



HAL
open science

Programmable electrical coupling between stochastic magnetic tunnel junctions

Sidra Gibeault, Temitayo Adeyeye, Liam Pocher, Daniel Lathrop, Matthew Daniels, Mark Stiles, Jabez Mclelland, William Borders, Jason Ryan, Philippe Talatchian, et al.

► **To cite this version:**

Sidra Gibeault, Temitayo Adeyeye, Liam Pocher, Daniel Lathrop, Matthew Daniels, et al.. Programmable electrical coupling between stochastic magnetic tunnel junctions. *Physical Review Applied*, 2024, 21 (3), pp.034064. 10.1103/PhysRevApplied.21.034064 . hal-04666736

HAL Id: hal-04666736

<https://hal.science/hal-04666736v1>

Submitted on 2 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Programmable electrical coupling between stochastic magnetic tunnel junctions

Sidra Gibeault,¹ Temitayo N. Adeyeye,¹ Liam A. Pocher,¹ Daniel P. Lathrop,¹
Matthew W. Daniels,² Mark D. Stiles,² Jabez J. McClelland,² William A. Borders,²
Jason T. Ryan,² Philippe Talatchian,³ Ursula Ebels,³ and Advait Madhavan^{1,2}

¹*Institute for Research in Electronics and Applied Physics,
University of Maryland, College Park, MD, USA*

²*Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA*

³*Univ. Grenoble Alpes, CEA, CNRS, Grenoble INP, SPINTEC, 38000 Grenoble, France*

Superparamagnetic tunnel junctions (SMTJs) are promising sources of randomness for compact and energy efficient implementations of probabilistic computing techniques. Augmenting an SMTJ with electronic circuits, to convert the random telegraph fluctuations of its resistance state to stochastic digital signals, gives a basic building block known as a probabilistic bit or p -bit. Though scalable probabilistic computing methods connecting p -bits have been proposed, practical implementations are limited by either minimal tunability or energy inefficient microprocessors-in-the-loop. In this work, we experimentally demonstrate the functionality of a scalable analog unit cell, namely a pair of p -bits with programmable electrical coupling. This tunable coupling is implemented with operational amplifier circuits that have a time constant of approximately 1 μ s, which is faster than the mean dwell times of the SMTJs over most of the operating range. Programmability enables flexibility, allowing both positive and negative couplings, as well as coupling devices with widely varying device properties. These tunable coupling circuits can achieve the whole range of correlations from -1 to 1 , for both devices with similar timescales, and devices whose time scales vary by an order of magnitude. This range of correlation allows such circuits to be used for scalable implementations of simulated annealing with probabilistic computing.

I. INTRODUCTION

Probabilistic computing is becoming increasingly popular given its applicability to many classes of optimization problems and its amenability to hardware acceleration [1]. The constraints of a given problem are encoded as interactions between Ising spins, which vary randomly to explore configuration space. Conventional implementations either generate random numbers in a centralized location and distribute them to the computational cores [2], or use microprocessors with digital-to-analog converters in the feedback loop to mediate interactions [3, 4], neither of which are conducive to scaling. In this work, we use low barrier magnetic tunnel junctions (MTJs) to implement the Ising spins and analog circuits to mediate direct interactions between them, leading to scalable analog implementations of probabilistic computing kernels.

The central objects in probabilistic computers are probabilistic bits, also known as p -bits [5]. Unlike a conventional Boolean bit, which takes only deterministic $\mathbf{1}$ and $\mathbf{0}$ values, a p -bit uses its room-temperature instability to encode a probability value. This is particularly useful in applications such as energy-based machine learning models [6, 7], combinatorial optimization [8], and quantum simulation [1, 9]. These applications map their problem statements into graphs of interconnected binary nodes with variable interaction strengths [8]. The parameters of the problem at hand are encoded in the interaction strengths, while the collective configuration of the binary nodes encodes the current proposed solution. The interaction strengths define an effective energy for each global configuration [10]. If the state of

each node is probabilistic, the configuration evolves with time and can explore configuration space according to a Boltzmann distribution, tending to spend more time in configurations with lower effective energy. Optimal solutions or solutions within the required error margin can be found using simulated annealing [11], in which the interaction strengths are collectively increased in a way analogous to lowering the temperature so that the configuration spends more time in the lowest energy state.

Implementing such calculations on conventional computing systems can be time-consuming and inefficient. For large problems, the interconnection weights are necessarily stored on off-chip memory, while iterative calculations of the solution are performed on chip [12]. Since conventional computing systems have separate locations for memory and computation, most of the performance limitations arise from having to move data between memory and computation centers [13] (the so-called von Neumann bottleneck). If stochastically fluctuating p -bits are used as the nodes, a large binary state space can be searched efficiently, in terms of both area and energy cost. Large arrays of low-barrier versions of MTJs would be particularly appealing implementations of p -bits provided they can be coupled by locally engineered, electronically tunable interactions between them. This would allow memory (interaction strengths) and computation (accumulation and randomness generation) to be co-located, promising significant performance boosts [14].

Magnetic tunnel junctions are nanoscale magnetic devices that consist of two ferromagnetic layers separated by a thin insulator [15]. The insulating layer is thin enough for electrons to tunnel from one ferromagnetic layer to another. The resistance seen by the tunneling

electrons depends on the relative orientation of the magnetizations of the ferromagnetic layers [16]. There are two stable configurations of the magnetizations, namely parallel and antiparallel (denoted as P and AP). The device resistance, also known as the tunnel magnetoresistance (TMR), is smaller if the magnetizations are parallel to each other than if they are antiparallel [17]. The relative orientations between layers can be changed by applying an external magnetic field across the device or passing a current through the device [18, 19].

MTJs have various properties that make them useful as memory devices when integrated into modern complementary metal oxide semiconductor (CMOS) chips [20, 21]. Their stable configurations are used to encode binary values: for example, the parallel state may encode a value of **0** and the antiparallel state a value of **1**. When used in storage applications [22, 23] requiring long retention times (≈ 10 years), the energy barrier between stable configurations is engineered to be $\geq 40 kT$, where k is Boltzmann's constant and $T \approx 300$ K is room temperature. While MTJ-based magnetic random access memory (MRAM) is a promising candidate for various memory and unconventional computing applications [23–27], reducing the energy barrier reduces exponentially the retention time of the MTJ [28, 29].

For energy barriers less than $14 kT$, thermal fluctuations at room temperature switch the device between its stable configurations [30] at time scales faster than a millisecond. In this regime, devices are said to be superparamagnetic tunnel junctions (SMTJs). A current applied across the device allows its fluctuating resistance state to be read out as a random telegraph voltage signal and simultaneously controls the relative time that the device spends in each of its stable states [29]. The ability to generate tunable random signals makes SMTJs useful for applications where cheap, energy-efficient randomness is required [31–34]. Conventional CMOS sources of randomness such as linear feedback shift registers are only pseudo-random and require comparatively large area and energy budgets [31].

Though proposals of large-scale probabilistic computing systems using SMTJs as p -bits have been reported in literature [3, 8, 9], practical implementations are still limited to a handful of devices. The largest demonstrations use up to 80 p -bits and perform tasks such as integer factorization [3], invertible logic gates [5], learning Boolean functions [6] and travelling salesman problems [4]. Larger scale demonstrations are limited by practical factors such as a lack of integrated platforms for exploration, yield and repeatability failures in the fabrication process, the need for an external magnetic field, and differences in time scales between memory and computation. External magnetic fields in particular, which are required to set devices in the superparamagnetic regime in many experiments, are inefficient and impractical in commercial applications.

Mismatch of timescales is an especially important issue. In probabilistic computing, the states of the nodes

are sampled and that set of sampled states determines the input controlling the state of each of the nodes. A key requirement for correct operation of probabilistic circuits is that determining the input for each node must complete before the next sample is taken, so that the subsequent sample is representative of the input [5]. For free running interactions, this requirement means that the time scales of the interaction circuits (which can become relatively complex), must be much smaller than the timescales of nodes themselves. Previous demonstrations [3, 4, 6] measure the device state and perform the calculations of the accumulated interactions with a microprocessor in the loop. This approach works because the devices have switching times in the tens of milliseconds or more, which is enough for thousands of microprocessor cycles to perform computations digitally.

Although using a microprocessor in the loop is useful to demonstrate the functionality and expressibility of probabilistic computing, it is impractical for scalable implementations. Recent papers report simple ways of directly coupling devices. When two SMTJs are connected in parallel [35, 36] or in series [37, 38] to a voltage source via a tunable resistor, the current flowing through one device depends on the state of the other. Though such approaches can scale theoretically [39], practical problems such as line resistance, device-to-device variability, and noise quickly overwhelm linear circuit solutions as the number of devices grows. Moreover, such simple connections between devices are not enough to implement the large scale interactions required to build practical systems.

The fastest and most energy-efficient implementations would consist of CMOS integrated with SMTJs in the back end of the line [25]. Such implementations would involve analog stages mediating interactions between SMTJ-based p -bit unit cells. In this work, we take a step in that direction with commercial off-the-shelf analog circuits. We experimentally demonstrate the functionality of a unit cell coupling circuit between p -bits, with direct analog interaction between field-free SMTJs that have operational time scales in the tens to hundreds of microseconds. The speeds of the analog computation circuits themselves are much faster, on the order of $1 \mu\text{s}$, allowing the devices to faithfully couple to each other. The degree of coupling can be tuned by varying the interaction strengths, while the uncoupled probabilities of the devices can be independently set by adjusting bias currents. These two features, coupled with a programmable gain circuit, provide the necessary ingredients to perform optimization based on simulated annealing. This unit cell can be readily adapted to couple more than two devices, making it a useful stepping stone towards integration of arrays of devices.

The rest of this paper is organized as follows. Section II describes the coupling circuit and discusses how this circuit can be used to perform simulated annealing. Section III contains a description of the experimental setup and the details of the specific SMTJs used. Section IV

presents the experimental results of coupling devices of various time scales. Sec. V describes a Markov model that accurately describes the switching time scales of the two-SMTJ coupled system, and Sec. VI discusses scalability of the coupling circuit.

II. SMTJ INTERACTIONS

A. Interaction circuit

Our coupling circuit allows two SMTJs to mutually influence the current through each other. The circuit senses the resistive state of one SMTJ and adds a differential current to a second SMTJ depending to the state of the first. The strength of the coupling can be tuned by adjusting the magnitude of the differential current. We make the coupling bi-directional, meaning that both SMTJs are influencing the current through each other, by using two independent, uni-directional interaction circuits. The uni-directional interaction circuit, shown in Fig. 1, consists of several operational amplifier stages that enable the raw, noisy SMTJ voltage to influence a second SMTJ. These signal conditioning stages, which can be classified as either input or output stages, have specific functions in the interaction pipeline.

The gain stage, which is the first stage in the input pipeline, receives its input from the last stage of the output pipeline of the previous device as shown in Fig 1. Its function is to set the strength of the coupling by adjusting the signal swing. This is done by setting the value of the R_{GS} resistor, which together with the R_{gain} resistor determines a programmable gain value $G = R_{gain}/R_{GS}$. The gain stage typically reduces the voltage spread of its digital input signal to a desired level in order to correctly influence the device it controls. In section VI, we describe how the gain stage can be augmented to support multiple SMTJ inputs for scaled implementations. The next stage in the input pipeline is the level shifting stage which shifts the mean voltage of the gain-stage signal to some pre-determined adjustable level. This level is typically appropriate for biasing the influenced device to the equal probability state. Therefore, when no input signal is received by the gain stage, a static current biases the influenced device to its maximum entropy state. The transconductance stage, which is the last stage of the input pipeline, turns the shifted voltage signal into a proportional two-level current, which is then input to the next SMTJ. Since the differential part of the current is proportional to G , the strength of the coupling from the previous SMTJ is adjusted through the gain.

The output pipeline consists of a single stage which is the threshold stage. The noisy signal from the SMTJ is converted into a digital output signal based on a programmable reference voltage which is controlled by I_{fixed} and R_{TH} . There is also a polarity selection in the threshold stage that can be used to invert the state of the device, effectively producing positive or negative coupling.

The gain, SMTJ bias and threshold reference voltage are programmed via digital potentiometers in this implementation.

B. SMTJ model and simulated annealing

The modified Néel-Brown model [29, 40] describes how the current flowing through an SMTJ affects its probability of being in either magnetoresistive state. The mean dwell time in each state at a current I is

$$\tau_{\pm}(I) = \tau_0 \exp \left[-\frac{\Delta E}{kT} \left(1 \mp \frac{I - I_0}{I_c} \right) \right] \quad (1)$$

where τ_0 is the characteristic time scale, ΔE is the energy barrier, T is the temperature, here nominally room temperature, k is the Boltzmann's constant, I_c is the critical current, and I_0 is the current at which the device spends equal time in each state. Here τ_+ is the mean dwell time in the antiparallel (AP) state and τ_- is the mean dwell time in the parallel (P) state.

The probability of finding the device in either state can be written as the fraction of total time spent in that state

$$\begin{aligned} P_{\pm}(I) &= \frac{\tau_{\pm}(I)}{\tau_{\pm}(I) + \tau_{\mp}(I)} \\ &= \frac{1}{1 + \exp \left[\mp \frac{2\Delta E}{kT} \left(\frac{I - I_0}{I_c} \right) \right]}. \end{aligned} \quad (2)$$

By changing the current applied to an SMTJ, the probability of being in either resistive state also changes.

Our implementation of two balanced unidirectional interaction circuits can be thought of as the simplest unit cell of an Ising machine [41] and is capable of performing simulated annealing. In simulated annealing, the objective is to find the ground state of spins in an interconnected graph that minimizes the global energy as described by the Ising Hamiltonian [1]. For our simple circuit, the Hamiltonian is determined by the weights of the two-edge graph (uni-directional interactions) connecting the SMTJs. The joint state of the SMTJ-pair that minimizes the energy is the desired solution. In this way, simulated annealing can be performed by starting with a gain G of 0 and gradually increasing it.

During simulated annealing, the probability of being in the i -th joint state, e.g. (P,P), (AP,AP), (P,AP), (AP,P), is determined by a Boltzmann distribution

$$P_i(T) = \frac{1}{Z(T)} \exp \left[-\frac{E_i(G)}{kT} \right], \quad (3)$$

where $E_i(G)$ is the effective energy of the i -th joint state at gain value G and $Z(T) = \sum_i \exp [-E_i(G)/kT]$. This effective energy can be linearly related to a model energy $E_i(G) = GE_i$. Here, $\bar{E}_i = J_{12}S_1^iS_2^i$ where S_n^i is the state of SMTJ n in configuration i and J_{12} is the nominal coupling strength for that pair.

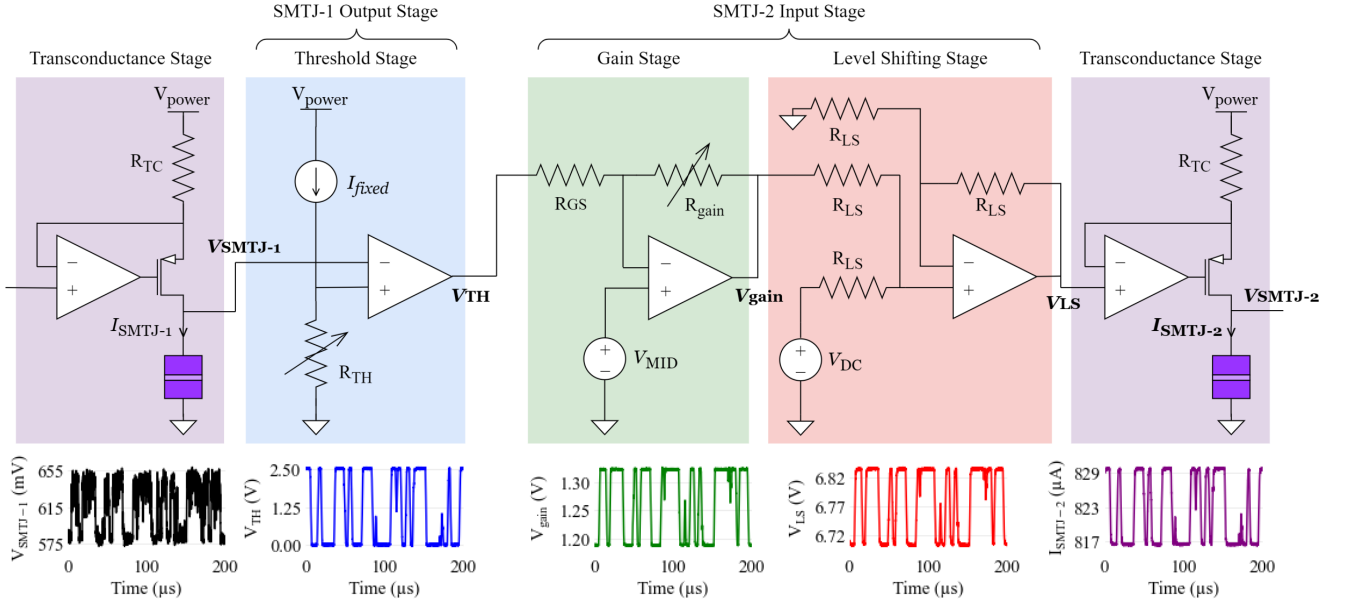


FIG. 1. Uni-directional SMTJ interaction circuit. Two copies of this circuit are used, one in each direction, to achieve bi-directional coupling. The SMTJ voltage gets sensed by a threshold stage, which converts the stochastic fluctuations to a digital output signal using a programmable reference voltage set by I_{fixed} and R_{TH} . This digital signal is passed through a gain stage, which adjusts its magnitude while retaining the mid voltage level. The gain $G = R_{gain}/R_{GS}$ is programmable. The level shifting stage adjusts the mid voltage level while retaining the voltage difference between high and low levels. V_{LS} is adjustable and controls the size of the shift. The shifted voltage is converted by a transconductance stage to a fluctuating current, which facilitates influencing the next SMTJ. In this example, the current in the next SMTJ is an inverted version of the SMTJ state, but both positive and negative coupling are possible with our circuit. In this example, $R_{TH} = 615 \Omega$, $I_{fixed} = 1 \text{ mA}$, $G = 0.05$, $V_{MID} = 1.25 \text{ V}$, and $V_{DC} = 5.5 \text{ V}$.

Rewriting the probability in terms of the model energy we get the following relation:

$$P_i(T_{\text{eff}}) = \frac{1}{Z} \exp \left[\frac{-\tilde{E}_i}{kT_{\text{eff}}} \right]. \quad (4)$$

Here, $T_{\text{eff}} = T/G$ is the effective temperature of the model.

The process of simulated annealing proceeds as follows. Initially, the two-device circuit, which is at room temperature, has gain equal to zero. Equation 4 shows that even though the system is at room temperature, the model temperature diverges in the sense that all joint SMTJ states are equally probable; the devices are effectively uncoupled. As the gain increases, the interaction circuit causes the devices to become correlated. The model temperature drops and the system spends more time in lower-energy joint states as prescribed by Boltzmann statistics. In conventional simulated annealing, the system would finally converge to a solution with the lowest energy; in our case, at maximal gain, the system still has enough energy from room temperature to explore low-energy degenerate states. SMTJs have a barrier breakdown current [42, 43] that imposes a limitation on how high the gain can be set without destroying the SMTJs. Higher breakdown currents would allow us to lower the temperature further for better convergence, but infinite gain (zero temperature) would still not be possible.

III. EXPERIMENTAL SETUP

To demonstrate the coupling of two SMTJs, we use perpendicular magnetic tunnel junctions with the following stack sequence: Si base / SiO₂ / TaN / [Co(0.5)/Pt(0.2)]₆ / Ru(0.8) / [Co(0.6)/Pt(0.2)]₃ / Ta(0.2) / Co(0.9) / W(0.25) / CoFeB(1) / MgO(0.8) / CoFeB(1.4) / W(0.3) / CoFeB(0.5) / MgO(0.75) / Ta(150) / Ru(8). The numbers in parentheses refer to layer thicknesses in nanometers, while the numbers beside square brackets show the bilayer repetitions. These devices, almost circular in shape, have approximate diameters of 150 nm. They exhibit a TMR near 120 % at room temperature and possess a resistance-area product of $10 \Omega \mu\text{m}^2$. These devices were initially developed as multifunctional magnetic tunnel junctions, useful for both non-volatile memory and radio frequency (RF) applications [42, 43]. Barrier breakdown has been observed in these devices at currents around 1 mA (corresponding to voltages around 0.7 V).

Field-free stochastic magnetization fluctuations have been observed in these perpendicular MTJ devices for a large range of diameters; for details see Ref. [44]. These devices allow for field-free fluctuations for two reasons. First, at the CoFeB thickness of 1.4 nm, the perpendicular magnetic anisotropy [45] nearly cancels the self-demagnetization field, leaving only a small energy bar-

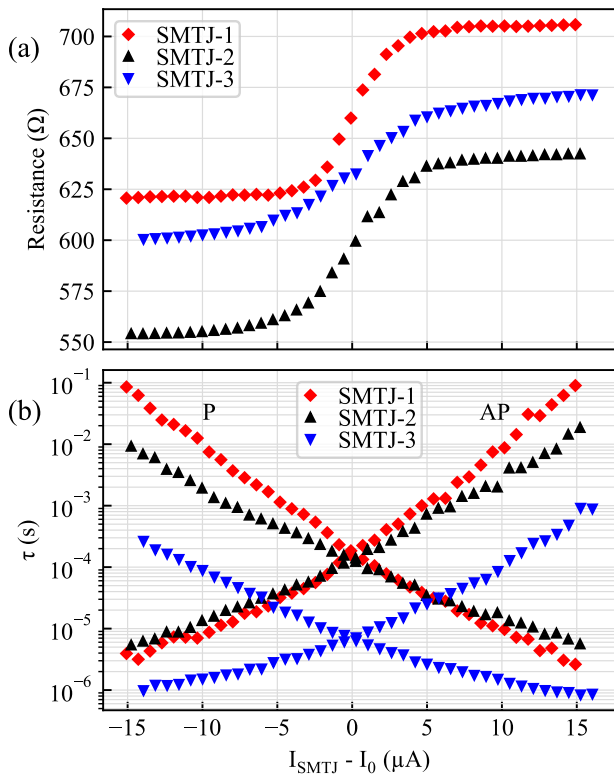


FIG. 2. Characteristics of the SMTJs used in our experiments. (a) Average resistance as a function of applied current. (b) Mean dwell time in the P (negative slope curves) and AP (positive slope) states as a function of applied current. SMTJ-1 and SMTJ-2 have similar timing characteristics, while SMTJ-3 switches an order of magnitude faster on average. The current at which the SMTJ is in the AP state 50 % of the time is denoted as I_0 , which is 0.95 mA, 0.90 mA, and 1.00 mA for SMTJ-1, SMTJ-2, and SMTJ-3, respectively. Statistical uncertainties are smaller than the plotting symbols.

rier. Second, the fringing field from the fixed layer is small enough that the energies of the parallel and antiparallel configurations can be aligned with relatively small external magnetic fields or moderate currents. Field-free operation of the coupling experiments ensures that operation remains practical and commercially viable. However, in the absence of a field, the current required to set these SMTJs to 50 % probability of being in each state is quite high (≈ 0.95 mA), necessitating operation close to barrier failure and suggesting refinement of the device design for future applications.

Figure 2(a) shows the average resistances of the SMTJs used in our experiments as a function of applied current. I_0 here is defined as the current at which the probabilities of being in either device state are equal. The resistances are low and almost independent of current for currents much below I_0 because the SMTJs are stable in the parallel state. As the current through the device approaches I_0 , the increasing spin-transfer torque causes the device to start switching stochastically, spending increased time

in the antiparallel state. This corresponds to the sharp increase in resistance between $-3 \mu\text{A}$ to $3 \mu\text{A}$. As currents are further increased, the rate of resistance increase slows down since the device spends all of its time in the antiparallel state.

Figure 2(b) shows the mean dwell times of the devices in the parallel and antiparallel states as a function of applied current. The parallel state dwell times are monotonically decreasing, while the antiparallel state dwell times are monotonically increasing. This is again by virtue of lower currents stabilizing the parallel state, while large currents stabilize the antiparallel state. Although the resistance values of SMTJ-1 and SMTJ-2 shown in Fig. 2(a) are considerably different, the difference in resistance does not affect the coupling dynamics, because the coupling circuit allows us to set the threshold voltage and I_0 separately for each SMTJ. We perform experiments with this pair of devices as the “ideal” case, treating them as identical due to their similar time scales over the current range of Fig. 2(b). On the other hand, SMTJ-3, whose resistance is between SMTJ-1 and SMTJ-2, is roughly an order of magnitude faster than the other two devices. Experiments done with SMTJ-2 and SMTJ-3 study the effect of coupling devices that have different time scales.

The printed circuit board (PCB) designed for this experiment performs multiple functions. It receives inputs from the probe station on which the SMTJ chip resides via coaxial cables connected to ground-signal (GS) probes. Each stage of the interaction circuit has outputs that can be connected to an oscilloscope for debugging and data collection. Peripheral bias and control inputs required to adjust the I_0 currents, threshold voltages, and interaction strengths are controlled by a microcontroller via a standard communication protocol. Both the microcontroller and the oscilloscope are connected to a host computer that performs all of the control operations, data collection, and analysis in software.

There is approximately $1 \mu\text{s}$ of propagation delay in our circuit; when one SMTJ switches, it takes $\approx 1 \mu\text{s}$ for the current through the other SMTJ to change in response. This delay makes our circuit sub-optimal for high-speed applications; however, more specialized amplifiers could be chosen to reduce the propagation delay. As discussed in Sec. VI, the delay would also be significantly reduced in an integrated circuit compared to a PCB.

We conduct the experiment as follows. First, both SMTJs are provided their respective I_0 currents to ensure that the probabilities of being in each state are equal. Then the gain is increased, increasing the effect that each SMTJ has on the other. Voltage-time traces are collected from each SMTJ at specific increments of gain. We then conduct statistical analyses on these voltage-time traces to demonstrate that increased gain concomitantly increases the degree of influence between the SMTJs.

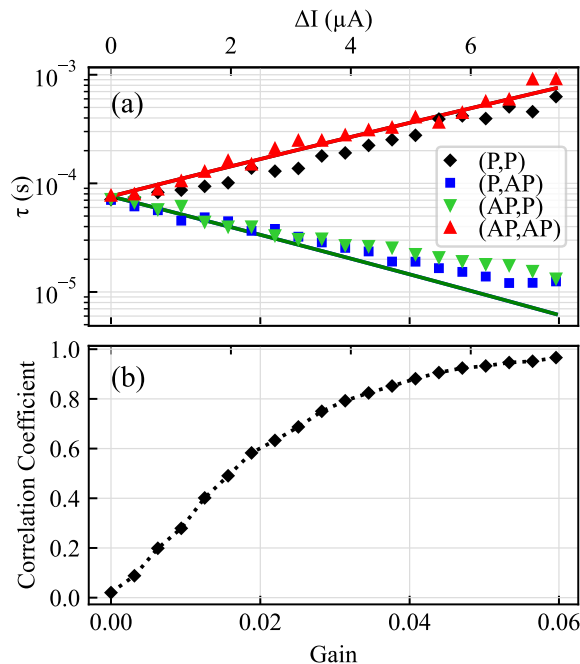


FIG. 3. Demonstration of positive coupling between SMTJ-1 and SMTJ-2, which have similar switching time scales. (a) Log-linear plot of mean dwell times in each of the four joint states the pair of SMTJs can be in. Positive coupling causes the SMTJs to spend more time in the same state (both in P or both in AP) and less time in opposite states with increasing gain. Solid lines are calculated mean joint dwell times using the Markov model described in Section V. (b) Pearson correlation coefficient between SMTJ voltage-time traces as gain is increased from 0 to 0.06. Both top and bottom axes apply to both panels. Statistical uncertainties are smaller than the plotting symbols.

IV. RESULTS

The SMTJs become more strongly coupled as the stochastic time between one SMTJ switching and the second SMTJ switching in response is reduced. If we consider the four joint states of the two-SMTJ system (P,P), (P,AP), (AP,P), and (AP,AP), we can define joint dwell times to be the mean dwell times in each of these four joint states. With positive coupling, when the system is in the (P,P) or (AP,AP) states, the currents through both SMTJs are such that there is a very low probability of either SMTJ switching out of its current state. We thus expect the dwell times in joint states (P,P) and (AP,AP) to increase with gain, and the dwell times in joint states (P,AP) and (AP,P) to decrease with gain, which is the behavior seen in Fig. 3(a). In the negatively coupled case, the joint states (P,AP) and (AP,P) are stabilized, so the dwell times in these states increase with gain as seen in Fig. 4(a).

Note that in Fig. 3(a), the (P,AP) and (AP,P) dwell times decrease linearly until around $G = 0.03$, where the

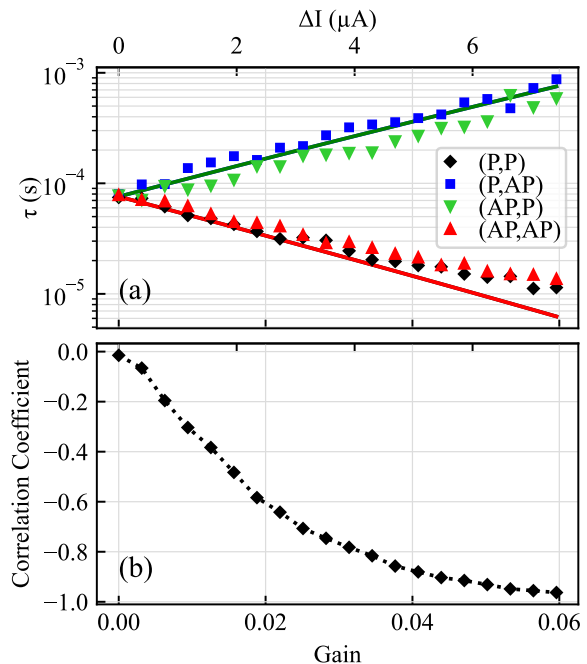


FIG. 4. Demonstration of negative coupling between SMTJ-1 and SMTJ-2, which have similar switching time scales. (a) Log-linear plot of mean dwell times in each of the four joint states the pair of SMTJs can be in. Negative coupling causes the SMTJs to spend more time in opposite states (one in P and the other in AP) and less time in the same state with increasing gain. Solid lines are calculated mean joint dwell times using the Markov model described in Section V. (b) Pearson correlation coefficient between SMTJ voltage-time traces as gain is increased from 0 to 0.06. Both top and bottom axes apply to both panels. Statistical uncertainties are smaller than the plotting symbols.

curve begins to flatten out. This is a sampling artifact. As the gain increases, the shorter dwell times start to approach the sampling times. As these times become closer some short-time events get missed, making the mean dwell times of the shorter events artificially larger, but also combining two of the long-time events, making those mean dwell times artificially longer as well. Computing dwell times requires saving entire time traces, making faster sampling prohibitive. However, these issues do not affect the calculation of the equal-time correlation coefficients between the two tunnel junctions, which is the main quantity of interest as a measure of the joint interaction between the two SMTJs.

The Pearson correlation coefficient can be used to quantify the similarity between the (digitized) SMTJ voltage-time traces for both positive and negative coupling (Fig. 3(b) and Fig. 4(b), respectively). A Pearson correlation coefficient of 1 indicates high similarity between the two voltage-time traces, namely that the SMTJs spend a lot of time in the correlated joint states (P,P) and (AP,AP), and very little time in the anti-

correlated joint states (P,AP) and (AP,P). A Pearson correlation coefficient of -1 indicates that the SMTJs are spending a lot of time in the joint states (P,AP) and (AP,P), and very little time in the joint states (P,P) and (AP,AP). A Pearson correlation coefficient of 0 indicates equal time in correlated and anti-correlated states. In the positive coupling experiment, the Pearson correlation starts at 0 and approaches 1 in the high gain limit. The negative coupling experiment has opposite dominant joint states to the positive coupling case, which results in the Pearson correlation approaching -1 .

Previous methods of coupling between SMTJs [35, 36, 38] which facilitate interaction via linear circuit elements have yielded a maximum Pearson correlation coefficient of 0.35. Such a low maximal correlation coefficient, coupled with the fact that these methods use a single resistor to control the interaction between their constituent SMTJs, make such approaches difficult to scale for simulated annealing tasks. On the other hand, our analog interaction circuit has a pair of independently programmable resistors for each pairwise set of SMTJs. It also features a global gain resistor that selects the effective temperature of the problem. Together, these features result in a Pearson correlation coefficient that approaches 1. This implies that our method is more viable for simulated annealing, where it is important to implement programmable pairwise interactions and be able to lower the effective temperature as much as possible.

The results discussed so far have demonstrated an analog circuit that facilitates tunable coupling between SMTJs with similar time scales. Figure 5 shows that the circuit successfully couples devices with very different time scales. When the positive coupling experiment is repeated with SMTJs of different time scales, the Pearson correlation and joint dwell times, plotted in Fig. 5, are very similar to the results in Fig. 3, aside from the joint state time scales being overall faster. In the next section, we examine how this difference in switching time scales affects the system's relaxation to equilibrium.

V. MARKOV MODELLING

Since the coupling between SMTJs in our experiments is facilitated entirely by analog circuitry, device properties such as TMR and I_0 have less of an impact on the coupling strength or on the behavior of the coupled system than in recent SMTJ coupling work [35, 38]. However, the difference in time scales of the coupled SMTJs does have a sizeable impact on the dynamics of the coupled system. To examine this effect, we can look at the eigenvalues of the Markov transition matrix for this system. This is a 4×4 matrix containing the transition rates between each of the four joint states (P,P), (P,AP), (AP,P), and (AP,AP), which we index as states 00, 01, 10, and 11, respectively.

We consider transitions where both SMTJs switch simultaneously (for instance, a switch from (P,P) to

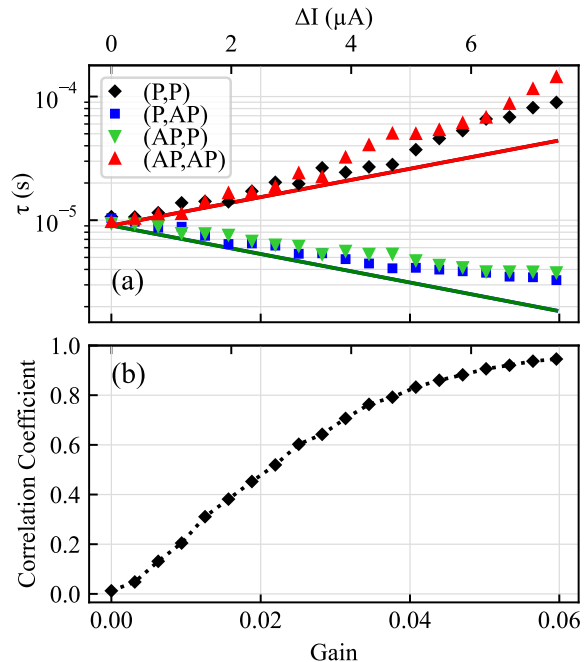


FIG. 5. Demonstration of positive coupling between SMTJ-2 and SMTJ-3, which have different switching time scales. (a) Log-linear plot of mean dwell times in each of the four joint states the pair of SMTJs can be in. Positive coupling causes the SMTJs to spend more time in the same state (both in P or both in AP) and less time in opposite states with increasing gain. Solid lines are calculated mean joint dwell times using the Markov model described in Section V. (b) Pearson correlation coefficient between SMTJ voltage-time traces as gain is increased from 0 to 0.06. Both top and bottom axes apply to both panels. Statistical uncertainties are smaller than the plotting symbols.

(AP,AP)) to have probability zero; all nonzero transition rates in our system represent single-SMTJ switching events. The switching rate of one SMTJ is dependent on the state of the other SMTJ due to the coupling; in the positive coupling case, when SMTJ-1 is in the P state, the current through SMTJ-2 will be $I_0 - \Delta I$. When SMTJ-1 is in the AP state, the current through SMTJ-2 will be $I_0 + \Delta I$. Referring to Fig. 6(a), we can then determine which inverse rate is the correct one for each of the eight possible transitions from one state into another. For example, $1/\tau_a^+$ is the rate for the first SMTJ to switch from the antiparallel state to the parallel state when the second SMTJ is in the antiparallel state, and $1/\tau_a^-$ is the rate for the first SMTJ to switch from the parallel state to the antiparallel state when the second SMTJ is in the parallel state.

The transition rate diagram is shown in Fig. 6(b), with Fig. 6(a) showing the relative locations of the inverse switching rates with respect to each other on a log-linear plot of mean dwell time against current. The notation in Fig. 6 can be simplified with the assumption that both

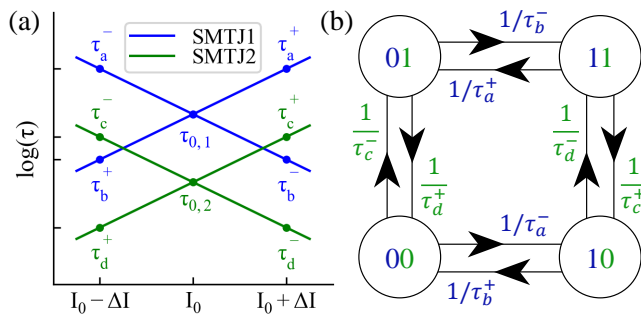


FIG. 6. (a) Conceptual mean dwell time plot (similar to Fig. 2(b)) showing the inverse switching rates in the two-SMTJ coupled system. Positive slope lines represent dwell times in the AP state, while negative slope lines represent dwell times in the P state. Note the log scale on the vertical axis. The mean dwell times in the P and AP states are equal at I_0 , and these dwell times are labeled as $\tau_{0,1}$ and $\tau_{0,2}$. The remaining labels are dwell times at either $I_0 + \Delta I$ or $I_0 - \Delta I$, which are used to write the transition rates from one state into another. (b) Markov transition diagram for the two-SMTJ coupled system. The switching rate out of each state is effectively a single-device switching rate, which is dependent on the state of the other SMTJ due to coupling.

devices have the same slopes for the current dependences, in which case $\tau_n^+ = \tau_n^- = \tau_n$, where $n = \{a, b, c, d\}$. Since we are interested in examining how the difference in time scales of the two devices impacts the coupled dynamics, we represent the difference in time scales by a scaling factor r on τ_0 via

$$\tau_{0,2} = \frac{1}{r} \tau_{0,1}. \quad (5)$$

Note that r will always be greater than 1 since we will always choose SMTJ1 to be the slower device. Rewriting the τ_n s in terms of $\tau_{0,1}$ yields

$$\begin{aligned} \tau_a &= g\tau_{0,1} & \tau_c &= \frac{g}{r}\tau_{0,1} \\ \tau_b &= \frac{1}{g}\tau_{0,1} & \tau_d &= \frac{1}{gr}\tau_{0,1}, \end{aligned} \quad (6)$$

where $g = \exp[B\Delta I]$ represents the change in the dwell times as a function of gain (G), since ΔI is linearly related to the gain. We would like to emphasize that although the gain G is proportional to the current increment ΔI , g is an exponential function of both G or equivalently ΔI . For the sake of clarity, we have chosen to make g for the two SMTJs equal. This implies that B for the two SMTJs are equal, which would make the slopes of the lines in Fig. 6(a) equal.

To generate a Markov model for our coupled system, we refer to Appendix B of Ref. [35]. The Markov transition rate matrix \mathbf{M} can be directly generated from Fig. 6(b) using the simplifications in Eq. 6. The steady state probability distribution is given by the eigenvector

corresponding to the zero eigenvalue of \mathbf{M} , namely

$$v_0 = \left(1, \frac{1}{g^2}, \frac{1}{g^2}, 1\right). \quad (7)$$

Note that this eigenvector, which encodes the (unnormalized) steady state distribution, is only a function of g , and the probability of being in the two disfavored states decreases quadratically with the value of g . Note that g itself depends exponentially on the gain.

In addition to this zero eigenvalue, there are three negative eigenvalues. The one closest to zero, provides the slowest rate of decay into the steady state distribution. In order to see how selecting SMTJs with different time scales affects the coupled dynamics, we can look at the eigenvalue that produces the term which decays most slowly,

$$\begin{aligned} \lambda_1 &= \frac{1}{\tau_{0,1}} \frac{-b + \sqrt{-16g^2r + b^2}}{2g} \\ b &= (g^2 + 1)(r + 1), \end{aligned} \quad (8)$$

which depends on both g and r . So although the final distribution is governed by g alone, the time scale of the coupled system's relaxation to equilibrium increases not only with gain but also with a higher ratio of SMTJ time scales.

We use this model to predict the mean joint dwell times during the coupled experiments. These model predictions are plotted as solid lines on the experimental dwell time plots in Section IV (Fig. 3(b), Fig. 4(b), and Fig. 5(b)). To apply the model to our experiment, we drop the assumption that B is the same for both SMTJs, reflecting the different slopes in Fig. 2. The result of this is that each SMTJ has its own g value, and the steady state distribution becomes dependent on r in addition to g . We fit lines to the characteristic dwell time plots of Fig. 2 to obtain B , which we use to calculate g for each SMTJ. $\tau_{0,1}$ and r were also determined from Fig. 2 for each pair of SMTJs. The negative reciprocals of the diagonal elements of \mathbf{M} give the dwell times in each of the four joint states for a given ΔI .

VI. DISCUSSION

The analog circuits used to facilitate coupling in this paper can be scaled up for larger bench-top experiments or integrated implementations with simple modifications. The threshold stage, used to convert noisy SMTJ signals to clean digital levels can be implemented by skewed inverter circuits with a digitally programmable threshold, or an analog latch-based sense amplifier design. Such circuits have operating speeds < 150 ps and are composed of < 10 transistors [46, 47]. Transconductance amplifiers for level shifting bias voltages and adding differential currents can also be practically implemented at nanosecond time scales. SMTJ switching speeds can be significantly

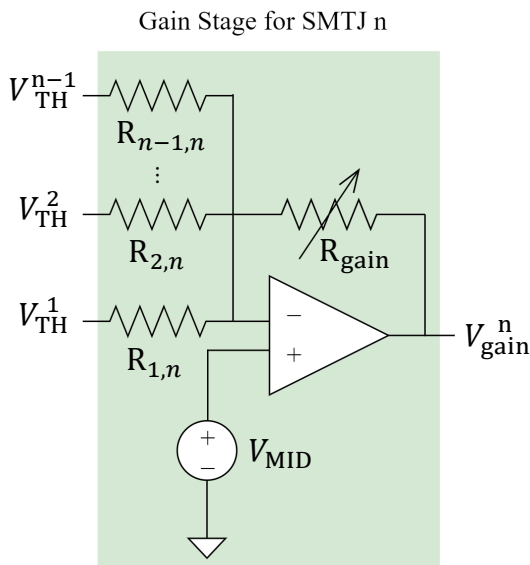


FIG. 7. Interaction circuit gain stage in a network of n SMTJs. Each SMTJ's gain stage receives voltage inputs from the threshold stages of the other $n - 1$ SMTJs. This way, the output of the gain stage contains the total influence from all other SMTJs in the network. This modification is a simple method of scaling the proposed coupling method up to large values of n . The number of operational amplifiers scales as n while the number of resistors scales as n^2 .

increased before CMOS input and output signal conditioning circuits become latency bottlenecks.

In a network with many SMTJ circuits such as the one described here, the gain stage of each interaction circuit can be modified to accept inputs from each of the other SMTJs in the network. The gain stage will then be a summing amplifier circuit, with rows of resistors feeding into the amplifier, as shown in Fig. 7. This scalability opens the possibility of creating an Ising machine to solve realistically large problems. The number of operational amplifiers in the interaction circuits grows linearly with the number of nodes in such an Ising machine, while the number of resistors needed for the modified gain stage grows quadratically if the nodes are all-to-all connected. These quadratic circuits can be implemented with crossbars of resistive switches. The push for building hardware platforms for matrix multiplication has led to the development of low-latency, energy-efficient large-scale crossbars. CMOS-integrated resistive RAM crossbars in sub-micron nodes (22 nm and 130 nm) have recently been demonstrated on such applications with arrays as large

as 1024×512 . When scaled to representative kernel sizes (256×256), such crossbars incur latencies on the order of $\approx 1 \mu\text{s}$ to $10 \mu\text{s}$ [48, 49]. The crossbar arrays implementing the all-to-all connections are the most area, energy, and latency-intensive aspects of larger-scale designs. Solving larger problems will require codesign of devices, architectures and algorithms. Using Hamiltonians with large connectivity matrices will be partitioned into crossbar kernels that solve a subset of the global problem. Additional nodes to reduce the connectivity requirements for particular problems [8] would be needed, augmented with specific circuits to handle communication between kernels.

Since p -bit based Ising machines operate on the principle that the programmable interaction circuits should operate at speeds faster than the nodes themselves, microsecond switching SMTJs may be sufficient for scaled integrated implementations. However, the SMTJs used in this study are far from optimal. Estimates of enhanced SMTJ properties for integrated designs involve lower operating currents [50], faster operating speeds [33], and larger TMR [50]. Larger TMR enables easier CMOS readout, while lower operating currents are essential for energy efficiency. The high energy cost of large static currents can be amortized by fast switching circuits that produce more random bits per unit time and hence current, but becomes difficult to justify when faster operation is not beneficial for directly coupled designs.

VII. ACKNOWLEDGEMENTS

This work was funded by the National Institute of Standards and Technology, National Science Foundation and Agence nationale de la recherche. SG, TA, LP, DPL and AM acknowledge support under NSF grant number CCF-CISE-ANR-FET-2121957. AM also acknowledges support under the NIST Cooperative Research Agreement Award No. 70NANB14H209 through the University of Maryland. PT and UE acknowledge support under the ANR StochNet Project award No. ANR-21-CE94-0002-01. The authors would like to thank Steve Moxim for help with experimental setup and Frank Mizrahi for valuable feedback. The authors acknowledge J. Langer and J. Wrona from Singulus Technologies for the MTJ stack deposition and N. Lamard, R. Sousa, L. Prejbeanu as well as the Upstream Technological platform PTA, Grenoble, France for the device nanofabrication.

-
- [1] A. Lucas, Ising formulations of many np problems, *Frontiers in physics* **2**, 5 (2014).
 [2] A. Grimaldi, L. Mazza, E. Raimondo, P. Tullio, D. Rodrigues, K. Y. Camsari, V. Crupi, M. Carpentieri, V. Puliafito, and G. Finocchio, Evaluating spintronics-

- compatible implementations of ising machines, arXiv preprint arXiv:2304.04177 (2023).
 [3] 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).
- [4] J. Si, S. Yang, Y. Cen, J. Chen, Z. Yao, D.-J. Kim, K. Cai, J. Yoo, X. Fong, and H. Yang, Energy-efficient superparamagnetic ising machine and its application to traveling salesman problems, arXiv preprint arXiv:2306.11572 (2023).
 - [5] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-bits for invertible logic, *Physical Review X* **7**, 031014 (2017).
 - [6] 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, *Physical Review Applied* **17**, 014016 (2022).
 - [7] N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, G. Finocchio, and K. Y. Camsari, Computing with invertible logic: Combinatorial optimization with probabilistic bits, in *2021 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2021) pp. 40–3.
 - [8] N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, J. M. Martinis, G. Finocchio, and K. Y. Camsari, Massively parallel probabilistic computing with sparse ising machines, *Nature Electronics* **5**, 460 (2022).
 - [9] S. Chowdhury, K. Y. Camsari, and S. Datta, Accelerated quantum monte carlo with probabilistic computers, *Communications Physics* **6**, 85 (2023).
 - [10] A. Z. Pervaiz, B. M. Sutton, L. A. Ghantasala, and K. Y. Camsari, Weighted p-bits for fpga implementation of probabilistic circuits, *IEEE transactions on neural networks and learning systems* **30**, 1920 (2018).
 - [11] D. Bertsimas and J. Tsitsiklis, Simulated annealing, *Statistical science* **8**, 10 (1993).
 - [12] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, *et al.*, In-datacenter performance analysis of a tensor processing unit, in *Proceedings of the 44th annual international symposium on computer architecture* (2017) pp. 1–12.
 - [13] W. A. Wulf and S. A. McKee, Hitting the memory wall: Implications of the obvious, *ACM SIGARCH computer architecture news* **23**, 20 (1995).
 - [14] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, and E. Eleftheriou, Memory devices and applications for in-memory computing, *Nature nanotechnology* **15**, 529 (2020).
 - [15] B. Jinnai, K. Watanabe, S. Fukami, and H. Ohno, Scaling magnetic tunnel junction down to single-digit nanometers—challenges and prospects, *Applied physics letters* **116** (2020).
 - [16] M. Julliere, Tunneling between ferromagnetic films, *Physics letters A* **54**, 225 (1975).
 - [17] J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meserve, Large magnetoresistance at room temperature in ferromagnetic thin film tunnel junctions, *Physical Review Letters* **74**, 3273 (1995).
 - [18] J. C. Slonczewski, Current-driven excitation of magnetic multilayers, *Journal of Magnetism and Magnetic Materials* **159**, L1 (1996).
 - [19] D. Ralph and M. Stiles, Spin transfer torques, *Journal of Magnetism and Magnetic Materials* **320**, 1190 (2008).
 - [20] D. C. Worledge, Spin-transfer-torque mram: the next revolution in memory, in *2022 IEEE international memory workshop (IMW)* (IEEE, 2022) pp. 1–4.
 - [21] K. Garello, F. Yasin, H. Hody, S. Couet, L. Souriau, S. H. Sharifi, J. Swerts, R. Carpenter, S. Rao, W. Kim, *et al.*, Manufacturable 300mm platform solution for field-free switching sot-mram, in *2019 Symposium on VLSI Circuits* (IEEE, 2019) pp. T194–T195.
 - [22] K. Lee, J. Bak, Y. Kim, C. Kim, A. Antonyan, D. Chang, S. Hwang, G. Lee, N. Ji, W. Kim, *et al.*, 1gbit high density embedded stt-mram in 28nm fdsoi technology, in *2019 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2019) pp. 2–2.
 - [23] S. Aggarwal, H. Almasi, M. DeHerrera, B. Hughes, S. Ikegawa, J. Janesky, H. Lee, H. Lu, F. Mancoff, K. Nagel, *et al.*, Demonstration of a reliable 1 gb standalone spin-transfer torque mram for industrial applications, in *2019 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2019) pp. 2–1.
 - [24] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Neuromorphic spintronics, *Nature Electronics* **3**, 360 (2020).
 - [25] S. Jung, H. Lee, S. Myung, H. Kim, S. K. Yoon, S.-W. Kwon, Y. Ju, M. Kim, W. Yi, S. Han, *et al.*, A crossbar array of magnetoresistive memory devices for in-memory computing, *Nature* **601**, 211 (2022).
 - [26] Q. Dong, Z. Wang, J. Lim, Y. Zhang, M. E. Sinangil, Y.-C. Shih, Y.-D. Chih, J. Chang, D. Blaauw, and D. Sylvester, A 1-mb 28-nm 1t1mtj stt-mram with single-cap offset-cancelled sense amplifier and in situ self-write-termination, *IEEE Journal of Solid-State Circuits* **54**, 231 (2018).
 - [27] O. Golonzka, J.-G. Alzate, U. Arslan, M. Bohr, P. Bai, J. Brockman, B. Buford, C. Connor, N. Das, B. Doyle, *et al.*, Mram as embedded non-volatile memory solution for 22ffl finfet technology, in *2018 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2018) pp. 18–1.
 - [28] S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, Theory of relaxation time of stochastic nanomagnets, *Physical Review B* **103**, 094423 (2021).
 - [29] W. Rippard, R. Heindl, M. Pufall, S. Russek, and A. Kos, Thermal relaxation rates of magnetic nanoparticles in the presence of magnetic fields and spin-transfer effects, *Physical Review B* **84**, 064439 (2011).
 - [30] M. Bapna and S. A. Majetich, Current control of time-averaged magnetization in superparamagnetic tunnel junctions, *Applied Physics Letters* **111** (2017).
 - [31] M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, Energy-efficient stochastic computing with superparamagnetic tunnel junctions, *Physical review applied* **13**, 034016 (2020).
 - [32] A. Shukla, L. Heller, M. G. Morshed, L. Rehm, A. W. Ghosh, A. D. Kent, and S. Rakheja, A true random number generator for probabilistic computing using stochastic magnetic actuated random transducer devices, in *2023 24th International Symposium on Quality Electronic Design (ISQED)* (2023) pp. 1–10.
 - [33] 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, arXiv preprint arXiv:2301.05694 (2023).
 - [34] D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. S. Friedman, A. F. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, S. Yuasa, S. Tiwari, J. Grollier, and D. Querlioz, Low-energy truly random number generation with superparamagnetic tunnel junctions for unconventional computing, *Physical Review Applied* **8**,

- 054045 (2017).
- [35] P. Talatchian, M. W. Daniels, A. Madhavan, M. R. Puffall, E. Jué, W. H. Rippard, J. J. McClelland, and M. D. Stiles, Mutual control of stochastic switching for two electrically coupled superparamagnetic tunnel junctions, *Physical Review B* **104**, 054427 (2021).
- [36] N.-T. Phan, L. Soumah, A. Sidi El Valli, L. Hutin, L. Anghel, U. Ebels, and P. Talatchian, Electrical coupling of perpendicular superparamagnetic tunnel junctions for probabilistic computing, in *Proceedings of the 17th ACM International Symposium on Nanoscale Architectures* (2022) pp. 1–6.
- [37] A. Mizrahi, N. Locatelli, J. Grollier, and D. Querlioz, Synchronization of electrically coupled stochastic magnetic oscillators induced by thermal and electrical noise, *Physical Review B* **94**, 054419 (2016).
- [38] L. Schnitzspan, M. Kläui, and G. Jakob, Electrical coupling of superparamagnetic tunnel junctions mediated by spin-transfer-torques, arXiv preprint arXiv:2307.15165 (2023).
- [39] M. W. Daniels, W. A. Borders, N. Prasad, A. Madhavan, S. Gibeault, T. Adeyeye, L. Pocher, L. Wan, M. Tran, J. A. Katine, *et al.*, Neural networks three ways: unlocking novel computing schemes using magnetic tunnel junction stochasticity, in *Spintronics XVI*, Vol. 12656 (SPIE, 2023) pp. 84–94.
- [40] W. F. Brown, Thermal fluctuations of a single-domain particle, *Phys. Rev.* **130**, 1677 (1963).
- [41] N. Mohseni, P. L. McMahon, and T. Byrnes, Ising machines as hardware solvers of combinatorial optimization problems, *Nature Reviews Physics* **4**, 363 (2022).
- [42] A. Chavent, V. Iurchuk, L. Tillie, Y. Bel, N. Lamard, L. Vila, U. Ebels, R. C. Sousa, B. Dieny, G. Di Pendina, *et al.*, A multifunctional standardized magnetic tunnel junction stack embedding sensor, memory and oscillator functionality, *Journal of Magnetism and Magnetic Materials* **505**, 166647 (2020).
- [43] R. Ma, A. Sidi El Valli, M. Kreißig, G. Di Pendina, F. Protze, U. Ebels, G. Prenat, A. Chavent, V. Iurchuk, R. Sousa, *et al.*, Microwave functionality of spintronic devices implemented in a hybrid complementary metal oxide semiconductor and magnetic tunnel junction technology, *Electronics Letters* **57**, 264 (2021).
- [44] A. S. El Valli, *Radiofrequency signal detection and random number generation with perpendicular magnetic tunnel junctions*, Ph.D. thesis, Université Grenoble Alpes [2020-....] (2022).
- [45] B. Dieny and M. Chshiev, Perpendicular magnetic anisotropy at transition metal/oxide interfaces and applications, *Reviews of Modern Physics* **89**, 025008 (2017).
- [46] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, Yield and speed optimization of a latch-type voltage sense amplifier, *IEEE journal of solid-state circuits* **39**, 1148 (2004).
- [47] S. Capra, F. Crescioli, L. Frontini, M. Garci, and V. Liberali, A digitally-controlled ring oscillator in 28 nm cmos technology, in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2018) pp. 1–5.
- [48] W. Wan, R. Kubendran, C. Schaefer, S. B. Eryilmaz, W. Zhang, D. Wu, S. Deiss, P. Raina, H. Qian, B. Gao, *et al.*, A compute-in-memory chip based on resistive random-access memory, *Nature* **608**, 504 (2022).
- [49] J.-M. Hung, C.-X. Xue, H.-Y. Kao, Y.-H. Huang, F.-C. Chang, S.-P. Huang, T.-W. Liu, C.-J. Jhang, C.-I. Su, W.-S. Khwa, *et al.*, A four-megabit compute-in-memory macro with eight-bit precision based on cmos and resistive random-access memory for ai edge devices, *Nature Electronics* **4**, 921 (2021).
- [50] G. Hu, G. Lauer, J. Sun, P. Hashemi, C. Safranski, S. Brown, L. Buzi, E. Edwards, C. D’Emic, E. Galligan, *et al.*, 2x reduction of stt-mram switching current using double spin-torque magnetic tunnel junction, in *2021 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2021) pp. 2–5.
- [51] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando, Giant room-temperature magnetoresistance in single-crystal fe/mgo/fe magnetic tunnel junctions, *Nature materials* **3**, 868 (2004).
- [52] M. D. Stiles and A. Zangwill, Anatomy of spin-transfer torque, *Physical Review B* **66**, 014407 (2002).
- [53] Z. Li and S. Zhang, Magnetization dynamics with a spin-transfer torque, *Physical Review B* **68**, 024404 (2003).