arXiv:2010.01152

Non-Markovian Momentum Computing: Thermodynamically Efficient and Computation Universal

Kyle J. Ray,[1,*] Alexander B. Boyd,[2,†] Gregory W. Wimsatt,[1,‡] and James P. Crutchfield[1,§]

[1] *Complexity Sciences Center and Physics Department, University of California at Davis, One Shields Avenue, Davis, CA 95616*

[2] *Complexity Institute, Nanyang Technological University, 3 Science Drive 2, Singapore 117543*

(Dated: March 5, 2021)

Practical, useful computations are instantiated via physical processes. Information must be stored and updated within a system's configurations, whose energetics determine a computation's cost. To describe thermodynamic and biological information processing, a growing body of results embraces rate equations as the underlying mechanics of computation. Strictly applying these continuous-time stochastic Markov dynamics, however, precludes a universe of natural computing. Within this framework, operations as simple as a NOT gate, flipping a bit, and swapping bits are inaccessible. We show that expanding the toolset to continuous-time *hidden* Markov dynamics substantially removes the constraints, by allowing information to be stored in a system's latent states. We demonstrate this by simulating computations that are impossible to implement without hidden states. We design and analyze a thermodynamically-costless bit flip, providing a counterexample to rate-equation modeling. We generalize this to a costless Fredkin gate—a key operation in reversible computing that is Turing complete (computational universal). Going beyond rate-equation dynamics is not only possible, but necessary if stochastic thermodynamics is to become part of the paradigm for physical information processing. The increased analytical challenges are readily addressed with recently-introduced spectral decomposition methods for nondiagnonalizable dynamics.

Keywords: rate equations, stochastic process, hidden Markov model, information processing, logical circuits, entropy production, reversibility

The burgeoning field of thermodynamic computing leverages recent progress in nonequilibrium thermodynamics and information and computation theories [1–5] to establish a new paradigm for physical information processing. It promises to increase computational power and efficiency and to reduce energy dissipation in a next generation of computers [6]. Thermodynamic computing is distinguished from alternative paradigms by its focus on an information-processing device's physical embedding; specifically, by constructively working with $k_{\mathrm{B}}T$-scale fluctuations that a thermal environment generates. More broadly, a general framework rooted in thermodynamics, as thermodynamic computing is, will provide the tools to understand the physics of computation in all its many forms. The following illustrates its breadth by introducing non-Markovian, momentum-based computing—a paradigm that is both computation universal and thermodynamically efficient.

Physically-embedded computation can be described as a stochastic mapping within a system's set $\mathcal{M}$ of memory states—Landauer's *information-bearing degrees of freedom* (IDoF) [7]. Carried out over time interval $t \in (0, \tau)$, the mapping is the conditional probability $\mathbf{p}$ of transitioning from an initial memory state $m(0) \in \mathcal{M}$ to a final state $m(\tau) \in \mathcal{M}$: $\mathbf{p}_{m(0)\to m(\tau)} = \Pr\left[m(\tau)|m(0)\right]$. The mapping $\mathbf{p}$ determines the probability of the final memory state given the initial memory state, and so updates the state distribution $\vec{p}(\tau) = \mathbf{p}\ \vec{p}(0)$. This describes the dynamics underlying a computation, but what of its thermodynamic consequences? To address this, we must first identify the constraints on dynamics that can implement computations.

Computing with Continuous-Time Markov Chains To date, proposed frameworks for the required mappings in thermodynamic computing assume that the memory state m obeys stochastic Markovian dynamics [8, 9, and references therein]. Taking time to be continuous, the dynamics are continuous-time Markov chains (CTMCs), where the state distribution changes continuously as a function of itself $\dot{\vec{p}}(t) = f(\vec{p}, t)$. The resulting dynamics is necessarily represented by a master equation over the memory-state distribution $\dot{\vec{p}}(t) = \mathbf{A}(t)\vec{p}(t)$ [8, 9]; that is, by *rate equations*. This is a powerful framework for stochastic thermodynamics [2, 10, 11] that yields insight into physical realizations of computations such as bit erasure and measurement [12].

The constraint that the computation $\mathbf{p}$ is generated by integrating continuous-time master equations comes at a substantial compromise—it limits the range of possible computations. For example, only input-output mappings whose determinants are positive are allowed when memory-state dynamics are restricted to obey CTMCs [9]. This eliminates many common and useful computations, including flipping a single bit of information. Reference [8] takes these restrictions as delineating the possible *physically realizable computations*. Given that any computation we can observe—a bit flip, to take one example—is necessarily physical, one must instead interpret the restrictions as a limitation of the CTMC framework, rather than of the physical world.

Understanding both the merits and limits of the CTMC framework requires a look at the physical mechanisms that underpin it. It might seem natural to say that the memory states $\mathcal{M}$ are microstates of a physical memory system $\mathcal{S}$, evolving under Hamiltonian dynamics. However, a physical computation device is typically coupled to an environment—which suggests treating $\mathcal{S}$ as a stochastic subsystem of a deterministic universe. If the environment is a large weakly-coupled heat bath, with degrees of freedom that relax sufficiently quickly, the effective dynamics for $\mathcal{S}$ are also Markov, and therefore CTMC [2, 13–16]. In essence, coarse-graining the thermal environment allows for accurate, probabilistic predictions about the memory system, while avoiding the task of tracking the full Hamiltonian dynamics of the joint system and bath.

This justification of the Markov evolution of a memory system recognizes an important fact: $\mathcal{S}$'s states are not themselves full descriptions of physical degrees of freedom. Instead, they are mesostates defined by a coarse-graining over the thermal environment's microstates. This coarse-graining is appropriate since the environment does not retain information about the past.

Computationally-useful memory states $\mathcal{M}$—Landauer's IDoF—are mesostates that also coarse-grain over $\mathcal{S}$'s CTMC-evolving states. It is possible, depending on the variables and timescales of interest, that this coarse-graining ignores only rapidly-relaxing subsystems of $\mathcal{S}$ and, then, the IDoF inherit the Markov property of the memory system. The result is a powerful and widely-used framework for thermodynamic computing in which the IDoF also obey CTMC dynamics. This case is typified by IDoF that are positional degrees of freedom and where $\mathcal{S}$ is described by overdamped Langevin dynamics.

Computing with Continuous-Time Hidden Markov Chains However, in the most general case, the IDoF coarse-grain over subsets of $\mathcal{S}$ that carry information relevant for predicting a computation's performance. That is, the states that $\mathcal{M}$ coarse-grains over are *hidden* in that they contain dynamically relevant information not determined from instantaneous realizations of the memory. The resulting memory dynamics are non-Markovian, since information is transmitted from past to future without ever appearing in the present memory state [17]. The sobering fact is that a general analytical treatment of partially-observed (and therefore non-Markovian) systems is highly nontrivial [5, 16, 18–20]. No matter, hidden states allow for more general forms of computation [9, 21], since non-Markov dynamics relax the constraints imposed by CTMCs. Following this argument to its conclusion, the following demonstrates that the appropriate setting for thermodynamic computing is continuous-time *hidden* Markov chains (CTHMCs), in which hidden variables store computationally-relevant information.

Moreover, when memory is stored in positional degrees of freedom, the conjugate momentum variables are particularly useful hidden variables for flexibly designing computations. We demonstrate this first by implementing a thermodynamically-costless bit flip—a simple computation that is explicitly forbidden by CTMCs. We then generalize this to a costless Fredkin gate—a key component in reversible computing that is also impossible to implement with CTMCs. This operation is computation universal (Turing complete), meaning that combinations of the Fredkin gate can implement any logical operation [22]. The implementation of this universal and reversible-logic gate via CTHMCs demonstrates that non-Markovian dynamics are essential to thermodynamic computing and that a new class of momentum-based computation is within reach.

Flipping a Physical Bit To execute a single bit flip over a time interval $t \in [0, \tau]$, the first step is to store a bit of information. One candidate is a particle with a single position dimension $x \in \mathbb{R}$ and corresponding momentum $p \in \mathbb{R}$ in an even potential energy landscape $V^{\text{store}}(x)$ containing two potential minima at $x = \pm x_0$ with an associated energy barrier between them equal to $\max\{V^{\text{store}}(x), x \in (-x_0, x_0)\}$. The particle's environment is a thermal bath at temperature T. As the height of the potential energy barrier rises relative to the bath energy scale $k_B T$, the probability that the particle transitions between left ($x < 0$) and right ($x \geq 0$) decreases exponentially. In this way, if we assign the left half of the position space to memory state 0 and the right half to memory state 1, the energy landscape is capable of metastably storing a bit $m \in \{0, 1\}$.

To execute a flip operation, we instantaneously reduce the coupling to the thermal reservoir to zero such that the memory system now follows dissipationless Hamiltonian dynamics. Simultaneously, the potential energy landscape changes to a positive quadratic well: $V^{\text{comp}}(x, t = 0^+) = kx^2/2$. The resulting particle motion is harmonic oscillation: $x(t) = x^* \cos(t\sqrt{k/m} + \phi)$, where x^* is the maximum distance from the cycle's origin and ϕ is the phase difference from maximum distance at the time $t = 0^+$. Maintaining the decoupled system in the quadratic potential energy landscape for half the oscillation period $t \in (0, \pi\sqrt{m/k})$, the particle's new position becomes: $x(\pi\sqrt{m/k}) = x^* \cos(\pi + \phi) = -x^* \cos(\phi) = -x(0)$. Thus, over the computation interval $\tau = \pi\sqrt{m/k}$, the position flipped sign so that the memory state has flipped as well: $m(\tau) = 1 - m(0)$. Finally, we instantaneously return the potential energy landscape to $V^{\text{store}}(x)$ and recouple to the thermal bath.

The work W involved is the time-integrated rate of potential energy change due to the change in the protocol parameter [23]: $W = \int dt \partial V(x, t')/\partial t'|_{x(t),t}$. The work cost for a particular trajectory is the instantaneous change in potential energy at $t = 0$ and $t = \tau$:

$W = V(x(0), 0^+) - V(x(0), 0) + V(x(\tau), \tau) - V(x(\tau), \tau^-)$, where 0^+ and τ^- are times immediately after and before $t = 0$ and $t = \tau$, respectively. Recall that the potential is time-symmetric ($V(x,t) = V(x, \tau - t)$)), that $x(\tau) = -x(0)$, and that the potential is even in x. These three qualities yield $-V(x(0), 0) + V(x(\tau), \tau) = V(x(0), 0^+) - V(x(\tau), \tau^-) = 0$. No net work is generated during the protocol.

Not only does this computation go beyond what is physically allowable according to rate-equation dynamics over the memory states, but the states only change while the Hamiltonian control is fixed. Thus, the computation is passive, meaning that it fits the information-ratchet framework introduced by Ref. [24].

A Physical Fredkin Gate The bit-flip implementation may seem obvious in its simplicity. However, sophisticated and functional computing can be built from a similar passive processes. Below we outline an implementation of the Fredkin gate, a reversible and computation universal logical gate [22], using the same strategy. This establishes that CTHMCs give straightforward access to complex and universal Turing thermodynamic computing.

The *Fredkin gate* operates on three bits $\mathcal{M} = \{0, 1\}^3$. That is, we encode the physical substrate as three particle-position variables (x, y, z) that are each separated into negative and positive memory-state regions, as above. This splits the memory states into eight respective octants: $(x < 0, y < 0, z < 0)$ corresponds memory state $m = 000$, $(x < 0, y \geq 0, z < 0)$ to $m = 010$, and so on. The information-storing Hamiltonian is a straightforward sum of bistable, even, one-dimensional storage potentials: $V^{\text{store}}(x, y, z) = V^{\text{store}}(x) + V^{\text{store}}(y) + V^{\text{store}}(z)$. This provides metastable regions corresponding to each memory state $m_x m_y m_z \in \{0, 1\}^3$.

Given this construction, we design physical transformations that implement the Fredkin gate with zero cost in finite time. The Fredkin gate is also known as the *controlled swap gate*, as it exchanges inputs m_y and m_z only if the control m_x is set to 1. In other words, the gate maps all inputs to themselves, excluding 101 and 110 that swap with each other. The implementation uses the bit-flip strategy of decoupling and adding a harmonic potential over the time interval $t \in (0, \tau)$, then recoupling and resetting the original information-storing Hamiltonian. The only difference is that the harmonic potential driving the computation is now embedded in the higher-dimensional space.

To execute the Fredkin gate, first note that the memory-state x-index must always be fixed: $m_x(\tau) = m_x(0)$. Moreover, behavior in the $y - z$ plane should only depend on x up to whether it is positive or negative. Thus, we first split the potential into two pieces: $V(x, y, z, t) = V^{\text{store}}(x) + V^{yz}(y, z, t)$. If $m_x(0) = 0$ then m_y and m_z must also not change. This suggests using the information-storing potential for this region of state space: $V(x < 0, y, z, t) = V^{\text{store}}(x, y, z)$ during the entire computation. For $m_x = 1$, however, we must nontrivially compute on m_y and m_z: $V^{yz}(x \geq 0, y, z, t \in (0, \tau)) = V^{\text{comp}}(y, z)$. Here, V^{comp} determines that part of the Hamiltonian which implements the switch $101 \to 110$ and $110 \to 101$ and remains unchanging over $t \in (0, \tau)$. Due to decoupling from the x-axis, particle behavior in either the positive or negative x regions can be considered as being purely the result of two-dimensional dynamics.

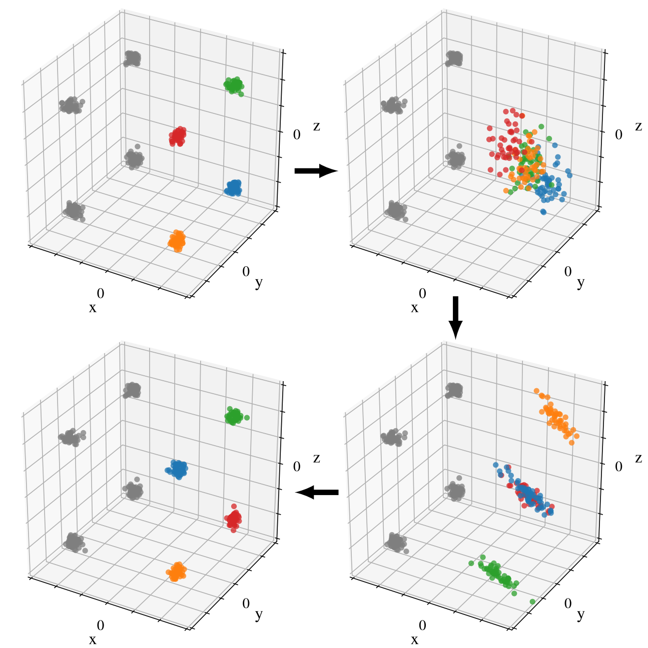

FIG. 1. Particle ensemble undergoing the Fredkin gate protocol with zero coupling to the thermal reservoir. Snapshots of the state evolution are given at times $0, \tau/4, \tau/2,$ and τ, with the black arrows indicating forward time. Color encodes in which informational state each trial begins: gray: 000, 001, 010, 011; orange: 100; red: 101; blue: 110; green: 111. Additional detail available in the Appendix. Animations available online.

To swap 101 and 110, while keeping 111 and 100 fixed, consider a new basis for the yz-space. Define new variables: $y' = (y - z)/\sqrt{2}$ and $z' = (y + z)/\sqrt{2}$, such that the local equilibrium distributions for states 110 and 101 are centered around $z' = 0$ and those for states 111 and 100 are centered around $y' = 0$. Thus, our goal is to swap the distributions in the y'-coordinate while preserving their z'-coordinate.

Given this, we split the computation Hamiltonian again into independent components: $V^{\text{comp}}(y, z) = V(y') + V(z')$. Flipping in the y'-coordinate employs the same Hamiltonian as for the previous bit-flip protocol: $V(y') = ky'^2/2$. As a result, when waiting half a period $\tau = \pi\sqrt{m/k}$, the y' coordinate changes sign $y'(\tau) = -y'(0)$, as does its momentum. We choose the z' coordinate's potential to be quadratic as well, but with an induced period of oscillation that is half as long: $V(z') = 2kz'^2$. z' then undergoes a full cycle after the duration τ, returning to its original value $z'(\tau) = z(0)$, as does its momentum. Over the control interval $t \in (0, \tau)$ the Hamiltonian operates piecewise with $V(x, y', z', t) = V^{\text{store}}(x) + V^{yz}(x, y', z', t)$, where:

$$V(x, y', z', t) = \begin{cases} V^{\text{store}}(x, y, z) & \text{if } x < 0 \\ V^{\text{store}}(x) + \frac{ky'^2}{2} + 2kz'^2 & \text{if } x \geq 0 \end{cases}.$$

In our original coordinates, this passive Hamiltonian transforms the particle's state by swapping y and z, but only when when $x > 0$ ($m_x = 1$); thus, it implements the Fredkin gate.

For a particular trajectory $(x, y, z)(t)$, the work invested only comes from the initial and final instantaneous changes in the energy landscape, as noted above. Recall that $x(t)$ is exponentially unlikely to change sign, since the energy barrier between states is much higher than the vast majority of thermal fluctuations can access. Thus, we assume that paths maintain a single sign for $x(t)$. If $x(t)$ is negative, then there is no instantaneous change, as the system is held in the same double-well potential, so $W = 0$.

That said, if $x(t)$ is positive, then the work invested also vanishes. For these trajectories, note that the $y - z$ subspace potential is symmetric with respect to exchange of the y and z coordinates and that the action of the map is to swap y and z. Thus, using the same arguments as for the bit flip, we see that the work production vanishes. The only work-producing trajectories are the exponentially suppressed barrier crossing events—so the average work production is nearly zero.

Figure 1 demonstrates the evolution of the phase space on an ensemble of initial conditions drawn from the equilibrium distribution of a quartic storage potential. As shown by the particle coloring, those that start in 110 and 101 swap while all others are fixed. Moreover, none of the particles' x-coordinates change informationally—confirming the effectiveness of the overall transformation.

Langevin Simulation The preceding stipulated that the logical system be isolated from its thermal environment during the swap. It might not seem surprising then, that we are able to accomplish a work-free bit flip, given that other classical implementations of efficient reversible computing—such as ballistic computing with billiards [22]—necessarily operate in a dissipationless environment.

However, a key and somewhat surprising point is that the Fredkin gate implemented above tolerates imperfect isolation from its thermal environment. The gate's robustness to fluctuations separates it from other implementations that are dynamically unstable, such as billiard computing.

To demonstrate this, we investigated how robust the operation is to thermal agitation by using *underdamped* Langevin dynamics. A simulation was carried out by initializing particles in the equilibrium distribution with a thermal reservoir under a quartic information-storing potential $V^{\text{store}}(x, y, z)$. Next, as described above, we exert work on the system by turning on the computational potential V^{comp} in the region $x > 0$. However, rather than reducing the thermal coupling to $\lambda = 0$, we drop the coupling coefficient to a nonzero value in the weak coupling regime. This coupling value and potential are held fixed for time $\tau = \pi\sqrt{m/k}$. (The Appendix provides additional detail.)

Thermodynamically Robust Fredkin Gate The particles experience thermal fluctuations as the weak coupling to the bath perturbs their trajectories from the otherwise expected harmonic motion. The work gained from shutting off the potential will not generally be the same as the work invested to turn it on (as in the idealized case of zero thermal coupling). In fact, the Second Law guarantees that, generally, positive work is invested for such cyclical transformations, because the net change in equilibrium free energy is zero. Nevertheless, one expects the behavior to approximate the desired Fredkin-gate dynamics if the coupling is sufficiently weak. While the energetic cost of implementing the gate does not remain zero as the coupling approaches zero, Fig. 2 shows that the logical fidelity approaches unity. And, it does so with *zero slope*, revealing that this Fredkin gate implementation is robust even in the presence of thermal fluctuations.

This also gives evidence that the implementation should have practical use for reversible universal computing in a thermal environment. This regime is particularly well suited, for example, to superconducting flux qubits working in the classical regime [25] in which a tunable resistance can act as a control parameter for damping.

As expected and shown in Fig. 2, the work invested approaches zero with decreasing coupling. However, as the coupling to the thermal reservoir increases, the average work required to compute increases to multiples of $k_B T$. This cost scaling is evidence that the thermal agitation has become significant enough to take the particles appreciably far from their ideal (costless) trajectories; nevertheless the protocol maintains high fidelity, even in this regime. Note that the work cost is exponentially unlikely to come from trajectories that "jump" over the $x = 0$ boundary, since the storage bits maintain perfect fidelity.

As a final note, the thermodynamic cost is much more than that predicted by the microscopic detailed-balance dynamics that underlie the Langevin simulation. This suggests the existence of a lower bound on entropy production—one that accounts for the course-graining, as Ref. [26] predicted.

Conclusion Rate-equation dynamics is certainly a venerable and powerful framework, central to reaction kinetics in chemistry [27, 28] and key to the master equations of applied statistical mechanics [2, 10, 11]. Due to the remarkable successes of continuous-time Markov chain

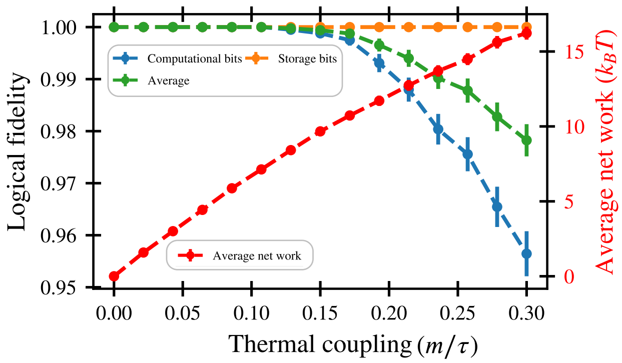

FIG. 2. Logical fidelity (successful trials/total trials) in the low-coupling Fredkin gate and the average net work required to implement it for different values of the thermal coupling constant λ, measured in units of m/τ. *Computational bits* refers to states that fall in the region $x > 0$, where the computational potential is in effect.

predictions of many thermodynamic behaviors, it might seem natural to claim that to be "physically realizable", thermodynamic computing and biological information processing should *only* be described and analyzed as rate-equation dynamics [8].

The results here demonstrated that this does not hold generally. And so, it cannot form a complete basis for thermodynamic computing. Moreover, it levies a heavy penalty, precluding engineering and analyzing Maxwellian information ratchets, which are the physical equivalent of Turing machines [24, 29–31]. The limits are especially draconian, since efficient time-symmetrically controlled general computations consist of involutions [26]—operations that are composed of bit swaps and identity maps in positional memory.

As a constructive alternative, we proposed employing continuous-time hidden Markov chains to realize *non-Markovian momentum computing*. We demonstrated it provides a more complete framework, using two explicit examples that are forbidden if one is restricted to rate equations to describe the evolution between memory states [8]. Additionally, we introduced explicit mechanisms for implementing both in finite time with zero work, proving them "physically realizable". However, we did fully acknowledge the increased analytical complexity posed by CTHMC dynamics. Fortunately, requisite tools have been developed that render the behaviors analytically tractable and in closed form [32, 33].

Given that convincing, physically-realizable implementations of the bit flip and Fredkin gate [22, 34, 35] have been known for some time, one can only conclude that computing devices already operate beyond the restrictions imposed by rate-equation dynamics. The examples presented here were intentionally couched in the thermodynamics of information to help bridge an apparent gap in understanding general computing. Most specifically, to fully realize the power and breadth of thermodynamic computing, the conception of memory must be expanded from being the realization of a microscopic physical state to being a mesoscopic coarse-graining, as Landauer emphasized half a century ago. Thus, CTHMCs and the momentum-based computing paradigm they inspire are invaluable tools, required for even the most basic computational tasks.

Acknowledgments. We thank Adam Kunesh, Thomas Ouldridge, and Mikhael Semaan for helpful discussions. JPC thanks the Santa Fe Institute and he and the other authors thank the Telluride Science Research Center for their hospitality during visits. This material is based upon work supported by, or in part by, FQXi Grant number FQXi-RFP-IPW-1902, the Templeton World Charity Foundation Power of Information fellowship TWCF0337, the U.S. Army Research Laboratory and the U. S. Army Research Office under grants W911NF-18-1-0028 and W911NF-21-100048.

* kylejray@gmail.com
† alecboy@gmail.com
‡ gwwimsatt@ucdavis.edu
§ chaos@ucdavis.edu

[1] T. Sagawa. Thermodynamics of information processing in small systems. *Prog. Theo. Phys.*, 127(1), 2012.

[2] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Physics*, 75(12), 2012.

[3] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa. Thermodynamics of information. *Nature Physics*, 11(2):131–139, 2015.

[4] Y. Hasegawa and T. Van Vu. Fluctuation theorem uncertainty relation. *Phys. Rev. Lett.*, 123(11):110602, 2019.

[5] U. Seifert. From stochastic thermodynamics to thermodynamic inference. *Ann. Rev. Cond. Mat. Physics*, 10:171–192, 2019.

[6] T. Conte et al. Thermodynamic computing. *arxiv:1911.01968*.

[7] R. Landauer. Information is physical. *Physics Today*, 44(5):23, 1991.

[8] E. Stopnitzky, S. Still, T. E. Ouldridge, and L. Altenberg. Physical limitations of work extraction from temporal correlations. *Phys. Rev. E*, 99:042115, 2019.

[9] J. A. Owen, A. Kolchinsky, and D. H. Wolpert. Number of hidden states needed to physically implement a given conditional distribution. *New J. Physics*, 21(1):013022, 2019.

[10] C. van den Broeck. Stochastic thermodynamics: A brief introduction. *Phys. Complex Colloids*, 2013.

[11] C. van den Broeck and M. Esposito. Ensemble and trajectory thermodynamics: A brief introduction. *Phyisca A*, 418:6–16, 2015.

[12] T. E. Ouldridge, C. C. Govern, and P. Rein ten Wolde. Thermodynamics of computational copying in biochemical systems. *Phys. Rev. X*, 7(2):021004, 2017.

[13] R. Alicki. The quantum open system as a model of the heat engine. *J. Phys. A*, 12(5), 1979.

[14] C. Jarzynski. Stochastic and macroscopic thermodynamics of strongly coupled systems. *Phys. Rev. X*, 7(011008), 2017.

[15] S. Deffner and E. Lutz. Nonequilibrium entropy production for open quantum systems. *Phys. Rev. Lett.*, 107(140404), 2011.

[16] P. Strasberg, G. Schaller, N. Lambert, and T. Brandes. Nonequilibrium thermodynamics in the strong coupling and non-Markovian regime based on a reaction coordinate mapping. *New J. Physics*, 18(073007), 2016.

[17] P. M. Ara, R. G. James, and J. P. Crutchfield. The elusive present: Hidden past and future dependence and why we build models. *Phys. Rev. E*, 93(2):022143, 2016.

[18] T. Koyuk and U. Seifert. Operationally accessible bounds on fluctuations and entropy production in periodically driven systems. *Phys. Rev. Lett.*, 122(23):230601, 2019.

[19] C. Maes. Frenetic bounds on the entropy production. *Phys. Rev. Lett.*, 119(16):160601, 2017.

[20] P. Strasberg and M. Esposito. Non-Markovianity and negative entropy production. *Phys. Rev. E*, 99(1)(012120), 2019.

[21] J. Bechhoefer. Hidden Markov models for stochastic thermodynamics. *New J. Physics*, 17(7):075003, 2015.

[22] E. Fredkin and T. Toffoli. Conservative logic. *Intl. J. Theo. Phys.*, 21(3-4):219–253, 1982.

[23] S. Deffner and C. Jarzynski. Information processing and the second law of thermodynamics: An inclusive, hamiltonian approach. *Phys. Rev. X*, 3(4):041003, 2013.

[24] A. B. Boyd, D. Mandal, and J. P. Crutchfield. Identifying functional thermodynamics in autonomous Maxwellian ratchets. *New J. Physics*, 18:023049, 2016.

[25] O.-P. Saira, M. H. Matheny, R. Katti, W. Fon, G. Wimsatt, J. P. Crutchfield, S. Han, and M. L. Roukes. Nonequilibrium thermodynamics of erasure with superconducting flux logic. *Phys. Rev. Res.*, 2(1):013249, 2020.

[26] P. M. Riechers, A. B. Boyd, G. W. Wimsatt, and J. P. Crutchfield. Balancing error and dissipation in computing. *Phys. Rev. Res.*, 2(3):033524, 2020.

[27] K. J. Laidler. The development of the Arrhenius equation. *J. Chem. Edu.*, 61.6(494), 1984.

[28] J. I. Steinfeld, J. S. Francisco, and W. L. Hase. *Chemical kinetics and dynamics.* Prentice Hall, 1999.

[29] A. B. Boyd, D. Mandal, P. M. Riechers, and J. P. Crutchfield. Transient dissipation and structural costs of physical information transduction. *Phys. Rev. Lett.*, 118:220602, 2017.

[30] A. Jurgens and J. P. Crutchfield. Functional thermodynamics of Maxwellian ratchets: Constructing and deconstructing patterns, randomizing and derandomizing behaviors. *Phys. Rev. Res.*, 2(3):033334, 2020.

[31] A. B. Boyd, D. Mandal, and J. P. Crutchfield. Thermodynamics of modularity: Structural costs beyond the Landauer bound. *Phys. Rev. X*, 8(3):031036, 2018.

[32] P. M. Riechers and J. P. Crutchfield. Fluctuations when driving between nonequilibrium steady states. *J. Stat. Phys.*, 168(4):873–918, 2017.

[33] P. M. Riechers and J. P. Crutchfield. Beyond the spectral theorem: Decomposing arbitrary functions of nondiagonalizable operators. *AIP Advances*, 8:065305, 2018.

[34] G. J. Milburn. Quantum optical Fredkin gate. *Phys. Rev. Lett.*, 62(18), 1989.

[35] J.-S. Wenzler, T. Dunn, T. Toffoli, and P. Mohanty. A nanomechanical Fredkin gate. *Nano Lett.*, 14(1):89–93, 2014.

Supplementary Materials

Non-Markovian Momentum Computing: Thermodynamically Efficient and Computation Universal

Kyle J. Ray, Alexander B. Boyd, Gregory W. Wimsatt, and James P. Crutchfield

Langevin Simulations

We simulated a system undergoing Langevin equations of motion:

$$\begin{aligned} dq &= v_q dt \\ m\, dv_q &= -\lambda v_q dt - \partial_q V dt + \sqrt{2 k_{\mathrm{B}} T \lambda}\, r(t) \sqrt{dt} \ , \end{aligned}$$

over three position coordinates $q = x$, y, or z. Here, v_q is the corresponding velocity, m is the mass, λ is the damping coefficient, and $r(t)$ is a memoryless Gaussian random variable with zero mean and unit variance. The one-dimensional storage potential $V^{\mathrm{store}}(q) = \alpha q^4 - \beta q^2$ (see Fig. S1) with coefficients α and β.

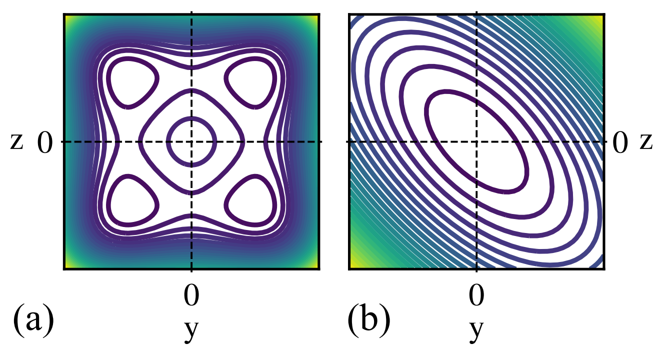

FIG. S1. Slice of the potential energy landscape $V(x, y, z)$ in the $y - z$ plane: (a) information-storing domain ($x = -x_0$) and (b) controlled-swap domain ($x = x_0$).

To simulate the above system, we first nondimensionalized the equations of motion:

$$\begin{aligned} d\widetilde{q} &= \widetilde{v}_q d\widetilde{t} \\ d\widetilde{v}_q &= -\gamma \widetilde{v}\, d\widetilde{t} + \partial_{\widetilde{q}} V d\widetilde{t} + \sqrt{2}\eta\, r(\widetilde{t}) \sqrt{d\widetilde{t}} \\ V^{\mathrm{store}}(\widetilde{q}) &= \widetilde{\alpha} \widetilde{q}^4 - \widetilde{\beta} \widetilde{q}^2 \\ V^{\mathrm{comp}}(\widetilde{y}', \widetilde{z}') &= \frac{\widetilde{k}}{2} \widetilde{y}'^2 + 2\widetilde{k} \widetilde{z}'^2 \ , \end{aligned}$$

where $\widetilde{\cdot}$ denotes a nondimensional quantity, and γ and η are two additional nondimensional parameters.

We conducted simulations over a range of values for $\gamma \in \{0.0, 0.5\}$, kept $\eta = \sqrt{\gamma}$, and held the remainder of the nondimensional parameters fixed: $\widetilde{\tau} = \pi$, $\widetilde{\alpha} = 2$, $\widetilde{\beta} = 16$, and $\widetilde{k} = 1$, where $\widetilde{\tau}$ is the nondimensional duration of the computation interval.

These choices are equivalent to four relationships between the dimensional parameters. The following three equalities held for all simulations: (i) $\tau = \pi\sqrt{m/k}$, (ii) $w = 2\sqrt{k_{\mathrm{B}} T / k}$, and (iii) $h = 32 k_{\mathrm{B}} T$, where τ is the dimensional duration of the computation interval, $w = \sqrt{\beta / 2\alpha}$ is the positional distance from the central maximum to the minima in the one-dimensional storage potential V^{store}, and $h = \beta^2 / 4\alpha$ is the energy difference between those points. The final dimensional relationship is an effective sweeping of the dimensional damping coefficient λ in $0 \leq \lambda \leq \frac{3}{10}\left(\frac{m}{\tau}\right)$.

The simulation employed the fourth-order Runge-Kutta method for the deterministic portion and Euler's method for the stochastic portion of the integration. (Python NumPy's Gaussian number generator was used to generate the memoryless Gaussian variable r(t).)

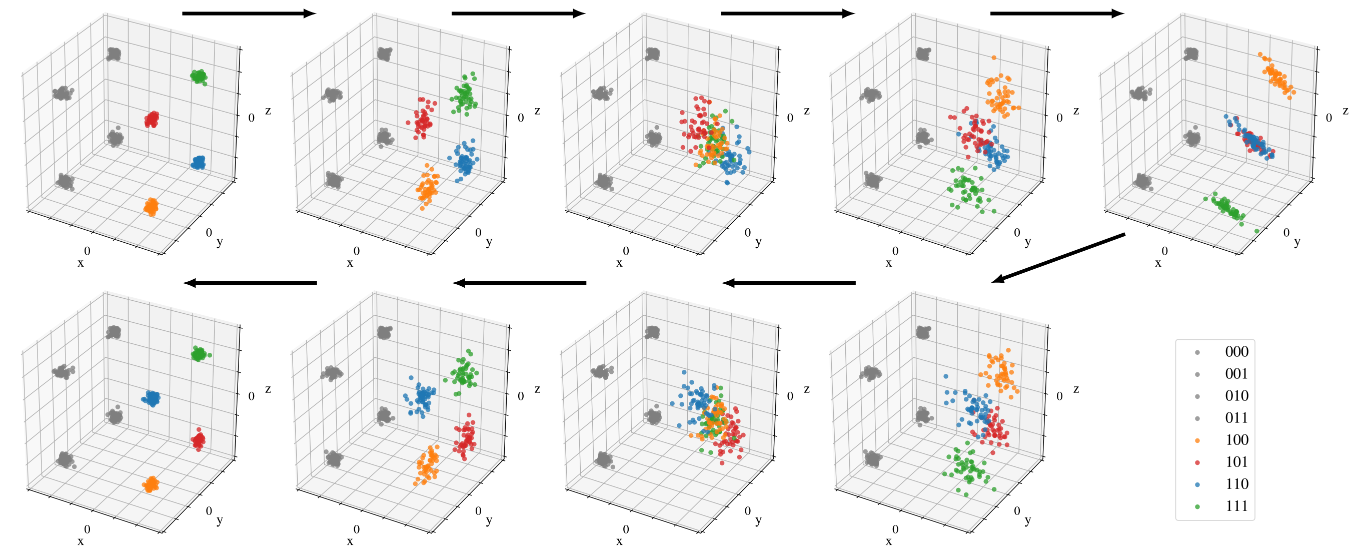

FIG. S2. Particle ensemble initialized in equilibrium with $V^{\text{store}}(x, y, z)$ undergoing the Fredkin gate protocol with zero coupling to the thermal reservoir. Each snapshot of the state evolution is separated by a time interval of $\tau/8$, with the black arrows indicating forward time. Color encodes in which informational state each trial begins. The 101 (red) and 110 (blue) states only oscillate by a quarter period in the time ($\tau/2$) it takes the 100 (yellow) and 111 (green) states to oscillate by a half cycle. As the 100 and 111 trials return to their initial positions, the 110 and 101 states approach their final positions: a half cycle from where they started (right). The states have been swapped. Animations available online.

Figure 2's plot displays 3σ error bars. The errors, though, are sufficiently small that they do not show up appreciably. Statistical errors were estimated using standard procedures for sample means and proportions.

Figure 2 was generated from simulation using the following procedure. First, an ensemble of 20,000trials were chosen from an approximate equilibrium distribution of $V^{\text{store}}(x, y, z)$ using the Monte Carlo algorithm. Second, this ensemble was thermalized while coupled to a bath ($\lambda = m/\tau$) until the ensemble energy changed by no more than 1 part in 1,000 over a unit time interval. Third, this ensemble was then used as the start state for the Fredkin gate operation. We then dropped λ down to a low coupling value and exposed the unit mass particles to the potential:

$$\begin{aligned} V(x, y', z', t) &= V^{\text{store}}(x) + V^{yz}(x, y', z', t) \\ &= \begin{cases} V^{\text{store}}(x, y, z) & \text{if } x < 0 \\ V^{\text{store}}(x) + \frac{ky'^2}{2} + 2kz'^2 & \text{if } x \geq 0 \end{cases} . \end{aligned} \tag{S1}$$

Fourth, we measured the work required to change the potential across our ensemble. Fifth, the potential was then held fixed for the computation duration τ using an integration step $dt \approx 0.0005\tau/\pi$. Finally, immediately following the computation interval, we measured the second work contribution—the work that would be harvested by dropping the potential back to V^{store}. The average net work is the ensemble average difference between the work invested when raising the potential and the work harvested when lowering it.

Figure 1 was generated by starting the particles in the equilibrium distribution described above and running the simulation with $\lambda = 0$, to simulate dissipationless oscillatory dynamics. For clarity, the plot shows a sample of 200 trials, rather than the full $20,000$. Figure S2 gives a more complete picture, with snapshots every $\tau/8$.