arxiv.org:2602.XXXXX

Momentum-Driven Reversible Logic Accelerates Efficient Irreversible Universal Computation

Kuen Wai Tang,* Kyle J. Ray,† and James P. Crutchfield‡
Department of Physics and Astronomy
University of California-Davis
One Shields Avenue, Davis, CA 95616
(Dated: February 7, 2026)

We present implementations of two physically-embedded computation-universal logical operations using a 2-bit logical unit composed of coupled quantum flux parametrons—Josephson-junction superconducting circuits. To illustrate universality, we investigate NAND gates built from these two distinct elementary operations. On the one hand, Controlled Erasure (CE) is designed using fixed-point analysis and assumes that information must be stored in locally-metastable distributions. On the other, Erasure-Flip (EF) leverages momentum as a computational resource and significantly outperforms the metastable approach, simultaneously achieving higher fidelity and faster computational speed without incurring any additional energetic cost. Notably, the momentum degree of freedom allows the EF to achieve universality by using both nontrivial reversible and irreversible logic simultaneously in different logical subspaces. These results not only provide a practical, high-performance protocol ripe for experimental realization but also underscore the broader potential of momentum-based computing paradigms.

Keywords: nonequilibrium thermodynamics, Landauer's Principle, Josephson junction, momentum computing, irreversible logic, reversible logic

I. INTRODUCTION

The information data industry is currently facing an unprecedented energy demand crisis. This challenge is driven by the rapid growth of data centers and the proliferation of digital services [1–3]. Historically, the industry's ability to consistently meet the growth in computational demand has been driven by the scaling achieved under Moore's Law [4, 5]. However, as conventional CMOS technology approaches fundamental physical limits during scaling-down [6–10], particularly regarding thermal dissipation, the need for sustainable, energy-efficient hardware is more critical than ever. Addressing this global energy cost problem requires the development of novel computing design paradigms capable of operating at or near the fundamental thermodynamic limits of computation.

One such effort targets low-energy circuits based on the quantum flux parametron (QFP), a superconducting logical element that operates in the classical regime [11–17] rather than the quantum regime. Various computational architectures based on QFPs have been proposed. For example, the single-flux quantum (SQF) [18] and rapid single-flux quantum (RSQF) [19–21] use voltage pulses to process information. Adiabatic QFPs [22–26] and reversible QFPs [27–29] store information in the components of a QFP's phase degrees of freedom and logic gate protocols are then designed assuming adiabatic operation—which assumes fast and reliable relaxation to a local steady-state relative to the protocol's timescale. Furthermore, other interesting reversible logic concepts utilizing superconducting circuits have been extensively studied [30–34].

In many operating regimes, the phase-space evolution of these devices are well modeled by underdamped Langevin dynamics [35]. This permits transient dynamics that can be dominated by the conjugate momentum coordinate in phase space instead of just the position. The attempt to harness these transient dynamics for useful computation provides a compelling alternative design principle for logic gates that we call *momentum computing* [36]. Single-bit momentum computing has been demonstrated using a feedback driven electromechanical cantilever [37], and a design has been proposed for an implementation based on a QFP-like circuit [38].

Coupling two QFP's inductively creates a device capable of operations on two logical bits instead of one. The following investigates the thermodynamics of universal computation via control protocols that implement a NAND operation using such a coupled quantum flux parametron (CQFP) device. Given that NAND is a universal logic gate, energy-efficient implementations position CQFPs as compelling candidates for next-generation, high-speed, low-power computational devices.

To this end, the following analyzes and compares

* tkwtang@ucdavis.edu
† kjray@ucdavis.edu
‡ chaos@ucdavis.edu

two distinct low-power protocols, each implementing NAND logic. First, we consider the Controlled Erasure (CE) protocol [39], which is designed using the principle of storing and processing information only in positional degrees of freedom. Next, we present a new universal momentum computing logic gate, the Erasure-Flip (EF). Remarkably, by leveraging the transient dynamics that conjugate momenta allow, the EF gate offers a NAND implementation that significantly increases both computational speed and fidelity without incurring additional energetic cost.

Section II A presents the CQFP circuit design and details the underlying principles for manipulating its potential landscape to perform logical computations. Section II B introduces the idea of implementing NAND operations in CQFPs by serially implementing more fundamental operations. Section II C then provides an in-depth analysis of the work cost, fidelity, and speed of the CE protocol, highlighting its inherent speed limitations. Addressing this, Section II E introduces a novel control scheme—the EF protocol—that leverages the system momentum to preserve information. We demonstrate through simulations that this protocol achieves significantly higher speed and fidelity than the CE while maintaining low work cost. Finally, Section II G shows the simulation results of using both the CE and EF subprotocols introduced in the preceding sections to perform the NAND operation. The development concludes in Section III with a summary of our findings.

II. RESULTS

A. Coupled Quantum Flux Parametron

1. Circuit

A CQFP is formed by inductively coupling two QFPs [40–43]. The coupling mechanism is often a third QFP, whose dynamics we represent with a time-dependent coupling parameter M_{12}. Figure 1 shows the CQFP circuit diagram. The Josephson junctions of the i-th QFP are denoted J_{ia} and J_{ib} with phase differences δ_{ia} and δ_{ib}, respectively. Each junction is modeled as a resistive capacitive shunted junction (RCSJ) [44, 45]. C_i and R_i represent the junctions' shunt capacitors and resistors. I^c_{ia} and I^c_{ib} are the junctions critical currents. Together, J_{ia} and J_{ib} form a dc-SQUID with inductance l_i in each branch. This dc-SQUID is then connected in a loop with a geometric inductance L_i, as in an rf-SQUID. For simplicity, we set $l_1 = l_2 = l$, $L_1 = L_2 = L$, $C_1 = C_2 = C$, and $R_1 = R_2 = R$. The device's characteristic timescale is given by $t_c = \sqrt{LC}$ which, for the specific parameters we consider, is 2.24ps.

The device's state is described by a four-dimensional vector: $(\phi_1, \phi_2, \phi_{1dc}, \phi_{2dc})$. Here, ϕ_1 and ϕ_2 represent the flux through the rf-SQUID of the two constituent QFPs, while ϕ_{1dc} and ϕ_{2dc} represent the flux through their small dc-SQUID loops. These dynamical coordinates can be expressed in terms of the JJ phases as:

$$\phi_i = \frac{\Phi_0}{2\pi}\frac{\delta_{ia}+\delta_{ib}}{2}$$
$$\phi_{idc} = \frac{\Phi_0}{2\pi}(\delta_{ia}-\delta_{ib})\ .$$

The shunt DC resistance R acts as a source of noise and damping, which allows us to model the dynamics of these coordinates using an underdamped Langevin equation. Since the system's state evolves according to Langevin dynamics (see Appendix A), we use the term "particle" as an analogy to describe the device's state and extend this metaphor by using the terms "position" and "velocity" to refer to values of the coordinate and its time derivative. This is a conceptual representation to motivate an intuitive understanding and is not meant to imply there is an actual physical particle.

The driving force in the Langevin equation that describes the motion of these particles can be written as the derivative of a potential landscape manipulated using five external parameters. These include external fluxes applied to the large loops (ϕ_{ix}) and the junction loops (ϕ_{ixdc}), as well as an inductive coupling between the two QFPs (M_{12}). The corresponding normalized external parameters are φ_{ix}, φ_{ixdc}, and m_{12} respectively. The relationships are given by:

$$\varphi_{ix} = \frac{2\pi\phi_{ix}}{\Phi_0}, \varphi_{ixdc} = \frac{2\pi\phi_{ix}}{\Phi_0}, \text{ and } m_{12} = \frac{M_{12}}{\sqrt{L_1 L_2}},$$

where Φ_0 is the magnetic flux quantum. The parameters introduced above and their corresponding values used in the simulations are summarized in Table I in Appendix C.

2. Potential

The potential energy function associated with the single QFP is well known [13, 15, 17, 22]. From it, one can find the CQFP's potential energy [46, 47], which we write in the following nondimensionalized form:

$$U' = U_1 + U_2 + U_{12}\ , \tag{1}$$

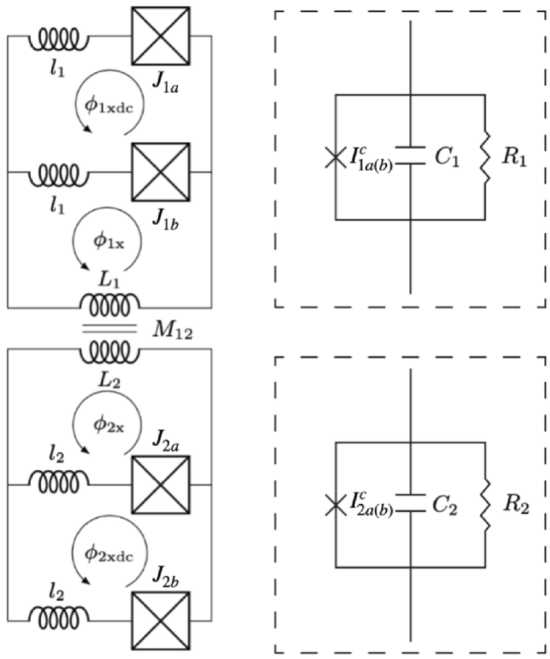


Figure 1. CQFP Circuit: Boxed diagrams (dashed lines) represent the Josephson junctions modeled as RCSJs. Table I in the Appendix explains the symbols.

with:

$$U_i = \frac{\xi}{2}(\varphi_i - \varphi_{ix})^2 + \frac{\gamma_i}{2}(\varphi_{idc} - \varphi_{ixdc})^2 + \beta_i \cos\varphi_i \cos\frac{\varphi_{idc}}{2} + \delta\beta_i \sin\varphi_i \sin\frac{\varphi_{idc}}{2}$$
$$U_{12} = m_{12}\xi(\varphi_1 - \varphi_{1x})(\varphi_2 - \varphi_{2x}),$$

where the index i = 1, 2.

Here, ξ is a modified coupling parameter $\xi \equiv 1/(1 - m_{12}^2)$ with $\xi \approx 1$ indicating the weak coupling limit. The parameter γ_i represents the ratio between the geometric inductances of the rf-SQUID loop and the dc-SQUID loop and is typically set to be $>> 1$. Due to this large ratio, any significant deviation of φ_{idc} from φ_{ixdc} incurs a large energy penalty, especially when compared to the deviation of φ_i from φ_{ix}. For this reason, it is often assumed that $\varphi_{idc} \approx \varphi_{ixdc}$, allowing simplification of the potential to a 2D landscape in φ_1 and φ_2 space. The parameters β_i and $\delta\beta_i$ are proportional to the sum and difference of the critical currents in the Josephson junctions, respectively. For simplification, the critical currents are assumed to be identical, resulting in $\delta\beta_i = 0$. However, in reality, it is unavoidable to have slightly asymmetric Josephson junctions, which produces an asymmetric potential that inherently favors some particular states. Fortunately, this asymmetric problem can be alleviated by calibration of the potential landscape; see Refs. [38, 47]. The full expressions for γ_i, β_i and $\delta\beta_i$ are provided in Table I.

When all external parameters are set to zero, the system exhibits a four-well potential within the φ_1-φ_2 state space, as visualized in the contour plot in Figure 3(a). Each of these potential minima is capable of storing information as a metastable state, because the energy scale of these wells is significantly higher than the thermal energy scale k_BT in typical operating regimes. The logical states are coarse-grained based on the sign of the individual phases: $\varphi_i < 0$ represents logical 0 and $\varphi_i > 0$ represents logical 1. Consequently, each well is labeled with its corresponding binary state: 00, 01, 10, and 11. The effects of the external control parameters on the system's potential energy landscape are further illustrated in Figure 13 of Appendix E.

The colored dots in the contour plot at $t = 0$ in Figure 3 represent samples from the initial equilibrium distribution. The coarse-grained initial position serves as the logical input for the operation, with color indicating the initial logical state of each sample. By dynamically manipulating the potential landscape, the particles are guided through the state space, representing the evolution of the system's microstates. The logical output of the operation is determined by a coarse-grained measurement of the final positions of the particles.

B. Physically-Embedded NAND

The NAND operation is a universal two-input logic gate whose output is zero if and only if both inputs are one; otherwise, the output is one. NAND gates are a favored logical element in modern general-purpose microprocessor architecture, and so the ability to simulate a NAND is a common standard candle for determining an operation's computation universality. Figure 2(a) gives the NAND truth table. Although the NAND operation is a 2-to-1 mapping, the CQFP produces a two-bit output. And, in fact, this will be the case for any physically embedded cyclic computation that stores 2-bit information in 2 degrees of freedom. If we call the input bits x ane y and the output bits x' and y', then an arbitrary operation is given by $(x', y') = (f(x, y), g(x, y))$ where $f(\cdot)$ and $g(\cdot)$ are functions of the input memory states.

A NAND gate, in this context, is an underdetermined concept. It could be the case that only $f(x, y) = \text{NAND}(x, y)$ or only $g(x, y) = \text{NAND}(x, y)$. Thus, when dealing with such physically embedded logic gates, it is more accurate to

(a)

input	NAND	CE	EF
x,y	NAND(x,y)	(AND(x,y),y)	(NAND(x,y),y)
0,0	1	(0,0)	(1,0)
0,1	1	(0,1)	(1,1)
1,0	1	(0,0)	(1,0)
1,1	0	(1,1)	(0,1)

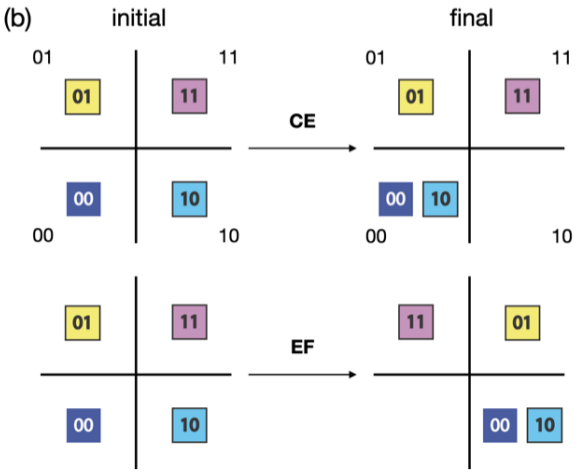


Figure 2. (a) Truth tables for the three operations: NAND, Controlled Erasure (CE), and Erasure-Flip (EF). (b) Schematic illustration of the initial and final logical-state distributions for CE (top) and EF (bottom) operations. The four 2-bit digits at the corners indicate the coarse-graining label for the potential wells and the colored boxes represent the particles. The initial state (left) represents the thermal equilibrium distribution of particles within the four-well potential, corresponding to the logical inputs. The final state (right) depicts the particle distribution after the respective operations, which should be compared with the truth table to determine the logical output.

say that an operation on the memory states of the device can "contain" or "implement" a NAND gate rather than "be" a NAND gate. We address this ambiguity by defining a "partial" NAND operation, which selects one of the two output bits as the result and is agnostic to the other one. (This terminology is common to the theory of computation [48].)

The EF operation in Figure 2(b) illustrates the memory state mapping in the CQFP state space as an example of a partial NAND. In this scenario, the initial equilibrium distributions representing inputs 00, 01, and 10 are mapped to a final logical outputs of either 10 or 11, while the input state 11 is mapped to 01. This partial NAND is described by $(x', y') = (\mathrm{NAND}(x, y), \mathrm{y})$. That being said, if a NAND gate is the only desired logic then we consider the second bit simply as an irrelevant bit, while the first is assumed to be the output.

A second approach—the "complete" NAND—removes ambiguity by requiring $f = g = \mathrm{NAND}(x, y)$. While there is no ambiguity in the map choice for this gate, our focus here is the partial NAND due to the flexibility it affords. Additional details regarding implementations of complete NAND operations are provided in Appendix F.

The absolute minimum possible work required for an operation is determined by the change in the system's entropy. Since we assume a cyclic protocol and the initial wells have identical local free energies, the change in the system's average internal energy is zero after local equilibration. According to the first law of thermodynamics $\Delta U = W + Q = 0$. That is, the work performed (W) is the negative of the heat dissipated (Q).

For a reversible, isothermal process, Q is given by the change in system entropy. This yields a minimal work value $W = -k_B T \Delta S$, which is often called Landauer's bound [49]. The fundamental thermodynamic cost of the partial NAND gate is 0.347 $k_B T$ (Appendix G). With this in mind, we note that the finite-time protocols we investigate fall typically between one and two orders of magnitude higher in work cost than this limit.

C. Mechanism of Controlled Erasure

This section describes the Controlled Erasure (CE) protocol, first proposed in Ref. [39], and, using extensive simulations, analyzes its work cost, fidelity, and speed. The key information processing of the CE protocol's is to erase one pair of inputs while keeping the other pair distinct. The truth table for one CE operation, along with the initial and final particle distributions, is illustrated in Figure 2. Under this protocol, both the 00 and 10 particles are mapped to the 00 well, while the 01 and 11 particles remain in their initial wells. Therefore, the protocol performs a conditional erasure: the first bit is erased to 0 if and only if the second bit is 0.

We refer to the 00 and 10 pair as the erasure pair and the 01 and 11 pair as the storage pair. In terms of binary logic, this CE returns $x' = \mathrm{AND}(x, y)$ and $y' = y$. It is worth noting that we cannot think of CEs in general as returning an AND and an identity. Another choice of CE that does the same information processing erases 00 and 10 to 10 (instead of 00), resulting in $(x', y') = (\mathrm{OR}(x, \mathrm{NOT}(y)), y)$. The full set of logical operations that can be implemented using CE protocols are detailed in Ref. [39].

Figure 3 displays snapshots of the potential landscape in the $\varphi_1 - \varphi_2$ space at four distinct times in the CE cycle. The graphs below the contour plots

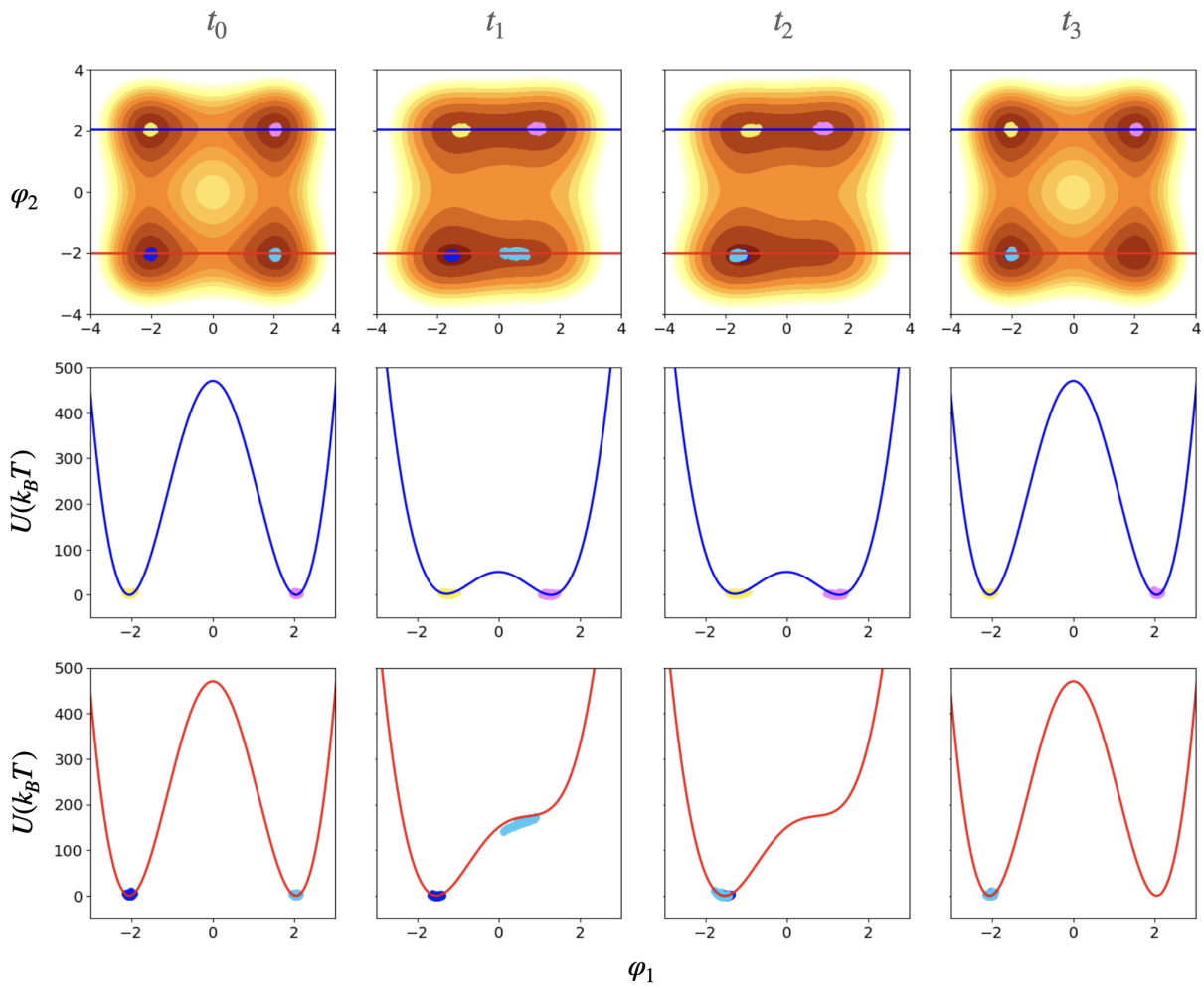


Figure 3. Snapshots illustrating the evolution of the CE protocol are presented at four distinct time points. The first row of the figure contains the contour plots of the potential energy landscape in the φ_1-φ_2 state space, where the colored circles mark the instantaneous distributions of the four particle types (00, 01, 10, and 11). The second and third rows depict the corresponding potential energy graphs along the blue and red cutlines, respectively, as indicated on the contour plots. (Animation available online).

illustrate the potential along the blue cutline (at $\varphi_2 = 2.3$) and the red cutline (at $\varphi_2 = -2.3$). The dots in the contour plots represent samples from the distribution of particles at that particular moment, with the color denoting the initial memory state of each sampled state.

The cycle begins at t_0 with a four-well potential. A detailed illustration of four-well potential is provided in Figure 4(a), where the numbered diamonds identify the critical points in the potential's energy landscape: Points 0 through 3 are local minima and Points 4 through 7 are saddle points.

At t_1, the potential is manipulated to achieve a saddle-node bifurcation. Figure 4(b) shows a detailed illustration of the potential at that particular time, showing how the critical points have evolved from t_0. The potential differences between these critical points are labeled as ΔU_{ab}, where the indices of the subscript indicate the points to which they refer. Figure 4(c) shows the potential along the blue and red cutlines. At this moment in the operation, the local minimum at Point 2 merges with the saddle point at Point 5. The elimination of ΔU_{25} removes the barrier that previously separated the 00 and 10 regions. As a consequence, the 10 particles (light blue) slide down a potential gradient defined by the height ΔU_{05}, gaining kinetic energy. Concurrently, the potential barriers ΔU_{17} and ΔU_{37} are maintained at approximately 50 k_BT. This barrier height is sufficient to ensure a negligible escape rate ($\approx 10^{-20}$ per operation) for the particles of the storage pair (01 and 11) during the protocol duration,

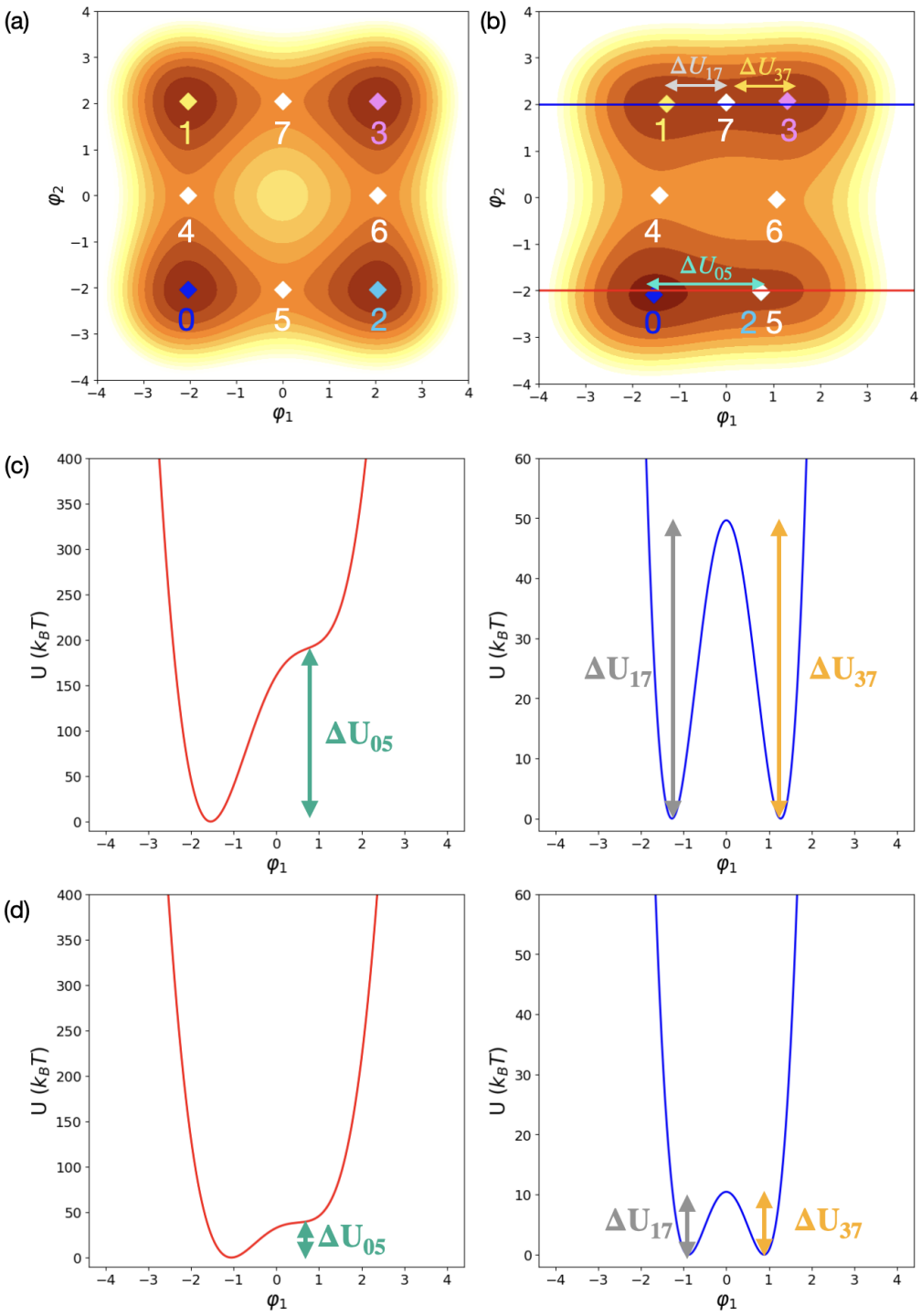


Figure 4. (a) Contour graph of the four-well potential. The diamonds indicate critical points of the potential. Number 0 to 3 are the local minima and number 4 to 7 are saddle points. (b) Potential contour at the moment of saddle point bifurcation happens. ΔU_{ij} indicate the potential differences between the critical points. The first index and second index in the subscript indicate the the local minimum and saddle point refering to. At the bifurcation point, stable point 2 and saddle point 5 merge together. (c) The left (right) potential plots trace the potential along the red (blue) line in (b). (d) Reducing the potential barrier height ΔU_{05} leads to a decrease in the overall work cost of the CE protocol. These two graphs illustrate the potential energy profile along the respective cutlines when a significantly lower barrier height is employed.

confining them to their initial wells.

The potential is held constant at this bifurcation point until t_2, allowing the 10 particles to fully transition into the 00 well. The kinetic energy gained during the descent is subsequently dissipated into the surrounding environment as heat, causing the particles to equilibrate. This process effectively erases the information that previously distinguished the 00 and 10 states. Throughout this process, the 01 and 11 particles remain isolated in their initial wells. This dissipative process dominates the protocol's energetic cost.

An alternative approach would use a pitchfork bifurcation, instead of a saddle-node bifurcation to merge the 00 and 10 wells. However, this is precluded by the necessity of maintaining high barriers to preserve the information stored in the 10 and 11 wells [39].

It is important to note that the motion of the 10 particles during the CE protocol is predominantly in the φ_1 direction, as the barriers separating the top and bottom wells are high enough (about 480 k_BT) to prevent movement over the barrier. This allows their motion to be effectively modeled as one-dimensional. Ref. [39] provides a range of circuit parameters that enable effective CE, establishing the device's robustness against logical errors.

Finally, from t_2 to t_3, the potential is restored to the initial four-well potential, completing one full cycle of the CE operation.

D. Simulation Results of CE

In this subsection, we present the simulation results for the CE protocol. The detailed equations and numerical methods used for these simulations are provided in Appendix C.

1. *Work Distributions*

Figure 5(a) shows the work distributions of all particles at t_3 for the protocol described in Figure 3. The total duration is 150 t_c (0.237ns). The full detailed protocol can be found in Figure 17(a) of Appendix E. Figure 5(b) shows the work distributions of each particle type. A distinguishing characteristic of CE work cost distributions is two peaks with a large distance between them. As expected, work is dominated by contributions from the 10 particles falling into the 00 well. Because only a quarter of the total particles undergo this transition, the average work for the ensemble is approximately one-fourth of the potential difference ΔU_{05} between the higher and lower wells. Meanwhile, the peak close to zero

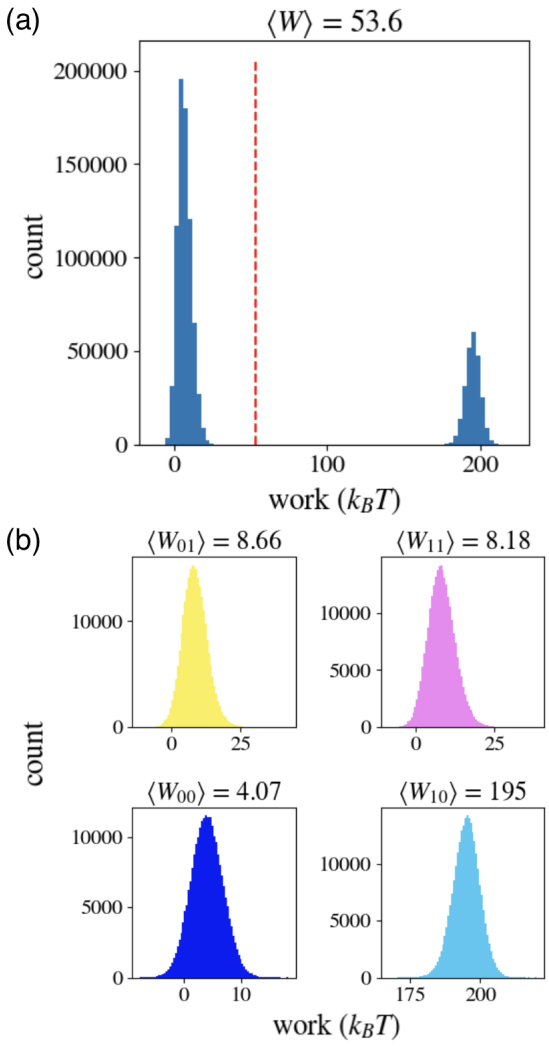


Figure 5. (a) Total work distribution of the CE protocol. The red dashed line indicates the average work, which is 53.6 k_BT. (b) Work distributions for each particle type: The average works for each particle type are also shown in the figure. A signature of this protocol is that the particles moving from a high potential well to a low potential well involve significantly more work compared to other particle types. Details of the protocol are shown in Figure 17(a) in Appendix E.

represents the 00, 01, and 11 particles, all of which remain in their initial local minima throughout the entire logic cycle. Although some work is done to change their potential energies during the protocol, this work is largely returned to the work reservoir upon completion of a full cycle. Consequently, these particles dissipate relatively minor amounts of energy.

2. Work, Fidelity, and Speed Trade-offs

An obvious approach to reducing work cost is to decrease ΔU_{05}, the potential difference between the higher and lower wells. This action directly reduces the potential energy converted into kinetic energy as the 10 particles transition down the potential gradient. This potential difference, however, is inherently constrained by the barrier height separating the 01 and 11 particles—ΔU_{17} and ΔU_{37}. Figures 4 (c) and (d) illustrate the potential differences in two distinct scenarios at t_1. This comes at the cost of increasing the escape rate for the storage pair and degrades computational fidelity. Under typical operating conditions, ΔU_{05} is approximately four times greater than both ΔU_{17} and ΔU_{37} at the bifurcation point. Given this substantial energy barrier, the probability of the 00 and 10 particles thermally exciting into the 10 region is strongly suppressed. Since they dominate the error, the overall CE fidelity can be accurately characterized by the escape rate of the storage pair, which is exponentially damped in the height of ΔU_{17} and ΔU_{37} barriers. Detailed calculations for the escape rate are provided in Appendix B.

This barrier height can be approximated as $\Delta E_B \equiv \beta_1 \cos\frac{\varphi_{1dc}}{2}(1-\cos\frac{\varphi_1^c}{2})$, where φ_1^c is the φ_1-coordinate of the 01 and 11 potential minima. (See derivation in Appendix B.) In many QFP studies [12, 13, 22, 43, 50, 51], it is standard practice to apply the simplification $\varphi_{idc} \approx \varphi_{idc}^c \approx \varphi_{ixdc}$, where φ_{idc}^c is the coordinate of the local minima along the φ_{idc} direction. This choice stems from the fact that the parameter γ is sufficiently large so that oscillations about the energy minima in φ_{dc} are several times smaller and faster than oscillations in φ. Under this assumption, ΔE_B is calculated by substituting $\varphi_{idc} = \varphi_{ixdc}$.

However, the exact relation between φ_{ixdc} and the local minima φ_{idc}^c can be obtained by taking the first derivative of U with respect to φ_{idc}. This yields

$$\varphi_{ixdc} = \varphi_{idc}^c - \frac{\beta_i}{2\gamma_i}\sin\frac{\varphi_{idc}^c}{2} \ . \tag{2}$$

The true minima φ_{idc}^c is somewhat displaced from φ_{ixdc}. The numerical solution of this relationship is shown in Figure 6(a), which plots the difference $\varphi_{idc}^c - \varphi_{ixdc}$, as a function of φ_{ixdc} for various combinations of $\beta = \{1.35, 2.3\}$ and $\gamma = \{9, 16\}$. From the plot, the difference is relatively small (no more than 6%), which appears to justify the assumption $\varphi_{ixdc} \approx \varphi_{idc}$.

However, this small difference in the coordinate can result in a large difference in the barrier height. We introduce two definitions to distinguish the two barrier-height calculations. *Intended barrier height* refers to that calculated under the assumption $\varphi_{idc}^c \approx \varphi_{ixdc}$. In contrast, *effective barrier height* is the actual barrier height, which is derived by first solving Eq.(2) to obtain φ_{idc}^c.

We extend these definitions to error rates. The error rate derived from the intended barrier height is defined as the *intended escape rate* and, similarly, the error rate derived from the effective barrier height is defined as the *effective escape rate*. Figure 6(b) illustrates the relationship between these two escape rates across different combinations of β and γ. The dashed line in the plot represents the ideal case where the intended and effective escape rates are equal.

Notably, for the case of $\beta = 2.3$, the effective escape rate is observed to be as much as 9 orders of magnitude higher than the intended escape rate. In contrast, the difference is significantly smaller when $\beta = 1.35$. This startling disparity highlights the fact that even minor variations between φ_{xdc} and φ_{dc}^c can substantially alter the system's energy landscape. This, in turn, leads to drastic implications for the storage pair's escape probability.

Figure 7 shows the work cost and fidelity as a function of the intended barrier height for different combinations of β and γ. One of the most striking features of the simulation results are the substantial differences in the observed error rates for different values of β and γ, even when the "intended" barrier heights are the same.

This can be explained by the difference between the intended barrier height and the effective barrier height, as described above. The nonmonotonic work cost for $\beta = 2.3$ reflects a failure of reliable information storage, also due to the inflated error rate. On the one hand, as the number of errors increases, more and more particles transit between the potential wells and thus do not stay at the potential minima. This increases the work cost. On the other hand, in the case of $\beta = 1.35$ (represented by the gray and red data points), the difference between the intended barrier height and effective barrier height is much smaller. Although errors are also observed for intended barrier height smaller than 20 k_BT, the error rate is orders of magnitude smaller.

Our analysis now shifts to the effect of operating speed on work cost and fidelity. Figure 8 illustrates the relationship between total protocol duration and performance, using a base case with $\beta = 1.35$ and $\gamma = 16$ and an intended barrier height of around 15 k_BT. Both work cost and the error rate increase as the total protocol duration is shortened from 150 t_c to 42 t_c. A faster protocol gives particles extra kinetic energy through sudden changes in potential energy. This excess energy increases the probability of the storage pair escaping its intended wells, leading to a higher error rate. Additionally,

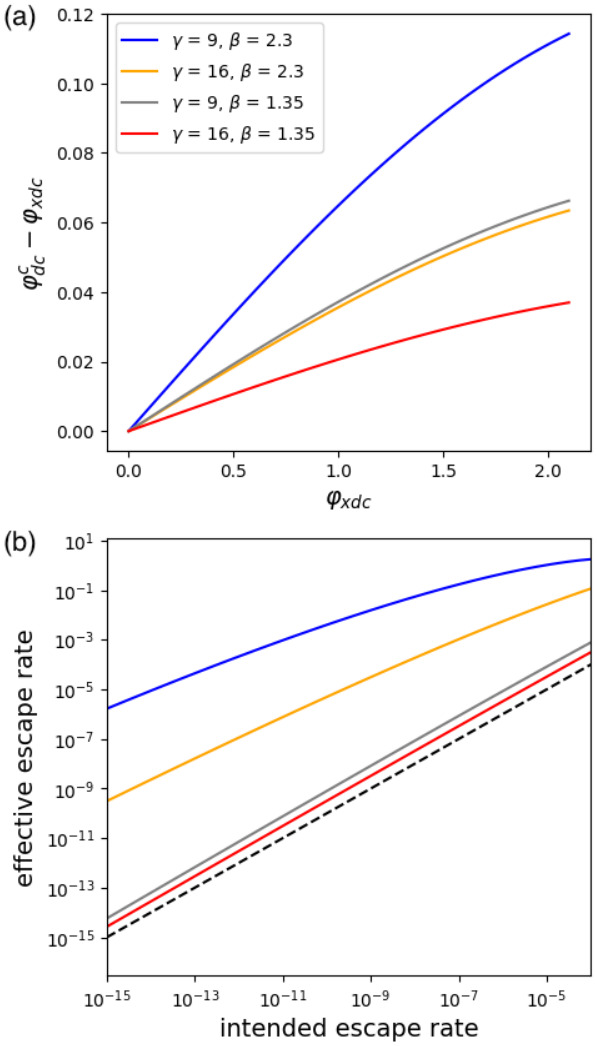


Figure 6. (a) This plot shows the difference between φ^c_{dc} and φ_{xdc} as φ_{xdc} increases from 0 to 2.1. The curves are generated by numerically solving Eq. (2) for four combinations of the coupling parameters, $\beta = \{1.35, 2.3\}$ and $\gamma = \{9, 16\}$. (b) This graph plots the intended escape rate against the effective escape rate per operation. The black dashed line represents the ideal case where the intended escape rate equals the effective escape rate. Note that the effective escape rate is always larger than the intended rate.

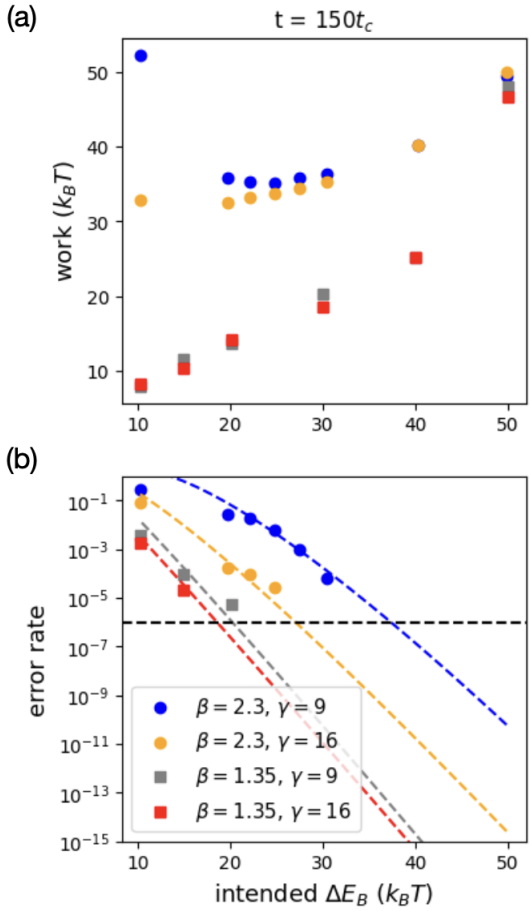


Figure 7. (a) Work cost for different combinations of β and γ with the time for duration with length of 150 t_c. The number of samples in each case is 10^6. When the intended barrier height decreases, the work cost for $\beta = 2.3$ first decreases but then increases afterward. However, for $\beta = 1.35$, the work costs keep decreasing. (b) Error rate against intended barrier height for different combinations of β and γ. The horizontal dashed line represents the threshold of observing errors given the sample size. The colored dashed lines represent the theoretical escape rates for the four cases. The observed error saturates at .25, as the particles become effectively randomized in the presence of vanishing barrier height

it increases the work cost because the excess kinetic energy is eventually dissipated into the environment. To maintain fidelity and reduce escape events as the barrier height decreases, the protocol must be executed slowly enough to ensure that the storage particles remain locally equilibrated to their local minima throughout the operation.

3. *Problem of Controlled Erasure*

In the CE protocol, maintaining a high potential barrier between the 01 and 11 wells is essential to prevent undesired particle transitions. First, this requirement—maintaining the storage pair within their own wells—led us to employ a nonadiabatic saddle-node bifurcation, rather than an adiabatic pitchfork bifurcation. Next, we observed another direct trade-off stemming from the storage bits: reducing the barrier height lowers the work cost but simul-

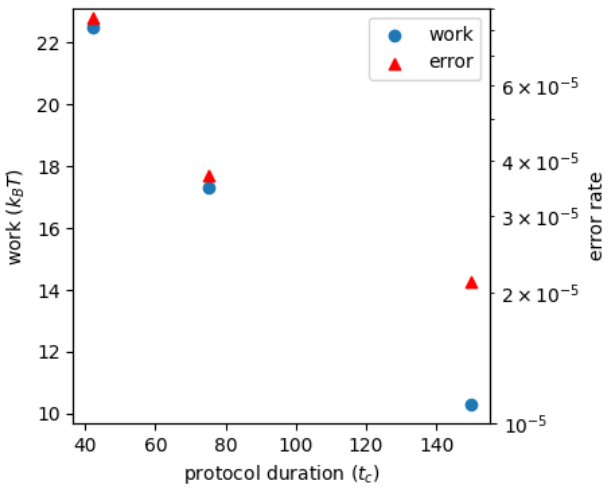


Figure 8. Work cost and error rate for a system with $\beta = 1.35$ and $\gamma = 16$ and with intended barrier height 15 k_BT. Simulated for three different protocol durations: 42, 75, and 150 t_c. As speed increases, the work cost and error rate increase.

taneously decreases fidelity by facilitating particle escape events. Another critical limitation we identified is that the operating speed of a high-fidelity CE is constrained by the storage pair. As the protocol's speed increased, these particles gain kinetic energy, causing them to escape their storage wells and compromise the outcome.

This highlights a fundamental constraint on CE: It relies on static particle positioning to store information during the computation. Theoretically, this should be a strength, since particles near their local equilibrium distributions dissipate energy very slowly. However, since the relevant barrier heights in this system are inescapably coupled during a CE, the result is a tension between the storage pair and the erasure pair. In short, the storage pair needs a constant passive barrier and the erasure pair needs a vanishing one.

The solution is to explore a method that does not require particles of the storage pair to be static. The next section introduces a new protocol, the Erasure-Flip protocol (EF), to solve this problems.

E. Mechanism of Erasure-Flip Protocol

In the momentum computing bit-flip protocol [38], the potential is abruptly changed from a double-well to a quasi-harmonic potential. This harmonic potential is maintained for half an oscillation period to allow the particles representing logical 0 and logical 1 to flip their positions. The potential then reverts to its original double-well form, trapping the flipped particles and preventing them from returning to their initial wells. A key feature of this process is that particles remain separable even when they are spatially mixed near the bottom of the harmonic potential, as their distinct start states are preserved in their opposing momenta. This ability to store and manipulate information in momentum space is a powerful feature not possible with traditional overdamped Langevin dynamics.

Inspired by this, the EF also utilizes a momentum-computing approach. Figure 2 displays the truth table for the EF protocol, along with the initial and final particle distributions for a full EF cycle. In contrast to the CE, the 00 and 10 particles merge within the 10 well, while the 01 and 11 particles flip their positions. Despite these differences, we maintain the nomenclature of calling the 00 and 10 particles the erasure pair and the 01 and 11 particles the storage pair. After all, the flip operation is logically reversible. And so, the logical information separating 01 and 11 particles is preserved.

Figure 9 shows the distribution of the particles and the potential energy of the system at four different times during an EF protocol. The protocol divides into three main stages. On the one hand, in the first stage (t_0 to t_2), the potential along the erasure pair is tilted toward 10 well so that 00 moves toward the 10 well. On the other hand, the barrier height between the storage pair is reduced to zero. The resulting quasi-harmonic potential allows the 01 and 11 particles to flip with each other. In the second stage (t_2 to t_3), the potential barrier is raised to trap these flipped particles, preventing them returning to the previous well. In the last stage (t_3 to t_4), the potential returns to the original four-well potential.

In this way, the EF's design addresses the CE's fundamental limitations. In stark contrast to CE, which must avoid the storage pair needlessly gaining kinetic energy, we intentionally impart kinetic energy to them and encourage these particles to move into the other well. Thus, the excess kinetic energy that was a source of error in CE is now a necessary condition for successful EF.

F. Simulation Results for EF

1. *Work distributions*

Figure 10(a) shows the total work for the EF protocol and individual work distributions for each particle type. In contrast to the CE, all four types move under EF. As none of them simply sit near local minima during the protocol, we expect for each particle

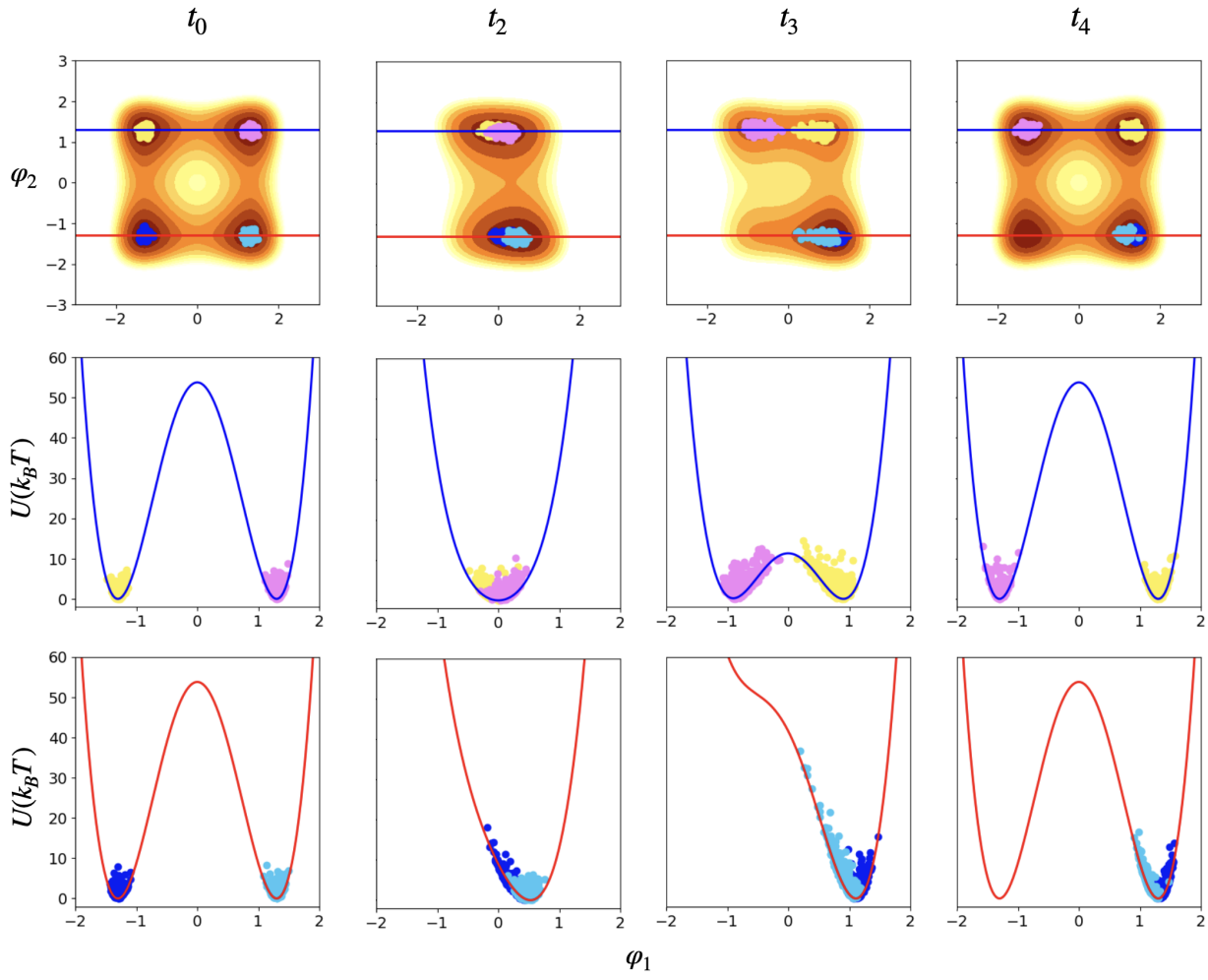


Figure 9. Snapshots of the evolution of the potential and particle distributions of the EF protocol at four important time steps: $t_0, t_2, t_3,$ and t_4. The partial barrier lowering at t_1 is not shown. (Animation available online).

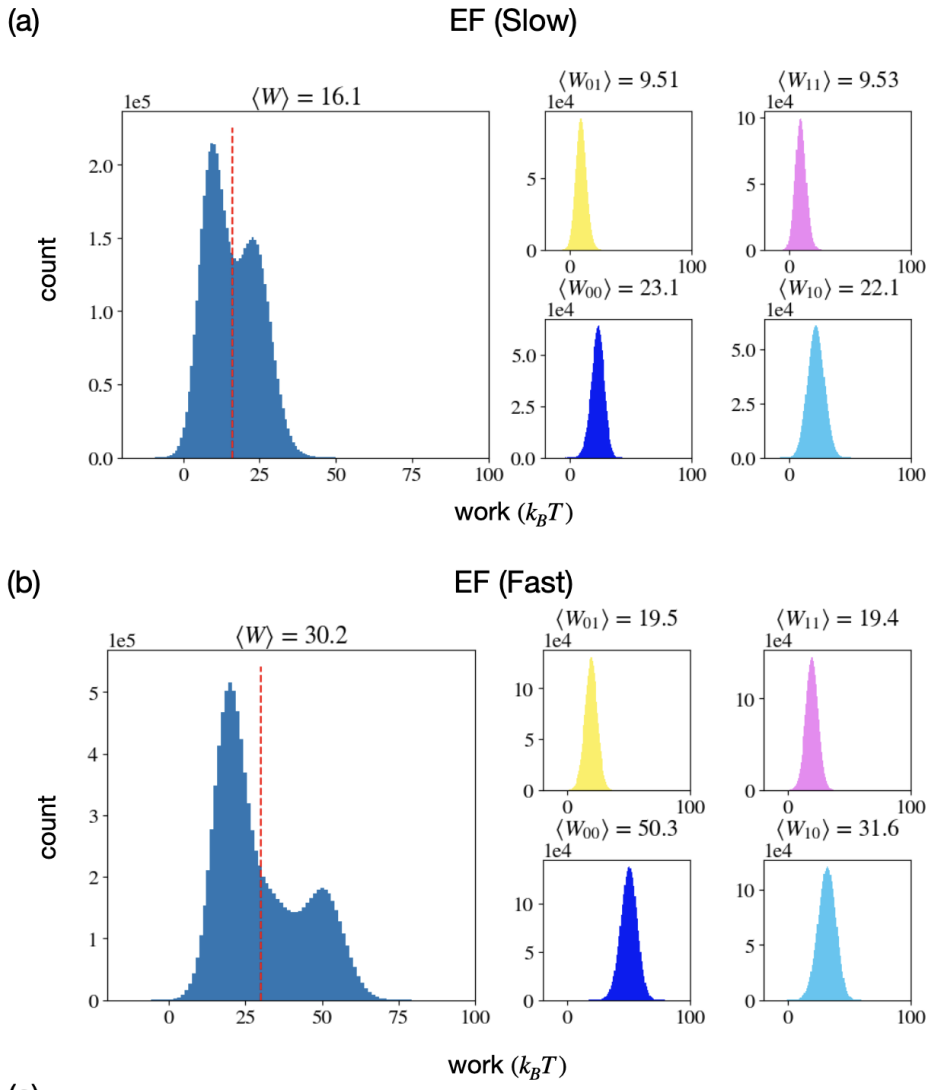
type that some input work will change the potential energy and so be converted to kinetic energy and dissipated to the surrounding. The average work cost is about 16 k_BT, with the average cost of the storage and erasure pairs being $\approx$ 9 k_BT and 23 k_BT, respectively. Here, the storage pair costs relatively less work since a significant portion of the kinetic energy gained from t_0 to t_2 is reinvested into potential energy from t_2 to t_3.

Section II C's CE analysis demonstrates that lowering the potential barrier height is beneficial for reducing the potential difference between the higher and lower wells of the erasure pair, thereby decreasing work cost. On the one hand, the EF takes this principle further by completely removing the potential barrier, allowing for a lowering the amount of potential energy converted to kinetic energy when 00 particles slide down the CE potential gradient. On the other hand, the EF protocol requires a rapid change in φ_{xdc} to trap the particles at just the right time. The work cost associated with this rapid change is directly related to the parameter γ, which can be conceptualized as the system's "stiffness" in the φ_{dc} space. Consequently, a smaller value of γ is crucial for reducing the work cost of EF.

2. *Trading-off work, fidelity, and speed*

With the potential barrier from the 10 well to the 00 well exceeding 50 k_BT at t_3 (Fig. 9), the probability of a particle transitioning back to the 00 well becomes exceedingly low. Consequently, the primary source of error still comes from the storage pair—specifically from the flipping between the 01 and 11 particles. The mechanism behind these errors is, however, quite different from CE.

The precise timing of the barrier-lowering phase



(c)

	EF (slow)	EF (fast)
duration (t_c)	42	34
work (k_BT)	16.1 ± 0.011	30.2 ± 0.0133
error rate	3.00×10^{-6}	2.00×10^{-7}
source of error	00 (0), 01 (12) 10 (0), 11 (18)	00 (2), 01 (0) 10 (0), 11 (0)

Figure 10. Work distributions for (a) EF (slow) and (b) EF (fast). For each protocol, the total work distribution across all particles and the work contribution from individual particle types are shown. The total number of particles simulated for both protocols is $N = 10^7$. (c) A summary table illustrating the duration, mean work cost, error rate, and source of error for the two protocols. For the error source, the number in the bracket is the number of observed errors for that particular particle type. Details of the protocol parameters are shown in Figure 17(b) and (c).

is critical to the EF protocol's success. If the barrier is lowered too slowly (t_0 to t_2), particles may not gain sufficient kinetic energy to transition to the target well. Instead, they could remain near the central point ($\varphi_1 = 0$) when the barrier starts to rise and end up in a random (indeed, wrong) well. Conversely, if the barrier is raised too quickly (t_2 to t_3), some particles may not have reached the target well. Then rising the barrier too slowly could allow particles in the target well to return to their original well.

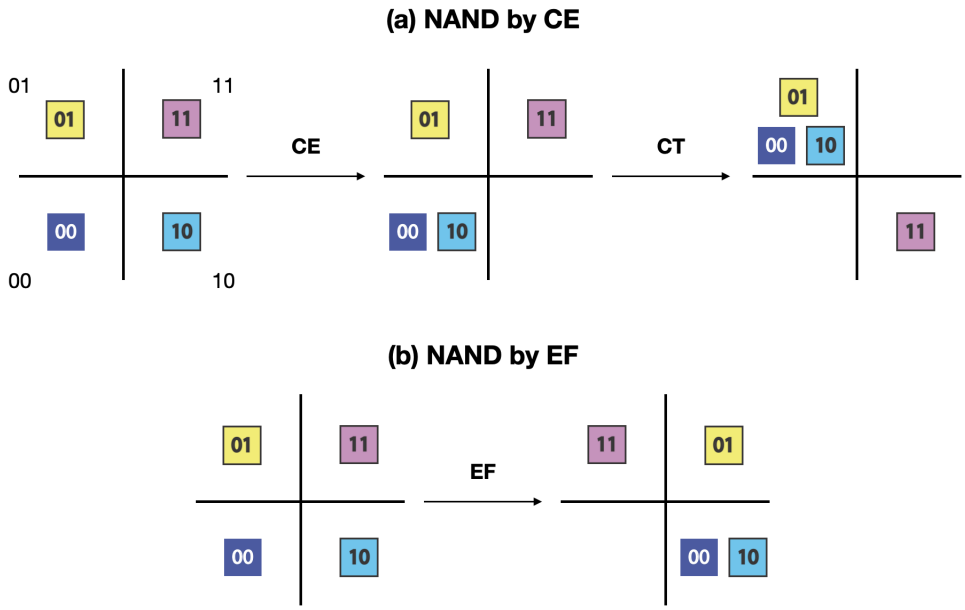


Figure 11. Schematic illustration comparing the partial NAND operation under (a) CE and (b) EF. After a CE operation, one more conditional tilt operation (CT) is required to finish the NAND operation. The first bit is the irrelevant and the second bit is the output bit. For the EF, the final distribution already satisfies the requirement of a partial NAND. The first bit is the output bit and the second bit is the irrelevant bit.

Therefore, a careful balance (and protocol timing) must be maintained to optimize both work cost and fidelity.

The simulation error rate and error source are summarized in the table of Figure 10. In 10^7 simulations, we observed only 30 instances of particles failing to reach the target wells, yielding an error rate of 3×10^{-6}. As expected, all of the failure particles are associated with the storage pair (12 trajectories from 01 particles and 18 trajectories from 11 particles) because these particles do not have sufficient speed to complete the transition.

Therefore, we simulated an alternative, faster protocol. Figure 10(b) shows the results. Across 10^7 trajectories, only two errors were observed, and these originated from the erasure pair rather than the storage pair. A detailed comparison of the average kinetic energy and examples of failure trajectories for the slow and fast EF are shown in Appendix D. This comparison suggests an intermediate protocol lying between these two might best minimize the error rate.

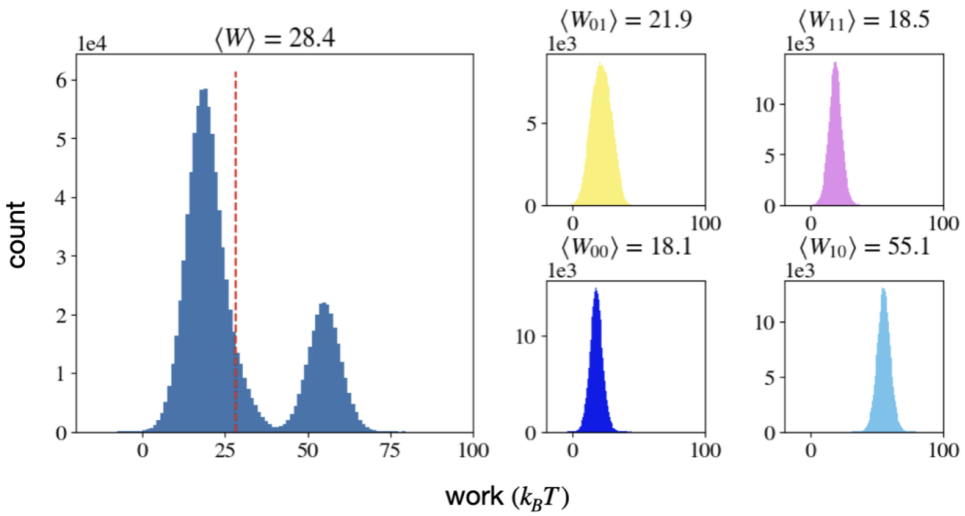
Taken together, these results highlight that a delicate balance must be struck between protocol parameters and duration to find optimal momentum-computing protocols. Finding an optimal protocol for EF that balances speed, work cost, and fidelity remains an open question. To address this, though, we have developed machine learning methods [52] that are a promising and well-behaved approach for optimization.

G. NAND operation by Controlled Erasure and Erasure-Flip

Finally, we discuss the costs of universal computation using both the adiabatic and the momentum computing approaches. Figure 11 shows the steps for achieving a partial NAND operation using either the CE or the EF protocol. A NAND based on the CE requires an additional operation. In order to stay within the adiabatic paradigm, we choose an operation that does not take advantage of momentum: the conditional tilt (CT). A CT is done first by lowering the potential barriers along a particular direction, and then changing the magnetic coupling. This protocol was investigated in detail in [47]; it can be implemented adiabatically—with work costs that approach Landauer's bound. While the energy cost of a CT can theoretically approach the Landauer limit with an infinitely long duration, we operate at a finite time scale, so the additional costs levied by this step are an important factor. Note also that, in this case, the first output bit is the irrelevant bit and the second output bit yields the NAND operation.

Figure 16 in Appendix E details the work cost of CT as a function of its duration. We find that a

(a)



(b)

	CE (fast)	EF(fast)	CE (slow)	EF(slow)
duration (t_c)	150 + 50	34	150 + 150	42
work (k_BT)	28.4 ± 0.0494	30.2 ± 0.0133	16.4 ± 0.0458	$16.1 \pm .011$
error rate	2.15×10^{-5}	2.00×10^{-7}	2.15×10^{-5}	3.00×10^{-6}

Figure 12. (a) This panel shows the total work distribution alongside the individual particle-type work distributions for the partial NAND protocol implemented via CE with CT duration 50 t_c (CE (fast)). Since the shape of the work distribution of CE (slow) is the same as CE (fast) with lower work cost, the work distribution for CE (slow) is not shown here. Animation for the protocol available via the links: NAND by CE (fast). (b) Table listing key performance metrics, including the duration, mean work cost, and error rate for CE (fast), EF (fast), CE (slow) and EF (slow) schemes.

CT with a duration of $50t_c$ incurs an energy cost of approximately $20k_BT$. The cost reduces to around $4k_BT$ when the duration is extended to $250t_c$. To facilitate a meaningful thermodynamic comparison, we select CT durations to ensure that the work cost of the NAND operations implemented by CE are comparable to the work costs of their respective counterparts, the EF (fast) and EF (slow) protocols. This results in CT durations of $50t_c$ and $150t_c$; we designate the respective CE protocols as CE (fast) and CE (slow), respectively.

On the other hand, a single EF cycle is sufficient to implement a partial NAND operation; the first output bit can be read directly to get the NAND result. The second output bit is simply a copy of the second input bit.

Figure 12 illustrates the work cost achieved when implementing the NAND operation using the CE (fast) protocol. The work distributions for the NAND operation implemented with the EF protocols are already presented in Figure 10, because no extra steps are required. The accompanying table summarizes the key comparative metrics—work cost, fidelity, and speed—for the four main protocols. Considering the comparison between NAND by CE (fast) and NAND by EF (fast), the advantage of EF protocol becomes apparent. While the mean work costs for both protocols are approximately 30 k_BT, EF (fast) protocol demonstrates vastly superior performance in both speed and fidelity. Specifically, EF (fast) implementation is six times faster and achieves an improvement in fidelity of two orders of magnitude compared to CE. Furthermore, comparing CE (slow) with EF (slow) also demonstrates EF's better overall performance. With work cost roughly 16 k_BT, EF (slow) is approximately seven times faster and the fidelity is ten times better than that of CE (slow).

As a final point of comparison, we speed up the CE-based NAND so that it is as fast as the fastest EF with a duration of 34 t_c. This induces a huge energy penalty with an average work of $\approx 108\ k_BT$ and the error rate spikes to an untenable $\approx 13\%$ (for comparison, randomization of unbiased inputs would achieve a partial NAND 50% of the time). At comparable operating speeds, it costs 3.6 times more energy while incurring more than $650,000$ errors for every error seen in the EF-based implementation. These comparisons underscore the foundational advantage of our hybrid reversible/irreversible

logic as a vehicle for general-purpose computing. The remarkable combination of speed and accuracy, achieved across different operating points without increasing thermodynamic costs, demonstrates the profound advantage of leveraging momentum in computation. And, this points to promising pathways for achieving dramatic improvements in speed, energy efficiency, and fidelity in digital computation.

III. DISCUSSION

This investigaton first pursued a detailed analysis of the work, fidelity, and speed of the Controlled Erasure (CE) protocol. The findings revealed an inherent lower bound on its work cost, when implemented with a CQFP. This cost arises from the unavoidable transfer of particles from higher to lower potential wells, due to an inescapable use of a saddle-node bifurcation. Furthermore, the protocol's speed was fundamentally limited by the storage pair, as rapid execution generates excessive kinetic energy, which in turn increases the probability of escape events.

To overcome these limitations, we introduced a logic gate that simultaneously performs reversible and irreversible logics on different memory subspaces. This universal operation, the Erasure-Flip (EF), transiently stores information in momentum space during the computational protocol, allowing logical states to depart from their local equilibria without losing information. This flexibility overcame the speed limit imposed by requiring information be stored in near-equilibrium quasi-static states. While the EF protocol does present greater control challenges, our results show the EF protocol significantly outperforms our adiabatic test operation in operating speed and in computational fidelity without leveraging additional energetic penalties. This successfully demonstrates the strength of the underdamped momentum computing paradigm when compared to the conventional overdamped approach. This finding opens the door to fundamentally different and more efficient computational architectures. This, in turn, paves the way for highly energy-efficient, accurate, and fast protocols that are not achievable with traditional overdamped perspective. Future work will focus on further exploring and building upon this promising new thermodynamic computing paradigm.

IV. METHOD

The simulation is written in Python using its NumPy library, and the figures and animations are generated by Matplotlib. To obtain the initial states, Monte Carlo sampling was performed. The system dynamics were then simulated using a hybrid integration scheme: a 4th-order Runge-Kutta method was applied to the deterministic portion, while the Euler-Maruyama method was used for the stochastic portion. This procedure generates an ensemble of trajectories for the system's initial state. Physical quantities, such as the mean work, are computed by taking the average across the trajectories produced by the simulator. A fixed time step of $dt = 1/1000$ was employed for all simulations. Statistics for physical observables, such as the mean work, are obtained by ensemble averaging over independent trajectories. Detailed derivations of the Langevin equation and additional simulation parameters are provided in Appendices A and C.

V. DATA AVAILABILITY

The data in the manuscript are provided in the main text, Supplementary Note, and available code repository. Additional data related to this study can be made available from the authors upon reasonable request.

VI. CODE AVAILABILITY

The simulator used is open-sourced and available at https://github.com/tkwtang/source/.

VII. AUTHOR CONTRIBUTIONS

KWT conceived the idea of the Erasure-Flip and was responsible for generating the simulation data, formal analysis, and visualization. KJR and JPC provided critical feedback, verified the theoretical framework, and were involved in supervision. JPC provided funding. All authors contributed to the editing of the manuscript and approved the final version for submission.

VIII. COMPETING INTERESTS

The authors declare no competing interests.

ACKNOWLEDGMENTS

We thank Christian Pratt and Sangbum Kim for helpful discussions. The authors thank the Telluride Science Research Center for its hospitality during

visits and the participants of the Information Engines workshop there for their valuable feedback. This material is based on work supported by, or in part by, the Art and Science Laboratory via a gift to UC Davis.

[1] C. Bogmans, P. Gomez-Gonzalez, G. Ganpurev, G. Melina, A. Pescatori, and S. Thube, Power hungry: How ai will drive energy demand, IMF Working Papers **81**, 2025 (2025).
[2] X. Chen, X. Wang, A. Colacelli, M. Lee, and L. Xie, Electricity demand and grid impacts of ai data centers: Challenges and prospects, arXiv preprint arXiv:2509.07218 (2025).
[3] E. Masanet, A. Shehabi, N. Lei, S. Smith, and J. Koomey, Recalibrating global data center energy-use estimates, Science **367**, 984 (2020).
[4] G. Moore, The future of integrated electronics, Fairchild Semiconductor internal publication **2** (1964).
[5] G. E. Moore, Cramming more components onto integrated circuits, Proceedings of the IEEE **86**, 82 (1998).
[6] H. Wong and H. Iwai, The road to miniaturization, Physics World **18**, 40 (2005).
[7] M. U. Khan, A. Arshad, S. Ume, R. Bukhari, A. Samer, and F. Tariq, Nanoscale CMOS scaling trends, challenges and their solutions, in *ICAME21, International Conference on Advances in Mechanical Engineering, Pakistan*, August (2021) pp. 1–10.
[8] H. H. Radamson, Y. Miao, Z. Zhou, Z. Wu, Z. Kong, J. Gao, H. Yang, Y. Ren, Y. Zhang, J. Shi, *et al.*, CMOS scaling for the 5 nm node and beyond: Device, process and technology, Nanomaterials **14**, 837 (2024).
[9] T. Kuroda, CMOS design challenges to power wall, in *Digest of Papers. Microprocesses and Nanotechnology 2001. 2001 International Microprocesses and Nanotechnology Conference (IEEE Cat. No. 01EX468)* (IEEE, 2001) pp. 6–7.
[10] M. Frank, Approaching the physical limits of computing, in *35th International Symposium on Multiple-Valued Logic (ISMVL'05)* (2005) pp. 168–185.
[11] K. Loe and E. Goto, Analysis of flux input and output josephson pair device, IEEE Transactions on Magnetics **21**, 884 (1985).
[12] R. Rouse, S. Han, and J. E. Lukens, Observation of resonant tunneling between macroscopically distinct quantum levels, Phys. Rev. Lett. **75**, 1614 (1995).
[13] S. Han, J. Lapointe, and J. E. Lukens, Variable ß RF SQUID, in *Single-Electron Tunneling and Mesoscopic Devices*, edited by H. Koch and H. Lübbig (Springer Berlin Heidelberg, Berlin, Heidelberg, 1992) pp. 219–222.
[14] K. K. Likharev, Classical and quantum limitations on energy consumption in computation, International Journal of Theoretical Physics **21**, 311 (1982).
[15] Y. Harada, E. Goto, and N. Miyamoto, Quantum flux parametron, in *1987 International Electron Devices Meeting* (1987) pp. 389–392.
[16] M. Hosoya, W. Hioe, J. Casas, R. Kamikawai, Y. Harada, Y. Wada, H. Nakane, R. Suda, and E. Goto, Quantum flux parametron: A single quantum flux device for Josephson supercomputer, IEEE Transactions on Applied Superconductivity **1**, 77 (1991).
[17] W. Hioe *et al.*, *Quantum flux parametron: a single quantum flux superconducting logic device*, Vol. 2 (World Scientific, 1991).
[18] K. Likharev, Dynamics of some single flux quantum devices: I. parametric quantron, IEEE Transactions on Magnetics **13**, 242 (1977).
[19] O. Mukhanov, V. Semenov, and K. Likharev, Ultimate performance of the RSFQ logic circuits, IEEE Transactions on Magnetics **23**, 759 (1987).
[20] W. Chen, A. Rylyakov, V. Patel, J. Lukens, and K. Likharev, Rapid single flux quantum T-flip flop operating up to 770 GHz, IEEE Transactions on Applied Superconductivity **9**, 3212 (1999).
[21] M. Tanaka, M. Ito, A. Kitayama, T. Kouketsu, and A. Fujimaki, 18-GHz, 4.0-aJ/bit operation of ultra-low-energy rapid single-flux-quantum shift registers, Japanese Journal of Applied Physics **51**, 053102 (2012).
[22] N. Takeuchi, T. Yamae, C. L. Ayala, H. Suzuki, and N. Yoshikawa, Adiabatic quantum-flux-parametron: A tutorial review, IEICE Transactions on Electronics **105**, 251 (2022).
[23] N. Takeuchi, T. Yamae, C. L. Ayala, H. Suzuki, and N. Yoshikawa, An adiabatic superconductor 8-bit adder with 24kBT energy dissipation per junction, Applied Physics Letters **114**, 042602 (2019).
[24] N. Takeuchi, Y. Yamanashi, and N. Yoshikawa, Thermodynamic study of energy dissipation in adiabatic superconductor logic, Phys. Rev. Appl. **4**, 034007 (2015).
[25] N. Takeuchi, Y. Yamanashi, and N. Yoshikawa, Measurement of 10 zJ energy dissipation of adiabatic quantum-flux-parametron logic using a superconducting resonator, Applied Physics Letters **102**, 052602 (2013).
[26] O. Chen, R. Cai, Y. Wang, F. Ke, T. Yamae, R. Saito, N. Takeuchi, and N. Yoshikawa, Adiabatic quantum-flux-parametron: Towards building extremely energy-efficient circuits and systems, Scientific Reports **9**, 10514 (2019).
[27] N. Takeuchi, Y. Yamanashi, and N. Yoshikawa, Reversible logic gate using adiabatic superconducting devices, Scientific Reports **4**, 6354 (2014).
[28] N. Takeuchi, Y. Yamanashi, and N. Yoshikawa, Reversibility and energy dissipation in adiabatic superconductor logic, Scientific Reports **7**, 75 (2017).

[29] T. Yamae, N. Takeuchi, and N. Yoshikawa, Minimum energy dissipation required for information processing using adiabatic quantum-flux-parametron circuits, Journal of Applied Physics **135** (2024).

[30] V. K. Semenov, G. V. Danilov, and D. V. Averin, Negative-inductance squid as the basic element of reversible josephson-junction circuits, IEEE transactions on applied superconductivity **13**, 938 (2003).

[31] V. K. Semenov, G. V. Danilov, and D. V. Averin, Classical and quantum operation modes of the reversible josephson-junction logic circuits, IEEE Transactions on Applied Superconductivity **17**, 455 (2007).

[32] K. D. Osborn and W. Wustmann, Reversible fluxon logic for future computing, in *2019 IEEE International Superconductive Electronics Conference (ISEC)* (IEEE, 2019) pp. 1–5.

[33] M. P. Frank, R. M. Lewis, N. A. Missert, M. A. Wolak, and M. D. Henry, Asynchronous ballistic reversible fluxon logic, IEEE Transactions on Applied Superconductivity **29**, 1 (2019).

[34] W. Wustmann and K. D. Osborn, Reversible fluxon logic: Topological particles allow ballistic gates along one-dimensional paths, Physical Review B **101**, 014516 (2020).

[35] Y. P. Kalmykov and W. T. Coffey, *Langevin Equation, The: With Applications To Stochastic Problems in Physics, Chemistry and Electrical Engineering*, Vol. 27 (World Scientific, 2012).

[36] K. J. Ray, A. B. Boyd, G. W. Wimsatt, and J. P. Crutchfield, Non-markovian momentum computing: Thermodynamically efficient and computation universal, Phys. Rev. Res. **3**, 023164 (2021).

[37] S. Dago and L. Bellon, Logical and thermodynamical reversibility: Optimized experimental implementation of the not operation, Physical Review E **108**, L022101 (2023).

[38] K. J. Ray and J. P. Crutchfield, Gigahertz sub-landauer momentum computing, Phys. Rev. Appl. **19**, 014049 (2023).

[39] C. Z. Pratt, K. J. Ray, and J. P. Crutchfield, Controlled erasure as a building block for universal thermodynamically robust superconducting computing, Chaos: An Interdisciplinary Journal of Nonlinear Science **35** (2025).

[40] I. Ozfidan, C. Deng, A. Smirnov, T. Lanting, R. Harris, L. Swenson, J. Whittaker, F. Altomare, M. Babcock, C. Baron, A. Berkley, K. Boothby, H. Christiani, P. Bunyk, C. Enderud, B. Evert, M. Hager, A. Hajda, J. Hilton, S. Huang, E. Hoskinson, M. Johnson, K. Jooya, E. Ladizinsky, N. Ladizinsky, R. Li, A. MacDonald, D. Marsden, G. Marsden, T. Medina, R. Molavi, R. Neufeld, M. Nissen, M. Norouzpour, T. Oh, I. Pavlov, I. Perminov, G. Poulin-Lamarre, M. Reis, T. Prescott, C. Rich, Y. Sato, G. Sterling, N. Tsai, M. Volkmann, W. Wilkinson, J. Yao, and M. Amin, Demonstration of a nonstoquastic Hamiltonian in coupled superconducting flux qubits, Phys. Rev. Appl. **13**, 034037 (2020).

[41] A. M. van den Brink, A. J. Berkley, and M. Yalowsky, Mediated tunable coupling of flux qubits, New Journal of Physics **7**, 230 (2005).

[42] R. Harris, A. J. Berkley, M. W. Johnson, P. Bunyk, S. Govorkov, M. C. Thom, S. Uchaikin, A. B. Wilson, J. Chung, E. Holtham, J. D. Biamonte, A. Y. Smirnov, M. H. S. Amin, and A. Maassen van den Brink, Sign- and magnitude-tunable coupler for superconducting flux qubits, Phys. Rev. Lett. **98**, 177001 (2007).

[43] R. Harris, T. Lanting, A. J. Berkley, J. Johansson, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han, Compound Josephson-junction coupler for flux qubits with minimal crosstalk, Phys. Rev. B **80**, 052506 (2009).

[44] W. C. Stewart, Current-Voltage Characteristics of Josephson Junctions, Applied Physics Letters **12**, 277 (1968).

[45] D. E. McCumber, Effect of AC impedance on DC voltage-current characteristics of superconductor weak-link junctions, Journal of Applied Physics **39**, 3113 (1968).

[46] C. Z. Pratt, K. J. Ray, and J. P. Crutchfield, Extracting equations of motion from superconducting circuits, Phys. Rev. Res. **7**, 013014 (2025).

[47] K. W. Tang, K. J. Ray, and J. P. Crutchfield, Nonequilibrium thermodynamics of a superconducting Szilard engine, New Journal of Physics **27**, 043011 (2025).

[48] H. R. Lewis and C. H. Papadimitriou, *Elements of the Theory of Computation*, 2nd ed. (Prentice-Hall, Englewood Cliffs, N.J., 1998).

[49] R. Landauer, Irreversibility and heat generation in the computing process, IBM Journal of Research and Development **5**, 183 (1961).

[50] S. Han, J. Lapointe, and J. E. Lukens, Effect of a two-dimensional potential on the rate of thermally induced escape over the potential barrier, Phys. Rev. B **46**, 6338 (1992).

[51] N. Takeuchi, D. Ozawa, Y. Yamanashi, and N. Yoshikawa, An adiabatic quantum flux parametron as an ultra-low-power logic device, Superconductor Science and Technology **26**, 035010 (2013).

[52] J. Lyu, K. J. Ray, and J. P. Crutchfield, Optimal computation from fluctuation responses https://doi.org/10.48550/arXiv.2510.03900 (2025).

[53] S. Han, J. Lapointe, and J. E. Lukens, Thermal activation in a two-dimensional potential, Phys. Rev. Lett. **63**, 1712 (1989).

[54] J. M. Martinis, M. H. Devoret, and J. Clarke, Experimental tests for the quantum behavior of a macroscopic degree of freedom: The phase difference across a Josephson junction, Phys. Rev. B **35**, 4682 (1987).

[55] A. B. Boyd, D. Mandal, and J. P. Crutchfield, Thermodynamics of modularity: Structural costs beyond the Landauer bound, Phys. Rev. X **8**, 031036 (2018).

[56] D. H. Wolpert and A. Kolchinsky, Thermodynamics of computing with circuits, New Journal of Physics **22**, 063047 (2020).

Appendix A: Langevin Dynamics for Josephson Circuits

Underdamped Langevin dynamics can often be used to simulate the behavior of the flux degrees of freedom when Josephson junctions (JJs) are modeled as resistively and capacitively shunted junction (RCSJ) model. The dimensional Langevin equation is:

$$m_i dv_i + \nu_i v_i dt = -\partial_{x_i} U(\vec{x}; \vec{\lambda}_{ext}(t))dt + r_i(t)\sqrt{2\nu_i \kappa dt}\,. \tag{A1}$$

Here, $U(\vec{x}; \vec{\lambda}_{ext}(t))$ is the driving potential and $\vec{\lambda}_{ext}(t)$ is the protocol of the external parameters. κ is equal to $k_B T$. The symbols x_i, v_i, m_i, ν_i, and r_i represent the i-th components of the $\vec{x}$ (position), $\vec{v}$ (velocity), $\vec{m}$ (mass), $\vec{\nu}$ (damping), and $\vec{r}$ (Gaussian random number), respectively In this, the $r_i(t)$ are the 4 independent memoryless Gaussian random variables with zero mean and unit variance. Note that $r_i(t)$ have the following correlation:

$$\langle r_i(t) r_j(t') \rangle = \delta_{ij}\delta(t - t'). \tag{A2}$$

The average $\langle \ldots \rangle$ is taken over the stochastic process trajectories.

The following shows how to obtain the dimensionless Langevin equations from the above dimensional equation. Rearranging terms of Eq. (A1) gives:

$$dv_i = -\frac{\nu_i}{m_i} v_i dt - \frac{1}{m_i}\partial_{x_i} U(\vec{x})dt + \frac{r_i(t)}{m_i}\sqrt{2\nu_i \kappa dt} \tag{A3}$$

These quantities can be written as a product of a constant with unit and a unitless quantity: $\vec{x} = x_c\vec{x'}, \vec{v} = v_c\vec{v'}, \vec{m} = m_c\vec{m'}, \vec{\nu_i} = \nu_c\vec{\nu_i'}, U = U_0 U', \kappa = \kappa_c\kappa'$ and $t = t_c t'$. Substitute these expression into Eq. (A3), it can be expressed as:

$$\begin{aligned} v_c dv' = &-\frac{\nu_c \nu'}{m_c m'} v_c v' t_c dt' - \frac{t_c}{m_c m'}\frac{U_0}{x_c}\partial_{x'} U' dt' \\ &+ \frac{1}{m_c m'} r(t')\sqrt{2\nu_c \nu' \kappa_c \kappa' t_c dt'}. \end{aligned} \tag{A4}$$

Note that subscript i is dropped to simplify the expression. Dividing by v_c on both sides and making use of $v_c = x_c/t_c$, the above can be simplified:

$$\begin{aligned} dv' = &-\frac{\nu_c t_c}{m_c}\frac{v'}{m'} v' dt' - \frac{1}{m'}\partial_{x'} U' dt' \\ &+ \frac{1}{m_c}\sqrt{\frac{\nu_c \kappa_c t_c^3}{x_c^2}}\frac{\sqrt{\nu'\kappa'}}{m'} r(t')\sqrt{2dt'}\,. \end{aligned} \tag{A5}$$

This can then be further simplified to:

$$dv' = -\lambda v' dt' - \theta \partial_{x'} U' dt' + \eta r(t')\sqrt{2dt'}\,. \tag{A6}$$

Here, λ is the thermal coupling coefficient, reflecting the strength of the damping force the particles experience from the thermal bath. θ is related to the strength of potential relative to the dynamic of the system. If θ is zero, the system is off from the potential. η determines noise strength. Returning subscript i gives the dynamics in the dimensionless form of the Langevin equations:

$$dv_i' = -\lambda_i v_i' dt' - \theta_i \partial_{x_i'} U' dt' + \eta_i r_i(t')\sqrt{2dt'}\,, \tag{A7}$$

where λ_i, θ_i, and η_i are defined as:

$$\begin{aligned} \lambda_i &= \frac{\nu_c t_c}{m_c}\frac{\nu_i'}{m_i'}, \\ \theta_i &= \frac{1}{m_i'}, \text{ and} \\ \eta_i &= \frac{1}{m_c}\sqrt{\frac{\nu_c \kappa_c t_c^3}{x_c^2}}\frac{\sqrt{\nu'\kappa'}}{m'} = \sqrt{\frac{\lambda_i \kappa'}{m'}}\,. \end{aligned}$$

In our specific circuit, with the assumption that $L = L_1 = L_2$, $C = C_1 = C_2$ and $R = R_1 = R_2$, the vectors can be written as:

$$\begin{aligned} \vec{x} &= (\phi_1, \phi_2, \phi_{1dc}, \phi_{2dc})\,, \\ \vec{v} &= (\frac{d\phi_1}{dt}, \frac{d\phi_2}{dt}, \frac{d\phi_{1dc}}{dt}, \frac{d\phi_{2dc}}{dt})\,, \\ \vec{m} &= (C, C, \frac{C}{4}, \frac{C}{4})\,, \\ \vec{\nu_i} &= (\frac{2}{R}, \frac{2}{R}, \frac{1}{2R}, \frac{1}{2R})\,, \text{ and} \\ \vec{r} &= (r_1, r_2, r_3, r_4)\,. \end{aligned}$$

Choosing the following constants,

$$\begin{aligned} x_c = \frac{\Phi_0}{2\pi}; v_c = \frac{x_c}{t_c}; t_c = \sqrt{LC}; \\ m_c = C; \nu_c = \frac{1}{R};\ \text{and} \\ \kappa_c = U_0 = \frac{m_c x_c^2}{t_c^2} = \frac{\Phi_0^2}{4\pi^2 L}, \end{aligned}$$

the vectors then become:

$$\begin{aligned} \vec{x} = x_c\vec{x'} &= \frac{\Phi_0}{2\pi}(\varphi_1, \varphi_2, \varphi_{1dc}, \varphi_{2dc}), \\ \vec{v} = v_c\vec{v'} &= \frac{x_c}{t_c}(\frac{d\varphi_1}{dt}, \frac{d\varphi_2}{dt}, \frac{d\varphi_{1dc}}{dt}, \frac{d\varphi_{2dc}}{dt}), \\ \vec{m} = m_c\vec{m'} &= C(1, 1, \frac{1}{4}, \frac{1}{4}),\ \text{and} \\ \vec{\nu_i} = \nu_c\vec{\nu_i'} &= \frac{1}{R}(2, 2, \frac{1}{2}, \frac{1}{2}), \end{aligned}$$

Therefore, $\vec{\lambda}, \vec{\theta}$ and $\vec{\eta}$ can be expressed:

$$\vec{\lambda} = \frac{2}{R}\sqrt{\frac{L}{C}}(1,1,1,1),$$
$$\vec{\theta} = (1,1,4,4), \text{ and}$$
$$\vec{\eta} = \sqrt{\frac{k_B T}{U_0}}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, 2\sqrt{\lambda_3}, 2\sqrt{\lambda_4}) \ .$$

Appendix B: Escape rate calculation

1. Escape rate

The escape rate controls the number of escape events in one second per particle. The escape rate for 2D [50, 53, 54] is related to the barrier height ΔE_B by the following equation:

$$\Gamma = a_t \frac{\Omega}{2\pi} \exp\left(-\frac{\Delta E_B}{k_B T}\right). \tag{B1}$$

Here, $\Omega = \frac{\omega_{lw}\omega_{tw}}{\omega_{ls}}$, where ω_{lw} and ω_{tw} are the angular frequencies of the oscillations in the longitudinal and transverse directions at the local minima, and ω_{ls} is the angular frequency of the oscillations in the longitudinal direction at the saddle point. Here, longitudinal (transverse) direction means the direction parallel (perpendicular) to the oscillation direction and transverse Using the second derivative approximation for oscillation frequency, these angular frequencies can be calculated using:

$$\omega = \frac{U_0}{x_c^2}\frac{d^2U'}{d\varphi^2}$$

at points of interest. ΔE_B is the barrier height separating the local minima and saddle point. The theoretical number of escape events per operation is:

$$\Gamma_{cycle} = \Gamma \times N \times t_{tot},$$

where N is the number of particles in the simulation and t_{tot} is the duration of the protocol.

2. Barrier height between 01 and 11 particles

The barrier height is calculated by finding the difference of the potential at the top of the barrier height, located at $(\varphi_1, \varphi_2) = (0, \varphi_2^c)$, and the potential at the bottom of the well, located at $(\varphi_1, \varphi_2) = (\pm\varphi_1^c, \varphi_2^c)$, where φ_1^c and φ_2^c are the absolute values of the critical points in φ_1 and φ_2 respectively. Thus, the barrier height between 01 and 11 particles (ΔE_B) can be approximated as:

$$\begin{aligned}\Delta E_B &= U(\varphi_1 = 0, \varphi_2 = \varphi_2^c) - U(\varphi_1 = \varphi_1^c, \varphi_2 = \varphi_2^c) \\ &= \beta_1 \cos\frac{\varphi_{1dc}}{2}(1 - \cos\frac{\varphi_1^c}{2}) + \frac{\xi}{2}[(\varphi_{1x})^2 - \\ &(\varphi_1^c - \varphi_{1x})^2] - m_{12}\xi\varphi_1^c(\varphi_2^c - \varphi_{2x}),\end{aligned}$$

where U is the potential U in Eq. (1). However, for the storage pairs, because the second and third terms balance out, ΔE_B is reduced to $\beta_1 \cos\frac{\varphi_{1dc}}{2}(1 - \cos\frac{\varphi_1^c}{2})$.

Appendix C: Simulation Details

1. Parameters

Table I summarizes the simulation parameters. Once the circuit is fabricated, the fabrication parameters are set. Consequently, the calculated parameters are entirely dependent upon these established fabrication parameters. The parameter $\delta\beta_i$ describes how asymmetric the potential is. This is set to be zero to keep the simulation simple, but in practice will be nonzero based on fabrication variability. Some amount of asymmetry can be compensated through calibrating offsets to the φ_{ix} control values, though an ideally symmetric device is preferable.

The external parameters are controls for modifying the potential landscape and interacting with the system dynamics. Figure 13 shows the effect of the external parameters on the potential landscape. When all external parameters are set to zero, the system exhibits four distinct local minima, forming a 4-well potential (Figure 13a). $\varphi_{1(2)xdc}$ adjust the potential barriers between top (left) and bottom (right) wells, while $\varphi_{1(2)x}$ tilts the potential landscape horizontally (vertically). For simplicity, only the horizontal cases are shown in (Figure 13b and c). m_{12} conditonally tilts the potential landscape (Figure 13d).

2. Work Done and Fidelity

The work done by the k-th particle is expressed as:

$$\begin{aligned}W_k = &\sum_{i=0}^{n}[U(\vec{x}_k(\tau_i), \tau_{i+1}) - U(\vec{x}_k(\tau_i), \tau_i)] \\ &- \sum_{i=0}^{n}[U_{\min}(\tau_{i+1}) - U_{\min}(\tau_i)] \ .\end{aligned}$$

Here, τ_i is the time at the i-th step, $\vec{x}_k$ is the state of the k-th particle, and U is the potential, which is

(a) Fabrication parameters

Symbol	Physical meaning	Value
R_1, R_2	Resistance of the JJs	100Ω
C_1, C_2	Capacitance of the JJs	1 pF
L_1, L_2	Inductance of the rf-SQUIDs	5 pH
l_1, l_2	Inductance in series with the JJs	0.16 - 0.5pH
I^c_{1a}, I^c_{1b}	Critical currents of J_{1a} and J_{1b}	$45, 75\ \mu A$
I^c_{2a}, I^c_{2b}	Critical currents of J_{2a} and J_{2b}	$45, 75\ \mu A$

(b) Calculated parameters

Symbol	Formula	Value
I_{+1}, I_{+2}	$I^c_{1a} + I^c_{1b}$, $I^c_{2a} + I^c_{2b}$	$90, 150\ \mu$A
I_{-1}, I_{-2}	$I^c_{1a} - I^c_{1b}, I^c_{2a} - I^c_{2b}$	$0\ \mu$A
$\beta_{1(2)}$	$2\pi L_{1(2)} I_{+1(2)}/\Phi_0$	1.35, 2.3
$\delta\beta_{1(2)}$	$2\pi L_{1(2)} I_{-1(2)}/\Phi_0$	0
$\gamma_{1(2)}$	$L_{1(2)}/2l_{1(2)}$	5, 9, 16

(c) External parameters

Symbol	meaning
$\varphi_{1(2)x}$	Dimensionless flux threading through the circuit loop containing J_a and L_1 (J_c and L_2)
$\varphi_{1(2)xdc}$	Dimensionless flux threading through the circuit loop containing J_a and J_b (J_c and J_d)
m_{12}	Coupling between L_1 and L_2

Table I. Tables listing device fabrication parameters and calculated parameters and external parameters of the circuit. With $C = 1000$ and $L = 5\ pH$, the time constant $t_c = 2.24\ ps$. The values of l_i are set at 0.16, 0.28, 0.5pH, which correspond to $\gamma_i = 16, 9, 5$. The values of $I^c_{ia(b)}$ are set at 45, $75\mu A$, which correspond to $\beta_i = 1.35, 2.3$. The temperature is set at 4.2 K, the normal boiling point of liquid helium.

a function of the particle state and time. The upper limit n of the summation is the total number of time steps.

The first summation term in the work done is the difference between the potential energy at τ_{i+1} of the i-th state and that at τ_{i+1} of the same state. The second summation is the difference of the minimum point of the potential at τ_{i+1} and τ_i. This term is used to eliminate the work cost for shifting the potential up and down, which costs zero net work after a complete cycle. The average work done is the average of the work done by all the particles in the ensemble.

Fidelity is defined as the ratio of particles that successfully reach their intended final well to the total number of particles in the ensemble. It serves as a direct measure of a protocol's performance in guiding particles to their correct logical state. The truth tables for NAND, CE, and EF are provided in Figure 2. Any particle whose final position is inconsistent with its corresponding truth table entry is classified as an error.

3. dt analysis

Figure 15 investigates the relationship between the numerical time step (dt) and the mean work cost. This analysis considers dt values across the range of $[1/50, 1/100, 1/500, 1/1000, 1/5000]$ with $10,000$ trajectories. Based on this investigation, we selected $dt = 1/1000$ for all subsequent simulations because it provides the optimal balance between numerical accuracy and computational efficiency.

4. Integration method

A 4th-order Runge-Kutta method was applied to the deterministic portion and the Euler-Maruyama method was used for the stochastic portion of the integration.

Appendix D: Comparing trajectories of slow and fast EF

This section illustrates the crucial role of kinetic energy and momentum in determining the fidelity of EF protocol. Figure 14(a) plots the average kinetic energy for the 10 and 11 particles in both the EF (slow) and EF (fast) protocols. In the EF (slow) protocol, during the critical barrier-raising substage (the region between the grey dashed lines), the error bars scatter across zero. This indicates that some particles may fail to complete the required logical flip due to lacking KE. The first column of Figure 14(b) illustrates a sample failure trajectory: during the barrier-raising substage, the failing particles remain clustered around $\varphi_1 = 0$. Since they do not acquire the necessary kinetic energy, they fail to complete their transition and consequently revert to their original wells.

The EF (fast) protocol successfully addresses this deficiency. Its average KE graphs show that the error bars are consistently far from zero during the barrier-raising substage, ensuring the 10 and 11 particles maintain high momentum and successfully transition into their target well. However, this substantial increase in speed simultaneously imparts excess kinetic energy to the 00 particles. This excess energy causes error in the erasure pair. The second column of Figure 14(b) illustrates a sample failure trajectory for EF (fast): the 00 particle, having

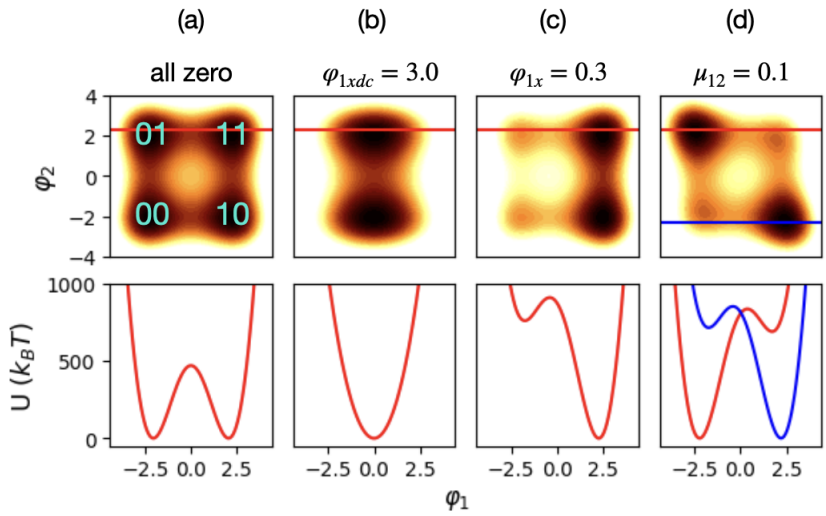


Figure 13. Potential landscape in the space of φ_1 and φ_2: The line graphs below show the energy profile along the red/orange line in the contour plots. (a) When all external parameters $\varphi_{1x}, \varphi_{2x}, \varphi_{1xdc}, \varphi_{2xdc}$, and m_{12} are zero, the energy landscape is a 4-well potential. The 2-digit numbers in the contour plot—(00, 01, 10, 11)—label the four wells. (b) φ_{ixdc} controls the barrier heights between potential wells. For simplicity, only the case of φ_{1xdc} is shown in this figure. (c) φ_{ix} tilts the potential. For simplicity, only the case of φ_{1x} is shown in this figure. (d) m_{12} conditionally tilts the potential landscape based on particle location in the left or right half-plane.

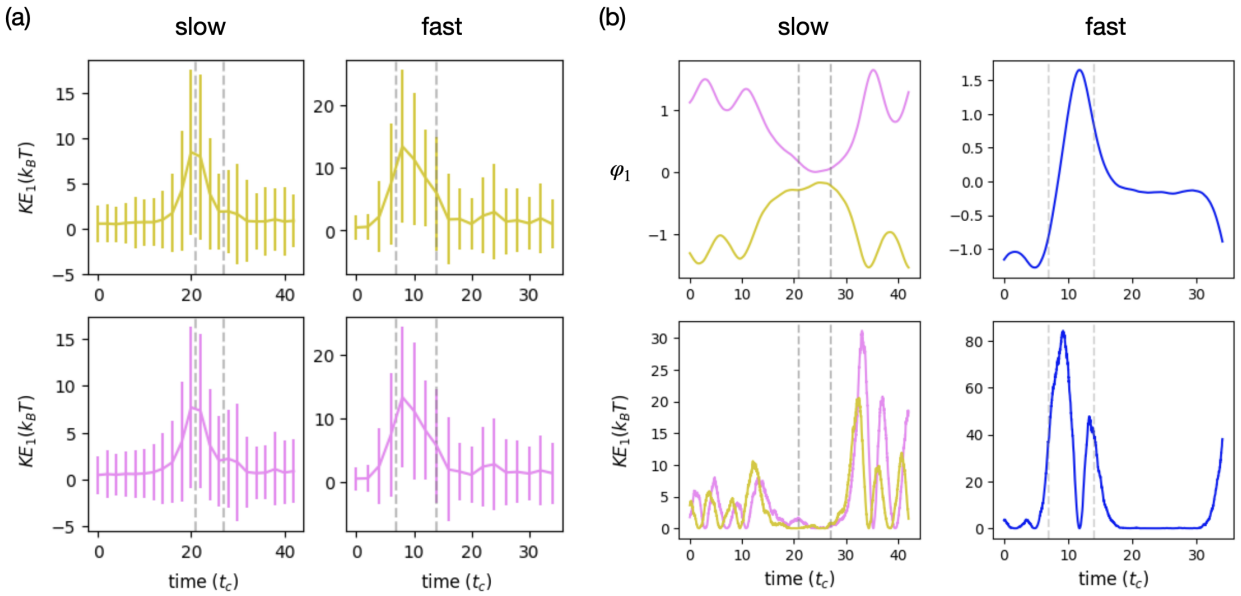


Figure 14. (a) Average kinetic energy of the 01 (yellow) and 11 (pink) particles as a function of time for EF (slow) and EF (fast): In the slower protocol, some particles have near-zero kinetic energy during the barrier-raising substage (region between the grey dashed lines). In contrast, in the faster version, the error bars no longer reach 0, indicating that the particles of the storage pair maintain a nonzero speed and can successfully transition to the target well. (b) Example trajectories of failure-prone particles in both the slow and fast versions of the EF protocol. In the slow version, the primary source of error is the storage pair. During the barrier-raising substage, these particles have near-zero kinetic energy and remain very close to the boundary, preventing them from successfully reaching their target wells. In EF (fast), the error comes from the erasure pair instead because too much kinetic energy is input to the system. The graph shows a failure trajectory from a 00 particle. (Animation: slow protocol, fast protocol)

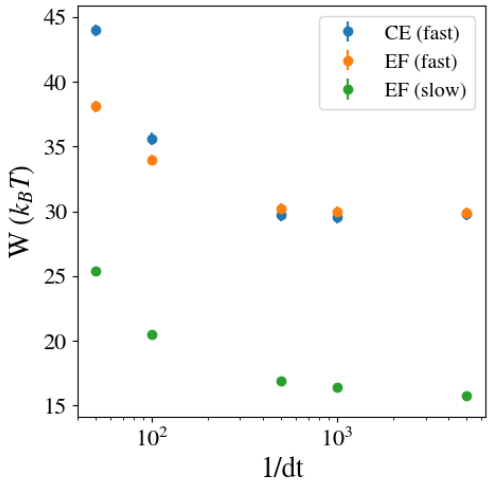


Figure 15. The graph plots the converged work cost for the three protocols (CE (fast), EF (slow), and EF (fast)) vs $1/dt$, where $dt = \{1/50, 1/100, 1/500, 1/1000, 1/5000\}$.

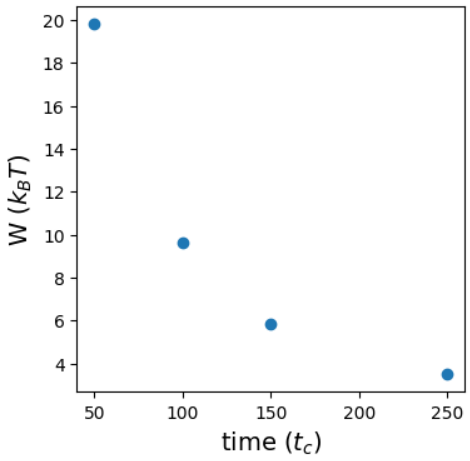


Figure 16. Work cost of conditional tilt for different durations

gained too much kinetic energy, fails to remain settled and reverts to its initial well, thereby causing a fidelity error in the erasure pair.

Appendix E: Protocol Detail

1. Work Cost of Conditional Tilt

The protocol for Conditional Tilt (CT) is defined by Substages 4 through 6 in Figure 17(d) and is assigned a base duration of $50t_c$. To analyze the work-cost-versus-speed trade-off of the CT itself, we simulated its performance using the same protocol but extending the duration by factors of two, three, and five times the base duration. The resulting work cost of the conditional tilt as a function of these different time lengths is presented in Figure 16. As expected theoretically, the work cost decreases as the duration increases. It would eventually approach Landauer's bound, but that may take an arbitrarily long protocol. The range displayed shows a reasonable swath of practical finite-time protocols with a reasonable balance of work cost and operating speed. There is not much energetic advantage in going beyond 250 t_c, when compared to the extra time spent processing. Conversely, protocols that are shorter than 50 t_c become more and more costly and quickly dominate the work cost of the whole protocol.

2. Protocols

Figure 17 presents the specific protocol tables used in the simulations. Figure 17(a) defines the CE protocol used in the simulation of Figure 5, which employed the parameters $\beta = 2.3$ and $\gamma = 9$. Figures 17(b), (c), and (d) define the protocols for EF (slow), EF (fast) and CE (fast) protocol, respectively. $\beta = 1.35$ in all the three cases, but $\gamma = 9$ for EF (slow), $\gamma = 5$ for EF (fast), and $\gamma = 16$ for CE (fast). All of the above simulations use the same capacitance, inductance, resistance, and critical currents as shown in Table I.

Appendix F: Alternative Protocols for NAND

1. Complete NAND

The analysis in the main text describes and presents simulation results for the partial NAND operation, implemented using both the CE and EF protocols. In this setting, the computation designates one bit as the primary output and the other as an irrelevant (garbage) bit. To eliminate the presence of the garbage bit and produce a clean, single-state result, one implements a complete NAND operation, which actively forces the logical states of the two output bits to be identical.

Figure 18 illustrates the conceptual implementation of the complete NAND using both CE and EF. For the CE-based implementation, the complete NAND is achieved by appending one additional CT substage to the partial NAND protocol. This final manipulation moves 00, 01, and 10 particles into the well 11, while moving the particle 11 into 00 well at

(a) Protocol of CE in Figure 5

substage	duration (τ_c)	φ_{1xdc}	φ_{2xdc}	m_{12}	φ_{1x}	φ_{2x}
1 (t_0 to t_1)	50	1.9	0	$0 \to -0.0523$	$0 \to -0.1045$	0
2 (t_1 to t_2)	50	$0 \to 1.9$	0	-0.0523	-0.1045	0
3 (t_2 to t_3)	50	$1.9 \to 0$	0	$-0.0523 \to 0$	$-0.1045 \to 0$	0

(b) Protocol of EF (slow)

substage	duration (τ_c)	φ_{1xdc}	φ_{2xdc}	m_{12}	φ_{1x}	φ_{2x}
1 (t_0 to t_1)	15	$0 \to 1.1$	0	0	0	0
2 (t_1 to t_2)	6	$1.1 \to 1.7$	0	$0 \to 0.0306$	$0 \to 0.04$	0
3 (t_2 to t_3)	6	$1.7 \to 1.1$	0	0.0306	0.04	0
4 (t_3 to t_4)	15	$1.1 \to 0$	0	$0.0306 \to 0$	$0.04 \to 0$	0

(c) Protocol of EF (fast)

substage	duration (τ_c)	φ_{1xdc}	φ_{2xdc}	m_{12}	φ_{1x}	φ_{2x}
1 (t_0 to t_1)	4	$0 \to 0.9$	0	0	0	0
2 (t_1 to t_2)	3	$0.9 \to 1.5$	0	$0 \to 0.0307$	$0 \to 0.04$	0
3 (t_2 to t_3)	3	1.5	0	0.0307	0.04	0
4 (t_3 to t_4)	4	$1.5 \to 1.1$	0	0.0307	0.04	0
5 (t_3 to t_4)	20	$1.1 \to 0$	0	$0.0307 \to 0$	$0.04 \to 0$	0

(d) Protocol of NAND by CE (fast)

substage	duration (τ_c)	φ_{1xdc}	φ_{2xdc}	m_{12}	φ_{1x}	φ_{2x}
1 (t_0 to t_1)	50	$0 \to 1.1$	0	$0 \to -0.0217$	$0 \to -0.0284$	0
2 (t_1 to t_2)	50	1.1	0	-0.0217	-0.0284	0
3 (t_2 to t_3)	50	$1.1 \to 0$	0	$-0.0217 \to 0$	$-0.0284 \to 0$	0
4 (t_3 to t_4)	20	$0 \to 1.5$	0	0	0	0
5 (t_4 to t_5)	10	1.5	0	$0 \to 0.06$	0	0
6 (t_5 to t_6)	20	$1.5 \to 0$	0	$0.06 \to 0$	0	0

Figure 17. The tables detail the specific time-dependent changes applied to the external control parameters for each simulation protocol: (a) Protocol of CE in Figure 5, (b) Protocol of EF (slow) in Figure 10(a), (c) Protocol of EF (fast) in Figure 10(b) and (d) Protocol of NAND by CE (fast) in Figure 12.

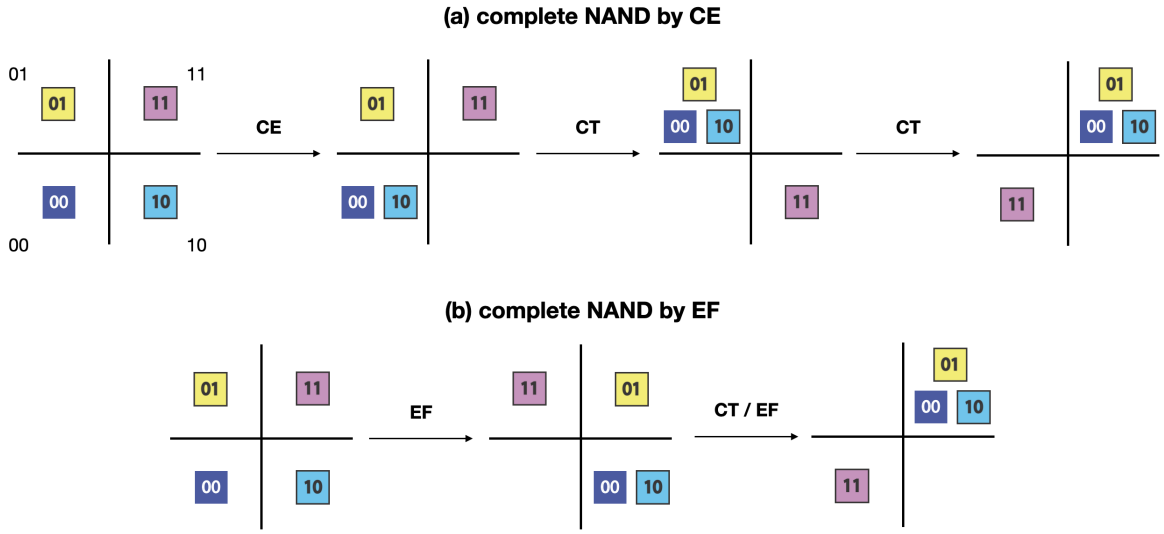


Figure 18. Schematic illustration showing the complete NAND protocol by (a) CE and (b) EF.

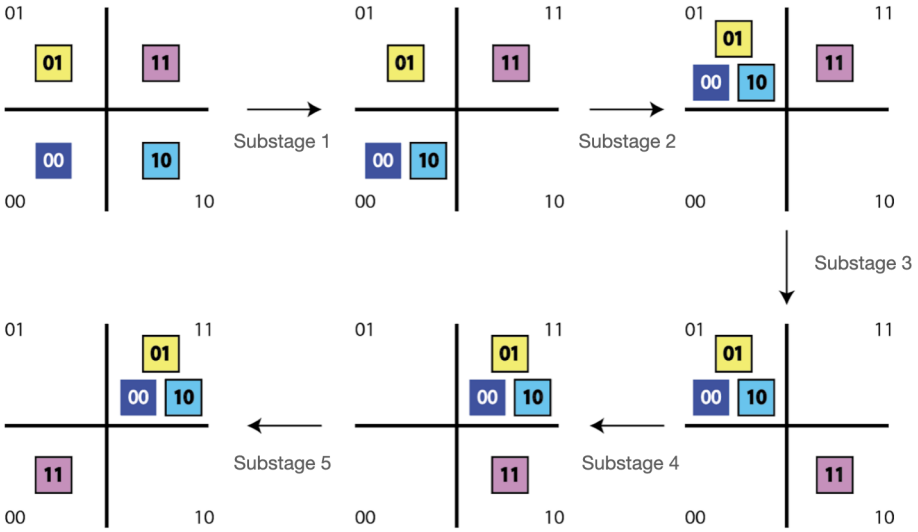


Figure 19. Performing controlled erasure five times for complete NAND operation.

the end of the operation.

For the partial NAND by EF, the complete NAND can be realized by applying either one additional CT or one additional EF substage. This final step moves the 11 particles to the 00 well, while concurrently steering the other three particle types (00, 01, and 10) to the 11 well, thereby ensuring both bits to be identical at the final state.

2. Complete NAND by CE only

An alternative complete NAND operation based off of CE only is described in Ref. [39]. As illustrated in Figure 19, this method involves performing CE protocol five times. The work cost for complete NAND by CE only (Figure 19) can be estimated by observing which particles types must move from a higher well to a lower well in each subprocess. The following table summarizes the particle types that must move from a higher well into a lower well. By summing the work cost column, the work cost is 8 times of the work cost of a single CE.

substage	type of particles	work cost
1	10	1
2	00, 10	2
3	11	1
4	00, 01, 10	3
5	11	1

3. Partial NAND by CE only

An alternative implementation of a partial NAND using three repeated CEs operations was also proposed in Ref. [39]. As it requires three serial CE operations the duration is three times that of a single CE. The error rate and work costs differ, however, as they were calculated assuming a uniform starting distribution over the four wells. When taking into account the change of distribution for each subsequent operation, we still expect the error rate to be triple that of the single CE performed on a uniform distribution. However, the work cost for this implementation is four times that of the single CE—a manifestation of the additional work costs associated with modularity and variations in initial distribution [55, 56]. In short, the CE-based NAND operation using a CT is simultaneously faster and more accurate, while also costing less work than one built completely out of CEs.

While these protocols incur higher work costs, lower fidelity, and longer durations, they may offer advantages in control complexity by exclusively using CE and forgoing the need for conditional tilts.

Appendix G: Entropy change of partial NAND and complete NAND

The fundamental work cost of a protocol is calculated through comparing the Shannon entropy difference between the initial and final distribution. This appendix shows how to calculate the fundamental work cost for a partial NAND and a com-

plete NAND. We start with a four-well potential and the Shannon entropy of the initial distribution is $-\frac{1}{4}\ln\frac{1}{4}\times 4 = 2\ln 2$.

The distribution after the first substage in Figure 18(b) corresponds to the final state of a partial NAND operation. In this state, two input states are mapped into a single well (total probability 1/2), while the other two states remain separate (each with probability 1/4). The Shannon entropy of this final distribution is: $S_{\text{Partial}} = -\left(\frac{1}{2}\ln\frac{1}{2} + \frac{1}{4}\ln\frac{1}{4} + \frac{1}{4}\ln\frac{1}{4}\right) = \frac{3}{2}\ln 2$. The minimum work cost for the partial NAND is calculated from the entropy change ΔS, which is about 0.347.

On the other hand, the distribution after the second substage in Figure 18(b) corresponds to the final state of a complete NAND operation, where three out of four inputs are mapped to the logical 0 state (total probability 3/4), and one input remains in the logical 1 state (probability 1/4). The Shannon entropy of this final distribution is: $S_{\text{Complete}} = -\left(\frac{3}{4}\ln\frac{3}{4} + \frac{1}{4}\ln\frac{1}{4}\right) = 2\ln 2 - \frac{3}{4}\ln 3$. The minimum work cost for the complete NAND is about 0.824.