Title

Non-Markovian Momentum Computing: Universal and Efficient

Permalink

https://escholarship.org/uc/item/89x804pp

Authors

Ray, Kyle J
Wimsatt, Gregory W
Boyd, Alexander B
et al.

Publication Date

2020-10-02

Peer reviewed

Non-Markovian Momentum Computing: Universal and Efficient

Kyle J. Ray,[*] Gregory W. Wimsatt,[†] and James P. Crutchfield[‡]
Complexity Sciences Center and Physics Department,
University of California at Davis, One Shields Avenue, Davis, CA 95616

Alexander B. Boyd[§]
Complexity Institute, Nanyang Technological University, 3 Science Drive 2, Singapore 117543

(Dated: October 6, 2020)

All computation is physically embedded. Reflecting this, a growing body of results embraces rate equations as the underlying mechanics of thermodynamic computation and biological information processing. Strictly applying the implied continuous-time Markov chains, however, excludes a universe of natural computing. We show that expanding the toolset to continuous-time *hidden* Markov chains substantially removes the constraints. The general point is made concrete by our analyzing two eminently-useful computations that are impossible to describe with a set of rate equations over the memory states. We design and analyze a thermodynamically-costless bit flip, providing a first counterexample to rate-equation modeling. We generalize this to a costless Fredkin gate—a key operation in reversible computing that is computation 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 burgeoning field of thermodynamic computing leverages recent progress in nonequilibrium thermodynamics and information and computation theories [1–5] to establish a new physical paradigm for computation [6]. Working upwards from fundamental laws of physics, it promises to increase computational efficiency and power and to reduce energy dissipation in a next generation of computers. 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. However, recent efforts inadvertently and unnecessarily limit the potential scope. The following illustrates the breadth of that scope by introducing non-Markovian momentum computing that is both computation universal and thermodynamically efficient.

A computation over a time interval $t \in (0, \tau)$ is described by the conditional input-output mapping between memory states $m(0), m(\tau) \in \mathcal{M}$: $\mathbf{p}_{m(0) \to m(\tau)} = \Pr[m(\tau)|m(0)]$. The mapping $\mathbf{p}$ characterizes the probability of the final memory state given the initial memory state.

Attempts to establish a general framework for the required mappings in thermodynamic computing have assumed that the memory state m is a physical state of a system—it obeys stochastic Markovian dynamics, with time-evolution depending only on the system's current state [7, 8, and references therein]. More broadly, weak coupling to a thermal bath and separation of time scales are often invoked to justify such Markovian thermodynamic behavior [2, 9–12].

Separating the memory system's time scale from that of the heat bath serves to eliminate any memory in the heat bath from the system's behavior, as the heat bath is assumed to rapidly relax to a local equilibrium. As a result, transitions among memory states are potentially stochastic, but Markovian—they only depend on the current state [2]. Thus, these systems obey continuous-time Markov chains (CTMCs), which are equivalently represented by rate equations [7, 8]. Heat bath interactions allow for a broader suite of behaviors among memory states than purely deterministic Hamiltonian evolution allows on its own. However, this framework is still restrictive. For example, only input-output mappings whose determinants are positive are allowed when memory-state dynamics are restricted to obey CTMCs. And this, for better or worse, eliminates a wide range of possible and common computations, including flipping a single bit of information [13].

And so, while CTMCs are a powerful framework for stochastic thermodynamics [2, 14, 15], they fail to capture a broad swath of physically-realizable computing. They neglect the possibility of physical variables that carry hidden memory of the past beyond the immediate computational memory configuration.

In a thermodynamic system consisting of a collection of particle positions and momenta $(\vec{x}, \vec{p})$ the dynamics are Langevin: stochastic, but Markovian [16]. The underlying dynamic that governs the combination of the system's full microstate and the thermal bath together, though, is deterministic and Hamiltonian. The system's stochastic evolution results from coarse-graining over the bath degrees of freedom. Focusing on the system's Markov dynamic alone is a conventional modeling strategy, especially since one is typically uninterested in the bath's details.

Similarly, if system memory-states are determined only by

particle position, that choice of state coarse-grains away an additional component of the system and bath: system momentum. However, *unlike* the microstates of the ideal thermal bath, system momentum may carry memory of past behavior and this contributes to memory-state dynamics. Soberingly, analytical treatment of partially-observed (and therefore non-Markov) systems is highly nontrivial [5, 12, 17–19]. That said, removing the CTMC restriction permits realizing a broader range of computations. Thus, despite the additional analytical burden, investigating systems that operate in the regime where hidden states carry computationally-useful memory is a topic of current focus [5, 20–24].

We argue that the appropriate setting for thermodynamic computing is continuous-time *hidden* Markov chains (CTHMCs), in which hidden variables may store computationally-relevant information. Recently, this was recognized as sufficient for a broad class of input-output mappings **p** by introducing ancillary hidden states which implement sequences of logical operations that individually obey CTMC dynamics [8, 25]. However, CTHMCs implement more general computations still.

The following first implements a thermodynamically-costless bit flip—a simple computation that is explicitly forbidden by CTMCs. It then generalizes this to a costless Fredkin gate [26]—a key component in reversible computing that is also impossible to implement with CTMCs. The implementation of this universal and reversible logic gate via CTHMCs demonstrates that non-Markov dynamics are essential to thermodynamic computing.

Bit Flip 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 $p \in \mathbb{R}$ with double-well potential energy landscape: $V^{\mathrm{DW}}(x) \equiv \alpha x^4 - \beta x^2$, where both α and β are positive and determine the location of the potential minima $x^* = \pm\sqrt{\beta/2\alpha}$. The parameters are set so that information is stored robustly at the beginning of the computation interval ($t = 0$).

The particle's environment is a thermal bath at temperature T. As the height of the potential energy barrier at $x = 0$ 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 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 it now follows dissipationless Hamiltonian dynamics. Simultaneously, the potential energy landscape changes to a positive quadratic well: $V^{\mathrm{DW}}(x, t = 0^+) = kx^2/2$. The resulting particle motion is harmonic oscillation: $x(t) = x^* \cos\left(t\sqrt{k/m} + \phi\right)$, where x^* is the maximum distance from the origin of the cycle and ϕ is the phase

difference from maximum distance at the time $t = 0^+$. If we maintain the decoupled system in the quadratic potential energy landscape for half the period of oscillation $t \in \left(0, \pi\sqrt{m}/\sqrt{k}\right)$, then the particle's new position becomes: $x\left(\frac{\pi\sqrt{m}}{\sqrt{k}}\right) = x^* \cos\left(\pi + \phi\right) = -x^* \cos\left(\phi\right) = -x(0)$. Thus, over the computation interval $\tau = \pi\sqrt{m}/\sqrt{k}$, 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 the original double well and recouple to the thermal bath.

The work cost comes from changes in the potential energy at $t = 0$ and $t = \tau$. However, since particle position simply flips sign between $t = 0$ and $t = \tau$ and the potential energy landscape is even in position, zero net work must be generated during this time-symmetric 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. [27].

Fredkin Gate The bit-flip implementation may seem obvious in its simplicity. Can sophisticated and functional computing, in fact, be built from such simple passive processes? We answer in the affirmative by showing that a similar strategy implements the Fredkin gate, a reversible and universal logical gate [26]. This straightforwardly establishes that the CTHMC framework for thermodynamic computing gives easy access to complex and universal Turing computing.

The *Fredkin gate* operates on three bits $\mathcal{M} = \{0, 1\}^3$. That is, we encode our physical system 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 double-wells in each dimension:

$$V^{\mathrm{store}}(x, y, z) = V^{\mathrm{DW}}(x) + V^{\mathrm{DW}}(y) + V^{\mathrm{DW}}(z)$$
$$= \alpha\left(x^4 + y^4 + z^4\right) - \beta\left(x^2 + y^2 + z^2\right) .$$

This provides metastable regions corresponding to each memory state $m_x m_y m_z \in \{0, 1\}^3$.

Within this framework, we consider physical transformations that implement the Fredkin gate and do so robustly. The Fredkin gate is also known as the *controlled swap gate*, as it swaps 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 which swap with each other. The implementation uses the same strategy of decoupling and adding a harmonic potential over the time
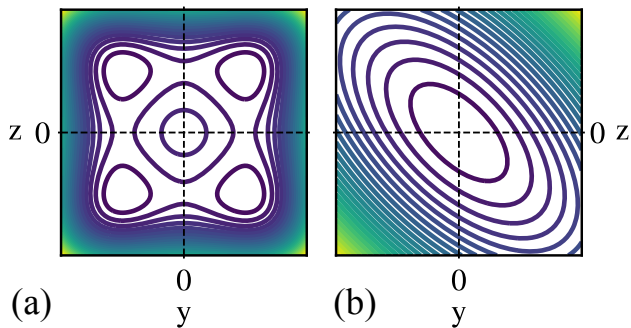
FIG. 1. Potential V^{yz} in the $y-z$ plane in the (a) information-storing ($x < 0$) and (b) controlled-swap ($x > 0$) domains.

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: $V(x, y, z, t) = V^{\mathrm{DW}}(x) + V^{yz}(x, y, z, t)$. 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: $V(x < 0, y, z, t) = V^{\mathrm{store}}(x, y, z)$, so that:

$$V^{yz}(x < 0, y, z, t) = V^{\mathrm{DW}}(y) + V^{\mathrm{DW}}(z) .$$

For $m_x = 1$, however, we must nontrivially compute on m_y and m_z:

$$V^{yz}(x \geq 0, y, z, t \in (0, \tau)) = V^{\mathrm{comp}}(y, z) .$$

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.

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: $V^{\mathrm{comp}}(y, z) = V(y') + V(z')$. 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 $\tau = \pi\sqrt{m}/\sqrt{k}$, 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 $\tau = \pi\sqrt{m}/\sqrt{k}$, returning to its original value $z'(\tau) = z(0)$, as does its momentum.

The resulting full Hamiltonian over the control interval operates piecewise. Figure 1 shows the potential in the $x < 0$ and $x > 0$ regions during the computation interval:

$$V(x, y', z', t) = V^{\mathrm{DW}}(x) + V^{yz}(x, y', z', t) \tag{1}$$

$$= \begin{cases} V^{\mathrm{store}} & \text{if } x < 0 \\ V^{\mathrm{DW}}(x) + \frac{ky'^2}{2} + 2kz'^2 & \text{if } x \geq 0 \end{cases}, \tag{2}$$

for $t \in (0, \tau)$. Translating back to the original coordinates $y = (y' + z')/\sqrt{2}$ and $z = (z' - y')/\sqrt{2}$, we find that for $x \geq 0$, this passive Hamiltonian transforms the particle's state by swapping y and z:

$$\begin{aligned}(y(\tau), z(\tau)) &= \left(\frac{y'(\tau) + z'(\tau)}{\sqrt{2}}, \frac{-y'(\tau) + z'(\tau)}{\sqrt{2}} \right) \\ &= \left(\frac{-y'(0) + z'(0)}{\sqrt{2}}, \frac{y'(0) + z'(0)}{\sqrt{2}} \right) \\ &= (z(0), y(0)) ,\end{aligned}$$

while it holds the other four quadrants where $m_x = 0$ in their respective potential minima. Thus, the transformation swapped y and z only when $m_x = 1$, implementing 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:

$$\begin{aligned}W &= V(x(0), y(0), z(0), 0^+) - V(x(0), y(0), z(0), 0) \\ &\quad + V(x(\tau), y(\tau), z(\tau), \tau) - V(x(\tau), y(\tau), z(\tau), \tau^-) .\end{aligned}$$

Recall the restriction that $x(t)$ is exponentially unlikely to change sign, because 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.

The x subspace of the potential decouples from the $y - z$ subspace and remains constant. Thus, there are no work contributions from the x-dependent terms. Additionally, the $y - z$ subspace potential is symmetric with respect to exchange of the y and z coordinates. So, the energy differences above will vanish for the y and z dependent terms as well. (Recall that the action of the potential over our interval is to swap the y and z coordinates so that $(y(\tau), z(\tau)) = ((z(0), y(0)))$.) And so, the average work production is nearly zero—only the exponentially
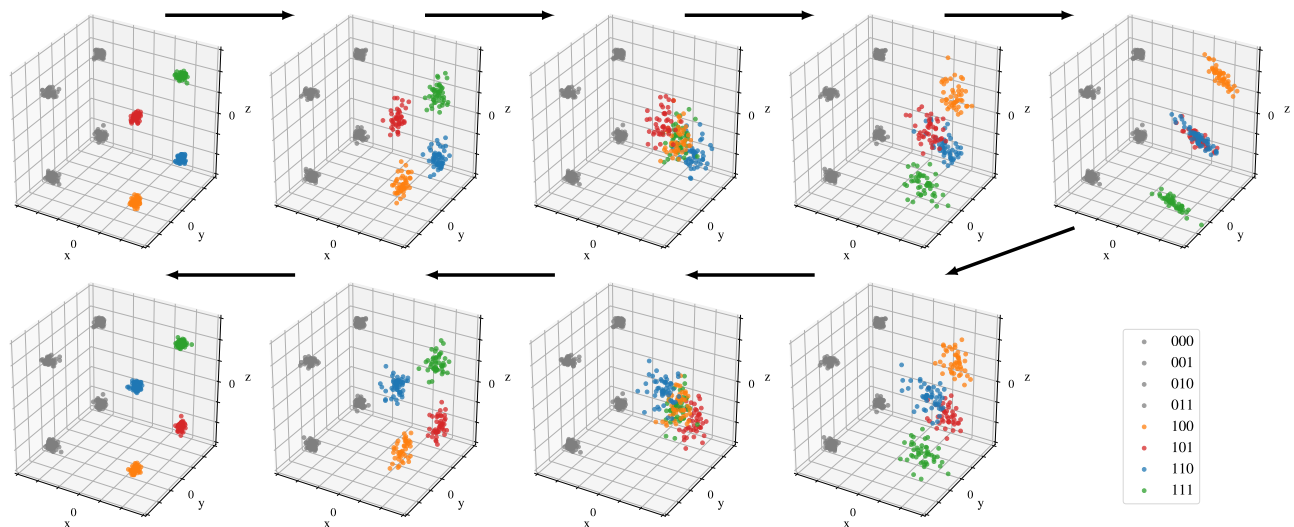
FIG. 2. Particle ensemble initialized in equilibrium with $V^{\mathrm{store}}(x, y, z)$ undergoing the Fredkin gate protocol with zero coupling to the thermal reservoir. Each snapshot of the state evolution is separated by a time interval of $\tau/8$, with the black arrows indicating forward time. Color 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.

suppressed barrier crossing events can contribute to non-zero work values.

Figure 2 demonstrates the evolution of the phase space on an ensemble of initial conditions drawn from the equilibrium distribution of $V^{\mathrm{store}}(x, y, z)$. 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-coordinate 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. However, the impact of adding thermal coupling is minimal. To demonstrate this, we investigated how robust the operation is to thermal coupling by using *underdamped* Langevin dynamics. A simulation was carried out by initializing 20000 particles in equilibrium with a thermal reservoir under the information-storing potential $V^{\mathrm{store}}(x, y, z)$. 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.)

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. 3 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. Thus, we see that the Fredkin gate (CTHMC) dynamics do not rely on removing the thermal reservoir.

As expected and shown in Fig. 3, 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 is much more than 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 predicted in Ref. [28].

Conclusion Rate equation dynamics is certainly a venerable and powerful framework, central to reaction kinetics in chemistry [29, 30] and key to the master equations of applied statistical mechanics [2, 14, 15]. In fact, perhaps due to the remarkable successes of continuous-time Markov chain predictions of many thermodynamic behaviors, it might seem natural to claim that in order to be "physically realizable", thermodynamic computing and biological information processing can *only* be described
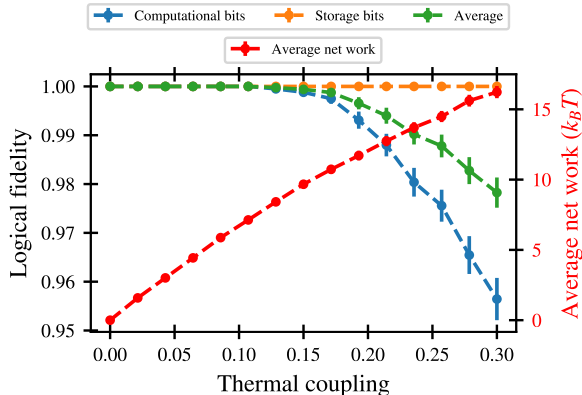
FIG. 3. Logical fidelity (successful trials/total trials) in the low-coupling Fredkin gate and the average net work required to implement it for different values of the thermal coupling constant λ. *Computational bits* refers to states that fall in the region $x > 0$, where the computational potential is in effect.

and analyzed as rate-equation dynamics [7].

However, we demonstrated this framework cannot form a complete basis for thermodynamic computing. Moreover, its strict application levies a penalty that precludes engineering and analyzing Maxwellian information ratchets, which are the physical equivalent of Turing machines [27, 31–33]. The limits are especially draconian, since efficient time-symmetrically controlled general computations consist of involutions [28]—operations that are composed of bit swaps and identity maps in positional memory (or any memory that is even under time reversal).

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 [7]. More and helpfully, we introduced explicit mechanisms for implementing both with zero work, proving that they are most certainly "physically realizable".

Not only are hidden Markov chains more general, but their added generality is critical in many circumstances. The fact that the Fredkin gate can be executed robustly, even when thermal fluctuations perturb the particle trajectories, suggests that this implementation will have practical use for reversible universal computing. The robustness of the gate to fluctuations separates it from other implementations of reversible computing—such as, ballistic computing with billiards—that are dynamically unstable [26].

We did, however, fully acknowledge the increased analytical complexity posed by CTHMC dynamics. Fortunately, the requisite tools have been developed that render the be-

haviors analytically tractable and in closed form [34, 35]. In short, there is little impediment to reaching the full generality of thermodynamic computing with CTHMCs.

Given that convincing physically-realizable implementations of the bit flip and and Fredkin gate [26, 36, 37] have been known for some time, one can only conclude that computing devices must be able to 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, the conception of memory must be modified, from being the realization of a microscopic physical state to being a mesoscopic coarse-graining, to fully realize the power and breadth of physical computations.

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

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

[8] 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, January 2019.

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

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

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

[12] 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.

[13] Such restrictions appear, for example, in Ref. [7]'s delineating what are "physically realizable" computations.

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

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

[16] R. Kubo, M. Toda, and N. Hashitsume. *Statistical physics II: Nonequilibrium statistical mechanics*, volume 31. Springer Science & Business Media, 2012.

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

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

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

[20] L. P. Fischer, P. Pietzonka, and U. Seifert. Large deviation function for a driven underdamped particle in a periodic potential. *Phys. Rev. E*, 97(2):022143, 2018.

[21] K. Brandner, T. Hanazato, and K. Saito. Thermodynamic bounds on precision in ballistic multiterminal transport. *Phys. Rev. Lett.*, 120(9):090601, 2018.

[22] P. S. Stumpf, R. C. G. Smith, M. Lenz, A. Schuppert, F-J. Müller, A. Babtie, T. E. Chan, M. P. H. Stumpf, C. P. Please, S. D. Howison, et al. Stem cell differentiation as a non-Markov stochastic process. *Cell Systems*, 5(3):268–282, 2017.

[23] L. Ferialdi. Exact closed master equation for Gaussian Non-Markovian dynamics. *Phys. Rev. Lett.*, 116:120402, Mar 2016.

[24] J. M. Horowitz and T. R. Gingrich. Proof of the finite-time thermodynamic uncertainty relation for steady-state currents. *Phys. Rev. E*, 96(2):020103, 2017.

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

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

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

[28] 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. arXiv:1909.06650.

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

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

[31] 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.

[32] 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.

[33] 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.

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

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

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

[37] 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: Universal and Efficient

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

Langevin Simulations

Simulations were carried out using Langevin equations of motion:

$$dx = vdt$$
$$dv = -\lambda vdt - \partial_x U(x,t)dt + \sqrt{2k_\mathrm{B}T\lambda}\, r(t)\sqrt{dt}\ ,$$

where $r(t)$ is a memoryless Gaussian random variable with zero mean and unit variance. Since we track behavior when sweeping the thermal coupling parameter λ only, it is convenient to consider a particle with unit mass ($m = 1$) and set $k_\mathrm{B}T = 1$. This yields a very simple dynamic that is readily interpreted:

$$dx = vdt$$
$$dv = -\lambda vdt - \partial_x U(x,t)dt + \sqrt{2\lambda}\, r(t)\sqrt{dt}\ .$$

The parameter λ (the *thermal coupling coefficient*) controls the damping force the particle experiences from the thermal bath when it has unit velocity. It is commonly called the damping coefficient or the inverse mobility.

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

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

Figure 3 was generated from simulation using the following procedure. First, an ensemble of 20000 trials were chosen from an approximate equilibrium distribution of $V^{\mathrm{store}}(x, y, z)$ with $\alpha = 2, \beta = 16$, using the Monte Carlo algorithm. Second, this ensemble was thermalized while coupled to a bath ($\lambda = 1$) until the ensemble energy changed by no more than 1 part in 1000 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 in Eq. (2) with $\alpha = 2$, $\beta = 16$, and $k = 1$. Fourth, at this point we measured the work required to change the potential across our ensemble. Fifth, the potential was then held fixed for a time $\tau = \pi\sqrt{k/m} = \pi$ using an integration step $dt \approx 0.0005$. 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 2 was generated by starting the particles in the equilibrium distribution described above, and running the simulation with $\lambda = 0$, to simulate dissipationless oscillatory dynamics. The plot shows a sample of 200 trials, rather than the full 20000, for clarity.