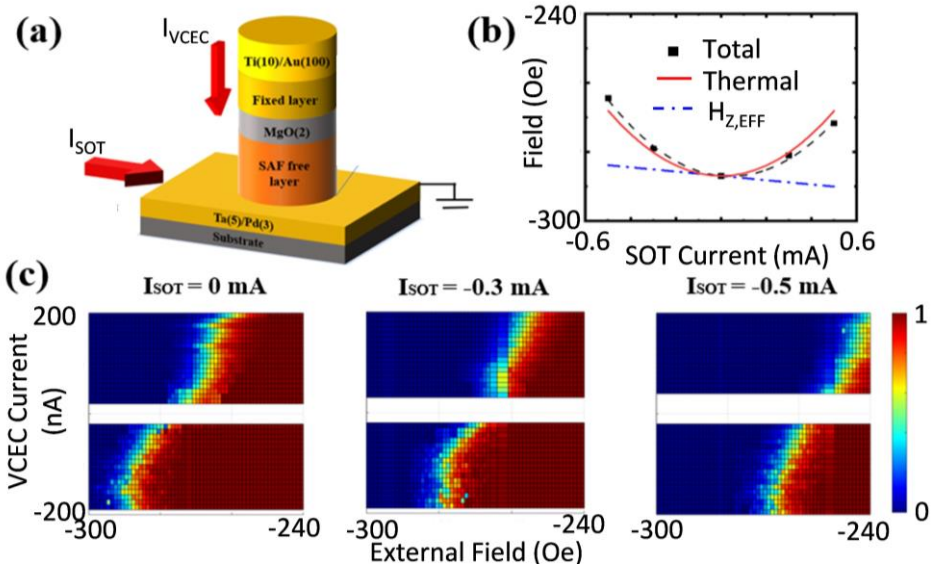

Fig. 15. Impact of SOT Current on the AP Rate. **(a)** Illustration of the VCEC structure and the testing circuit. **(b)** Effective field induced by the SOT current. **(c)** Phase diagram of the time-averaged resistance (or AP rate) at SOT currents of 0, -0.3 and -0.5 mA, illustrating the AP rate as a function of magnetic field and applied current. Each data point is extracted from a 40-second waveform. Reprinted with permission from [44] Copyright 2025 American Chemical Society.

noticeable temperature change, the SOT-induced effects remain unaffected even when VCEC is simultaneously applied. Conversely, the strength of VCEC remains nearly constant with temperature, further confirming the independent nature of these two effects. However, as with STT + SOT dual-biasing, it is unclear if two-degrees of tunability is achieved with VCEC + SOT dual-biasing. Using a voltage source for the VCEC bias rather than a current source would be one way to generate stochastic switching signals with two-degrees of tunability.

As stated previously, there is very little work done on VCEC driven p-bits. Since the effective field induced by the SOT current on the SAF free layer is primarily dominated by thermal effects, which reduces the free layer anisotropy, combining VCEC with SOT is one possible method of improving the switching speed in VCEC p-bits. However, one potential shortcoming of this strategy is it can increase the overall energy consumption of VCEC driven p-bits. In most cases, the write energy in SOT MTJs can be reduced by engineering the spin Hall channel to achieve maximum charge-to-spin conversion, thus reducing the current density needed for magnetization switching. However, if thermal effects are the dominating force, then improving charge-to-spin conversion will not change the effective field induced by the SOT current significantly and large current densities are still required. One possible solution in minimizing energy consumption in SOT + VCEC driven p-bits is to use a topological insulator (such as BiSe [68]) to increase the resistance of the SOT channel.

E. Comparison Between Tuning/Biasing Schemes

An overview of the key experimental and theoretical studies for each p-bit biasing mechanism discussed is shown in Table I. For the analysis of STT biased p-bits, we categorized them into 3 types: perpendicular magnetic anisotropy (PMA), in-plane anisotropy (IMA), and double free layer (DFL). It should be noted that a DFL MTJ can have perpendicular or in-plane anisotropy, however, for simplicity, we will assume that DFL MTJs have in-plane anisotropy since a majority of the studies on DFL MTJs are in-plane. Furthermore, the theoretical studies listed under IMA and PMA MTJs are also applicable to DFL MTJs In fact, some of these studies may be more applicable to DFL MTJs because they use an sMTJ connected to a transistor as a p-bit and assume that the switching probability of the sMTJ does not change as the gate voltage of the transistor varies. The only way to achieve this behavior is using a DFL MTJ similar to the ones described in refs. [99–101] where the output probability is constant with bias voltage. Otherwise, the output probability of the sMTJ will be modified via STT rather than through electrical modulation. SOT-

Table I. Overview of key experimental and theoretical studies for each tuning/biasing mechanism.

Biasing mechanism		Experimental work	Theory/simulation work
STT	**PMA**	Device-level [102] Electrical coupling [110] Integer factorization [93,182] Hardware-aware machine learning [14]	Correlation time prediction [90]
	IMA	Device-level verification [41] Nanosecond operation [103–105] Random number generation [105,106] Coupling of sMTJs [107–109]	Correlation time prediction [90] Bayesian Network [11,20] Invertible Logic [15,20,40] Integer factorization [15] Boolean satisfiability [15] Accelerated Machine Learning [20] Combinatorial Optimization [20]
	DFL	Device-level experiments IMA based DFL [100] SAF based DFL [101]	Concept prediction [99]
SOT	**Type I**	Device-level [119] Random number generation [42,118] Frequency synchronization [117] Integer factorization using time division multiplexing [120]	Performance comparison [90] Invertible Logic [92] Bayesian Network [13] Boltzmann machine [13] Deep Belief Networks [91] Invertible Boolean Logic [92] Combinatorial Optimization [17]
	Type II and III	Device-level [97,124] Random number generator [128] Integer factorization [126] Gibbs sampling for restricted Boltzmann machine [127]	Device-level analysis [97]
VCMA	**Short Pulse Method**	Random number generator [134]	Boltzmann machine [131,132] Bayesian Network [133]
	Long pulse Method	Integer factorization [43,130] **These studies used a spintronic/CMOS hybrid design.*	Random bit-stream generator [82,135]
VCEC		Device-level [44]	None
STT+ Field		Device-level [139–141] Bipolar random number generator [144] Invertible AND gate [143]	None
STT+ME coupling		None	Proposed ME for p-bits [147] Tune switching rate in sMTJs [151] Tune coupling between sMTJs [150] Predict two-degrees of tunability with STT+ME coupling [142]
STT+SOT		Device-level and random number generator [165]	Random bit-stream generator [135] Combinatorial Optimization [159]
VCEC+SOT		Device-level [44]	None

driven p-bits are divided into two categories: Type I and Types II and III. Types II and III are placed in the same category since there are similarities in how these types operate as well as their performance metrics.

Lastly, VCMA p-bits are separated into two different methods (recall section 3(c)) and the remaining biasing methods are placed in their own category.

There are several notable trends presented in Table I. One is that STT and SOT driven p-bits have received significantly more attention than any other biasing mechanism. On the other hand, only one paper has studied VCEC driven p-bits. This is not surprising since STT and SOT are also the most common write mechanisms for MRAM over the last decade, whereas VCEC is a relatively newly discovered write mechanism. For STT driven p-bits, there exists a good balance of experimental and theory work, however, the large-scale experimental demonstrations have been done using PMA sMTJs, whereas the experimental work on IMA and DFL MTJs are currently limited to small-scale (1-2 MTJs) experiments. Conversely, most of the theoretical work on STT p-bits use IMA MTJs for their study. SOT and VCMA driven p-bits follow similar trends. Most of the theory work on SOT p-bits use Type I devices whereas the largest scale experiments are done using Type II and III devices. Note that integer factorization has been done on type I SOT devices [120], however this was done using a single SOT MTJ along with additional circuitry and algorithm realizing time division multiplexing. Similarly, most of the theory work on VCMA p-bits use the short pulse method whereas the experimental work on VCMA p-bits use the long pulse method. Significantly less work has been conducted on dual-biased p-bits, but the experimental work tends to focus on STT + field dual-biasing whereas a large portion of the theory work is focused on STT + magnetoelectric (ME) coupling driven p-bits.

The trends observed in Table I where some biasing schemes receive more attention in theoretical studies than experimental studies or vice versa can be explained by the performance metrics of each p-bit biasing mechanism, which are shown in Tables II and III. Table II shows the switching speed and energy consumption (per random bit) for both experimentally determined and theoretically predicted values. Table III provides a qualitative assessment of the susceptibility for device variations for each biasing mechanism. These tables divided the biasing mechanisms into the same groups as in table I. The data presented in these tables for STT driven MTJs explain why PMA sMTJs are often used in experiments whereas IMA MTJs are used in theoretical studies. P-bits that use IMA sMTJs have much lower correlation times (as predicted in [90]) than p-bits using PMA sMTJs. Both PMA and IMA sMTJs are very susceptible to variations in switching speed, however, the variations in the sigmoid curve are much smaller for PMA sMTJs than for IMA sMTJs. The data in Tables II and III show that DFL sMTJs are the most promising option for p-bit implementation for STT-driven p-bits due to their potential for high switching speed and low energy consumption. The reason that the energy consumption is much lower for DFL sMTJs than IMA sMTJs is due to the reduction in current needed for operation. For IMA sMTJs, the switching probability is tuned via STT whereas the p-bit output for DFL sMTJs is tuned by the signal offset relative to the threshold of the inverter caused by the transistor, details of which are described in section 5(a) (also see ref. [40]). This means that the current change required in IMA sMTJs to produce a swing of switching probability from 0 % to 100 % is 10X higher than for DFL sMTJs. Furthermore, unlike IMA sMTJs, DFL sMTJs are not susceptible to variations in the sigmoid characteristics since it has been observed that their switching probability is fixed at 50 % regardless of bias voltage.

SOT driven p-bits using type I SOT MTJs (thermally unstable) are susceptible to the same variations in switching speed as STT driven p-bits. However, type I SOT p-bits have a few advantages over STT driven p-bits in terms of their susceptibility to variations. One is that the center of the sigmoid curves in type I SOT sMTJs only vary by around 10 μA [120], meaning there are smaller variations in sigmoid curves in type I SOT p-bits compared to IMA sMTJs. The other is that the coupling strength in SOT oscillators is much stronger. In ref. [119], SOT driven stochastic oscillators would still synchronize to signals one order of magnitude higher than the natural frequency. This means that SOT driven p-bits could vary by one order of magnitude and still have strong correlations. Table II shows that SOT driven p-bits (type I) also have the potential to achieve GHz switching rates and energy consumption on the order of 1 fJ/bit, however, these predictions are significantly different than what is achieved experimentally. For the performance of type I SOT p-bits to reach their full theoretical potential, the thermal stability of the sMTJ will need to decrease to near 0 k_BT and the switching current density will need to reduce below 10^6 A/cm^2.

Table III shows that SOT driven p-bits with type II and III methods are the least susceptible to variations. This is because these p-bits operate in a synchronous manner at a speed determined by the clocking cycle of the write-read operations. Therefore, variations in switching speeds is not a significant issue. Furthermore, the SOT current just needs to be strong enough to set the magnetization of the MTJ in the in-plane direction so variations in write current are also not a factor. While this robustness in variations has made these types of SOT driven p-bits an attractive method for carrying out high-level experimental demonstrations [126,127], the data in table II suggest they have limitations in terms of their energy

Table II. Comparison of performance metrics and overview of circuit-level demonstration for each tuning/biasing mechanism.

Biasing mechanism		Switching speed (Hz)		Energy per bit (fJ/bit)	
		Experimental	Predicted	Experiment	Predicted
STT	PMA	1.5 x 10^5 [110]	10^7 [90]	> 300 [110]	2 [93]
	IMA	5 x 10^8 [103]	10^9 [90]	>200 [103,105]	>100 *Assuming 10^9 Hz switching rate.
	DFL	5 x 10^7 [100]	3 x 10^9 [101]	>200 [100]	3.6 [101]
SOT	Type I	2.5 x 10^5 [120]	1 x 10^{10} [90]	>7000 [120]	3 [90]
	Type II and III	20 x 10^6 (maximum) [127,128]		>100,000 [116,127,128]	
VCMA	Short Pulse Method	2 x 10^5 [134]	166 x 10^6 [132]	1.1 [134]	1 [131]
	Long Pulse Method	100 x 10^6 (maximum) [43]		311 [82]	
VCEC		10^2 [44]	Unknown	>500,000 [44] *40 nW power consumption	< 0.1 *If sub 10 ns switching is assumed
STT+ Field		>10^7 (tunable over 3 orders of magnitude) [139–141]	STT switching rate possible (10^9)	>10,000 *Assumes field generated by Cu wire and STT voltages used in [141]	
STT+ME coupling		Unknown	10^3 to 10^6 (tunable) [151] *STT switching rate possible*	Unknown	STT energy + 0.26 [150,151]
STT+SOT		10^2 to 10^3 [165] (tunability unknown)	SOT switching rate possible	8000 [165]	3X lower than STT p-bits [135]
VCEC+SOT		10^2 [44](tunability unknown)	Unknown.	> 500 x 10^6 [44]	> 100 *Assumes SOT current densities larger than 10^6 A/cm^2

consumption and speed. Since they operate on synchronous write and read pulses, their speed is typically limited to 20 MHz (refs. [127,128] used 50 ns pulses) and their energy consumption is likely to exceed 100 pJ/bit due to the large current densities needed for hard-axis initialization (≈10^8 A/cm^2).

Similar trade-offs between performance and robustness exist between short and long pulse methods in VCMA driven p-bits, as seen in Tables II and III. In short, the short pulse method benefits from higher speed and low energy consumption (theoretically predicted) whereas the long pulse method is much more robust against device variations, thus making it an ideal solution for prototyping of MTJ based probabilistic Ising machines [43,130]. Since both methods have synchronous write and read pulses, the maximum bit rate is still below 1 GHz for both methods. The key difference is that the short pulse method uses write pulses that are on the order of ≈100 ps, meaning that both the experimental and predicted energy consumption is around 1 fJ/bit whereas the long pulse method requires pulse widths

≥ 10 ns, meaning that the leakage energy consumption is at least 100X larger. However, VCMA driven p-bits using the short pulse method is potentially the most susceptible to device variations among all the biasing mechanisms (see table III), making it an unfeasible solution for large-scale p-bit circuits. It should be noted that there are no studies that make any claims to device variations for VCMA p-bits using the short pulse method, however, we can still claim that they have high susceptibility to device variations based on the fact that the switching probability depends entirely on the precession frequency of the free layer. Previous studies on switching in the precessional regime [166] show that the switching probability is proportional to $\sin^2\varphi$ where φ is proportional to $\exp(J_{C0})$. Any variation in resistance or anisotropy changes φ exponentially, meaning that the switching probability with pulse width is very sensitive to intrinsic device properties. On the other hand, the long pulse method uses pulse widths > 10 ns to ensure 50 % switching probability regardless of variations in device properties making it a much more reliable method.

Among all of the biasing methods, VCEC is the least studied and therefore, has several remaining unknowns regarding its overall performance. Only one experimental study has been published on VCEC p-bits [44]. Because of this lack of research, the remaining issues that need to be studied are maximum switching speed and susceptibility to variations. Furthermore, the energy per bit in ref. [44] exceeds 500 pJ/bit. Despite this issue, the existing research makes VCEC an attractive biasing method for p-bits for three reasons. One is that VCEC switching does not require the thickness of the MgO tunneling barrier to be less than 1 nm. This makes VCEC MTJs less conductive, have better endurance [167], and reduce variations in resistance and TMR between MTJs [168]. The second reason is that ref. [44] showed that VCEC driven p-bits operate at currents below 100 nA. This means that the projected energy consumption for VCEC driven p-bits is less than 0.1 fJ/bit if we assume switching rates below 10 ns can be achieved. Note that this is a fair estimation considering that ref. [169] suggested a write speed of below 10 ns for VCEC MRAM and ref. [136] even demonstrated sub-ns switching with VCEC. For this reason, future experimental work on p-bit hardware should investigate VCEC as a biasing mechanism.

Table III shows that circuits that use dual-biased p-bits are potentially much less susceptible to device variations as single-biased p-bits. There are four key reasons for this. One is that employing two biases allows for the possibility of defect-aware training, where the secondary bias can be used to correct for any variations in the output probability of the p-bit compared to the desired ones. It should be noted that this strategy will likely involve large overhead circuitry to have dual-control over every single p-bit, so it may not be feasible in large-scale circuits. However, there are still three additional benefits of dual-biasing. The second benefit is that manipulation of two biases can reduces the effective thermal stability of the MTJ and can even generate tunable telegraphic switching signals in intrinsically thermally stable MTJs. This relaxes the criteria for the MTJs to be superparamagnetic and makes them more compatible with modern state-of-the-art MRAM technologies. The third benefit is that the average dwell times can be tuned by over three orders of magnitude. This flexibility in switching rate can be used to modify the resistance distributions of the MTJs from bimodal to continuous (recall Fig. 10), thus ensuring that all p-bits have the same resistance distribution behavior. The fourth benefit of dual-biasing is that some combinations allow for a unique capability called two-degrees of tunability, where the AP- and P-state dwell times can be controlled separately. This novel feature can double the information capacity in signals generated from p-bits and provide the possibility of utilizing temporal encoding in stochastic signals [142], further increasing the information capacity.

Previous studies show that STT + field and STT + ME coupling can generate stochastic signals with two-degrees of tunability. Table I shows that the key studies on STT + field dual-biasing are all experimental with no notable theory/simulation studies. This is simply because of the large energy consumption required to generate a tunable magnetic field at each p-bit (see energy consumption in Table II) as well as inability for a field-switchable MRAM cell to be scaled below 90 nm [64]. While refs. [139–141] showed that STT + field can achieve two-degrees of tunability, which was very useful for small-scale experimental demonstrations [143,144], the limitations in energy efficiency and cell size make it an unfeasible solution for large-scale p-bit circuits. STT + ME coupling is a promising alternative approach to dual-biasing in p-bits. It has many of the same benefits as STT + field dual-biasing such as the capability of tuning the switching rates to change the resistance distributions [151] and two degrees of tunability [142], but with significantly lower energy consumption (see predicted values in table II). The key challenge for STT + ME coupling in p-bits is the lack of experimental work that has been done (see Table I) leads to a lot of uncertainty if the performance metrics shown in Table II can actually be achieved. The experimental work that has been done pertaining to ME coupling (using strain) for MRAM require large

Table III. Qualitative assessment of the susceptibility to device-to-device variations for each tuning/biasing mechanism.

Biasing mechanism		Overall susceptibility	Explanation
STT	**PMA**	Medium	• Switching speed can vary by over 2 orders of magnitude [93]. • Very little variation in sigmoid curve (width only varies by 2 – 4 μA) [102].
	IMA	High	• Switching speed varies by over 2 orders of magnitude, sigmoid curve shifts by over 30 Oe, and resistance varies by over 20% [104].
	DFL	Medium	• Subject to the same variations in speed and resistance as IMA MTJs. • Probability is fixed at 50% regardless of voltage bias, therefore no variations in sigmoid curve.
SOT	**Type I**	Medium	• SOT p-bits are more robust against variations in switching speed than STT MTJs (better correlation strength between different speeds) [119]. • Still susceptible to effects of variation for sMTJs with different resistance distributions.
	Type II and III	Very low	• SOT current needs to be strong enough to initialize MTJ to the in-plane direction. • Variation in switching probability from secondary source is only 1.9% [126].
VCMA	**Short Pulse Method**	High	• Any variation in precession frequency drastically influences the switching probability distribution.
	Long Pulse Method	Very low	• 50 % switching probability is ensured with long write pulse and p-bit functions can be implemented in CMOS [43,130].
VCEC		Unknown	• True influence of variations is unknown. • Thick MgO tunnel barriers makes VCEC MTJs less susceptible to resistance and TMR variations [168].
STT+ Field		Low	• No variations in resistance distributions (increases average switching rate by three orders of magnitude). • Relaxes criteria for MTJs to be superparamagnetic. • Two-degrees of tunability achieved [139–142].
STT+ME coupling		Low	
STT+SOT		Medium	• Changes in switching speed is minimal, not clear if this method can solve variations in resistance distributions.
VCEC+SOT		Unknown	• True influence of variations is unknown. • Benefits from thick MgO tunnel barrier.

voltages for switching (>100 V) [152–154]. The low energy calculations in [152,154] assume that strain voltage will scale with piezo size, however, as stated earlier, these calculations do not account for the reduction in piezoelectric coefficient as the piezo substrate reaches sub-μm thicknesses [155]. Therefore, more experimental work is needed to confirm that STT + ME coupling can achieve the predicted energy efficiency presented in Table II.

When evaluating the performance of STT + SOT dual-biasing, Table II shows that it has the possibility of reducing the energy consumption per bit 3X when compared to STT driven p-bits. However, work in ref. [165] indicates that the combination of STT and SOT does not change the average switching rate by

several orders of magnitude, meaning that it cannot be used as a method for modifying the resistance distributions in p-bits. This means that STT + SOT is slightly more susceptible to device variations than STT + field and STT + ME coupling dual-biasing. A simple way to fix this is to use a voltage source for the STT bias rather than a current source, thus making the strength of the STT state dependent. This may also allow for two degrees of tunability to be achieved with STT + SOT dual-biasing, but this requires more research to confirm. One major advantage to STT + SOT dual-biasing is that it does reduce the overall energy consumption which is seen in ref. [135] without affecting the switching probability distribution curve [165].VCEC+SOT also has a lower range of tunability of the switching rate when compared to STT+field or STT+ME coupling dual-biasing does, making it potentially more susceptible to device variations. As stated earlier, this could be resolved by using a voltage source rather than a current source for the VCEC bias, but more research is needed to confirm this. Unlike STT + SOT, it is not clear that combining VCEC and SOT will increase the overall energy consumption. This is because the biggest influence that the SOT current has on VCEC switching is thermal effects. This means that large current densities ($> 10^6$ A/cm^2) are required to have an influence on the effective thermal stability of the MTJ, which will likely not be reduced with improved design of the spin Hall channel. Therefore, VCEC may be the only biasing mechanism that is not improved when combined with other mechanisms, however, this conclusion is based on limited research.

Among various tuning mechanisms explored for p-bit implementations, techniques based on spin-transfer torque (STT) and spin-orbit torque (SOT) might appear slightly more advantageous, creating an impression of superior scalability compared to alternative approaches. This perceived advantage primarily arises because the perpendicular-MTJ stacks used for STT-based MRAM are already qualified in high-volume manufacturing environments, offering established CMOS co-integration, mature fabrication protocols such as etching and chemical-mechanical polishing (CMP), and well-characterized defect models. Hence, existing industrial fabrication lines can seemingly scale these torque-based methods to mega-scale probabilistic circuits with minimal modifications, typically involving only incremental cell-area increases due to additional in-plane SOT current paths. Conversely, voltage-controlled methods such as voltage-controlled magnetic anisotropy (VCMA), voltage-controlled exchange coupling (VCEC), or strain-mediated magnetoelectric biasing—despite their superior energy efficiency at the device level—often require ultra-thin high-quality oxide barriers, specialized multilayer stacks involving rare-earth ferromagnets or piezoelectric materials, or sophisticated 3-D heterogeneous integration steps incompatible with conventional foundry design rules. These intricate material and integration requirements inherently limit wafer-scale yields and achievable circuit densities. Additionally, academia's limited resources and the absence of standardized test vehicles further constrain extensive studies of these promising alternative methods to small-scale demonstrators. In contrast, STT and SOT approaches are supported by mature toolchains, robust industry investment, and reliable supply chains for CoFeB/MgO-based materials. Consequently, the existing literature and industry research trends may disproportionately reflect the scalability and manufacturability of torque-driven techniques, not necessarily due to an intrinsic superiority of these methods, but largely because current ecosystem readiness and fabrication maturity heavily favor their rapid deployment for large-scale probabilistic computing applications.

V. FUTURE OF MTJ-BASED PROBABILISTIC COMPUTERS

A. From Devices to Circuits

The architecture of p-computers can vary significantly to adapt to various intended applications and tasks. However, in a more granular perspective, synapses, interconnects, and p-bits are the key building blocks. The synapses and interconnects are essentially the same as those intensively studied by the neuromorphic computing community [170]. Although typically the synapses significantly outnumber the p-bits, it is the adoption of p-bits, instead of neurons in more traditional definitions, that introduces stochasticity to p-computers, differentiating them from other neuromorphic computing paradigms. Therefore, in the rest of this section, we will focus on the designs and working principles of various MTJ-based p-bit hardware implementations.

A p-bit is fundamentally a unit that generates a stochastic signal, where the statistical mean of the output varies non-linearly with the input signal level. The hardware design of a p-bit integrates spintronic devices alongside passive and active CMOS components, including resistor networks, voltage dividers, Analog-to-Digital Converters (ADC), Digital-to-Analog Converters (DAC), comparators,

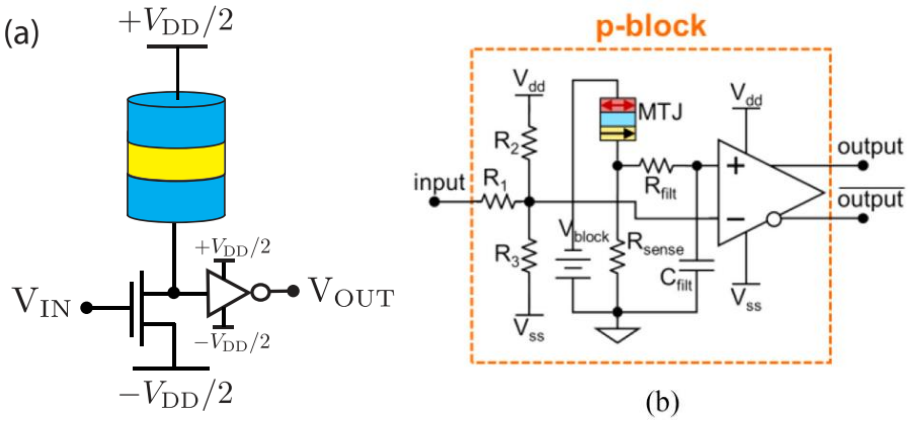

Fig. 16. P-bit circuits with electrical modulation. **(a)** input-dependent electrical bias on MTJ (reproduced from ref. [40] © 2017 IEEE), **(b)** input-independent electrical bias on MTJ (reproduced from ref. [143], Reprinted with permission from IEEE).

and amplifiers. These auxiliary circuits remain consistent across various probabilistic computing architectures and applications, regardless of the specific implementation of the underlying stochastic unit. Consequently, the functional realization of p-bit hardware extends beyond the intrinsic stochastic properties of spintronic devices, incorporating dependencies on additional circuit elements responsible for signal conditioning, amplification, and control.

Generally, spintronic devices provide stochasticity while the auxiliary components and circuits provide working conditions to the spintronic devices and conditions the input and output signals. The key to a qualitatively functioning design of p-bit is to implement a mechanism that modulates the statistical mean value of the output stochastic signal based on the signal level input. Existing approaches to this modulation can be broadly classified into electrical and physical (spintronic) strategies. The electrical modulation approach primarily utilizes circuit-level transformations to process the stochastic signal generated by the Magnetic Tunnel Junction (MTJ) and establish its dependence on the input signal. This is achieved through the peripheral CMOS components mentioned earlier, which regulate and refine the MTJ output. In contrast, the physical modulation mechanism leverages the intrinsic material properties and magnetization dynamics of the MTJ itself to establish input-output dependence, directly modulating stochasticity at the device level. A prime example of electrical modulation is the implementation of an embedded MTJ-based p-bit as shown in Fig. 16(a), where a conventional 1T-1MTJ cell, commonly found in MRAM architectures, is modified by replacing a thermally stable MTJ with a low-barrier nanomagnet (LBM). Additionally, an inverter is incorporated to amplify and sense the state of the p-bit. This modified structure is often referred to as a three-terminal p-bit cell, though it deviates from the standard MTJ-based nomenclature. Unlike configurations such as spin-Hall effect (SHE) + spin-orbit torque (SOT) [92] MTJs, where the MTJ itself serves as a three-terminal device, the p-bit cell here retains a two-terminal MTJ while utilizing an access transistor as the third terminal. In Fig. 16(a), the source transistor offsets the stochastic signal relatively to the internal threshold of the inverter based on the input signal level by Vin controlling the resistance of the transistor, resulting in an input-dependent stochastic output signal. While this design offers simplicity, it comes with a notable drawback—the bias voltage and current can influence the switching or pinning behavior of the MTJ, inadvertently introducing an additional dependence beyond that of the source transistor. This added variability reduces the robustness of the system, limits the possibility of incorporating a more diverse range of spintronic or other novel devices, making it more challenging to precisely control the output probability as a function of the input signal.

An alternative electrical modulation strategy, illustrated in Fig. 16(b) [143], mitigates this issue by decoupling the MTJ signal path from the input signal. In this architecture, both signals are fed into a comparator, which, within an MRAM setting, is

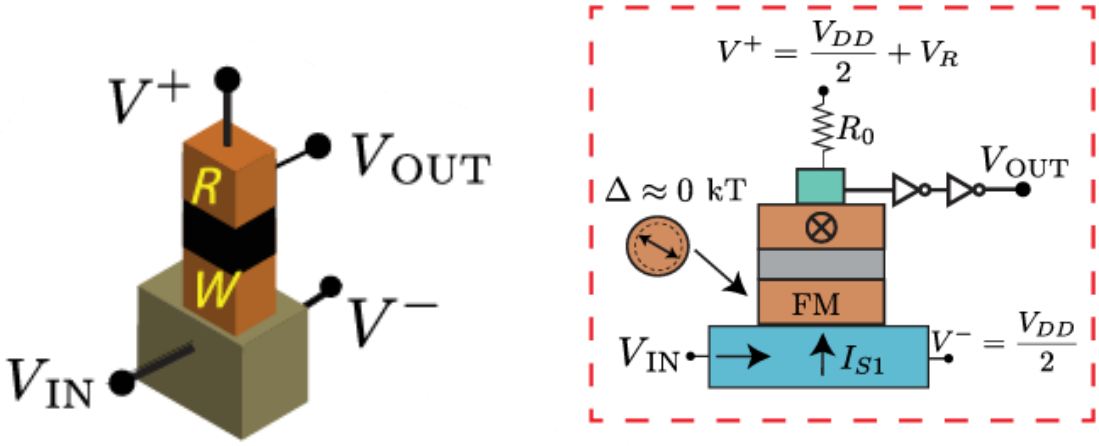

Fig. 17. A p-bit circuit utilizing physical modulation (reproduced from ref. [92], https://creativecommons.org/).

largely similar to a sense amplifier, as demonstrated in ref. [38]. This approach offers several advantages: it maintains a stable and independent electrical bias on the MTJ, preventing unwanted fluctuations, and additionally, it supports thermally stable MTJs through a dual-biasing mechanism. However, the increased circuit complexity and the requirement for additional components pose potential area and power trade-offs. Despite these challenges, the improved stability and greater control over the MTJ switching dynamics make this approach a compelling alternative for probabilistic computing applications.

Fig. 17 illustrates a Magnetic Tunnel Junction (MTJ) integrated atop a Spin-Hall Effect (SHE) channel, alongside a corresponding p-bit circuit design. In this architecture, an applied input voltage (Vin) drives a charge current through the underlying SHE channel, inducing a spin current that influences the magnetic dynamics of the nanomagnet within the MTJ. Consequently, the input voltage directly modulates the statistical behavior of the MTJ's magnetic state, thereby shaping the output characteristics of the p-bit. Notably, this modulation mechanism is inherently embedded within the spintronic device itself, reducing the need for external control elements.

This design offers several advantages. Since STT switching can cause breakdown in the MgO tunnel barrier [171], the separation of read and write current paths enhances device endurance, potentially extending its operational lifespan. Furthermore, the circuit architecture remains relatively simple, requiring fewer components compared to alternative designs, which contributes to improved scalability. However, certain limitations persist. While three-terminal devices such as SHE+STT and SOT may still lag slightly behind conventional two-terminal STT-MTJs in maturity, the gap is narrowing as major industry players like TSMC have begun developing fabrication processes for these technologies, thereby reducing the differences in CMOS integration advantages between STT and SOT-based approaches. Additionally, static power consumption is highly device-dependent, as it is intrinsically tied to design parameters, making it particularly challenging to optimize for low-power operation. Despite these trade-offs, this spintronic-based modulation approach holds promise for advancing energy-efficient and compact probabilistic computing architectures.

Despite the promise of p-bit-based probabilistic computing, existing designs impose several architectural constraints. A key challenge lies in ensuring that the transistor conductance is precisely matched to the average conductance of the MTJ ($(G_{AP}+G_P)/2$), while also optimizing inverter transistor sizing to deliver adequate gain. Additionally, precise biasing of inverters to their midpoint in the DC operating range is essential to achieve a reliable and broad output probability distribution [14]. These design complexities become even more pronounced when scaling from circuit-level implementations to full system architectures across various applications.

For instance, in the weighted array structure of a Restricted Boltzmann Machine (RBM) [91], a chain of inverters, where each stage is twice the size of its predecessor, is necessary to ensure accurate output sensing and maintain proper synaptic functionality. However, this approach introduces substantial area and power overheads, limiting scalability. Similarly,

in synaptic networks [40], offline resistor network tuning is required to match the current and impedance requirements of p-bit neurons, thereby hindering dynamic learning capabilities. Alternative synapse and weight-update implementations using FPGAs or microcontrollers [14] suffer from additional inefficiencies, as bulky DACs and control modules further contribute to increased power consumption and area constraints.

Several studies [13,40,88,91,92] have highlighted the energy and area benefits of device-level innovations, demonstrating that MTJ-based p-bit architectures have numerous advantages over their software counterparts, FPGA implementations, and stochastic CMOS circuits such as Linear Feedback Shift Registers (LFSRs), RRAM, and PCM-based designs. CMOS based true random number generators consume over 4 pJ/bit at rates of 100 MHz [172], which underperforms relative to most MTJ based p-bit solutions (recall Table II). RRAM and PCM based random number generators have been proposed [173–175], however, these solutions rely on statistical variations among multiple cells to provide random bits, whereas MTJs can provide tunable random numbers with a single device. Photonic devices can be used as p-bits [176], however, they require vacuum conditions for operation. Another potential alternative to spintronic devices for p-bit implementation are FeFETs [177,178], however, there is a lack of experimental work on FeFET p-bits and the modelling work appears to indicate that FeFET p-bits switch at a time scale of μs rather than ns. Several studies have also used other emerging technologies for Ising machines. For example, Chiang et al proposed a RRAM based Ising machine that leverages noise inherent to RRAM devices to avoid local minimum issues [179]. Ouyang et al demonstrated a photonic Ising machine where spins were encoded via phase delay and a series of beam splitters are used to update the pattern of spins [180]. Xu et al demonstrated an FeFET based Ising machine that program and reconfigure the weight elements between spins without additional configuration of overhead circuitry [181]. None of these solutions benefit from the nonlinearity and tunable randomness inherent to p-bits. The common bottleneck across all these implementations remains the bulky analog peripheral circuitry, which hinders scalability and restricts real-world deployment of probabilistic computing hardware. Addressing this critical limitation is imperative to unlock the full potential of p-bit-based computing paradigms.

Alternatively, synchronous approaches can be employed on MTJ/CMOS hybrid systems, which may reduce the analog peripheral circuitry. Recently Yang, Shuhan, et al. demonstrated first probabilistic Ising machine (PIM) that enables 250 STT-MTJs with an FPGA back-end [182]. Each MTJ acts as a tunable true-random p-bit, while 16 on-board DAC/ADC banks and an FPGA scheduler distribute inputs and collect states. The process of solving combinatorial optimization problems is similar to that used in [130], where the MTJs provide the entropy source and calculations for the Ising Hamiltonian is done via CMOS. Note that even though the work in [182] still uses DACs and ADCs, the work in [130] showed that lookup tables, which are much less area-intensive than DACs, can be used instead. At this time, these MTJ/CMOS hybrid systems are the largest scale probabilistic Ising machines demonstrated and are the closest to reaching prototyping stage.

At the device level, this work advocates for a dual-biasing approach, particularly leveraging Voltage-Controlled Magnetic Anisotropy (VCMA)-based tuning mechanisms, as a future roadmap for MTJ-based p-bit implementations [130]. Unlike conventional Low-Barrier Magnet (LBM) engineering, which demands intricate energy barrier optimizations, VCMA tuning facilitates the use of thermally stable, well-matured MRAM cells for p-bit applications, mitigating device-to-device variations. Moreover, fluctuation rate variations in LBM-based MTJs can be improved through multi-tunability strategies. By adopting dual-biasing, higher MTJ resistance values can be achieved, enabling a broader voltage range at inverters for precise state sensing while simultaneously allowing the use of thicker barrier layers. This not only enhances device endurance and reduces on-current losses but also relaxes fabrication constraints, thereby enabling further scaling of MTJs to seamlessly integrate with advanced CMOS technologies.

By following this strategic device roadmap, the field can overcome key scaling challenges and fully harness the capabilities of probabilistic computing, ultimately paving the way for energy-efficient, high-performance, and scalable p-bit architectures.

B. Next Generation MTJs

Significant developments have been made to optimize MTJ stacks to address the demand for p-bit computing, aiming for optimized TMR ratios, improved resistance distributions, and integration with

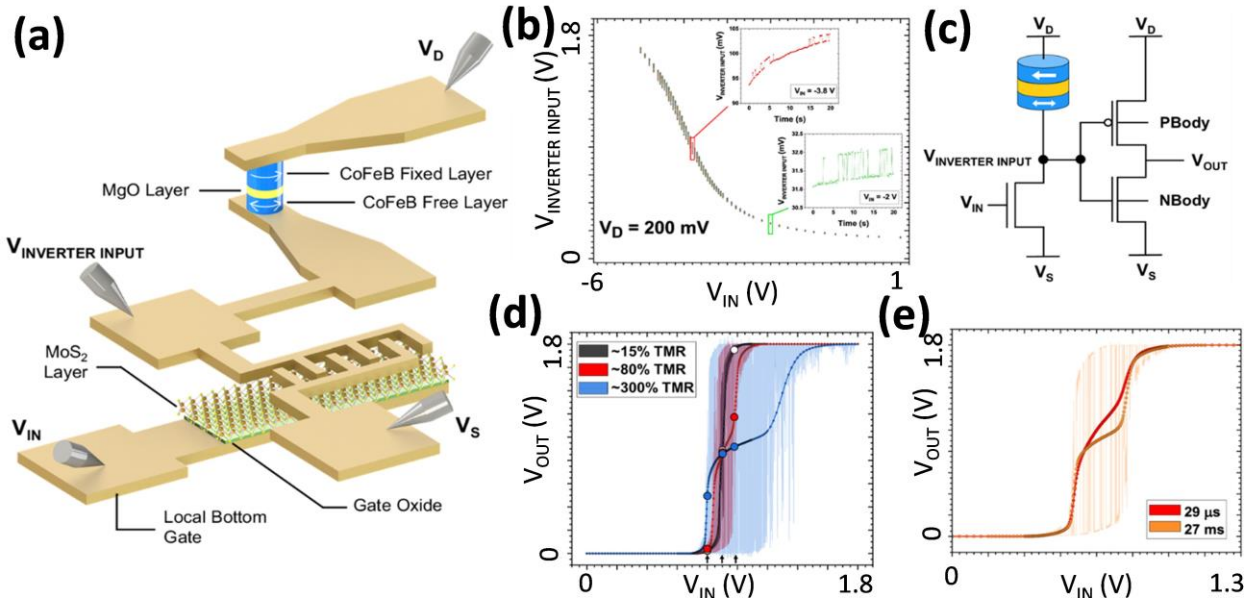

Fig. 18. Illustration depicting the integration scheme of a stochastic MTJ and a monolayer MoS_2 transistor for realizing an on-chip probabilistic bit (p-bit). **(b)** Depiction of the functionality of an integrated p-bit core, highlighting stochastic variations in the output voltage ($V_{INVERTER\ INPUT}$), which can be controlled by adjusting the transistor gate voltage (V_{IN}). The displayed fluctuations are captured through multiple measurements of $V_{INVERTER\ INPUT}$ at various V_{IN} values. Insets present detailed time-series examples of these fluctuations for two gate voltages: $V_{IN} = -3.8$ V (red inset) and $V_{IN} = -2$ V (green inset). **(c)** Schematic representation of the p-bit circuit designed in Cadence software for performing circuit simulations, including extra body bias terminals (denoted as PBody and NBody) to enable tuning of the PMOS and NMOS transistors within the inverter. **(d)** Simulation results comparing p-bit outputs with different TMR ratios, indicating that excessively large TMRs create undesirable plateaus in the stochastic output, while moderate (~80%) TMR provides smoother and more tunable probabilistic behavior. **(e)** Effect of MTJ dwell times on p-bit performance. Faster stochastic MTJs, with shorter dwell times and more continuous resistance distributions, significantly improve the uniformity and tunability of the p-bit's output compared to slower, more bimodal MTJs. Reproduced from ref. [114] (https://creativecommons.org/licenses/by/4.0/).

present CMOS platforms. Liu, Samuel et al. has reported that the conductance noise generated by MTJs which is important in p-bit computing has dependency on MTJ's TMR ratio [183]. It is reported that higher TMR ratios lead to larger differences between parallel and anti-parallel conductance states leads to a broader range noise level. For Bayesian neural networks (BNNs) this broader range is important for encoding both narrow which corresponds to low-noise and broad which corresponds to high-noise probabilistic distributions. Furthermore, this study highlights the modulation of TMR through VCMA effect providing control over conductance noise magnitude, shows that a high TMR ratio allows encoding weight distributions in BNNs ranging from highly noisy to nearly deterministic. The TMR ratio modulation with the application of voltage using VCMA effect has been reported in both theoretical studies and experimental results [184–187].

However, Daniel, John, et al. recently demonstrated an on-chip p-bit core utilizing stochastic spin-transfer torque magnetic tunnel junctions (s-MTJs) integrated with monolayer MoS_2 FETs, as depicted in Fig. 18 [114] This integration facilitates efficient, voltage-controllable stochasticity, high transistor performance with currents around 0.6 mA at V_{DS} = 0.1 V, high on/off ratios (~10^{10}), and subthreshold slopes of approximately 94 mV/dec. Evaluation of s-MTJs with varying TMR ratios (15%, 80%, and 300%) revealed that a high TMR ratio of 300% led to substantial fluctuations but introduced undesirable plateau effects in the stochastic region, negatively affecting output tunability and smoothness. A medium TMR ratio of 80% offered balanced properties, demonstrating smooth transitions and an appropriate range of fluctuations. Additionally, dwell time and distribution factors significantly impact p-bit operational optimization. Shorter (faster) dwell times result in lower distribution factors, indicating a more uniform state distribution, essential for smooth p-bit outputs.

The motivation behind integrating s-MTJs with 2D FETs includes enabling low thermal budget processing (beneficial for MTJ-CMOS interfaces and avoid tunnel barrier degradation), achieving low contact resistance, and leveraging the large bandgap of 2D materials.

Therefore, for optimized p-bits computing, MTJs are supposed to have optimal TMR levels as high TMR ratios could lead to undesirable plateaus in p-bit outputs, hindering performance. For reliable p-bit operation the MTJ's TMR must be neither too small nor too large. A very large TMR causes the inverter input to straddle both noise-margin regions, producing a flat 'plateau' in the time-averaged output that makes the output probability insensitive to its input. Conversely, a very small TMR yields a narrow fluctuation window that cannot drive subsequent stages. Experiments show that plateaus slow network convergence by a factor of two in a 5-p-bit full-adder [14]. It should also have uniform MTJ resistance distribution (less bimodal) that tend to produce smoother sigmoidal outputs, which are desirable for p-bit computation.

Further, MTJs incorporating 3D materials have experienced significant advancements, particularly regarding integration with CMOS technology, tolerance of the high thermal budget during CMOS fabrication processes, improved TMR values, and enhanced uniformity in MTJ resistance distribution. However, several challenges persist, primarily related to the quality of the interface between the ferromagnetic (FM) layer and the barrier layer [188]. Interface bonding plays a crucial role in determining the effectiveness of these advancements, directly influencing spin transport and magnetoresistive properties. Traditional chemical bonding methods often lead to issues such as spin scattering and electron transmission inefficiencies, which ultimately affect conductance [189]. One promising solution is the integration of layered two-dimensional (2D) materials, which utilize van der Waals (vdW) forces rather than chemical bonding, potentially mitigating these interface-related problems [190].

Due to the intrinsic electronic structure, atomically flat interfaces, and precise layer-by-layer control achievable with 2D materials like graphene, hBN, MoS_2, and WS_2, vdW-based MTJs typically exhibit narrower and more uniform resistance distributions. These materials allow for unique opportunities in magneto-electronics [191–193], where electrical properties can be manipulated for device engineering, and in developing interfacial phenomena through vdW engineering [194–198]. The uniform and defect-free tunnel barriers of hBN enable stable [199], well-defined resistance states, while graphene-based MTJs offer moderate spin polarization leading to smaller but highly uniform TMR ratios. These attributes of 2D vdw based MTJs enables them for potential incorporation in p-bit computing. Nevertheless, despite these promising attributes, achieving uniformity across large-scale arrays remains a practical challenge due to complexities involved in 2D vdw material synthesis and device fabrication.

VI. CONCLUSION

In this article, we reviewed the recent progress of MTJ based probabilistic bits and provided an overview of the prospects and challenges of each biasing method. The biasing methods that we focused on were spin transfer torque (STT), spin orbit torque (SOT), voltage controlled magnetic anisotropy (VCMA), and voltage-controlled exchange coupling (VCEC). While STT and SOT are the most commonly studied methods, they require large current densities to operate which limits their energy efficiency. Using a superparamagnetic MTJ (sMTJ) with a double free layer is one way to avoid these large current densities as their output probability is not determined by the current through the device. Alternatively, electric field control over MTJs via VCMA or VCEC offers a much more energy efficient solution as magnetization switching can be achieved with these methods on MTJs with thick MgO barriers, thus reducing the conductance of the device. All of these biasing methods are susceptible to device-to-device variations in switching speed and transfer response due to defects that occur during MTJ fabrication. One way to mitigate the impacts of these defects is to change the operation of the p-bits to a synchronous approach with clocked write and read pulses. This can be done either through long VCMA pulses or through hard-axis initialization with an SOT current. While both of these strategies successfully mitigate the effects of device variation, they significantly reduce the operation speed and increase the energy consumption. A more straight-forward approach to addressing the issue of device variations is to use two biasing methods to control the output of p-bits. Therefore, we expanded our review to include various dual-biasing strategies, namely, STT + field, STT + magnetoelectric (ME) coupling, STT + SOT, and STT + VCEC. However, there are several trade-offs when combining different biases to p-bits. For example, STT + field and STT + ME coupling offer tremendous flexibility in control over each p-bit. The average switching rate can be tuned over three orders of magnitude, which can be used to fix variations in resistance distributions between p-bits. Furthermore, both of these combinations offer a

unique capability called two-degrees of tunability where the average high-state and low-state dwell times can be tuned separately. Yet these combinations may increase the overall energy consumption. On the other hand, STT + SOT will reduce the overall energy consumption when compared to single-biased p-bits that operate using either STT or SOT effects, but the limited number of experimental studies on STT + SOT dual-biasing shows that this combination does not offer the same range of control over the MTJs switching speed as STT + field or STT + ME coupling offers. Lastly, the least studied biasing method for p-bit operation is VCEC. Nevertheless, the few studies on VCEC driven p-bits show that they can operate at currents less than 100 nA, making VCEC a promising solution for developing p-bits that operate at energies less than 1 fJ/bit. These results warrant further investigation into the speed and robustness of VCEC driven p-bits.

ACKNOWLEDGMENTS

This work was supported, in part, by SMART, one of the seven centers of nCORE, a Semiconductor Research Corporation program, sponsored by the National Institute of Standards and Technology (NIST), by the Defense Advanced Research Projects Agency (DARPA) (Advanced MTJs for computation in and near random access memory) under Grant HR001117S0056-FP-042, by the Global Research Collaboration (GRC) Logic and Memory program, sponsored by Semiconductor Research Corporation (SRC), NSF ASCENT program TUNA: No. 2230963, NSF MEITY Program No: ECCS 2415836. and NSF CBET-2226579. Parts of this work were carried out in the Characterization Facility, University of Minnesota, which receives partial support from the NSF through the MRSEC (Award Number DMR-2011401) and the NNCI (Award Number ECCS-2025124) programs. Portions of this work were conducted in the Minnesota Nano Center, which is supported by the National Science Foundation through the NNCI under Award Number ECCS-2025124. We would also like to thank Jabez McClelland at the Alternative Computing Group at NIST for supporting Brandon R. Zink during the preparation of this manuscript.

[1] F. Zhu, P. Xu, and J. Zong, Moore's Law: The potential, limits, and breakthroughs, Proceedings of the 2023 International Conference on Mechatronics and Smart Systems, **10**, 307 (2023).

[2] R. G. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, and T. Mudge, Near Threshold Computing: Overcoming Performance Degradation from Aggressive Voltage Scaling, Workshop on Engineering Efficient Design held in conjunction with ISCA 36, pp. 44 – 49 (2009).

[3] K. Thiagarajan, C. K. Dixit, M. Panneerselvam, C. A. Madhuvappan, S. Gadde, and J. N. Shrote, Analysis on the Growth of Artificial Intelligence for Application Security in Internet of Things, in Proceedings of the 2nd International Conference on Artificial Intelligence and Smart Energy (ICAIS), pp. 6–12 (2022).

[4] M. Soori, B. Arezoo, and R. Dastres, Artificial Intelligence, Machine Learning and Deep Learning in Advanced Robotics, a Review, Cognitive Robotics, **3**, 54 (2023).

[5] M. F. Safitra, M. Lubis, T. F. Kusumasari, and D. P. Putri, Advancements in Artificial Intelligence and Data Science: Models, Applications, and Challenges, Procedia Computer Science, **234**, 381 (2024).

[6] S. Chowdhury, A. Grimaldi, N. A. Aadit, S. Niazi, M. Mohseni, S. Kanai, H. Ohno, S. Fukami, L. Theogarajan, G. Finocchio, S. Datta, and K. Y. Camsari, A Full-Stack View of Probabilistic Computing With p-Bits: Devices, Architectures, and Algorithms, IEEE J. Explor. Solid-State Computational Devices and Circuits **9**, 1 (2023).

[7] S. Misra, L. C. Bland, S. G. Cardwell, J. A. C. Incorvia, C. D. James, A. D. Kent, C. D. Schuman, J. D. Smith, and J. B. Aimone, Probabilistic Neural Computing with Stochastic Devices Adv. Mater., **35**, 2204569 (2023).

[8] K. Y. Camsari, Probabilistic Computing with P-Bits: Optimization, Machine Learning and Quantum Simulation, in 2024 IEEE International Magnetic Conference – Short papers (INTERMAG short papers), pp. 1-2 (2024).

[9] N. A. Aadit, A. Grimaldi, G. Finocchio, and K. Y. Camsari, Physics-Inspired Ising Computing with Ring Oscillator Activated p-Bits, 22nd IEEE Int. Conf. Nanotechnology (NANO) pp. 393-396 (2022).

[10] J. Kaiser and S. Datta, Probabilistic Computing with P-Bits, Appl. Phys. Lett., **119**, 150503 (2021).

[11] R. Faria, K. Y. Camsari, and S. Datta, Implementing Bayesian networks with embedded stochastic MRAM, AIP Adv., **8**, 045101 (2018).

[12] R. Faria, J. Kaiser, K. Y. Camsari, and S. Datta, Hardware Design for Autonomous Bayesian Networks, Front. Comput. Neurosci., **15**, 584797 (2021).

[13] B. Behin-Aein, V. Diep, and S. Datta, A building block for hardware belief networks, Sci. Rep., **6**, 29893 (2016).

[14] 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).

[15] N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, G. Finocchio, and K. Y. Camsari, Computing with Invertible Logic: Combinatorial Optimization with Probabilistic Bits, Int. Electron Devices Meeting (IEDM), pp. 40.3.1-40.3.4 (2021).

[16] N. Onizawa and T. Hanyu, Enhanced convergence in p-bit based simulated annealing with partial deactivation for large-scale combinatorial optimization problems, Sci. Rep., **14**, 1339 (2024).

[17] B. Sutton, K. Y. Camsari, B. Behin-Aein, and S. Datta, Intrinsic optimization using stochastic nanomagnets, Sci. Rep., **7**, 44370 (2017).

[18] M. Yamaoka, C. Yoshimura, M. Hayashi, T. Okuyama, H. Aoki, and H. Mizuno, A 20k-spin Ising chip to solve combinatorial optimization problems with CMOS annealing, IEEE J. Solid-State Circuits, **51**, 303 (2016).

[19] W. Moy, I. Ahmed, P. W. Chiu, J. Moy, S. S. Sapatnekar, and C. H. Kim, A 1,968-node coupled ring oscillator circuit for combinatorial optimization problem solving, Nat. Electron, **5**, 310 (2022).

[20] K. Y. Camsari, B. M. Sutton, and S. Datta, P-Bits for Probabilistic Spin Logic, Appl. Phys. Rev., **6**, 011305 (2019).

[21] O. Hassan, S. Datta, and K. Y. Camsari, Quantitative Evaluation of Hardware Binary Stochastic Neurons, Phys. Rev. Appl., **15**, 064046 (2021).

[22] W. H. Choi, Y. Lv, J. Kim, A. Deshpande, G. Kang, J.-P. Wang, and C. H. Kim, A Magnetic Tunnel Junction Based True Random Number Generator with conditional perturb and real-time output probability tracking, Int. Electron Devices Meeting (IEDM), pp. 12.5.1 – 12.5.4 (2014).

[23] 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).

[24] J. G. Zhu and C. Park, Magnetic tunnel junctions, Materials Today, **9**, 36 (2006).

[25] J. C. Slonczewski, Current-Driven Excitation of Magnetic Multilayers, J. Magn. Magn. Mater., **159**, L1 – L7 (1996).

[26] L. Berger, Emission of Spin Waves by a Magnetic Multilayer Traversed by a Current, Phys. Rev. B, **54**, 9353 (1996).

[27] L. Liu, T. Moriyama, D. C. Ralph, and R. A. Buhrman, Spin-torque ferromagnetic resonance induced by the spin Hall effect, Phys. Rev. Lett., **106**, 036601 (2011).

[28] I. M. Miron, K. Garello, G. Gaudin, P. J. Zermatten, M. V. Costache, S. Auffret, S. Bandiera, B. Rodmacq, A. Schuhl, and P. Gambardella, Perpendicular Switching of a Single Ferromagnetic Layer Induced by In-Plane Current Injection, Nature, **476**, 189 (2011).

[29] L. Liu, O. J. Lee, T. J. Gudmundsen, D. C. Ralph, and R. A. Buhrman, Current-induced switching of perpendicularly magnetized magnetic layers using spin torque from the spin hall effect, Phys. Rev. Lett., **109**, 096602 (2012).

[30] L. Liu, C.-F. Pai, Y. Li, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, Spin-Torque Switching with the Giant Spin Hall Effect of Tantalum, Science, **336**, 555 (2012).

[31] T. Nozaki, T. Yamamoto, S. Miwa, M. Tsujikawa, M. Shirai, S. Yuasa, and Y. Suzuki, Recent Progress in the Voltage-Controlled Magnetic Anisotropy Effect and the Challenges Faced in Developing Voltage-Torque MRAM, Micromachines, **10**, 327 (2019).

[32] C. Grezes, F. Ebrahimi, J. G. Alzate, X. Cai, J. A. Katine, J. Langer, B. Ocker, P. Khalili Amiri, and K. L. Wang, Ultra-low switching energy and scaling in electric-field-controlled nanoscale magnetic tunnel junctions with high resistance-area product, Appl. Phys. Lett., **108**, 012403 (2016).

[33] W. G. Wang, M. Li, S. Hageman, and C. L. Chien, Electric-field-assisted switching in magnetic tunnel junctions, Nat. Mater., **11**, 64 (2012).

[34] Y.-C. Chen, T. Peterson, Q. Jia, Y. Yang, S. Liang, B. R. Zink, Y. H. Huang, D. Lyu, B. Dixit, and J.-P. Wang, Large and Tunable Electron-Depletion-Based Voltage-Controlled

Magnetic Anisotropy in the CoFeB/MgO System via Work-Function-Engineered Pt_xW_{1-x} Underlayers, ACS Nano, **19**, 15953 (2025).

[35] D. Zhang, M. Bapna, W. Jiang, D. Sousa, Y.-C. Liao, Z. Zhao, Y. Lv, P. Sahu, D. Lyu, A. Naeemi et al., Bipolar Electric Field Switching of Perpendicular Magnetic Tunnel Junctions Through Voltage - Controlled Exchange Coupling, Nano Lett., **22**, 622 (2021).

[36] S. Ikegawa, F. B. Mancoff, J. Janesky, and S. Aggarwal, Magnetoresistive Random Access Memory: Present and Future, IEEE Trans. Electron Devices **67**, 1407 (2020).

[37] 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).

[38] D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. S. Friedman, A. F. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, S. Yuasa et al., Low-Energy Truly Random Number Generation with Superparamagnetic Tunnel Junctions for Unconventional Computing, Phys. Rev. Appl., **8**, 054045 (2017).

[39] K. H. Han, Y. J. Kim, H. C. Koo, O. J. Lee, and S. Hong, Probabilistic computing enabled by continuous random numbers sampled from in-plane magnetized stochastic magnetic tunnel junctions, Appl. Phys. Lett., **125**, 142402 (2024).

[40] K. Y. Camsari, S. Salahuddin, and S. Datta, Implementing p-bits with Embedded MTJ, IEEE Electron Device Letters, **38**, 1767 (2017).

[41] M. Bapna and S. A. Majetich, Current control of time-averaged magnetization in superparamagnetic tunnel junctions, Appl. Phys. Lett., **111**, 243107 (2017).

[42] 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**, 4503305 (2019).

[43] Y. Shao, C. Duffee, E. Raimondo, N. Davila, V. Lopez-Dominguez, J. A. Katine, G. Finocchio, and P. Khalili Amiri, Probabilistic computing with voltage-controlled dynamics in magnetic tunnel junctions, Nanotechnology **34**, 495203 (2023).

[44] Q. Jia et al., Energy-Efficient Stochastic Signal Manipulation in Superparamagnetic Tunnel Junctions via Voltage-Controlled Exchange Coupling, Nano Lett., **25**, 9181 (2025).

[45] W. Cai. K. Shi, Y. Zhuo, D. Zhu, Y. Huang, J. Yin, K. Cao, Z. Wang, Z. Guo, Z. Wang, et al, Sub-ns Field-Free Switching in Perpendicular Magnetic Tunnel Junctions by the Interplay of Spin Transfer and Orbit Torques, IEEE Electron Device Letters **42**, 704 (2021).

[46] M. Cubukcu, O. Boulle, N. Mikuszeit, C. Hamelin, T. Brächer, N. Lamard, M.-C. Cyrille, L. Buda-Prejbeanu, K. Garello, I. M. Miron et al., Ultra-Fast Perpendicular Spin-Orbit Torque MRAM, IEEE Trans. Magn. **54**, 9300294 (2018).

[47] V. D. Nguyen, S. Rao, K. Wostyn, and S. Couet, Recent progress in spin-orbit torque magnetic random-access memory, npj Spintronics **2**, 48 (2024).

[48] E. Grimaldi, V. Krizakova, G. Sala, F. Yasin, S. Couet, G. Sankar Kar, K. Garello, and P. Gambardella, Single-shot dynamics of spin–orbit torque and spin transfer torque switching in three-terminal magnetic tunnel junctions, Nat. Nanotechnol. **15**, 111 (2020).

[49] B. R. Zink, D. Zhang, H. Li, O. J. Benally, Y. Lv, D. Lyu, and J. P. Wang, Ultralow Current Switching of Synthetic-Antiferromagnetic Magnetic Tunnel Junctions Via Electric-Field Assisted by Spin–Orbit Torque, Adv. Electron Mater., **8**, 2200382 (2022).

[50] Md. A. Abeed and S. Bandyopadhyay, Low Energy Barrier Nanomagnet Design for Binary Stochastic Neurons: Design Challenges for Real Nanomagnets With Fabrication Defects, IEEE Magn Lett **10**, 4504405 (2019).

[51] N. Onizawa and T. Hanyu, GPU-accelerated simulated annealing based on p-bits with real-world device-variability modeling, Sci. Rep., **15**, 6118 (2025).

[52] A. Abeed and S. Bandyopadhyay, Sensitivity of the Power Spectra of Magnetization Fluctuations in Low Barrier Nanomagnets to Barrier Height Modulation and Defects, Spin, **10**, 2050001 (2020).

[53] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen, V. Dau, F. Petroff, P. Eitenne, G. Creuzet, A. Friederich, and J. Chazelas, Giant Magnetoresistance of (001)Fe/(001) Cr Magnetic Superlattices, Phys. Rev. Lett., **61**, 2472 (1988).

[54] G. Binasch, P. Grűnberg, F. Saurenbach, and W. Zinn, Enhanced Magnetoresistance in Layered Magnetic Structures with Antiferromagnetic Interlayer Exchange, Phys. Rev. B, **39**, 4828 (1989).

[55] C. Chappert, A. Fert, and F. Nguyen Van Dau, The emergence of spin electronics in data electronics, Nat. Mater., **6**, 813 (2007).

[56] M. Julliere, TUNNELING BETWEEN FERROMAGNETIC FILMS, Phys. Lett. A, 54, 225 (1975).

[57] J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, Large Magnetoresistance at Room Temperature in Ferromagnetic Thin Film Tunnel Junctions, Phys. Rev. Lett., **74**, 3273 (1995).

[58] T. Miyazaki and N. Tezuka, Giant Magnetic Tunneling Effect in Fe/AlzO3/Fe Junction, J. Magn. Magn. Mater., **139**, pp. L231 – L234 (1995).

[59] W. H. Butler, X. G. Zhang, T. C. Schulthess, and J. M. MacLaren, Spin-dependent tunneling conductance of Fe/MgO/Fe sandwiches, Phys. Rev. B, **63**, 544161 (2001).

[60] J. Mathon and A. Umerski, Theory of tunneling magnetoresistance of an epitaxial Fe/MgO/Fe(001) junction, Phys. Rev. B, **63**, 220403(R) (2001).

[61] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando, Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions, Nat. Mater., **3**, 868 (2004).

[62] S. S. P. Parkin, C. Kaiser, A. Panchula, P. M. Rice, B. Hughes, M. Samant, and S. H. Yang, Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers, Nat. Mater., **3**, 862 (2004).

[63] T. Scheike, Z. Wen, H. Sukegawa, and S. Mitani, 631% room temperature tunnel magnetoresistance with large oscillation effect in CoFe/MgO/CoFe(001) junctions, Appl. Phys. Lett., **122**, 112404 (2023).

[64] R. Sbiaa, H. Meng, and S. N. Piramanayagam, Materials with Perpendicular Magnetic Anisotropy for Magnetic Random Access Memory, Phys. Status Solidi RRL, **5**, 413 (2011).

[65] S. Ikegawa, K. Nagel, F. B. Mancoff, S. M. Alam, M. Arora, M. DeHerrera, H. K. Lee, S. Mukherjee G. Shimon, J. J. Sun et al., High-Speed (400MB/s) and Low-BER STT-MRAM Technology for Industrial Applications, Int. Electron Devices Meeting (IEDM), pp. 1041–1044 (2022).

[66] L. Liu, O. J. Lee, T. J. Gudmundsen, D. C. Ralph, and R. A. Buhrman, Current-induced switching of perpendicularly magnetized magnetic layers using spin torque from the spin hall effect, Phys. Rev. Lett., **109**, 096602 (2012).

[67] C. F. Pai, L. Liu, Y. Li, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, Spin transfer torque devices utilizing the giant spin Hall effect of tungsten, Appl. Phys. Lett., **101**, 122404 (2012).

[68] M. Dc, R. Grassi, J.-Y. Chen, M. Jamali, D. R. Hickey, D. Zhang, Z. Zhao, H. Li, P. Quarterman, Y. Lv et al., Room-temperature high spin–orbit torque due to quantum confinement in sputtered $Bi_xSe_{(1-x)}$ films, Nat. Mater., **17**, 800 (2018).

[69] N. H. D. Khang, Y. Ueda, and P. N. Hai, A conductive topological insulator with large spin Hall effect for ultralow power spin–orbit torque switching, Nat. Mater., **17**, 808 (2018).

[70] D. Zhang, W. Jiang, H. Yun, O. J. Benally, T. Peterson, Z. Cresswell, Y. Fan, Y. Lv, G. Yu, J. G. Barriocanal et al., Robust negative longitudinal magnetoresistance and spin–orbit torque in sputtered Pt_3Sn and $Pt_3Sn_xFe_{1-x}$ topological semimetal, Nat. Commun., **14**, 4151 (2023).

[71] S. Bhatti, R. Sbiaa, A. Hirohata, H. Ohno, S. Fukami, and S. N. Piramanayagam, Spintronics Based Random Access Memory: A Review, Materials Today, **20** 530 (2017).

[72] A. Manchon, J. Železný, I. M. Miron, T. Jungwirth, J. Sinova, A. Thiaville, K. Garello, and P. Gambardella, Current-induced spin-orbit torques in ferromagnetic and antiferromagnetic systems, Rev. Mod. Phys., **91**, 035004 (2019).

[73] Q. Shao, P. Li, L. Liu, H. Yang, S. Fukami, A. Razavi, H. Wu, K. Wang, Y. Mokrousov et al., Roadmap of Spin-Orbit Torques, IEEE Trans. Magn., **57**, 800439 (2021).

[74] J. E. Hirsch, Spin Hall Effect, Phys. Rev. Lett., **83**, 1834 (1999).

[75] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, Spin Hall effects, Rev. Mod. Phys., **87**, 1213 (2015).

[76] Q. Shao, G. Yu, Y. W. Lan, Y. Shi, M. Y. Li, C. Zheng, X. Zhu, L. J. Li, P. K. Amiri, and K. L. Wang, Strong Rashba-Edelstein Effect-Induced Spin-Orbit Torques in Monolayer Transition Metal Dichalcogenide/ Ferromagnet Bilayers, Nano Lett., **16**, 7514 (2016).

[77] A. D. Caviglia, M. Gabay, S. Gariglio, N. Reyren, C. Cancellieri, and J. M. Triscone, Tunable Rashba Spin-Orbit Interaction at Oxide Interfaces, Phys. Rev. Lett., **104**, 126803 (2010).

[78] I. M. Miron, G. Gaudin, S. Auffret, B. Rodmacq, A. Schuhl, S. Pizzini, J. Vogel, and P. Gambardella, Current-driven spin torque induced by the Rashba effect in a ferromagnetic metal layer, Nat. Mater., **9**, 230 (2010).

[79] H. Wu, P. Zhang, P. Deng, Q. Pan, S. A. Razavi, X. Che, L. Huang, B. Dai, K. Wong et al., Room-Temperature Spin-Orbit Torque from Topological Surface States, Phys. Rev. Lett., **123**, 207205 (2019).

[80] Y. Wang, P. Deorani, K. Banerjee, N. Koirala, M. Brahlek, S. Oh, and H. Yang, Topological surface states originated spin-orbit torques in Bi_2Se_3, Phys. Rev. Lett., **114**, 257202 (2015).

[81] P. Sahu, Y. Yang, Y. Fan, H. Jaffrès, J.-Y. Chen, X. Devaux, Y. Fagot-Revurat, S. Migot, E. Rongione, T. Chen et al., Room Temperature Spin-to-Charge Conversion in Amorphous Topological Insulating Gd-Alloyed Bi_xSe_{1-x}/CoFeB Bilayers, ACS Appl. Mater. Interfaces, **15**, 38592 (2023).

[82] H. Lee, F. Ebrahimi, P. K. Amiri, and K. L. Wang, Design of high-throughput and low-power true random number generator utilizing perpendicularly magnetized voltage-controlled magnetic tunnel junction, AIP Adv., **7**, 055934 (2017).

[83] Y. Shao, V. Lopez-Dominguez, N. Davila, Q. Sun, N. Kioussis, J. A. Katine, and P. Khalili Amiri, Sub-volt switching of nanoscale voltage-controlled perpendicular magnetic tunnel junctions, Commun. Mater., **3**, 87 (2022).

[84] W. Kang, Y. Ran, Y. Zhang, W. Lv, and W. Zhao, Modeling and Exploration of the Voltage-Controlled Magnetic Anisotropy Effect for the Next-Generation Low-Power and High-Speed MRAM Applications, IEEE Trans. Nanotechnol., **16**, 387 (2017).

[85] Y. C. Wu, K. Garello, W. Kim, M. Gupta, M. Perumkunnil, V. Kateel, S. Couet, R. Carpenter, S. Rao, S. Van Beek et al., Voltage-Gate-Assisted Spin-Orbit-Torque Magnetic Random-Access Memory for High-Density and Low-Power Embedded Applications, Phys. Rev. Appl. **15**, 064015 (2021).

[86] Y. C. Chen, Q. Jia, Y. Yang, Y. H. Huang, D. Lyu, T. J. Peterson, and J. P. Wang, Enhanced Voltage-Controlled Magnetic Anisotropy and Field-Free Magnetization Switching Achieved with High Work Function and Opposite Spin Hall Angles in W/Pt/W SOT Tri-Layers, Adv. Funct. Mater., **35**, 2416570 (2024).

[87] P. Bruno, Theory of interlayer magnetic coupling, Phys. Rev. B, **52**, 411 (1995).

[88] A. Z. Pervaiz, B. M. Sutton, L. A. Ghantasala, and K. Y. Camsari, Weighted p-bits for FPGA implementation of probabilistic circuits, IEEE Trans. Neur. Netw. Learning Syst., **30**, 1920 (2019).

[89] R. Rahman and S. Bandyopadhyay, Robustness of Binary Stochastic Neurons Implemented with Low Barrier Nanomagnets Made of Dilute Magnetic Semiconductors, IEEE Magn. Lett., **13**, 4505104 (2022).

[90] 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**, 4502805 (2019).

[91] R. Zand, K. Y. Camsari, S. D. Pyle, I. Ahmed, C. H. Kim, and R. F. DeMara, Low-Energy Deep Belief Networks Using Intrinsic Sigmoidal Spintronic-Based Probabilistic Neurons, in Proceedings of the ACM Great Lakes Symposium on VLSI (GLSVLSI), pp. 15 – 20 (2018).

[92] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-bits for invertible logic, Phys. Rev. X, **7**, 031014 (2017).

[93] 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).

[94] A. Sengupta, C. M. Liyanagedera, B. Jung, and K. Roy, Magnetic Tunnel Junction as an On-Chip Temperature Sensor, Sci. Rep., **7**, 11764 (2017).

[95] S. Wang, H. Lee, C. Grezes, P. K. Amiri, K. L. Wang, and P. Gupta, Adaptive MRAM Write and Read with MTJ Variation Monitor, IEEE Trans. Emerg. Top Comput., **9**, 402 (2021).

[96] J. L. Drobitch and S. Bandyopadhyay, Robustness and Scalability of p-bits Implemented with Low Energy Barrier Nanomagnets, IEEE Trans. Magn., **10**, 4510404 (2019).

[97] P. Debashis, R. Faria, K. Y. Camsari, and Z. Chen, Design of Stochastic Nanomagnets for Probabilistic Spin Logic, IEEE Magn. Lett., **9**, 4305205 (2018).

[98] S. Z. Peng, Y. Zhang, M. X. Wang, Y. G. Zhang, and W. Zhao, Magnetic Tunnel Junctions for Spintronics: Principles and Applications, in Wiley Encyclopedia of Electrical and Electronics Engineering (Wiley, 2014), pp. 1–16.

[99] K. Y. Camsari, M. M. Torunbalci, W. A. Borders, H. Ohno, and S. Fukami, Double-Free-Layer Magnetic Tunnel Junctions for Probabilistic Bits, Phys. Rev. Appl., **15**, 044049 (2021).

[100] R. Ota, K. Kobayashi, K. Hayakawa, S. Kanai, K. Y. Çamsarı, H. Ohno, and S. Fukami, Voltage-insensitive stochastic magnetic tunnel

junctions with double free layers, Appl. Phys. Lett., **125**, 022406 (2024).
[101] K. Selcuk, S. Kanai, R. Ota, H. Ohno, S. Fukami, and K. Y. Camsari, Double-free-layer stochastic magnetic tunnel junctions with synthetic antiferromagnets, Phys. Rev. Appl., **21**, 054002 (2024).
[102] 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).
[103] 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).
[104] K. Hayakawa, S. Kanai, T. Funatsu, J. Igarashi, B. Jinnai, W. A. Borders, H. Ohno, and S. Fukami, Nanosecond Random Telegraph Noise in In-Plane Magnetic Tunnel Junctions, Phys Rev Lett **126**, 117202 (2021).
[105] L. Schnitzspan, M. Kläui, and G. Jakob, Nanosecond True-Random-Number Generation with Superparamagnetic Tunnel Junctions: Identification of Joule Heating and Spin-Transfer-Torque Effects, Phys. Rev. Appl., **20**, 024002 (2023).
[106] J. Z. Sun, C. Safranski, P. Trouilloud, C. D'Emic, P. Hashemi, and G. Hu, Easy-plane dominant stochastic magnetic tunnel junction with synthetic antiferromagnetic layers, Phys. Rev. B, **108**, 064418 (2023).
[107] A. Mizrahi, N. Locatelli, J. Grollier, and D. Querlioz, Synchronization of electrically coupled stochastic magnetic oscillators induced by thermal and electrical noise, Phys. Rev. B, **94**, 054419 (2016).
[108] L. Schnitzspan, M. Kläui, and G. Jakob, Electrical coupling of superparamagnetic tunnel junctions mediated by spin-transfer-torques, Appl. Phys. Lett., **123**, 232403 (2023).
[109] P. Talatchian, M. W. Daniels, A. Madhavan, M. R. Pufall, E. Jué, W. H. Rippard, J. J. McClelland, and M. D. Stiles, Mutual control of stochastic switching for two electrically coupled superparamagnetic tunnel junctions, Phys. Rev. B, **104**, 054427 (2021).
[110] S. Gibeault, T. N. Adeyeye, L. A. Pocher, D. P. Lathrop, M. W. Daniels, M. D. Stiles, J. J. McClelland, W. A. Borders, J. T. Ryan, P. Talatchian et al., Programmable electrical coupling between stochastic magnetic tunnel junctions, Phys. Rev. Appl., **21**, 034064 (2024).
[111] H. Kaneko, R. Ota, K. Kobayashi, S. Kanai, M. Elyasi, G. E. W. Bauer, H. Ohno, and S. Fukami, Temperature dependence of the properties of stochastic magnetic tunnel junction with perpendicular magnetization, Appl. Phys. Express, **17**, 053001 (2024).
[112] T. Funatsu, S. Kanai, J. Ieda, S. Fukami, and H. Ohno, Local bifurcation with spin-transfer torque in superparamagnetic tunnel junctions, Nat. Commun., **13**, 4079 (2022).
[113] N. S. Singh, K. Kobayashi, Q. Cao, K. Selcuk. T. Hu, S. Niazi, N. A. Aadit, S. Kanai. H. Ohno, S. Fukami, and K. Y. Camsari, CMOS plus stochastic nanomagnets enabling heterogeneous computers for probabilistic inference and learning, Nat. Commun., **15**, 2685 (2024).
[114] J. Daniel, Z. Sun, X. Zhang, Y. Tan, N. Dilley, Z. Chen, and J. Appenzeller, Experimental demonstration of an on-chip p-bit core based on stochastic magnetic tunnel junctions and 2D MoS2 transistors, Nat. Commun., **15**, 4098 (2024).
[115] K. Danouchi, G. Prenat, P. Talatchian, L. Hutin, and L. Anghel, Robustness and Power Efficiency in Spin-Orbit Torque-Based Probabilistic Logic Circuits, in Proceedings of IEEE Computer Society Annual Symposium on VLSI, (ISVLSI), pp. 1 -6 (2023).
[116] R. Zhang, X. Li, M. Zhao, C. Wan, X. Luo, S. Liu, Y. Zhang, Y. Wang, G. Yu, and X. Han, Probability-Distribution-Configurable True Random Number Generators Based on Spin-Orbit Torque Magnetic Tunnel Junctions, Adv. Sci., **11**, 2402182 (2024).
[117] P. Debashis, A. K. Maskay, P. Upadhyaya, and Z. Chen, Spin-orbit torque controlled stochastic oscillators with synchronization and frequency tunability, J. Appl. Phys., **131**, 123901 (2022).
[118] P.B. Alisha and Dr. T. S. Warrier, Optimizing free layer of Magnetic Tunnel Junction for true random number generator, Memories - Materials, Devices, Circuits and Systems, **5**, 100075 (2023).
[119] P. Debashis, R. Faria, K. Y. Camsari, S. Datta, and Z. Chen, Correlated fluctuations in spin orbit torque coupled perpendicular nanomagnets, Phys. Rev. B, **101**, 094405 (2020).
[120] J. Yin, Y. Liu, B. Zhang, A. Du, T. Gao, X. Ma, Y. Dong, Y. Bai, S. Lu, Y. Zhuo et al., Scalable Ising Computer Based on Ultra-Fast Field-Free Spin Orbit Torque Stochastic Device with Extreme 1-Bit Quantization, Int.

Electron Devices Meeting (IEDM), pp. 3611–3614 (2022).
[121] D. J. P. De Sousa, P. M. Haney, J. P. Wang, and T. Low, Field-Free-Switching State Diagram of Perpendicular Magnetization Subjected to Conventional and Unconventional Spin-Orbit Torques, Phys. Rev. Appl., **18**, 054020 (2022).
[122] A. Roy, M. H. D. Guimarães, and J. Sławińska, Unconventional spin Hall effects in nonmagnetic solids, Phys. Rev. Mater., **6**, 045004 (2022).
[123] Y. Yang, S. Lee, Y.-C. Chen. Q. Jia, D. Sousa, M. Odlyzko, J. Garcia-Barriocanal, G. Yu, G. Haugstad, Y. Fan et al., Giant spin Hall effect with multi-directional spin components in Ni4W, arXiv:2411.05682 (2024).
[124] V. Ostwal, P. Debashis, R. Faria, Z. Chen, and J. Appenzeller, Spin-torque devices with hard axis initialization as Stochastic Binary Neurons, Sci. Rep., **8**, 16689 (2018).
[125] Y. Shim, S. Chen, A. Sengupta, and K. Roy, Stochastic Spin-Orbit Torque Devices as Elements for Bayesian Inference, Sci. Rep., **7**, 14101 (2017).
[126] R. Ren, Y. Cao, C. Wang, Y. Guan, S. Liu, L. Wang, Z. Du, C. Feng, Z. A. Bekele, X. Lan et al., Initialization-Free and Magnetic Field-Free Spin-Orbit p-Bits with Backhopping-like Magnetization Switching for Probabilistic Applications, Nano Lett., **24**, 10072 (2024).
[127] X. Li, C. Wan, R. Zhang, M. Zhao, S. Xiong, D. Kong, X. Luo, B. He, S. Liu, J. Xia et al., Restricted Boltzmann Machines Implemented by Spin-Orbit Torque Magnetic Tunnel Junctions, Nano Lett., **24**, 5420 (2024).
[128] Y. Q. Xu, X. H. Li, R. Zhang, C. H. Wan, Y. Z. Wang, S. Q. Liu, X. M. Luo, G. B. Lan, J. H. Xia, G. Q. Yu et al., Self-stabilized true random number generator based on spin–orbit torque magnetic tunnel junctions without calibration, Appl. Phys. Lett., **125**, 132403 (2024).
[129] W. Kang, L. Chang, Y. Zhang, and W. Zhao, Voltage-Controlled MRAM for Working Memory: Perspectives and Challenges, Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 542-547 (2017).
[130] C. Duffee, J. Athas, Y. Shao, N. Davila Melendez, E. Raimondo, J. A. Katine, K. Y. Camsari, G. Finocchio, and P. K. Amiri, Integrated Probabilistic Computer Using Voltage-Controlled Magnetic Tunnel Junctions as Its Entropy Source, arXiv:241208617 (2024).
[131] J. Deng, V. P. K. Miriyala, Z. Zhu, X. Fong, and G. Liang, Voltage-Controlled Spintronic Stochastic Neuron for Restricted Boltzmann Machine with Weight Sparsity, IEEE Electron Device Letters, **41**, 1102 (2020).
[132] S. Nasrin, J. L. Drobitch, S. Bandyopadhyay, and A. R. Trivedi, Low Power Restricted Boltzmann Machine Using Mixed-Mode Magneto-Tunneling Junctions, IEEE Electron Device Letters, **40**, 345 (2019).
[133] S. Nasrin, J. Drobitch, P. Shukla, T. Tulabandhula, S. Bandyopadhyay, and A. R. Trivedi, Bayesian reasoning machine on a magneto-tunneling junction network, Nanotechnology, **31**, 484001 (2020).
[134] A. Fukushima, T. Yamamoto, T. Nozaki, K. Yakushiji, H. Kubota, and S. Yuasa, Recent Progress in Random Number Generator Using Voltage Pulse-Induced Switching of Nano-Magnet: A Perspective, APL Mater., **9**, 030905 (2021).
[135] S. Liu, J. Kwon, P. W. Bessler, S. G. Cardwell, C. Schuman, J. D. Smith, J. B. Aimone, S. Misra, and J. A. C. Incorvia, Random Bitstream Generation Using Voltage-Controlled Magnetic Anisotropy and Spin Orbit Torque Magnetic Tunnel Junctions, IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, **8**, 194 (2022).
[136] Q. Jia, Y.-C. Chen, D. Zhang, Y. Lv, S. Liang, O. J. Benally, Y. Yang, B. Dixit, D. Lyu, B. Zink, and J.-P. Wang, Ultrafast and Directional Magnetization Control via Voltage-Controlled Exchange Coupling, arXiv: 2504:06509 (2025).
[137] Y. Lv, B. Dixit, and J.-P. Wang, Modulation of Switching Dynamics in Magnetic Tunnel Junctions for Low-Error-Rate Computational Random-Access Memory, arXiv:2505.14829 (2025).
[138] Z. Zhao, A. K. Smith, M. Jamali, and J. Wang, External-Field-Free Spin Hall Switching of Perpendicular Magnetic Nanopillar with a Dipole-Coupled Composite Structure, Adv. Electron Mater., **6**, 1901368 (2020).
[139] 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).
[140] 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 Trans. Electron Devices, **66**, 5353 (2019).
[141] B. R. Zink and J.-P. Wang, Influence of Intrinsic Thermal Stability on Switching Rate and Tunability of Dual-Biased Magnetic Tunnel Junctions for Probabilistic Bits, IEEE Magn. Lett., **12**, 4501405 (2021).
[142] K. Yang and A. Sengupta, Stochastic magnetoelectric neuron for temporal information encoding, Appl. Phys. Lett., **116**, 043701 (2020).
[143] Y. Lv, R. P. Bloom, and J. P. Wang, Experimental Demonstration of Probabilistic Spin Logic by Magnetic Tunnel Junctions, IEEE Magn. Lett., **10**, 4510905 (2019).
[144] Y. Lv, B. R. Zink, and J. P. Wang, Bipolar Random Spike and Bipolar Random Number Generation by Two Magnetic Tunnel Junctions, IEEE Trans Electron Devices **69**, 1582 (2022).
[145] Z. Zhao, M. Jamali, N. D'Souza, D. Zhang, S. Bandyopadhyay, J. Atulasimha, and J. P. Wang, Giant voltage manipulation of MgO-based magnetic tunnel junctions via localized anisotropic strain: A potential pathway to ultra-energy-efficient memory technology, Appl. Phys. Lett., **109**, 092403 (2016).
[146] A. Chen, Y. Wen, B. Fang, Y. Zhao, Q. Zhang, Y. Chang, P. Li, H. Wu, H. Huang, Y. Lu et al., Giant nonvolatile manipulation of magnetoresistance in magnetic tunnel junctions by electric fields via magnetoelectric coupling, Nat. Commun., **10**, 243 (2019).
[147] K. Y. Camsari, R. Faria, O. Hassan, B. M. Sutton, and S. Datta, Equivalent Circuit for Magnetoelectric Read and Write Operations, Phys. Rev. Appl., **9**, 044020 (2018).
[148] N. Tiercelin, Y. Dusch, A. Klimov, S. Giordano, V. Preobrazhensky, and P. Pernod, Room temperature magnetoelectric memory cell using stress-mediated magnetoelastic switching in nanostructured multilayers, Appl. Phys. Lett., **99**, 192507 (2011).
[149] A. Klimov, N. Tiercelin, Y. Dusch, S. Giordano, T. Mathurin, P. Pernod, V. Preobrazhensky, A. Churbanov, and S. Nikitov, Magnetoelectric write and read operations in a stress-mediated multiferroic memory cell, Appl. Phys. Lett., **110**, 222401 (2017).
[150] M. T. McCray, M. A. Abeed, and S. Bandyopadhyay, Electrically programmable probabilistic bit anti-correlator on a nanomagnetic platform, Sci. Rep., **10**, 12361 (2020).
[151] R. Rahman, S. Ganguly, and S. Bandyopadhyay, Reconfigurable stochastic neurons based on strain engineered low barrier nanomagnets, Nanotechnology, **35**, 325205 (2024).
[152] N. D'Souza, M. Salehi Fashami, S. Bandyopadhyay, and J. Atulasimha, Experimental Clocking of Nanomagnets with Strain for Ultralow Power Boolean Logic, Nano Lett., **16**, 1069 (2016).
[153] A. K. Biswas, H. Ahmad, J. Atulasimha, and S. Bandyopadhyay, Experimental Demonstration of Complete 180° Reversal of Magnetization in Isolated Co Nanomagnets on a PMN-PT Substrate with Voltage Generated Strain, Nano Lett., **17**, 3478 (2017).
[154] A. K. Biswas, S. Bandyopadhyay, and J. Atulasimha, Complete magnetization reversal in a magnetostrictive nanomagnet with voltage-generated stress: A reliable energy-efficient non-volatile magneto-elastic memory, Appl. Phys. Lett., **105**, 072408 (2014).
[155] K. P. Kelley, D. E. Yilmaz, L. Collins, Y. Sharma, H. N. Lee, D. Akbarian, A. C. T. Van Duin, P. Ganesh, and R. K. Vasudevan, Thickness and strain dependence of piezoelectric coefficient in BaTiO3 thin films, Phys. Rev. Mater., **4**, 024407 (2020).
[156] R. Gupta and R. K. Kotnala, A Review on Current Status and Mechanisms of Room-Temperature Magnetoelectric Coupling in Multiferroics for Device Applications, J. Mater. Sci., **57**, 12710 (2022).
[157] B. Zink, B. Ma, D. Zhang, D. Bhattacharya, M. A. Abeed, S. Bandyopadhyay, J. Atulasimha, and J. P. Wang, Influence of surface acoustic wave (SAW) on nanoscale in-plane magnetic tunnel junctions, AIP Adv., **14**, 025104 (2024).
[158] D. Bhattacharya, P. Sheng, M. A. Abeed, Z. Zhao, H. Li, J. P. Wang, S. Bandyopadhyay, B. Ma, and J. Atulasimha, Surface acoustic wave induced modulation of tunneling magnetoresistance in magnetic tunnel junctions, J. Appl. Phys., **130**, 033901 (2021).
[159] A. Grimaldi, L. Mazza, E. Raimondo, P. Tullo, D. Rodrigues, K. Y. Camsari, V. Crupi, M. Carpentieri, V. Puliafito, and G. Finocchio, Evaluating Spintronics-Compatible Implementations of Ising Machines, Phys. Rev. Appl., **20**, 024005 (2023).
[160] M. Wang, W. Cai, D. Zhu, Z. Wang, J. Kan, Z. Zhao, K. Cao, Z. Wang, Y. Zhang, T. Zhang et al., Field-free switching of a perpendicular magnetic tunnel junction through the interplay

of spin–orbit and spin-transfer torques, Nat. Electron, **1**, 582 (2018).

[161] S. Fukami, C. Zhang, S. Duttagupta, A. Kurenkov, and H. Ohno, Magnetization switching by spin-orbit torque in an antiferromagnet-ferromagnet bilayer system, Nat. Mater., **15**, 535 (2016).

[162] G. Yu, P. Upadhyaya, Y. Fan, J. G. Alzate, W. Jiang, K. L. Wong, S. Takei, S. A. Bender, L.-T. Chang, Y. Jiang et al., Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields, Nat. Nanotechnol., **9**, 548 (2014).

[163] L. You, O. Lee, D. Bhowmik, D. Labanowski, J. Hong, J. Bokor, and S. Salahuddin, Switching of perpendicularly polarized nanomagnets with spin orbit torque without an external magnetic field by engineering a tilted anisotropy, Proc. Natl. Acad. Sci. U.S.A., **112**, 10310 (2015).

[164] Q. Ma, Y. Li, D. B. Gopman, Y. P. Kabanov, R. D. Shull, and C. L. Chien, Switching a Perpendicular Ferromagnetic Layer by Competing Spin Currents, Phys. Rev. Lett., **120**, 117703 (2018).

[165] D. Koh, Q. Wang, B. C. McGoldrick, C. T. Chou, L. Liu, and M. A. Baldo, Closed Loop Superparamagnetic Tunnel Junctions for Reliable True Randomness and Generative Artificial Intelligence, Nano Lett., **25**, 3799 (2025).

[166] Z. Diao, Z. Li, S. Wang, Y. Ding, A. Panchula, E. Chen, L.-C. Wang, and Y. Huai, Spin-transfer torque switching in magnetic tunnel junctions and spin-transfer torque random access memory, J. Physics: Cond. Matter **19**, 165209 (2007).

[167] Q. Chen, T. Min, T. Torng, C. Horng, D. Tang, and P. Wang, Study of dielectric breakdown distributions in magnetic tunneling junction with MgO barrier, J. Appl. Phys., **105**, 07C931 (2009).

[168] Dipanjan Mazumdar, Coherent Magnetotunneling Based on (001) Magnesium Oxide Barrier, PhD Dissertation, Brown University, 2007.

[169] Y. C. Liao, C. Pan, and A. Naeemi, Benchmarking and Optimization of Spintronic Memory Arrays, IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, **6**, 9 (2020).

[170] 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**, 1533 (2018).

[171] H. Yun, D. Lyu, Y. Lv, B. R. Zink, P. Khanal, B. Zhou, W. G. Wang, J. P. Wang, and K. A. Mkhoyan, Uncovering Atomic Migrations Behind Magnetic Tunnel Junction Breakdown, ACS Nano, **18**, 25708 (2024).

[172] N. Nguyen, G. Kaddoum, F. Pareschi, R. Rovatti, and G. Setti, A fully CMOS true random number generator based on hidden attractor hyperchaotic system, Nonlinear Dyn., **102**, 2887 (2020).

[173] R. Govindaraj, S. Ghosh, and S. Katkoori, CSRO-Based Reconfigurable True Random Number Generator Using RRAM, IEEE Trans. Very Large Scale Integr. (VLSI) Syst., **26**, 2661 (2018).

[174] E. Piccinini, R. Brunetti, and M. Rudan, Self-Heating Phase-Change Memory-Array Demonstrator for True Random Number Generation, IEEE Trans Electron Devices, **64**, 2185 (2017).

[175] Y. Liu, Q. Hu, Q. Wu, X. Liu, Y. Zhao, D. Zhang, Z. Han, J. Cheng, Q. Ding, Y. Han et al., Probabilistic Circuit Implementation Based on P-Bits Using the Intrinsic Random Property of RRAM and P-Bit Multiplexing Strategy, Micromachines, **13**, 924 (2022).

[176] C. Roques-Carmes, Y. Salamin, J. Sloan, S. Choi, G. Velez, E. Koskas, N. Rivera, S. E. Kooi, J. D. Joannopoulos, and M. Soljačić, Biasing the Quantum Vacuum to Control Macroscopic Probability Distributions, Science, **381**, 205 (2023).

[177] S. Luo, Y. He, B. Cai, X. Gong, and G. Liang, Probabilistic-Bits Based on Ferroelectric Field-Effect Transistors for Probabilistic Computing, IEEE Electron Device Letters, 44, 1356 (2023).

[178] S. Luo, Y. He, B. Cai, X. Gong, and G. Liang, Ferroelectric Probabilistic Bits Based on Thermal Noise Induced Randomness for Stochastic Computing, 7th IEEE Electron Devices Technology & Manufacturing Conference (EDTM), pp. 1 – 3 (2023).

[179] H. W. Chiang, C. F. Nien, H. Y. Cheng, and K. P. Huang, ReAIM: A ReRAM-Based Adaptive Ising Machine for Solving Combinatorial Optimization Problems, ACM/IEEE 51st Annual International Symposium on Computer Architecture (ISCA), pp. 58–72 (2024).

[180] J. Ouyang, Y. Liao, Z. Ma, D. Kong, X. Feng, X. Zhang, X. Dong, K. Cui, F. Liu, W. Zhang et al., On-demand photonic Ising machine with simplified Hamiltonian calculation by phase encoding and intensity detection, Commun. Phys., **7**, 168 (2024).

[181] W. Xu, J. Luo, Z. Fu, R. Han, S. Bao, K. Wang, Q. Huang, and R. Huang, Novel Ferroelectric-Based Ising Machine Featuring Reconfigurable Arbitrary Ising Graph and Controllable Annealing through Device-Algorithm Co-Optimization, Int. Electron Device Meeting (IEDM) pp. (2024). doi: 10.1109/IEDM50854.2024.10873528.

[182] S. Yang, A. Grimaldi, Y. Bao, E. Raimondo, J. Si, G. Finocchio, and H. Yang, 250 Magnetic Tunnel Junctions-Based Probabilistic Ising Machine, arXiv: 2506.14590 (2025).

[183] S. Liu, T. P. Xiao, J. Kwon, B. J. Debusschere, S. Agarwal, J. A. C. Incorvia, and C. H. Bennett, Bayesian neural networks using magnetic tunnel junction-based probabilistic in-memory computing, Frontiers in Nanotechnology **4**, 1021943 (2022).

[184] K. Zhang, D. Zhang, C. Wang, L. Zeng, Y. Wang, and W. Zhao, Compact Modeling and Analysis of Voltage-Gated Spin-Orbit Torque Magnetic Tunnel Junction, IEEE Access, **8**, 50792 (2020).

[185] V. Krizakova, E. Grimaldi, K. Garello, G. Sala, S. Couet, G. S. Kar, and P. Gambardella, Interplay of Voltage Control of Magnetic Anisotropy, Spin-Transfer Torque, and Heat in the Spin-Orbit-Torque Switching of Three-Terminal Magnetic Tunnel Junctions, Phys. Rev. Appl., **15**, 054055 (2021).

[186] P. Li, A. Chen, D. Li, Y. Zhao, S. Zhang, L. Yang, Y. Liu, M. Zhu, H. Zhang, and X. Han, Electric field manipulation of magnetization rotation and tunneling magnetoresistance of magnetic tunnel junctions at room temperature, Adv. Mater., **26**, 4320 (2014).

[187] Y. Shiota, S. Murakami, F. Bonell, T. Nozaki, T. Shinjo, and Y. Suzuki, Quantitative evaluation of voltage-induced magnetic anisotropy change by magnetoresistance measurement, Appl. Phys. Express, **4**, 043005 (2011).

[188] J. Maria De Teresa, A. Barthélémy, A. Fert, J. Pierre Contour, F. Montaigne, and P. Seneor, Role of Metal-Oxide Interface in Determining the Spin Polarization of Magnetic Tunnel Junctions, Science, **286**, 507 (1999).

[189] O. Wunnicke, N. Papanikolaou, R. Zeller, P. H. Dederichs, V. Drchal, and J. Kudrnovský, Effects of resonant interface states on tunneling magnetoresistance, Phys. Rev. B, **65**, 064425 (2002).

[190] W. Echtenkamp, B. Dixit, Y. Yang, D. Lyu, T. Peterson, Q. Jia, Y. C. Chen, and J. P. Wang, Prospects of Electric Field Control in Perpendicular Magnetic Tunnel Junctions and Emerging 2D Spintronics for Ultralow Energy Memory and Logic Devices, Adv. Funct. Mater., 2505426 (2025) https://doi.org/10.1002/adfm.202505426.

[191] S. Liang, T. Xie, N. A. Blumenschein, T. Zhou, T. Ersevium, Z. Song, J. Liang, M. A. Susner, B. S. Conner, S.-J. Gong et al., Small-voltage multiferroic control of two-dimensional magnetic insulators, Nat. Electron., **6**, 199 (2023).

[192] Q. Song, C. A. Occhialini, E. Ergeçen, B. Ilyas, D. Amoroso, P. Barone, J. Kapeghian, K. Watanabe, T. Taniguchi, A. S. Botana et al., Evidence for a single-layer van der Waals multiferroic, Nature, **602**, 601 (2022).

[193] C. Gong, E. M. Kim, Y. Wang, G. Lee, and X. Zhang, Multiferroicity in atomic van der Waals heterostructures, Nat. Commun., **10**, 2657 (2019).

[194] J. Eom, I. H. Lee, J. Y. Kee, M. Cho, J. Seo, H. Suh, H.-J. Choi, Y. Sim, S. Chen, H. J. Chang et al., Voltage control of magnetism in $Fe_{3-x}GeTe_2/In_2Se_3$ van der Waals ferromagnetic/ferroelectric heterostructures, Nat. Commun., **14**, 5605 (2023).

[195] X. Huang, L. Zhang, L. Tong, Z. Li, Z. Peng, R. Lin, W. Shi, K.-H. Xue, H. Dai, H. Cheng et al., Manipulating exchange bias in 2D magnetic heterojunction for high-performance robust memory applications, Nat. Commun., **14**, 2190 (2023).

[196] Y. Ou, W. Yanez, R. Xiao, M. Stanley, S. Ghosh, B. Zheng, W. Jiang, Y.-S. Huang, T. Pillsbury, A. Richardella et al., $ZrTe_2/CrTe_2$: an epitaxial van der Waals platform for spintronics, Nat. Commun., **13**, 2972 (2022).

[197] D. Zhong, K. L. Seyler, X. Linpeng, R. Cheng, N. Sivadas, B. Huang, E. Schmidgall, T. Taniguchi, K. Watanabe, M. A. McGuire et al., Van Der Waals Engineering of Ferromagnetic Semiconductor Heterostructures for Spin and Valleytronics, Science Advances, **3**, e1603113 (2017).

[198] C. Gong and X. Zhang, Two-dimensional magnetic crystals and emergent heterostructure devices, Science, **363**, eaav4450 (2019).

[199] T. Yamaguchi, Y. Inoue, S. Masubuchi, S. Morikawa, M. Onuki, K. Watanabe, T. Taniguchi, R. Moriya, and T. Machida, Electrical spin injection into graphene through monolayer hexagonal boron nitride, Appl. Phys. Express, **6**, 073001 (2013).