Applied Physics Reviews

RESEARCH ARTICLE | NOVEMBER 02 2026

Robust power-law computing for edge intelligence using spin-orbit torque magnetic tunnel junctions

Shilong Xiong; Caihua Wan ; Xiaohan Li ; Yingqian Xu ; Dehao Kong; Shengkai Xia ; Tianyi Zhang; Aleksandr V. Davydenko ; Guanxiang Du ; Alexander S. Samardak ; Guoqiang Yu ; Xiufeng Han

Check for updates

Appl. Phys. Rev. 13, 041413 (2026)
https://doi.org/10.1063/5.0352180

View Online
Export Citation

Articles You May Be Interested In

AVS Quantum Science publishes original research and more in 2021

AVS Quantum Sci. (March 2021)

Seven nonstandard models coupling quantum matter and gravity

AVS Quantum Sci. (May 2022)

03 November 2026 01:40:15

AIP Publishing

Special Topics Open for Submissions

Learn More

AIP Publishing

Applied Physics Reviews ARTICLE pubs.aip.org/aip/are

Robust power-law computing for edge intelligence using spin-orbit torque magnetic tunnel junctions

Cite as: Appl. Phys. Rev. **13**, 041413 (2026); doi: 10.1063/5.0352180
Submitted: 12 July 2026 · Accepted: 14 October 2026 ·
Published Online: 2 November 2026

View Online Export Citation CrossMark

Shilong Xiong,[1,2] Caihua Wan,[1,2,a)] Xiaohan Li,[1] Yingqian Xu,[1,2] Dehao Kong,[1,2] Shengkai Xia,[1,2] Tianyi Zhang,[1,2] Aleksandr V. Davydenko,[3] Guanxiang Du,[4] Alexander S. Samardak,[3,a)] Guoqiang Yu,[1,2,3] and Xiufeng Han[1,2,a)]

AFFILIATIONS

[1]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China
[2]Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
[3]Laboratory of Spin-Orbitronics, Institute of High Technologies and Advanced Materials, Far Eastern Federal University, Vladivostok 690922, Russia
[4]Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China

[a)]**Authors to whom correspondence should be addressed:** wancaihua@iphy.ac.cn; samardak.as@dvfu.ru; and xfhan@iphy.ac.cn

ABSTRACT

Power-law transformations are widely used in edge perception, where compact implementations must balance numerical precision, circuit resources, and device variability. Here, we demonstrate a stochastic power-law architecture based on spin–orbit-torque magnetic tunnel junctions (SOT-MTJs). An event-driven feedback loop adjusts the write voltage to regulate the probability of an all-1 event in an n-bit block. Under the fixed-point sampling model, this establishes a switching probability $p = b^{1/n}$, from which m-bit event statistics yield the power-law target $b^{m/n}$. Online probability locking removes the need for offline per-device probability-voltage characterization or stored calibration tables. Measurements on W-based SOT-MTJs, combined with digital analysis of the acquired bitstreams, demonstrate representative power-law operations with absolute deviations of 0.003 or less at the reported precision, as well as parallel decoding of multiple exponents from one source. A 4-bit gamma-correction example preserves dark-region image detail and reaches approximately 35 dB peak signal-to-noise ratio relative to the theoretical gamma-domain reference near 1000 raw samples. A 5 Oe device-level field perturbation and a joint ±90% simulation centered on the measured parameter means of 25 devices examine adaptation and the safe voltage operating boundary under the tested conditions. These results establish a nonlinear computing primitive with accuracy controlled by sampling depth and an operating point maintained through feedback.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0352180

INTRODUCTION

Power-law operations, $x^{m/n}$, form a recurring class of nonlinear primitives in edge intelligence and sensory hardware.[1,2] They appear in normalization and feature-scaling operations in neural networks,[3] perceptual and sensory response models,[4] dynamic-range compression,[5] and empirical scaling laws in physical systems.[6] Although these operations often require only moderate precision at the edge, conventional implementations based on lookup tables, piecewise-linear approximations, or polynomial datapaths can still require substantial memory, multipliers, and parameter-specific calibration. This mismatch motivates compact hardware primitives whose native physics can operate directly in the probability domain. Ideally, such a primitive should compute tunable power-law mappings from local stochastic state, allow accuracy to be adjusted by sample count, and reuse the same physical circuit and digital decoder across multiple nonlinear workloads.

Stochastic computing offers a different route: a number is represented by the probability of 1s in a bitstream, and nonlinear functions can be estimated by event statistics rather than by deterministic arithmetic.[7,8] Magnetic tunnel junctions are attractive physical random sources because their thermally assisted switching can be tuned electrically and integrated with compact readout circuitry.[9–13] Spin-orbit-torque

03 November 2026 01:40:15

magnetic tunnel junctions (SOT-MTJs), in particular, separate write and read paths, enabling repeated probabilistic sampling without destructive sensing and making them natural candidates for stochastic nonlinear hardware.[14–26] Their stochasticity is not emulated by pseudo-random logic; it is generated by nanoscale magnetization dynamics, so the random source and the memory element can be physically collocated. Because probabilities are accumulated over time, the same primitive can expose a simple precision-latency trade-off that is difficult to obtain from fixed deterministic approximations in compact edge deployment systems.

The same device physics that creates useful randomness also creates a practical obstacle. Nominally identical MTJs rarely share the same probability–voltage (P–V) curve; threshold voltage, slope, thermal environment, and aging all shift the mapping from voltage to switching probability.[27,28] Conventional stochastic hardware must either calibrate each element or tolerate biased outputs. For edge systems deployed outside controlled laboratory conditions, this calibration burden can be as important as the arithmetic cost itself. The challenge is especially acute for nonlinear operations because a small probability bias before exponentiation can become a systematic output error after repeated event counting.

We address this requirement with an event-driven feedback loop that regulates the occurrence of all-1 blocks. By adjusting the SOT-MTJ write voltage online, the loop establishes and maintains the prescribed probability despite variations in the underlying P–V characteristics, from which several rational powers of the same input can be decoded. Because the proposed scheme requires neither offline extraction of device-specific P–V characteristics nor stored calibration tables, we refer to this mode of operation as calibration-free throughout this work. We first characterize the stochastic device response and then demonstrate probability locking and shared-source power-law decoding. A gamma-correction application connects these operations to image processing, while a field-perturbation experiment and parameterized simulations examine their response to changes in the device characteristics.

SOT-MTJ STOCHASTICITY AND DEVICE VARIABILITY

The stochastic source is a three-terminal SOT-MTJ with a W write channel and an elliptical junction pillar approximately $300 \times 100\,\mathrm{nm}^2$ in size [Fig. 1(a)]. Current through W exerts spin–orbit torque on the free layer, and the junction resistance distinguishes the parallel and antiparallel states. The separated paths enable the repeated reset-write-read cycles required for Bernoulli sampling. The type-Y geometry also supports field-free switching.[15–17] For computation, we operate the device near its stochastic switching threshold, where individual switching outcomes vary while their probability remains voltage-dependent. Fabrication details and the measurement setup are given in supplementary material Note 1.

At room temperature, the parallel and antiparallel resistances are approximately 36 and 100 kΩ, respectively, corresponding to a tunneling magnetoresistance ratio of about 175% [Fig. 1(b)]. This contrast allows each trial to be classified as a binary state. Pulses above the switching threshold produce deterministic reversal, whereas subcritical pulses generate probabilistic switching [Figs. 1(c) and 1(d)]. Thermal fluctuations in this regime give a sigmoid-like dependence of switching probability on voltage,

$$P(V) = \frac{1}{1 + \exp[-k_s(V - V_c)]},$$

where V_c is the 50% switching voltage and k_s describes the transition steepness. The model provides a compact description of the measured curves and the parameter variations considered later; its fitted parameters are not inputs to the feedback algorithm.

Figures 1(e)–1(g) demonstrate voltage-controlled probability encoding. Using voltages selected from the measured P–V curve, a single device generated bitstreams at target probabilities of 0.3, 0.5, and 0.8, with measured averages following the programmed values. These measurements establish the tunability needed for computation. In the feedback experiments that follow, the voltage is instead adjusted from the observed binary events. Power-law outputs and their statistics are obtained by digital processing of the acquired SOT-MTJ state sequences.

The 25 junctions in Fig. 1(d) show device-to-device variation in both V_c and k_s. These differences explain why a common fixed voltage cannot generally encode the same probability across devices. Independent sigmoid fits give $V_c = 1.881 \pm 0.151$ V (mean ± sample standard deviation), with an observed range of 1.603–2.236 V, corresponding to −14.78% to +18.86% relative to the measured mean and a coefficient of variation of 8.02%. The fitted k_s values are $7.543 \pm 1.883\ \mathrm{V}^{-1}$, with an observed range of 3.595–11.818 V^{-1}, corresponding to −52.34% to +56.68% and a coefficient of variation of 24.96%. This variation is typically treated as a fabrication problem to be minimized or calibrated away. For large arrays, however, per-device calibration would introduce storage, time, and update overhead that directly conflicts with edge deployment. Here, it motivates a probability-locking strategy in which voltage becomes an internal control variable rather than a calibrated input supplied by the user.

FEEDBACK-LOCKED POWER-LAW COMPUTATION

The feedback algorithm uses event statistics to hide the unknown P–V curve. For an input b in [0,1] and a denominator n, the controller applies a voltage V_k and samples the SOT-MTJ n times. It then checks the all-1 event $A_k = (x_1 = x_2 = \ldots = x_n = 1)$. If the instantaneous switching probability is $\sigma(V_k)$, the event probability is thus $[\sigma(V_k)]^n$. The controller updates voltage from this binary event only, without estimating $\sigma(V)$ itself. This is important for hardware because the loop only needs a shift register, an event detector, and a small voltage-update rule; it does not need an analog-to-probability conversion table.

When A_k occurs, the observed event rate is likely above the target b, and the controller applies negative feedback:

$$V_{k+1} = V_k - l \cdot (1 - b). \tag{1a}$$

When A_k does not occur, the controller applies positive feedback

$$V_{k+1} = V_k + l \cdot b, \tag{1b}$$

where l is the voltage learning rate. The asymmetry of the two steps gives the conditional mean drift $E[\Delta V_k | V_k] = l[b - [\sigma(V_k)]^n]$. On a monotonic switching branch, after convergence, the voltage therefore moves toward a fixed point satisfying $\sigma(V^*) = b^{1/n}$. This is the basis for obtaining a root through probability locking. The update acts as stochastic integral feedback driven by a binary event error; it shares target-tracking and disturbance-compensation principles with

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

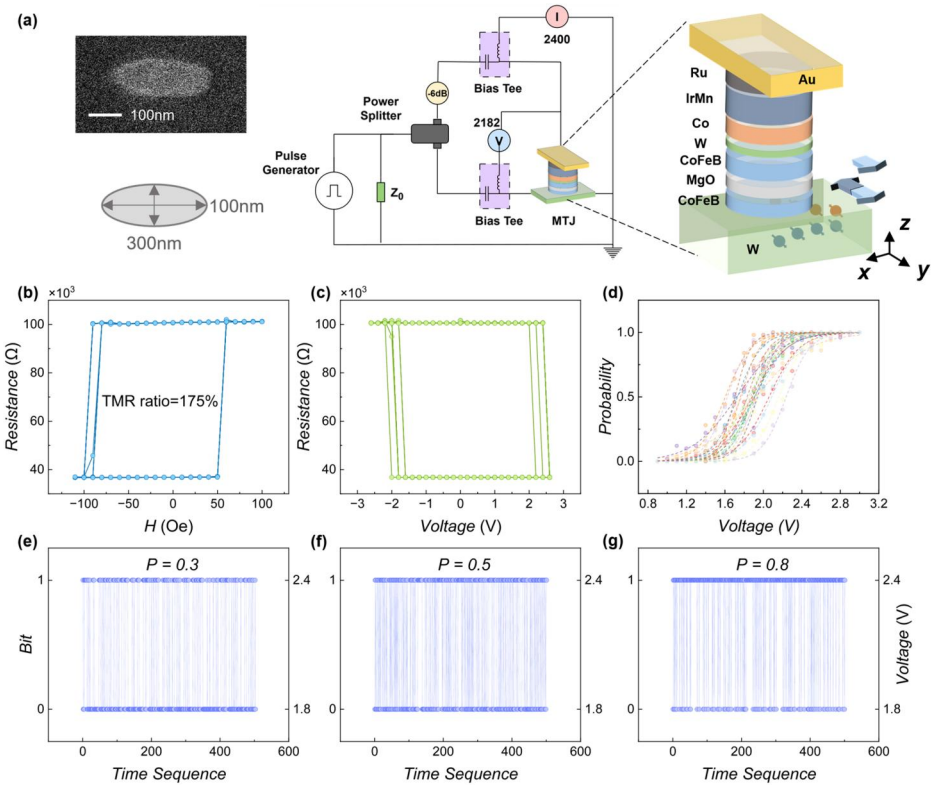


FIG. 1. SOT-MTJ structure and stochastic switching. (a) Device structure, four-terminal measurement geometry, and SEM image of the junction. (b) and (c) Measured TMR loop and deterministic switching characteristics: $R_P \approx 36\,k\Omega$, $R_{AP} \approx 100\,k\Omega$, TMR $\approx 175\%$, a loop offset of about -25 Oe, and switching thresholds near -1.8 and $+2.4$ V under 50 ns pulses. (d) Measured P–V curves for 25 nominally identical junctions, with sigmoid fits. (e)–(g) State sequences acquired from one SOT-MTJ at target probabilities of 0.3, 0.5, and 0.8, together with digitally evaluated Bernoulli statistics. Panels (b)–(d) present direct device characterization; panels (e)–(g) establish voltage tunability before closed-loop compensation.

classical feedback control but has no explicit proportional or derivative term. Supplementary material Note 2 derives the drift and compares the rule with proportional-integral-derivative (PID) control.

The fixed point describes the mean update law. With a finite learning rate, the voltage continues to fluctuate, and a finite bitstream yields an estimate of the target rather than an exact deterministic value. A larger l can shorten acquisition while increasing steady-state fluctuations; the local response slope determines the appropriate step size. Operation also requires the target to remain reachable on a monotonic branch within a safe voltage interval. A bounded controller can enforce that interval, while the endpoints $b = 0$ and $b = 1$ require accessible deterministic regimes. Specific voltage-clipping and end point protocols remain to be established for an integrated implementation. Supplementary material Note 2 relates these operating conditions to local stability and finite-step error.

Once the base probability is stabilized, counting all-1 events in m-bit groups gives the power-law target $b^{m/n}$ under the fixed-probability sampling model. Figure 2 summarizes the event-driven feedback loop and the subsequent power-law readout. Separate counters can decode several numerators from the same source when the input and denominator are shared. To improve statistical efficiency under the finite available bitstream length, we use sliding groups for point estimates and distribution plots, thereby exploiting all admissible local windows rather than only a disjoint subset. The resulting overlapping observations reuse raw bits and are therefore correlated, so we consequently quantify uncertainty using a separate partition into nonoverlapping outer blocks. This distinction, together with checks of raw-bit temporal dependence, is developed in supplementary material Note 3. Inputs requiring normalization must be mapped into the probability interval with the corresponding output transformation specified, as discussed in supplementary material Note 2.

POWER-LAW DEMONSTRATION AND SHARED-SOURCE OUTPUTS

We tested the feedback rule with three representative operations: $0.5^{3/2}$, $0.5^{1/3}$, and 0.5^3. These cover a fractional power above 1, a root operation, and an integer power. In each case, the measured SOT-MTJ state sequence drove the voltage update, and the acquired bits were processed digitally to evaluate the output. All measurements used a cold start at $V_0 = 0$ V and a fixed learning rate $l = 50$ mV, without a pre-stored operating voltage.

The trajectories in Fig. 3(a) show how the loop acquires that fixed point. For $0.5^{1/3}$, the initially low switching probability produces few all-1 events, so upward updates bring the device into its stochastic switching regime. The output approaches the theoretical value after roughly 200 samples and subsequently fluctuates around it. Here, a sample is one physical SOT-MTJ trial. The other two tasks show comparable acquisition behavior with the same update rule. The voltage traces represent the evolving controller state, including the initialization cost of probability locking.

After locking, the point estimates are 0.351, 0.793, and 0.127, compared with theoretical values of 0.354, 0.794, and 0.125 for the three tasks, respectively. Their absolute deviations at the displayed precision are 0.003, 0.001, and 0.002. Figure 3(c) displays the corresponding distributions of sliding-window estimates. For the

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

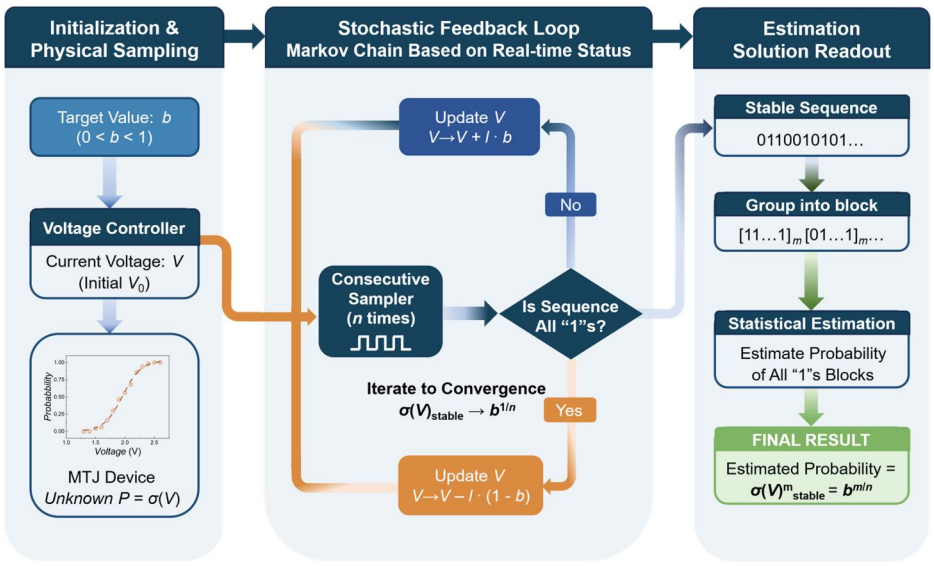


FIG. 2. Event-driven feedback and power-law decoding. The controller samples n-bit blocks and adjusts the voltage according to whether each block is all 1s. Under the fixed-point sampling model, the regulated event probability b corresponds to a device switching probability $b^{1/n}$. Digital counters then estimate $b^{m/n}$ from m-bit groups, allowing several numerators to share one source and feedback loop. Changes in the device response can be accommodated by a new operating voltage, provided the target remains reachable within the safe range. No offline device-specific P–V table is used by the loop.

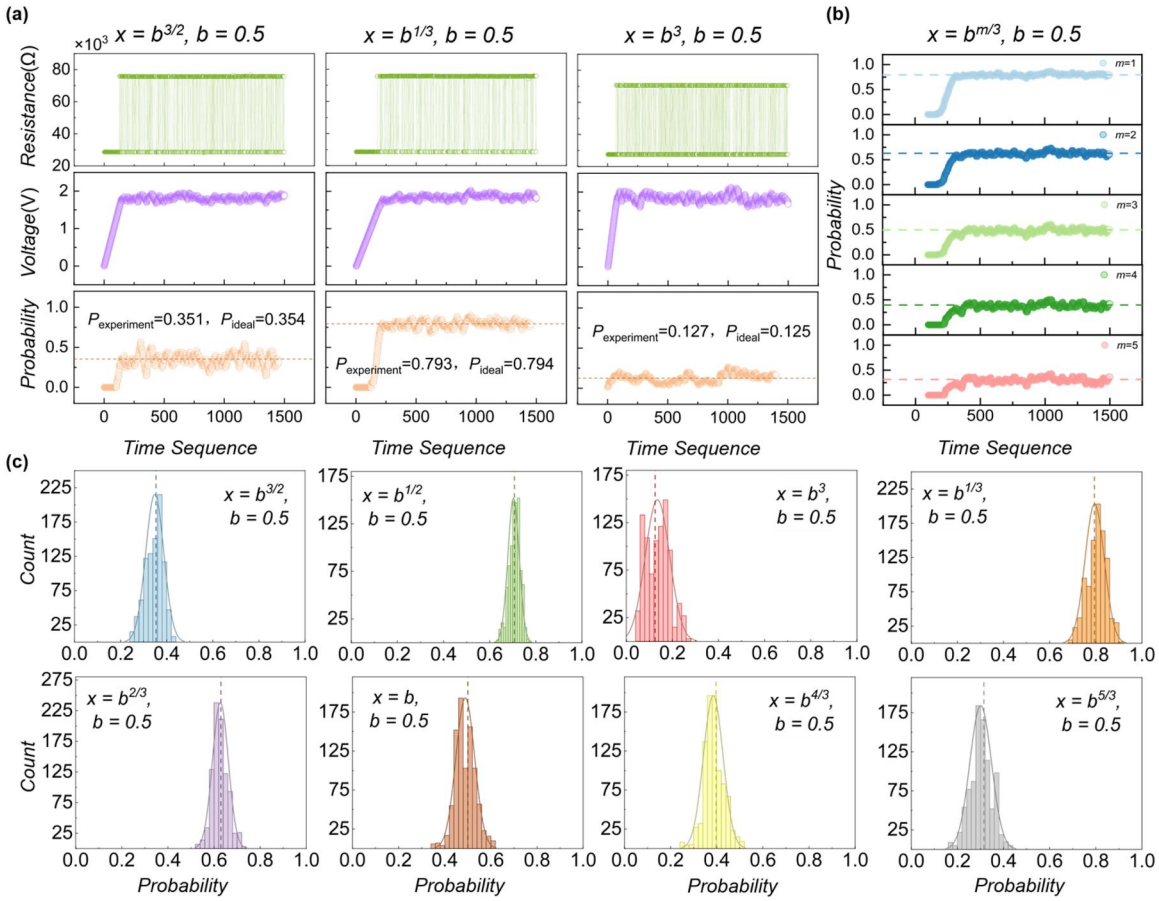


FIG. 3. Power-law evaluation using acquired SOT-MTJ state sequences and digital analysis. (a) Resistance state, feedback voltage, and estimated output for $0.5^{3/2}$, $0.5^{1/3}$, and 0.5^{3}, starting at $V_0 = 0$ V with $l = 50$ mV. Dashed lines indicate theoretical values. (b) Parallel decoding of $0.5^{m/3}$ for $m = 1$–5 from a shared stream. (c) Post-lock distributions of sliding-window estimates, adjacent windows share raw bits. All panels derive from measured device sequences rather than parameterized device simulations.

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

uncertainty analysis, the stationary sequences are instead partitioned into nonoverlapping outer blocks of the respective lengths. Table S1 gives the block-level standard deviations and 95% intervals, each of which contains the theoretical target. The temporal-dependence checks and the available statistical-method details are provided in supplementary material Note 3.

Probability locking also enables several outputs to reuse the same physical source. With $n = 3$ and $b = 0.5$, digital counters with grouping lengths $m = 1$–5 decode $0.5^{1/3}$ through $0.5^{5/3}$ from one stabilized stream [Fig. 3(b)]. For a family of powers sharing the same operand and denominator, the additional hardware is primarily the grouping and counting logic. This single-stream, multiple-output mode is useful when several powers of the same operand are needed because it avoids replicating the MTJ source and feedback circuit for every exponent.

The sampling cost has two components: acquiring the fixed point and estimating the output. If the loop makes K updates using n trials per update, acquisition consumes nK physical trials. Output precision then depends on the available post-lock samples and their grouping. Supplementary material Note 7 distinguishes these counts and presents the parameterized Monte Carlo results in Fig. S5. This sampling framework also guides the image-processing example below, in which stream length controls fidelity to a gamma mapping.

ROBUST GAMMA-CORRECTION APPLICATION

Gamma correction provides a direct application of rational-power decoding. For a normalized intensity x and nominal display gamma $\gamma_{\mathrm{disp}} = 2.2$, the encoding is $y = x^{1/2.2} = x^{5/11}$. The exponent is approximately 0.4545, with 0.45 used only as a rounded representation. This mapping allocates more of the available quantization levels to dark regions before low-bit storage.[29] It is therefore useful for examining whether a stochastic power-law output can retain image structure that direct 4-bit linear quantization suppresses.

We normalized the original 8-bit grayscale image to [0,1] and obtained gamma-domain values for the discrete input codes from SOT-MTJ bitstreams using $m = 5$ and $n = 11$. Pixels sharing an input code used the same experimentally derived gamma value, so the image was processed through a code-to-output mapping rather than separate device sampling for every pixel. The gamma-domain values were represented with 4 bits, decoded, and transformed back to the original-intensity domain using $x' = y'^{11/5}$. The code values were thus derived from device sampling and digital processing, with no device-specific calibration curve supplied to the feedback computation.

The two stages of this process require different comparisons. Peak signal-to-noise ratio (PSNR) and SSIM compare the normalized gamma-domain output with the theoretical gamma-transformed image, assessing fidelity to the intended nonlinear function. MAE, RMSE, and error heatmaps compare the inverse-transformed reconstruction with the original grayscale image, assessing the complete quantization and reconstruction process. Supplementary material Note 6 defines the references, intensity ranges, and units for each metric.

Figure 4(c) shows that direct 4-bit linear quantization removes shadow contours and texture, while the reconstruction based on SOT-MTJ gamma values retains more dark-region details and follows the theoretical gamma-corrected reference. Fidelity to the gamma mapping improves with the raw bitstream length N_{bit} [Fig. 4(b)], reaching approximately 35 dB PSNR near 1000 samples. The 35 dB value characterizes this image comparison; it is not a universal perceptual threshold. Longer streams improve the statistical estimate at the cost of additional sampling time and energy.

We next examined whether feedback could maintain the mapping when the device response changed. In the dynamic experiment, a 5 Oe field was applied at the 500th sample [Fig. 4(d)]. The change in switching probability altered the observed event statistics, prompting the loop to shift the write voltage toward a new operating point within dozens of update cycles. The recovery of the target probability and image quality illustrates compensation of this perturbation during active computation.

To relate the feedback simulations to the measured device variability, we varied V_c and k_s simultaneously from −90% to +90% about their measured means, using a 5% grid spacing. At every grid point, the same closed-loop feedback method was applied with the confirmed safe drive-voltage window of −3.5 to +3.5 V. Figure 4(e) displays the mean gamma-domain image PSNR relative to the ideal gamma output for the same quantized input codes. The fitted parameter pairs of the 25 measured devices are overlaid on the simulated map.

The measured mean parameters give 34.83 dB, and the 25 fitted device pairs give 31.75–36.83 dB. Across the deliberately broad ±90% domain, the mean PSNR ranges from 4.61 to 37.51 dB. The low-PSNR region occurs primarily for high V_c combined with low k_s, where some target probabilities require voltages above +3.5 V. The controller then saturates at the safety limit and cannot support the full computational target range. Finite sampling and feedback fluctuations remain sources of error within the voltage-accessible region. This map is a parameterized model study and predictions outside the measured 0.9–3.0 V interval involve model extrapolation.

HARDWARE RESOURCES AND FURTHER DEVELOPMENT

The 4-bit example demonstrates an adaptive nonlinear mapping and its sampling-accuracy trade-off. A fixed 4-bit lookup table contains only 16 entries, so this example alone does not establish a storage-area advantage. The broader resource question concerns higher input precision, multiple selectable mappings, or families of powers that share a source. A direct table requires 2^B entries per mapping for a B-bit input. The SOT-MTJ architecture instead assigns resources to a stochastic source, write/read and sensing circuits, event detection, counters, and voltage feedback. Supplementary material Note 4 compares these requirements at the level of circuit primitives.

Repeated physical sampling is central to that resource balance. The complete trial includes reset, write, read, and control operations, so a 50 ns write pulse alone does not determine output latency. Acquisition and post-lock estimation must both be included in a computation budget. Endurance and bit-rate results from earlier MTJ systems provide context,[23,30] but their translation to this architecture depends on the pulse sequence, peripheral circuitry, and workload. Supplementary material Note 4 therefore treats them as conditional estimates and distinguishes hypothetical per-pixel sampling from the discrete-code mapping used here. Circuit integration will need to address these costs alongside the probability-control accuracy.

Shared-source decoding offers a further direction for development. With $n = 1$, estimates of successive integer powers could form a polynomial basis for functions beyond the power laws demonstrated here. Exponential and logarithmic approximations would require

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

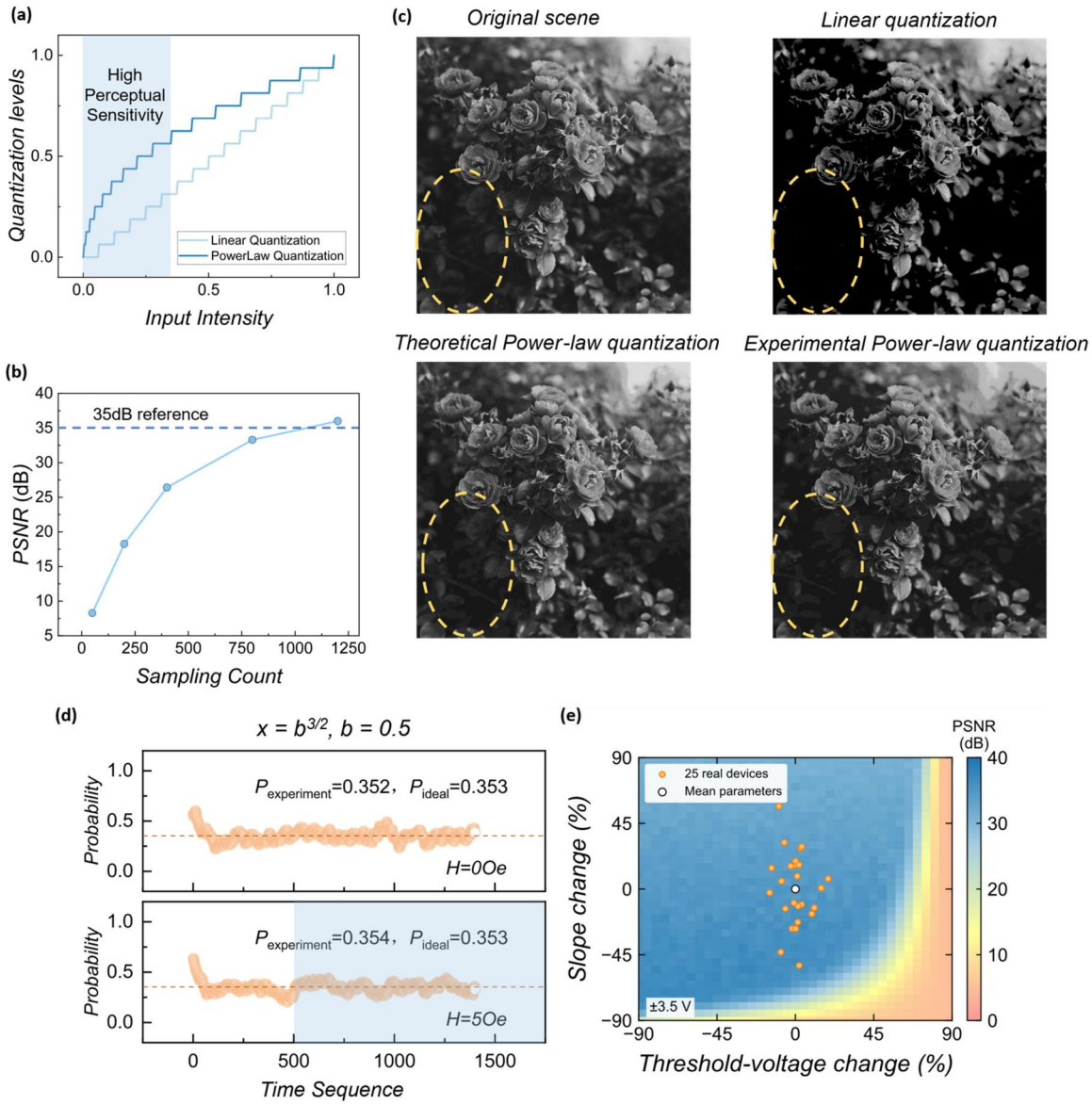


FIG. 4. Gamma correction and response to device perturbations. (a) The mapping $y = x^{5/11}$ before 4-bit quantization. (b) Gamma-domain PSNR relative to the theoretical gamma image vs sample count, reaching about 35 dB near 1000 samples. The horizontal marker denotes a 35 dB reference. (c) Original image, linear 4-bit quantization, theoretical gamma correction, and gamma correction derived from measured SOT-MTJ bitstreams. (d) Device-level recovery following a 5 Oe field perturbation during iteration. (e) Mean gamma-domain PSNR from joint V_c and k_s feedback simulations over ±90% about the 25-device measured means, with a safe voltage clamp of ±3.5 V. Orange circles mark the 25 measured device parameter pairs, and the white circle marks their mean.

coefficient storage and weighted accumulation, including signed-value handling where needed. A logarithmic approximation must also be restricted away from zero or use range reduction. Supplementary material Note 2.7 describes these possible extensions and their additional costs. Retaining the controller voltage for a warm start at a nearby target and adapting the learning rate are likewise prospective optimizations; the present measurements use cold starts and a fixed learning rate.

CONCLUSION

We have demonstrated a feedback-locked SOT-MTJ architecture for calibration-free stochastic power-law computing. By regulating the all-1 event probability, the system computes $x^{m/n}$ without knowing the device P–V curve, achieves absolute errors below 0.003 in representative operations, and supports multiple outputs from one stochastic stream. A 4-bit gamma example retains dark-region detail with

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

about 35 dB PSNR relative to the theoretical gamma-domain reference. The field-perturbation experiment and separate parameter sweeps show adaptation within the tested conditions. This device-algorithm co-design suggests a compact route to robust nonlinear functions for edge hardware, where modest precision, adaptability, and low calibration overhead are more valuable than deterministic arithmetic generality. Future integration can further reduce latency by parallelizing MTJ sampling and co-designing the voltage driver with the event detector.

SUPPLEMENTARY MATERIAL

See the supplementary material for device fabrication and measurement details, feedback pseudocode and theory, statistical analysis, hardware-resource and sampling-cost models, parameterized variation studies, and image-quality definitions and results.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (MOST) (Grant No. 2022YFA1402800), the National Natural Science Foundation of China (NSFC) (Grant Nos. T2495210, 12134017, W2412079, and 12374131), and the Chinese Academy of Sciences President's International Fellowship Initiative (PIFI Grant No. 2025PG0006). The work of A. V. Davydenko, A. S. Samardak, and G. Yu was supported by the Russian Science Foundation, Grant No. 26-72-31015 (https://rscf.ru/project/26-72-31015/).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Shilong Xiong: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Caihua Wan:** Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **Xiaohan Li:** Investigation (equal). **Yingqian Xu:** Investigation (equal). **Dehao Kong:** Investigation (equal). **Shengkai Xia:** Validation (equal). **Tianyi Zhang:** Validation (equal). **Aleksandr V. Davydenko:** Formal analysis (equal); Writing – review & editing (equal). **Guanxiang Du:** Validation (equal). **Alexander S. Samardak:** Formal analysis (equal); Writing – review & editing (equal). **Guoqiang Yu:** Formal analysis (equal); Writing – review & editing (equal). **Xiufeng Han:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material and from the corresponding authors upon reasonable request.

REFERENCES

[1]M. Horowitz, "1.1 computing's energy problem (and what we can do about it)," in *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)* (IEEE, San Francisco, CA, 2014), pp. 10–14.
[2]V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," Proc. IEEE **105**, 2295–2329 (2017).
[3]B. Zhang and R. Sennrich, "Root mean square layer normalization," arXiv:1910.07467 (2019).
[4]S. S. Stevens, "To honor Fechner and repeal his law," Science **133**, 80–86 (1961).
[5]J. M. Kates, "Principles of digital dynamic-range compression," Trends Amplif. **9**, 45–76 (2005).
[6]M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," Contemp. Phys. **46**, 323–351 (2005).
[7]B. R. Gaines, "Stochastic computing systems," in *Advances in Information Systems Science*, edited by J. T. Tou (Springer, Boston, MA, 1969), Vol. 2, pp. 37–172.
[8]A. Alaghi, W. Qian, and J. P. Hayes, "The promise and challenge of stochastic computing," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **37**, 1515–1531 (2018).
[9]J. Grollier *et al.*, "Neuromorphic spintronics," Nat. Electron. **3**, 360–370 (2020).
[10]K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, "Stochastic p-bits for invertible logic," Phys. Rev. X **7**, 031014 (2017).
[11]W. A. Borders *et al.*, "Integer factorization using stochastic magnetic tunnel junctions," Nature **573**, 390–393 (2019).
[12]S. Niazi *et al.*, "Training deep Boltzmann networks with sparse Ising machines," Nat. Electron. **7**, 610–619 (2024).
[13]C. Duffee *et al.*, "An integrated-circuit-based probabilistic computer that uses voltage-controlled magnetic tunnel junctions as its entropy source," Nat. Electron. **8**, 784–793 (2025).
[14]Q. Shao *et al.*, "Roadmap of spin-orbit torques," IEEE Trans. Magn. **57**, 1–39 (2021).
[15]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).
[16]L. Liu *et al.*, "Spin-torque switching with the giant spin Hall effect of tantalum," Science **336**, 555–558 (2012).
[17]C.-F. Pai *et al.*, "Spin transfer torque devices utilizing the giant spin Hall effect of tungsten," Appl. Phys. Lett. **101**, 122404 (2012).
[18]X. H. Li *et al.*, "True random number generator based on spin-orbit torque magnetic tunnel junctions," Appl. Phys. Lett. **123**, 142403 (2023).
[19]H. Chen *et al.*, "Binary and ternary true random number generators based on spin orbit torque," in *2018 IEEE International Electron Devices Meeting (IEDM)* (IEEE, San Francisco, CA, 2018), pp 36.5.1–36.5.4.
[20]R. Ren *et al.*, "Initialization-free and magnetic field-free spin-orbit p-bits with backhopping-like magnetization switching for probabilistic applications," Nano Lett. **24**, 10072–10080 (2024).
[21]X. Li *et al.*, "Restricted Boltzmann machines implemented by spin-orbit torque magnetic tunnel junctions," Nano Lett. **24**, 5420–5428 (2024).
[22]R. Zhang *et al.*, "Probabilistic greedy algorithm solver using magnetic tunneling junctions for traveling salesman problem," Nat. Commun. **17**, 189 (2025).
[23]R. Zhang *et al.*, "Probability-distribution-configurable true random number generators based on spin-orbit torque magnetic tunnel junctions," Adv. Sci. **11**, 2402182 (2024).
[24]Y. Q. Xu *et al.*, "Self-stabilized true random number generator based on spin-orbit torque magnetic tunnel junctions without calibration," Appl. Phys. Lett. **125**, 132403 (2024).
[25]Y. Lv, R. P. Bloom, and J.-P. Wang, "Experimental demonstration of probabilistic spin logic by magnetic tunnel junctions," IEEE Magn. Lett. **10**, 1–5 (2019).
[26]V. B. Adishesha, T. Kim, J. Park, and H. Yang, "PRIME: Probabilistic reconfigurable Ising memory engine, an 8000 p-bit Ising machine with 28 nm multi-function STT-MRAM," in *2025 IEEE International Electron Devices Meeting (IEDM)* (IEEE, San Francisco, CA, 2025), pp. 1–4.

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing

[27]B. R. Zink *et al.*, "Tunable spintronic devices with different switching mechanisms for probabilistic and stochastic computing," Phys. Rev. Appl. **24**, 047001 (2025).
[28]B. Zhang *et al.*, "Automatic extraction and compensation of P-bit device variations in large array utilizing Boltzmann machine training," Micromachines **16**, 133 (2025).
[29]S. P. Vimal and P. K. Thiruvikraman, "Automated image enhancement using power law transformations," Sadhana **37**, 739–745 (2012).
[30]S. Zou, X. Zhao, Y. Xue, J. Gao, Y. Li, Y. Cui, and J. Luo, "A high bit rate and self-stable probability true random number generator based on magnetic tunnel junction for IoT security applications," IEICE Electron. Express **22**, 20250084 (2025).

03 November 2026 01:40:15

Published under an exclusive license by AIP Publishing