Shortcuts to Thermodynamic Computing:
The Cost of Fast and Faithful Erasure

Alexander B. Boyd,[1, *] Ayoti Patra,[2, †] Christopher Jarzynski,[2, 3, 4, ‡] and James P. Crutchfield[1, §]

[1] *Complexity Sciences Center and Physics Department,*
University of California at Davis, One Shields Avenue, Davis, CA 95616
[2] *Department of Physics, University of Maryland, College Park, Maryland 20742, USA*
[3] *Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, USA*
[4] *Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA*
(Dated: December 28, 2018)

Landauer's Principle states that the energy cost of information processing must exceed the product of the temperature and the change in Shannon entropy of the information-bearing degrees of freedom. However, this lower bound is achievable only for quasistatic, near-equilibrium computations—that is, only over infinite time. In practice, information processing takes place in finite time, resulting in dissipation and potentially unreliable logical outcomes. For overdamped Langevin dynamics, we show that counterdiabatic potentials can be crafted to guide systems rapidly and accurately along desired computational paths, providing shortcuts that allows for the precise design of finite-time computations. Such shortcuts require additional work, beyond Landauer's bound, that is irretrievably dissipated into the environment. We show that this dissipated work is proportional to the computation rate as well as the square of the information-storing system's length scale. As a paradigmatic example, we design shortcuts to erase a bit of information metastably stored in a double-well potential. Though dissipated work generally increases with erasure fidelity, we show that it is possible perform perfect erasure in finite time with finite work. We also show that the robustness of information storage affects the energetic cost of erasure—specifically, the dissipated work scales as the information lifetime of the bistable system. Our analysis exposes a rich and nuanced relationship between work, speed, size of the information-bearing degrees of freedom, storage robustness, and the difference between initial and final informational statistics.

I. INTRODUCTION

Information processing requires work. For example, no less than $k_\mathrm{B} T \ln 2$ of work must be supplied in order to erase a single bit of information at temperature T [1]. More generally, Landauer's Principle bounds the work investment by the change in the memory's Shannon entropy [2]:

$$\langle W \rangle \geq k_\mathrm{B} T \ln 2 \left(\mathrm{H}[Y_0] - \mathrm{H}[Y_\tau] \right) . \tag{1}$$

Here, Y_0 and Y_τ are random variables describing initial and final memory states, and $\mathrm{H}[Y] = -\sum_y \Pr(Y = y) \log_2 \Pr(Y = y)$ denotes the uncertainty in bits of a random variable Y.

Mathematically, information processing is described by a communication channel [3] that maps an initial distribution $\Pr(Y_0)$ to a final distribution $\Pr(Y_\tau)$. Physically, a memory is realized by a system whose thermodynamically-metastable states encode logical

states $\{y\}$. The simplest example is a Brownian particle in a double-well potential, with two deep wells representing the $y = 0$ and $y = 1$ states of a single bit of information. More generally, the collection of all possible memory states $\mathcal{Y} = \{y\}$ represents a mesoscopic coarse-graining of the space of explicit physical microstates $\mathcal{X} = \{x\}$ of the memory device. Information processing is implemented by varying the system's energy landscape so as to drive the flow of probability between memory states in a controlled fashion, to achieve a desired computation.

A computation can be implemented to achieve the Landauer bound, Eq. (1), by varying the energy landscape infinitely slowly, so that the system remains in metastable equilibrium from beginning to end [4, 5]. Such quasistatic computations, however, take infinitely long to implement. For computations performed in finite time the underlying physical system is driven out of equilibrium, resulting in the irretrievable dissipation of energy into thermal surroundings. The problem of minimizing this dissipation has recently been explored in the linear response regime, using the tools of geometric thermodynamic control [6–9].

In the present work, we consider the separate problem of how to implement a computation rapidly and reliably. That is, we study how to design protocols for varying

---
* abboyd@ucdavis.edu
† ayotipatra@gmail.com
‡ cjarzyns@umd.edu
§ chaos@ucdavis.edu

the energy landscape of a system, so as to produce a desired computation in a given time interval, no matter how short its duration τ. In effect, we place a premium on speed of computation rather than energy efficiency, although we then proceed to analyze the energetic costs of rapid computation.

To achieve rapid and precisely-controlled information processing, we use recently developed tools from the field of *shortcuts to adiabaticity* [10]. Specifically, we apply generic methods of counterdiabatic control of classical overdamped systems [11], which were inspired by pioneering experiments on the engineered swift equilibration of a Brownian particle [12]. The results we obtain are not limited to the linear response regime—they remain valid even when the system is driven far from equilibrium during the information processing.

For concreteness, we show how to apply counterdiabatic control to erase a single bit of information rapidly and accurately, but our approach generalizes to other cases of information processing. To embed the memory states $\mathcal{Y}$ physically, we consider a one-dimensional position space $\mathcal{X}$ governed by overdamped Fokker-Planck dynamics. The energy landscape at the beginning and end of the protocol is the double-well potential shown in Fig. 2, with a barrier sufficiently high to prevent the leakage of probability between the two wells. Thus, the landscape provides a means of storing information in metastable mesoscopic states. As we will show, counterdiabatic control of the potential can be used to drive any initial distribution over the memory states to any desired final distribution in finite time—in fact, arbitrarily rapidly. Mirroring results in geometric control, we show that the work required to perform this counterdiabatic process decomposes into a change in free energy, which captures Landauer's change in state space, plus an additional contribution that scales as the rate of computation and the square of the length scale of the information-storing potential [7]. This additional work is proportional to the global entropy production and so quantifies thermodynamic inefficiency.

Our approach reveals additional trade-offs beyond that between computation rate, length scale, and thermodynamic efficiency. We show that dissipation also increases with the difference between initial and final bit distributions of the computation and with the robustness of information storage. In this way, we give a more complete picture of information processing beyond Landauer's bound. Rather than a tradeoff between information and energy, more complex tradeoffs are revealed between information, energy, statistical bit-bias difference, speed, size of the memory states, and information robustness.

II. THERMODYNAMIC COMPUTING

What is physical computing? At the outset, information must be encoded in collections of microscopic states $\mathcal{X}$ of a physical system. Let $\mathcal{Y}$ denote these information-containing microstate groups—the accessible *memory states* [13, 14]. By manipulating the physical system, a microstate collection evolves, transforming the information it contains. Generally, an information processor has only partial control over and knowledge of the underlying microstates of its physical implementation. We now consider how such information processing can be modeled by stochastic dynamics governed by a controlled potential.

A. Memory States and Symbolic Dynamics

There are many ways to form memory states out of physical microstates. Here, we choose a framework for information erasure and general information processing in which the physical degrees of freedom $\mathcal{X}$ participate in metastable equilibria. Each metastable equilibrium is a microstate distribution that corresponds to a memory state $y \in \mathcal{Y}$. For example, we can have memory states $\mathcal{Y} = \{0, 1\}$, such that they are stable for intermediate, if not asymptotically long, time scales. The coarse-graining $c : \mathcal{X} \rightarrow \mathcal{Y}$ of physical states to form the informational states specifies the memory alphabet $\mathcal{Y} = \{c(x) | x \in \mathcal{X}\}$. This translates a distribution $\Pr(X_t)$ over physical microstates $x \in \mathcal{X}$ to a distribution $\Pr(Y_t)$ over informational states $y \in \mathcal{Y}$. In this way, controlling a physical system determines not only its raw physical dynamics, but also the *symbolic dynamics* of the informational states [15].

We use random variable notation, $\Pr(X_t) = \{(x, \Pr(X_t = x)), x \in \mathcal{X}\}$, common in symbolic dynamics [16], rather than $\rho(x, t)$, which is more standard in stochastic thermodynamics, due to its specificity and flexibility. The probability of being in microstate x at time t is expressible in both notations $\Pr(X_t = x) = \rho(x, t)$, but the random variable notation works with many different distributions over the same microstate space $\mathcal{X}$. And so, rather than specify many different probability functions, we specify their random variables. Other advantages of this choice is that it readily expresses joint probabilities, such as $\Pr(X_t = x, X_{t+\tau} = x')$, and entropies:

$$\mathrm{H}[X_t] = - \sum_{x \in \mathcal{X}} \Pr(X_t = x) \log_2 \Pr(X_t = x) . \quad (2)$$

While not all of the potential functionality is used in the following, a number of benefits will follow in due course.

B. Overdamped Fokker-Planck Dynamics

The first challenge of thermodynamic computing is to control a system's Hamiltonian over the physical degrees of freedom $\mathcal{X}$ such that the induced microstate distribution $\Pr(X_t)$ at time t matches a desired distribution $\Pr(X_t^d)$, where X_t and X_t^d are the random variables for the actual physical distribution and desired physical distribution, respectively, at time t, each realizing states $x \in \mathcal{X}$. The second challenge, which we come to later, is to associate the microstate distributions with mesostate distributions that support the desired information-storing and -processing.

We consider a Hamiltonian controlled via a potential energy landscape $V(x, t)$ over the time interval $t \in (0, \tau)$, where $x \in \mathcal{X}$. We will demonstrate that one can exactly guide an overdamped Fokker-Planck dynamics in one dimension along the desired time sequence of distributions $\Pr(X_t^d = x)$, resulting in a powerful tool for thermodynamic control and information processing.

In fact, overdamped stochastic systems are a promising and now common framework for elementary thermodynamic information processing [17, 18]. With a single physical degree of freedom $\mathcal{X} = \mathbb{R}$, one information processing task is to change the initial distribution to a final distribution in finite time. The actual microstate distribution $\Pr(X_t)$ obeys the Fokker-Planck equation:

$$\frac{\partial \Pr(X_t = x)}{\partial t} = \mu \frac{\partial}{\partial x} \left(\Pr(X_t = x) \frac{\partial V(x, t)}{\partial x} \right) + \mu k_{\mathrm{B}} T \frac{\partial^2 \Pr(X_t = x)}{\partial x^2} , \quad (3)$$

where $V(x, t)$ is the potential energy landscape at time t, T is the temperature of the thermal environment, and μ is the inverse friction coefficient. Recall that the Boltzmann equilibrium distribution:

$$\Pr(X_t^{\mathrm{eq}} = x) = \frac{e^{-V(x,t)/k_{\mathrm{B}}T}}{Z(t)} , \quad (4)$$

is a stationary distribution for the Fokker-Planck equation if the potential is held fixed at time t. That is, substituting into the righthand side of Eq. (3) yields:

$$\frac{\partial \Pr(X_t^{\mathrm{eq}} = x)}{\partial t} = 0 .$$

III. WORK PRODUCTION DURING COUNTERDIABATIC PROTOCOLS

Next, we identify how the evolution of the physical distribution yields useful changes in memory states

that robustly store a computation's result. We break the development into two parts. This section considers counterdiabatic Hamiltonian control of the physical states $x \in \mathcal{X}$ such that they follow specified distributions $\Pr(X_t^d)$ over the time interval $t \in (0, \tau)$ [11]. For the resulting finite-time protocol, we determine the work production and show that it increases with both the size of the memory states and the speed of operation, if the overall computational task is fixed. This holds for any counterdiabatically-controlled computation. The subsequent section then addresses the particular computational task of information erasure in a bistable potential well. While the analytical and numerical results there do not explicitly generalize to other computational tasks, they introduce general relationships between dissipated work, information storage robustness, and computation fidelity that hold broadly.

A. Inverse Problem for Thermodynamic Control

For a specified potential $V(x, t)$, the Fokker-Planck equation Eq. (3) evolves an initial distribution $\Pr(X_0)$ to a density $\Pr(X_t)$ at any later time in the control interval $t \in (0, \tau)$. Together, the probability density and potential determine the average energy expended as work $\langle W \rangle$ by the protocol on the physical system [14]:

$$\langle W \rangle = \int_0^\tau dt \int_{-\infty}^\infty dx \, \Pr(X_t = x) \partial_t V(x, t) . \quad (5)$$

What if, rather than starting with an initial distribution and control protocol, we are given a desired trajectory of probability distributions $\Pr(X_t^d)$ over some time interval $t \in [0, \tau]$, a *distribution trajectory*, and are tasked to determine the control protocol that yields the trajectory? This challenge—the *inverse problem* of reconstructing dynamical equations of motion from distributions over trajectories—falls within purview of state-space reconstruction [19, 20] and computational mechanics [21] which provide principled approaches for inferring generators of observed time series. Broadly speaking, our challenge here is to reconstruct dynamical equations of motion for evolving distributions that perform computations and, then, to show how the work cost relates to the computation's effectiveness. The setting here is both more constrained and more challenging than state-space reconstruction.

Generally, as with most inverse problems, determining the control protocol from a distribution trajectory does not lead to a unique solution. Many alternate dynamics can generate the same observed distributions [22]. However, Appendix A shows, for the specific case of overdamped Fokker-Planck dynamics in a single dimension

$\mathcal{X} = \mathbb{R}$, that the distribution trajectory $\Pr(X_t^d)$ uniquely determines the control protocol $V(x, t)$ up to a baseline energy $E(t)$ that is constant in position and so adds no force. Moreover, if $\Pr(X_t^d)$ characterizes our desired computation then, up to a readily-recovered change in baseline energy $E(\tau) - E(0)$, the work is uniquely determined for that computation. Thus, by designing a single protocol that guides the system along the desired distribution trajectory, we find both the unique protocol and the unique work investment required for that trajectory.

When τ is much larger than the system's relaxation timescale τ^{eq}, a control protocol can be determined by assuming the system remains approximately in equilibrium at all times: $\Pr(X_t = x) \approx \Pr(X_t^{\mathrm{eq}} = x)$. This *quasistatic* (adiabatic) control protocol is determined from the *quasistatic potential*:

$$V^Q(x, t) \equiv F^{\mathrm{eq}}(t) - k_{\mathrm{B}} T \ln \Pr(X_t^d = x) , \qquad (6)$$

where the equilibrium free energy:

$$\begin{aligned} F^{\mathrm{eq}}(t) &= -k_{\mathrm{B}} T \ln Z(t) \\ &= -k_{\mathrm{B}} T \ln \int_{-\infty}^{\infty} dx e^{-V^Q(x,t)/k_{\mathrm{B}} T} \end{aligned} \qquad (7)$$

is the baseline energy $E(t)$. Note that $\Pr(X_t^d)$ is the equilibrium distribution corresponding to $V^Q(x, t)$, see Eq. (4). In the large-τ case, the system follows this equilibrium distribution, as shown in Fig. 1, and the quasistatic protocol provides the unique solution to our control problem. Moreover, the work invested is the change in equilibrium free energy:

$$\langle W^Q \rangle = \Delta F^{\mathrm{eq}} . \qquad (8)$$

If, however, τ is *not* much larger than τ^{eq}, then evolution under the quasistatic potential $V^Q(x, t)$, defined by Eq. (6), does not drive the system along the desired trajectory distribution $\Pr(X_t^d)$. Rather, the actual distribution $\Pr(X_t)$ deviates from the desired distribution as the system is pushed away from equilibrium. Recent work [11] describes how to construct a *counterdiabatic protocol* that achieves the desired evolution, $\Pr(X_t) = \Pr(X_t^d)$, for all $t \in (0, \tau)$. In this approach the overdamped system evolves under a potential:

$$V(x, t) = V^Q(x, t) + V^{CD}(x, t) , \qquad (9)$$

which consists of both the quasistatic term $V^Q(x, t)$ and a *counterdiabatic potential* $V^{CD}(x, t)$, constructed to guarantee that the actual distribution tracks the desired distribution, $\Pr(X_t) = \Pr(X_t^d)$, as illustrated in Fig. 1.

By Eq. (6) $\Pr(X_t^d)$ is the equilibrium distribution corresponding to the quasistatic potential $V^Q(x, t)$, but it is not the equilibrium distribution corresponding to the total potential $V(x, t)$ given by Eq. (9). Thus, when the system evolves under the counterdiabatic protocol, it is out of equilibrium with respect to the instantaneous potential $V(x, t)$ at intermediate times $t \in (0, \tau)$. However, to ensure that the system starts and ends in equilibrium, we choose $\Pr(X_t^d)$ such that $\partial_t \Pr(X_t^d)$ vanishes at the start and end of the protocol. This way the counterdiabatic potential vanishes at the endpoints of the protocol: $V^{CD}(x, t = 0, \tau) = 0$. And so, the potential energy becomes the quasistatic potential $V(x, t = 0, \tau) = V^Q(x, t = 0, \tau)$, as shown in Fig. 1.

B. Counterdiabatic Control of Stochastic Systems

Reference [11] showed that the counterdiabatic potential $V^{CD}(x, t)$ is constructed from the desired distribution $\Pr(X_t^d)$ by integrating a velocity *flow field* $v(x, t)$, defined shortly:

$$V^{CD}(x, t) = -\frac{1}{\mu} \int_0^x v(x', t) dx' .$$

The lower limit of integration is set to 0 for convenience. In fact, it may take any value, as the physics is unchanged by the addition of an arbitrary function $f(t)$ to the potential. The velocity flow field:

$$v(x, t) = \left. \frac{\partial x}{\partial t} \right|_{C = \mathrm{const}} = -\frac{\partial_t C}{\partial_x C}$$

is the velocity of constant values of the cumulative distribution function:

$$C(x, t) = \int_{-\infty}^x \Pr(X_t^d = x') dx' .$$

Combining results, we have, explicitly:

$$V^{CD}(x, t) = \frac{1}{\mu} \int_0^x \int_{-\infty}^{x'} \frac{\partial_t \Pr(X_t^d = x'')}{\Pr(X_t^d = x')} dx'' dx' , \quad (10)$$

for $t \in (0, \tau)$. For $t \notin (0, \tau)$ we set $V^{CD}(x, t) = 0$, hence $V(x, t) = V^Q(x, t)$ outside of the control interval. As a result, the system begins in the equilibrium distribution at $t = 0$ and it ends (and subsequently remains) in equilibrium at $t \geq \tau$.

Since the potential energy $V(x, t)$ consists of quasistatic and counterdiabatic terms, we can similarly de-
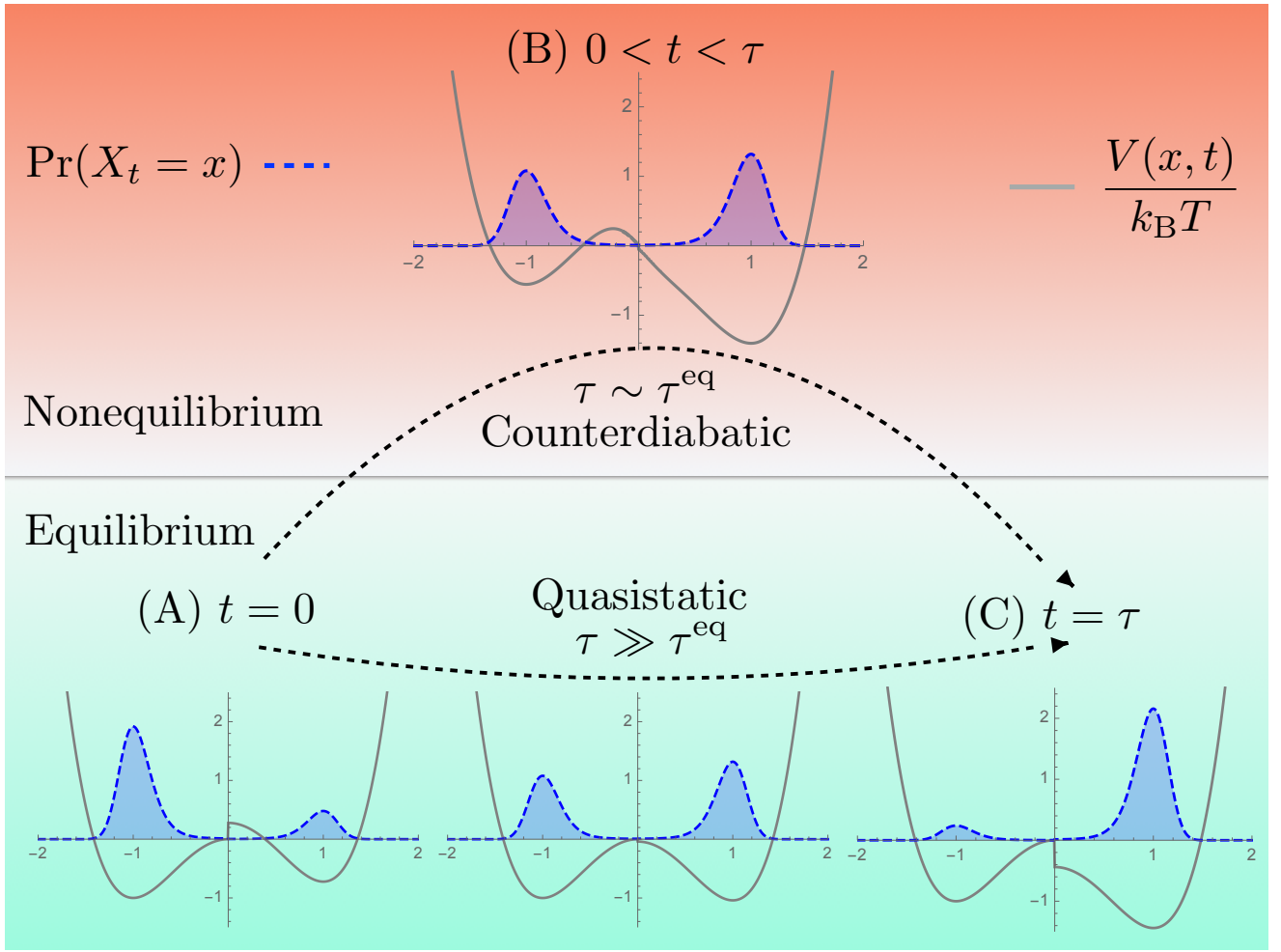
FIG. 1. Counterdiabatic control of the energy landscape $V(x,t)$ (solid gray curve) at times along the interval $t \in [0,\tau]$ guides the probability distribution $\Pr(X_t = x)$ (dashed blue curve) along a desired trajectory $\Pr(X_t^d = x)$ in finite time τ. The system starts in equilibrium in stage (A) and ends in equilibrium at stage (C), meaning that $V(x,0) = V^Q(x,0)$ and $V(x,\tau) = V^Q(x,\tau)$ are the quasistatic potentials for the initial and final distributions, respectively. However, at intermediate times, in stage (B), the necessary control protocol $V(x,t)$ required to guide the system along the desired distribution changes as we change the speed of the protocol. If the timescale of equilibration is relatively very small $\tau^{\text{eq}} \ll \tau$, then the control protocol must be in equilibrium with the desired distribution, such that the control potential is described by the quasistatic potential $V(x,t) = V^Q(x,t)$, as shown in the lower half of stage (B). Otherwise, an additional counterdiabatic term $V^{CD}(x,t)$ is added to the potential energy, which pushes the system out of equilibrium, as shown in the upper half of stage (B).

compose the work in Eq. (5) into two contributions:

$$\langle W \rangle = \int_0^\tau dt \int_{-\infty}^\infty dx \, \Pr(X_t = x) \partial_t V^Q(x,t)$$
$$+ \int_0^\tau dt \int_{-\infty}^\infty dx \, \Pr(X_t = x) \partial_t V^{CD}(x,t)$$
$$= \langle W^Q \rangle + \langle W^{CD} \rangle$$
$$= \Delta F^{\text{eq}} + \langle W^{CD} \rangle . \tag{11}$$

The first term $\langle W^Q \rangle$ is the amount of work that would be performed *if* the protocol were executed quasistatically, i.e., reversibly. This *quasistatic work* is simply the change in equilibrium free energy, as follows by direct substitution of Eq. (6) into the first line above. This contribution depends only on the initial and final potential and not on either (i) the sequence of intermediate distributions or (ii) the duration of the protocol.

The second contribution $\langle W^{CD} \rangle$ is the *counterdiabatic work*, and it is proportional to the global entropy production $\langle \Sigma \rangle$. Specifically, when the system begins and ends in equilibrium we have [23]:

$$T\langle \Sigma \rangle = \langle W \rangle - \Delta F^{\text{eq}} = \langle W^{CD} \rangle , \tag{12}$$

where $\langle \Sigma \rangle \geq 0$ quantifies the net change in the entropy of the system and its thermal surroundings.

In Eq. (11), the quasistatic work is fixed and the coun-

terdiabatic work gives the path-dependent dissipated work required to complete the transformation in finite time. Thus, all dependence on intermediate details is captured by $\langle W^{CD} \rangle$. This quantity is our principal focus and, as we now show, it scales particularly simply with system size and computation time.

We note that Eqs. (11) and (12), along with the inequality $\langle \Sigma \rangle \geq 0$, generalize to transformations between nonequilibrium states, with ΔF^{eq} replaced by the recoverable nonequilibrium free energy, ΔF^{neq}; see Refs. [2, 24, 25] for details. We will use this generalized result in Sec. IV when discussing counterdiabatic erasure.

C. System-Size and Computation-Rate Dependence

A protocol's *duration* τ is the time over which the Hamiltonian varies. For our one-dimensional system, we define a characteristic *system length* L reflecting the extent of the desired probability distribution's support. Since we wish to capture only the distribution's bulk and not the support's absolute extent, there are many ways to define this length. A candidate is the initial variance:

$$L = \sqrt{\int_{-\infty}^{\infty} dx \Pr(X_0^d = x)x^2 - \left(\int_{-\infty}^{\infty} dx \Pr(X_0^d = x)x \right)^2} \ .$$

The particular form is somewhat arbitrary. All we ask is that L scale appropriately when transforming the distribution. With these definitions in hand, we can analyze how the protocol and dissipation change under rescalings.

Consider the probability trajectory $\{\Pr(X_t^d = x) : t \in (0, \tau)\}$ and a system of length L, yielding the control protocol $V(x,t) = V^Q(x,t) + V^{CD}(x,t)$. To preserve the probability trajectory shape while changing the duration to τ' and length to L', we introduce a new desired trajectory:

$$\Pr(X_t^{d\prime} = x) = \Pr(X_{\tau t/\tau'}^d = Lx/L') \frac{L}{L'} \ .$$

This stretches the original distribution's support by a factor L'/L and increases the computation rate by a factor τ/τ'.

In the expression for the resulting counterdiabatic control protocol:

$$V'(x,t) = V'^Q(x,t) + V'^{CD}(x,t) \ ,$$

we define a new quasistatic potential as the similarly-scaled version of the original:

$$V^{Q\prime}(x,t) = V^Q(Lx/L', \tau t/\tau') \ .$$

The associated equilibrium free energy is expressed in terms of the original free energy:

$$
\begin{aligned}
F^{\mathrm{eq}\prime}(t) &= -k_{\mathrm{B}}T \ln Z'(t) \\
&= -k_{\mathrm{B}}T \ln \int_{-\infty}^{\infty} dx e^{-V^{Q\prime}(x,t)/k_{\mathrm{B}}T} \\
&= -k_{\mathrm{B}}T \ln \int_{-\infty}^{\infty} dx' \frac{L'}{L} e^{-V^Q(x', \tau t/\tau')/k_{\mathrm{B}}T} \\
&= k_{\mathrm{B}}T \ln \frac{L}{L'} + F^{\mathrm{eq}}(\tau t/\tau') \ , \quad (13)
\end{aligned}
$$

where the third line comes from substituting $x = x'L'/L$. Equation (13) implies:

$$\Delta F^{\mathrm{eq}\prime} = F^{\mathrm{eq}\prime}(\tau') - F^{\mathrm{eq}\prime}(0) = \Delta F^{\mathrm{eq}} \ ,$$

hence the quasistatic work is the same for protocols with different durations and lengths:

$$\langle W^{Q\prime} \rangle = \langle W^Q \rangle \ .$$

The counterdiabatic contributions, however, yield meaningful differences when changing system length or protocol duration. Substituting the rescaled probability trajectory into the expression for counterdiabatic potential in Eq. (10), we find:

$$
\begin{aligned}
V^{CD\prime}(x,t) &= \frac{1}{\mu} \int_0^x \int_{-\infty}^{x'} \frac{\partial_t \Pr(X_t^{d\prime} = x'')}{\Pr(X_t^{d\prime} = x')} dx'' dx' \\
&= \frac{1}{\mu} \int_0^x \int_{-\infty}^{x'} \frac{\partial_t \Pr(X_{\tau t/\tau'}^d = Lx''/L')}{\Pr(X_{\tau t/\tau'}^d = Lx'/L')} dx'' dx' \\
&= \frac{1}{\mu} \frac{L'^2}{L^2} \int_0^{Lx/L'} \int_{-\infty}^{x'''} \frac{\partial_{t'} \Pr(X_{t'}^d = x'''')\partial_t t'}{\Pr(X_{t'}^d = x''')} dx'''' dx''' \\
&= \frac{\tau L'^2}{\tau' L^2} V^{CD}(Lx/L', \tau t/\tau') \ ,
\end{aligned}
$$

using the substitutions $t' = \tau t/\tau'$, $x''' = Lx'/L'$, and $x'''' = Lx''/L'$. Thus, the counterdiabatic potential scales as the square of the length of the information storage device and as the inverse of the protocol duration. Equivalently, the additional nonequilibrium force $F^{CD}(x,t) = -\partial_x V^{CD}(x,t)$ applied to the system scales as the computation rate and square of the system size.

For the counterdiabatic work we similarly find:

$$\langle W^{CD\prime}\rangle = \int_0^{\tau'} dt \int_{-\infty}^{\infty} dx \Pr(X_t^{d\prime} = x)\partial_t V^{CD\prime}(x,t)$$

$$= \frac{\tau L'^2}{\tau' L^2} \int_0^{\tau'} dt \int_{-\infty}^{\infty} dx \frac{L}{L'} \Pr\left(X_{\frac{\tau}{\tau'}t}^d = \frac{L}{L'}x\right) \partial_t V^{CD}\left(\frac{L}{L'}x, \frac{\tau}{\tau'}t\right)$$

$$= \frac{\tau L'^2}{\tau' L^2} \int_0^{\tau} \frac{\tau'}{\tau} dt' \int_{-\infty}^{\infty} \frac{L'}{L} dx' \frac{L}{L'} \Pr(X_{t'}^d = x')\partial_t V^{CD}(x',t')$$

$$= \frac{\tau L'^2}{\tau' L^2} \frac{\tau'}{\tau}(\partial_t t') \int_0^{\tau} dt' \int_{-\infty}^{\infty} dx' \Pr(X_{t'}^d = x')\partial_{t'} V^{CD}(x',t')$$

$$= \frac{\tau L'^2}{\tau' L^2} \langle W^{CD}\rangle .$$

And so, too, the dissipated counterdiabatic work scales as system length squared and linearly with computation rate. This work, in turn, is proportional to the entropy production, so we find that the entropy production obeys a similar scaling:

$$\langle \Sigma'\rangle = \frac{\langle W^{CD\prime}\rangle}{T}$$

$$= \frac{\tau L'^2}{\tau' L^2} \langle \Sigma\rangle .$$

D. Efficient Protocols

When changing the protocol duration $\tau \to \tau'$ and system length $L \to L'$ of a desired distribution trajectory $\{\Pr(X_t^d)\}$, the counterdiabatic control becomes:

$$V'(x,t) = V^Q\left(\frac{L}{L'}x, \frac{\tau}{\tau'}t\right) + \frac{\tau L'^2}{\tau' L^2} V^{CD}\left(\frac{L}{L'}x, \frac{\tau}{\tau'}t\right) ,$$

where $V^Q(x,t)$ and $V^{CD}(x,t)$ are the original quasistatic and counterdiabatic potential energies. This leads to the work investment:

$$\langle W'\rangle = \Delta F^{\mathrm{eq}} + \frac{\tau L'^2}{\tau' L^2} \langle W^{CD}\rangle ,$$

where ΔF^{eq} is the original change in free energy and $\langle W^{CD}\rangle$ is the original nonequilibrium addition to work.

Since protocols are uniquely determined by the distribution trajectory, the above scaling relations apply directly to maximally efficient computations as well. Since a computation maps an initial equilibrium distribution $\Pr(X_0)$ to a final one $\Pr(X_\tau)$, there are many compatible distribution trajectories that evolve continuously from the initial to the final distribution. A minimally dissipative distribution trajectory $\Pr(X_{t,\min})$ has a corresponding $V_{\min}(x,t) = V_{\min}^Q(x,t) + V_{\min}^{CD}(x,t)$ that yields the

minimum work:

$$\langle W^{CD}\rangle_{\min} = \min\{\langle W^{CD}\rangle : \Pr(X_{0,\tau}^d) = \Pr(X_{0,\tau})\} .$$

Since quasistatic work is identical for all such protocols, up to an instantly recoverable additional energy, this condition also minimizes invested work.

Changing protocol duration $\tau \to \tau'$ and initial and final system length—viz., $\Pr(X_0' = x) = \Pr(X_0 = Lx/L')$ and $\Pr(X_{\tau'}' = x) = \Pr(X_\tau = Lx/L')$—we can determine how the minimally dissipative distribution trajectory changes, as well the minimum dissipation. A natural guess for the minimally dissipative trajectory is to take the scaled minimal distribution:

$$\Pr(X_t' = x) = \Pr(X_{\tau t/\tau',\min} = Lx/L')\frac{L}{L'} ,$$

which satisfies:

$$\langle W^{CD\prime}\rangle = \frac{\tau L'^2}{\tau' L^2} \langle W^{CD}\rangle_{\min} .$$

(See Sec. III C.)

If this proposed trajectory is not minimally dissipative, then there is another trajectory $\{\Pr(X_{t,\min}')\}$ that dissipates work $\langle W^{CD\prime}\rangle_{\min} < \langle W^{CD\prime}\rangle$. However, if this were the case, then we could reverse the duration and size scalings $\tau' \to \tau$ and $L' \to L$ on that trajectory to generate the dissipation:

$$\frac{\tau' L^2}{\tau L'^2}\langle W^{CD\prime}\rangle_{\min} < \frac{\tau' L^2}{\tau L'^2}\langle W^{CD\prime}\rangle$$

$$= \frac{\tau' L^2}{\tau L'^2}\frac{\tau L'^2}{\tau' L^2}\langle W^{CD}\rangle_{\min}$$

$$= \langle W^{CD}\rangle_{\min} .$$

This is a contradiction, since it states that it is possible to dissipate less than the minimal dissipation for the original computation that evolves the distribution between $\Pr(X_0)$ and $\Pr(X_\tau)$. We conclude that the spatially- and temporally-scaled minimally dissipative distribution trajectories are themselves minimally dissipative.

While this does not provide a method for discovering the minimally dissipative protocol, paralleling other approaches in the restricted linear regime [7, 26], it shows how to achieve maximal efficiency given other constraints on space and time, if one finds the minimally dissipative protocol in one setting. Moreover, it gives the temporal scaling of the minimally dissipative control protocol:

$$V_{\min}'(x,t) = V_{\min}^Q(Lx/L', \tau t/\tau')$$

$$+ \frac{\tau L'^2}{\tau' L^2} V_{\min}^{CD}(Lx/L', \tau t/\tau') ,$$

and of the minimum work production:

$$\langle W'\rangle_{\min} = \Delta F^{\mathrm{eq}} + \frac{\tau L'^2}{\tau' L^2}\langle W^{CD}\rangle_{\min} \ .$$

This matches independent analyses on the scaling of dissipated work for optimal control [7]. In contrast, as we see, the present results apply more generally: without restricting control parameters—all potential landscapes are allowed—and beyond linear response. This perhaps explains the puzzle that the results derived assuming linear response [6, 7] appeared to work outside of those constraints. Moreover and constructively, counterdiabatic protocols allow a control engineer to specify exact start and ending distributions. The latter must be inferred from dynamics in other treatments.

In short, counterdiabatic control of Fokker-Planck dynamics in one dimension gives precise control over distributions and yields constructive methods for designing control protocols. The resulting energetic costs depend simply on overall system temporal and spatial scales, revealing a tradeoff beyond that between a computation's information processing and energy cost.

IV. COUNTERDIABATIC ERASURE

We now apply the counterdiabatic approach to the paradigmatic example of erasing a bit of information in a metastable system. The analysis exposes new elements in the resource tradeoffs that arise in thermodynamic computing.

A. Nonequilibrium Information Storage

The ability to quickly shift probability distributions in one-dimensional nonlinear Langevin systems gives a physical way to implement fast logical operations. For instance, erasure is an essential part of most computations and can be achieved by controlling a double-well potential landscape [27, 28]. Landauer stated that erasure requires dissipating at least $k_{\mathrm{B}} T \ln 2$ of work—a cost deriving from the microstate space contraction induced by the logically irreversible operation [1]. This bound is indeed achievable in the present setting, but only in the limit of quasistatic operations, where zero entropy is produced globally. That is, it is achievable only in infinite time. For finite-time processes, there is additional dissipation and, thus, additional work required for erasing a bit of information [7, 27, 28]. We will now derive the same additional cost for finite-time erasure but, departing from prior treatments, we determine the initial and final distributions and thus exactly specify the fidelity of

erasure, instead of merely recreating it. This provides a detailed analysis of thermodynamic resources for a given accuracy level of information processing. Additionally, we will design a protocol that gives perfect erasure in finite time and at finite cost.

The expression for the counterdiabatic potential Eq. (10) specifies how to design a protocol $V(x,t)$ that maintains the distribution $\Pr(X_t)$ exactly in a desired distribution $\Pr(X_t^d)$ over the interval $t \in (0, \tau)$. However, we must also consider how to use the map to informational states $c : \mathcal{X} \to \mathcal{Y}$ to perform useful and robust computation. One strategy is to design the energy landscape such that physical states $x \in \mathcal{X}$ in one informational state $y \in \mathcal{Y}$ rarely transition to different informational states $y' \neq y$. This allows the information processing device to remain in a passive "default" state while retaining the information of its computation for long times, regardless of the outcome. In contrast, if a computation is designed such that the equilibrium distribution $\Pr(X_t^{\mathrm{eq}})$ exactly matches the desired distribution $\Pr(X_t^d)$ after the computation, with $t > \tau$, then the energy landscape is given by $V(x,t) = F^{\mathrm{eq}}(t) - k_{\mathrm{B}} T \ln \Pr(X_t^d)$ for $t \geq \tau$. The potential energy characterizes the external configuration of our memory storage device, and it is through control of this external configuration that $V(x,t)$ has time dependence.

Thus, the relevant information about the computation is stored in the external configuration of the memory device $V(x,t)$ (our control), rather than the actual physical distribution $\Pr(X_t)$. By choosing a default energy landscape that stores metastable physical distributions, the computational device can robustly store information without explicitly encoding the outcome distribution of the computation in the energy landscape and thus the external configuration. With metastable information, a memory device can store information about its input that is not explicitly encoded in the control parameters.

To experimentally test Landauer's prediction [1], Ref. [28] employed a protocol that starts and ends in a symmetric double-well potential, where each well is interpreted as a distinct mesoscopic informational state: $Y_t = 0$ or $Y_t = 1$. Such a potential stores informational states determined by the probability $\Pr(Y_t = 0)$ to realize the informational state 0. Following this setup, if we start and end in a symmetric bistable potential:

$$V(x,0) = V(x,\tau) = \alpha x^4 - \beta x^2 \ , \qquad (14)$$

then, at a temperature T, the equilibrium distribution:

$$p(x) \equiv \Pr(X_{\{0,\tau\}}^{\mathrm{eq}} = x) = \frac{e^{-V(x,0)/k_{\mathrm{B}} T}}{Z} \ , \qquad (15)$$

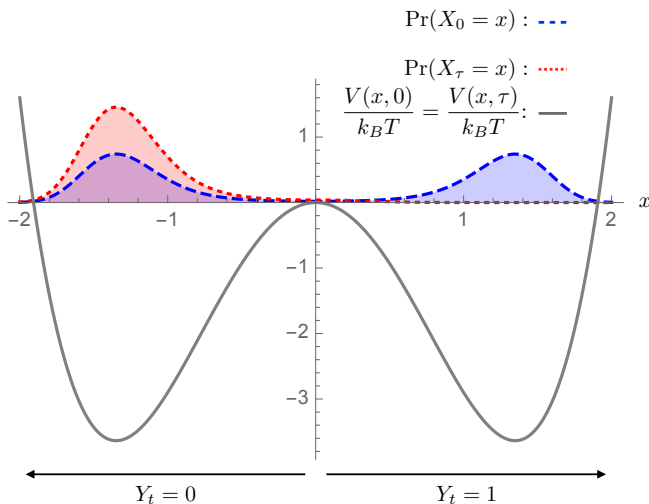is bimodal; see the blue dashed curve in Fig. 2. While

FIG. 2. Default energy landscape: A double-well potential that stores many different distributions $\Pr(Y_t)$ over the informational states $Y_t \in \{0, 1\}$. For protocols that process the information in the distributions over the times $t \in (0, \tau)$, the energy landscape is set to be the same at the beginning and end, shown by the gray curve $V(x, 0) = V(x, \tau)$. The equilibrium distribution, delineated by the dashed blue curve, gives equal probabilities for the informational states: $\Pr(Y_0 = 0) = \Pr(Y_0 = 1) = 1/2$. This is the initial distribution for the system $\Pr(X_0)$ in this case. It stores $\mathrm{H}[\Pr(Y_0)] = 1$ bit of information, where $\mathrm{H}[Z]$ is the Shannon information of random variable Z [29]. The red curve $\Pr(X_\tau = x)$ gives unit probability of informational state 0 ($\Pr(Y_\tau = 0) = 1$) and is the distribution of the system after an effective erasure protocol. Its Shannon information vanishes and so the initial and final distributions represent bit erasure. The final distribution $\Pr(X_\tau)$ is out of equilibrium, but the energy barrier between the two informational states keeps it nearly fixed for short times. This distribution, as well as many other nonequilibrium distributions, are metastable, and will eventually, slowly relax to equilibrium.

this distribution is exactly stationary (when the potential is held fixed), we can construct other distributions that are (temporarily) effectively stationary, such as that given by the dotted red curve shown in Fig. 2. This has the same shape as the equilibrium distribution over the interval $(-\infty, 0)$, but vanishes outside. By specifying a time-dependent *bit bias* probability $\Pr(Y_t = 0) = b(t)$, we fully specify a metastable physical distribution [30]:

$$\Pr(X_t^{\mathrm{met}} = x) = \begin{cases} p(x) \cdot 2b(t) & \text{if } x \leq 0 \\ p(x) \cdot 2(1 - b(t)) & \text{if } x > 0 \end{cases} . \quad (16)$$

We take this distribution to be our desired distribution $\Pr(X_t^d)$, which in turn defines the quasistatic potential $V^Q(x, t)$.

Figure 2 shows the metastable distributions before (blue dashed curve) and after (red dotted curve) an erasure protocol, where the initial distribution is unbiased

$b(0) = 1/2$, and the final distribution is totally biased $b(\tau) = 1$. The energy barrier between informational states 0 and 1 inhibits large probability flow between the two local equilibria. That is, these distributions relax to a global equilibrium very slowly, depending on barrier height relative to $k_B T$ [30]. Thus, these metastable states robustly store nonequilibrium informational states and provide a basis for information processing by a controlled double-well potential.

B. Counterdiabatic Information Processing

We now consider how to use counterdiabatic driving to steer the system through a sequence of metastable states specified by a given time-dependent bit bias $b(t)$, with the symmetric initial and final energy landscape of Eq. (14). Despite the symmetric initial and final configurations of the memory device, this modified counterdiabatic control allows for $b(0) \neq 1/2$ and $b(\tau) \neq 1/2$.

Since the initial and final metastable states are out of equilibrium with respect to the symmetric potential $V(x, 0) = V(x, \tau) = \alpha x^4 - \beta x^2$ (Fig. 2), we must modify the counterdiabatic protocol described in Sec. III B, which was developed for transitions between initial and final equilibrium distributions. Two additional steps are needed, each a quench, as shown in Fig. 3. (*Quench* here means a nearly instantaneous change in the Hamiltonian [31], as opposed to a nearly instantaneous change in temperature, as often intended.) These quenches are added to make the quasistatic potential $V^Q(x, t)$ match the equilibrium distribution of the desired metastable distribution $\Pr(X_t^{\mathrm{met}})$ over the open time interval $t \in (0, \tau)$.

Specifically, for $t \in (0, \tau)$ we set:

$$V^Q(x, t) = F^{\mathrm{eq}}(t) - k_B T \ln \Pr(X_t^{\mathrm{met}} = x) .$$

Hence, at $t = 0$ the energy landscape undergoes a quench from the symmetric potential $V(x, 0)$ to the asymmetric potential $V^Q(x, 0)$. We then add the counterdiabatic term:

$$V^{CD}(x, t) = \frac{1}{\mu} \int_0^x \int_{-\infty}^{x'} \frac{\partial_t \Pr(X_t^{\mathrm{met}} = x'')}{\Pr(X_t^{\mathrm{met}} = x')} dx'' dx',$$

such that the overall potential becomes:

$$V(x, t) = V^Q(x, t) + V^{CD}(x, t) .$$

For $t \in (0, \tau)$ the system evolves through the desired sequence $\Pr(X_t^{\mathrm{met}} = x)$, corresponding to the equilibrium states of $V^Q(x, t)$. At the end of the protocol the system undergoes another quench, from the asymmetric potential $V^Q(x, \tau^-)$ to the symmetric potential $V(x, \tau)$.
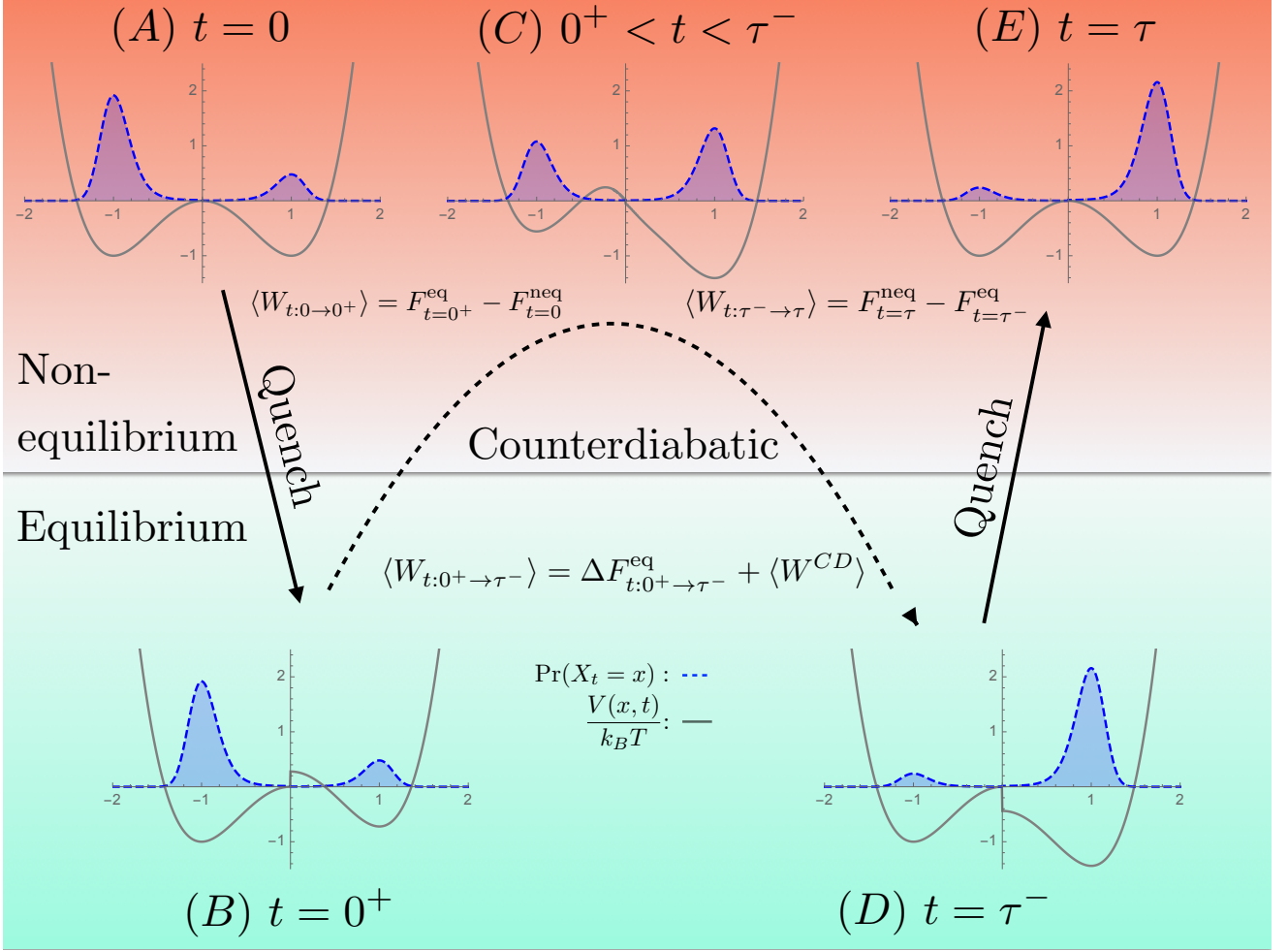
FIG. 3. Counterdiabatic information processing in three steps: Distribution $\Pr(X_t = x)$ shown with blue dashed curves and energy landscapes $V(x, t)$ shown by the gray curves. First, the information landscape is instantaneously changed to fit the starting distribution from stage (A) to stage (B). These share the same distribution but have different energy landscapes. Second, the counterdiabatic protocol is applied to take the system from the equilibrium distribution at stage (B) to that shown in stage (D), passing through nonequilibrium distributions driven from equilibrium by the counterdiabatic potential; such as that shown in stage (C). Third, the last quench step instantaneously takes the system from equilibrium stage (D) to the nonequilibrium metastable stage (E). All three transitions are labeled with the associated work investment.

In this way, we drive the system through a sequence of metastable distributions with precise control of the bit bias $b(t)$.

Although the protocol just described pertains to the specific case of a double well, the procedure of quenching, controlling counterdiabatically, and then quenching again is a general technique for evolving between nonequilibrium distributions in finite time. For such a computation, the total work simplifies to the net change in nonequilibrium free energy plus the counterdiabatic work:

$$\langle W \rangle = \Delta F^{\mathrm{neq}} + \langle W^{CD} \rangle \,, \qquad (17)$$

as shown in Fig. 3. The change in nonequilibrium free energy is given by the sum of the quasistatic work and the

quenching work [2, 24]. For the metastable distributions we discussed, where each informational state contributes the same local free energy, ΔF^{neq} reduces to the change in the Shannon entropy of the information variable [2]:

$$\Delta F^{\mathrm{neq}} = k_{\mathrm{B}} T \ln 2 \left(\mathrm{H}[Y_0] - \mathrm{H}[Y_\tau] \right) \,. \qquad (18)$$

Since $\langle W^{CD} \rangle = T \langle \Sigma \rangle \geq 0$ (see Sec. III B), Eqs. (17) and (18) lead immediately to a generalized form of Landauer's bound:

$$\langle W \rangle \geq k_{\mathrm{B}} T \ln 2 \left(\mathrm{H}[Y_0] - \mathrm{H}[Y_\tau] \right) \,, \qquad (19)$$

which takes on the familiar form, $\langle W \rangle \geq k_{\mathrm{B}} T \ln 2$, when starting with fully randomized bits, $b(0) = 1/2$ and when

the operation's fidelity is perfect, $b(\tau) = 1$. As we shall see, while the Landauer bound cannot be achieved in finite time, perfect fidelity can be achieved in finite time with finite work.

The amount of entropy produced, $\langle \Sigma \rangle = \langle W^{CD} \rangle / T$, reflects the additional cost beyond Landauer's bound of implementing a computation in finite time. For metastable erasure in a symmetric double well, this additional cost depends on duration, system length scale, bit bias difference, and information lifetime—a measure of information storage robustness. We have already seen that the value of $\langle \Sigma \rangle$ scales as the inverse of the protocol duration τ and the square of the system characteristic length scale L (Sec. III C). We now show how bit bias difference and information lifetime lead to additional energy costs.

Metastability simplifies the expression for the counterdiabatic potential, leading to simple relationships between the work, bit bias difference, and robustness of information storage. As shown in Appendix B, the counterdiabatic potential can be expressed as a product of a piecewise-continuous function and a function that depends only on the equilibrium distribution:

$$V^{CD}(x,t) = h(x) \times \begin{cases} -\partial_t \ln b(t) & \text{if } x \leq 0 \\ -\partial_t \ln(1 - b(t)) & \text{if } x > 0 \end{cases}, \quad (20)$$

where:

$$h(x) = \frac{1}{\mu} \int_0^{|x|} dx' \frac{1}{p(x')} \int_{-\infty}^{-|x'|} dx'' p(x'')$$

and $p(x) = \Pr(X_0^{\text{eq}} = x)$ is the equilibrium distribution for the symmetric, bistable potential of Eq. (15). This result allows us to design protocols for evolving a metastable distribution from an initial bit bias $b(0) = b_i$ to any final bit bias $b(\tau) = b_f$, given a bistable potential. For instance, the choices $b_i = 1/2$ and $b_f = 1$ correspond to perfect erasure, where the system ends entirely in the left well.

C. Tradeoffs in Metastable Symmetric Erasure

As discussed above, the equilibrium distribution $p(x)$ and bit bias $b(t)$ determine both the desired metastable distribution trajectory of Eq. (16) and the counterdiabatic potential of Eq. (20) that generates this evolution. In Appendix B we show that the functions $p(x)$ and $b(t)$ are multiplicatively separable in the expression for counterdiabatic work. Specifically:

$$\langle W^{CD} \rangle = f_1[p(\cdot)] f_2[b(\cdot)] ,$$

where:

$$f_1[p(\cdot)] = \frac{1}{\mu} \int_0^\infty dx\, p(x) \int_0^x dx' \frac{1}{p(x')} \int_{-\infty}^{-x'} dx'' p(x'')$$

$$f_2[b(\cdot)] = \int_0^1 dt\, \frac{(\partial_t b(t))^2}{b(t) - b(t)^2} . \quad (21)$$

This separability entails additional tradeoffs between dissipation, bit bias difference, and information lifetime.

Functional f_1 depends on the equilibrium distribution $p(x)$ which in turn is determined by the bistable potential $V(x,0)$. The shape of this potential (e.g., the height of the barrier relative to the left and right minima) determines the expected "lifetime" of a stored bit, in the absence of external driving. Thus, f_1 contains all the dependence of the counterdiabatic work on information storage robustness.

Functional f_2 depends on the bit bias trajectory $b(t)$. One can now entertain a variety of bias trajectories, using this functional to determine how the counterdiabatic work changes. However, note that one must restrict to paths for which the initial and final time-derivative vanishes $\partial_t b(t)|_{t \in \{0, \tau\}} = 0$, since $\partial_t \Pr(X_t = x)_{t \in \{0, \tau\}}$ must vanish for the counterdiabatic potential itself to be zero initially and finally.

Note too that f_1 and f_2 contain the system length and protocol duration dependence, respectively. If we rescale the system spatially and the protocol temporally, we obtain the new equilibrium distribution and bias trajectory:

$$p'(x) = \frac{L}{L'} p(Lx/L') \text{ and}$$
$$b'(t) = b(\tau t / \tau') .$$

Plugging these in, we find the new functionals:

$$f^1[p'(\cdot)] = \frac{L'^2}{L^2} f^1[p(\cdot)] \text{ and}$$
$$f^2[b'(\cdot)] = \frac{\tau}{\tau'} f^2[b(\cdot)] .$$

To further separate dependencies, we introduce unitless functionals of both bias and the default equilibrium distribution:

$$F_1[p(\cdot)] = f_1[p(\cdot)]/L^2 \text{ and}$$
$$F_2[b(\cdot)] = f_2[b(\cdot)]\tau .$$

F_2 captures the difference between initial and final bias without dependence on computation rate. F_1 captures the depth between left and right wells without dependence on the spatial scale, which also affects how long a bistable system can robustly store information. In short, the counterdiabatic work is expressed as the product of
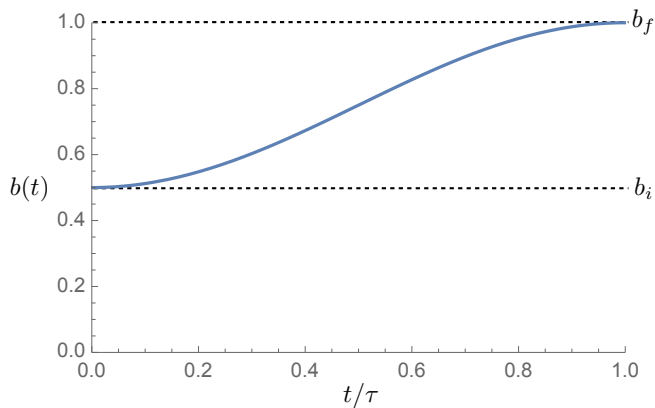
FIG. 4. Nonlinear protocol for driving between initial bit bias $b(0) = b_i$ and final bias $b(\tau) = b_f$. The nonlinear protocol $b(t)$ (blue curve) has zero slope initially and finally such that the counterdiabatic potential vanishes at the protocol's beginning and end.
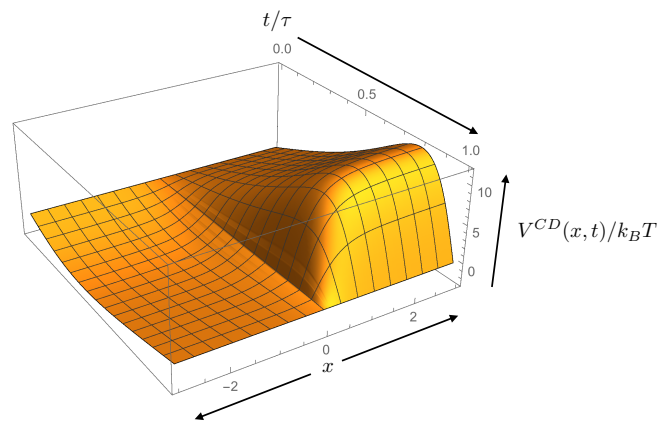

FIG. 5. Counterdiabatic potential for the nonlinear erasure protocol of Fig. 4 which takes a bistable potential well from an initial bias $b(0) = 0.5$ to a final bias $b(\tau) = 0.95$. For this protocol we set $\tau = 1$, $\mu = 1$, $k_B T = 1$, $\alpha = 1$, and $\beta = 2$. The counterdiabatic potential vanishes at the beginning and end, so that the system begins and ends in equilibrium.

four terms:

$$\langle W^{CD} \rangle = \frac{L^2}{\tau} F_1[p(\cdot)]F_2[b(\cdot)] .$$

Since F_1 and F_2 are independent of duration and system length, this implies a five-way tradeoff between the main dependencies of computation: dissipation, duration, length, F_1, and F_2. We next study how F_1 and F_2 depend on properties of the erasure protocol, leading to a practical consequence.

D. Perfect Erasure in Finite Time with Finite Work

Let us consider control protocols for which the bit bias trajectory is given by:

$$b(t) = b_i \cos^2(\pi t/2\tau) + b_f \sin^2(\pi t/2\tau) . \qquad (22)$$

This schedule takes the system monotonically from $b(0) = b_i$ to $b(\tau) = b_f$, as shown in Fig. 4. Since $\partial_t b = 0$ at $t = 0$ and $t = \tau$, the counterdiabatic potential vanishes at the initial and final times, except in the special cases that b_i or b_f are either 0 or 1.

The counterdiabatic potential in this case is:

$$V^{CD}(x,t) = \frac{h(x)}{2\tau} \qquad (23)$$

$$\times \begin{cases} -\frac{(b_f - b_i)\pi \sin(\pi t/\tau)}{b_i \cos(\pi t/2\tau)^2 + b_f \sin(\pi t/2\tau)^2} & \text{if } x \leq 0 \\ \frac{(b_f - b_i)\pi \sin(\pi t/\tau)}{1 - b_i \cos(\pi t/2\tau)^2 - b_f \sin(\pi t/2\tau)^2} & \text{if } x > 0 \end{cases} .$$

Note that the explicit dependence on duration factors out, yielding the prefactor τ^{-1}, as expected. Calculating

$h(x)$ numerically, Fig. 5 plots the counterdiabatic potential $V^{CD}(x,t)$. The nonlinear protocol begins and ends with zero counterdiabatic potential, hence the distribution begins and ends in equilibrium. This guarantees that when instantaneously changing back to the default bistable potential landscape, the work investment beyond the counterdiabatic work equals the change in nonequilibrium free energy.

Equation (20) indicates that any protocol ending with all probability in a single well, such that $b_f = 0$ or $b_f = 1$, has divergent counterdiabatic potentials, since either $b(t)$ or $1 - b(t)$ vanishes. A vanishing numerator $\dot{b}(t)$ is no compensation, since under any number of applications of L'Hopital's rule to evaluate convergence the numerator becomes nonzero first; it is the derivative of the denominator. Despite this, through numerical calculations, we find that a counterdiabatic potential $V^{CD}(x,t)$ that starts and ends at zero can perform perfect erasure in finite time with finite work. If the thermodynamic-computing designer wishes to avoid a divergent final potential, they can approach perfectly faithful erasure asymptotically while keeping the final state in equilibrium, because $V^{CD}(x,\tau) = 0$ for all $b_f \neq 0, 1$. As the final bias b_f approaches 1, the resulting work approaches a constant value but the system approaches perfect erasure, as shown in the rightmost plot of Fig. 6.

To study the dependence of dissipated work on initial and final bias, b_i and b_f, we substitute the nonlinear bias function, Eq. (22), into functional $F_2[b(\cdot)]$. This functional is proportional the dissipated work with the default distribution $p(x)$ and duration fixed. As Fig. 6 shows, starting with an unbiased state $b_i = 0.5$, then increasing the final bias towards $b_f = 1.0$, the required
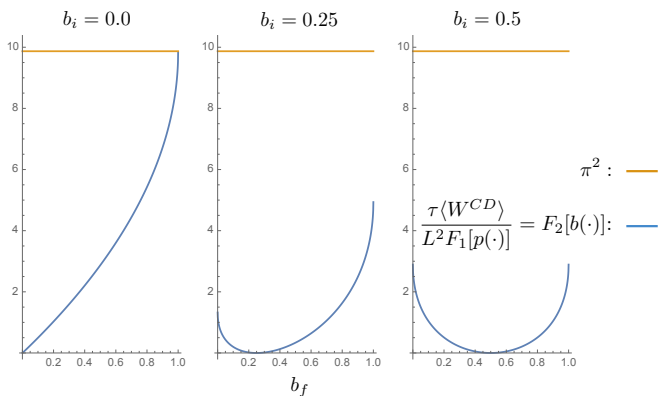
FIG. 6. Dissipated work to execute a logical operation changes with initial bit bias b_i and final bias b_f: Dissipated work is proportional to $F_2[b(\cdot)]$ when the duration τ and equilibrium distribution $p(\cdot)$ are held fixed. (Left) Initial bit bias $b_i = 0.0$: as the probability b_f of informational state 0 increases the cost of erasure increases steadily to a maximum at $b_f = 1.0$. (Center) Similar behavior for an initial bias $b_i = 0.25$. (Right) Fair initial distribution $b_i = 0.5$, demonstrating a finite cost for perfect erasure ($b_f = 1.0$) in finite time.

work increases, but not indefinitely. $F^2[b(\cdot)]$ converges to approximately 0.293, meaning that the protocol can perform perfect erasure in finite time with finite work.

More generally, Fig. 6 shows for three different initial biases $b_i = 0.0$, 0.25, and 0.5 that the dissipation increases with the magnitude of the bias difference $|b_i - b_f|$. However, the contribution of the bias difference is bounded by the case where the bias difference is maximal $|b_i - b_f| = 1$, for which $F_2[b(\cdot)] = \pi^2$.

Note that these plots are intentionally designed in a way similar to Fig. 3 of Ref. [7] and reveal similar dependence on initial and final bias. Quantitatively, the values are proportional. Reference [7] showed that optimal control in the linear response regime requires dissipated heat proportional to twice the square of the Hellinger distance:

$$K(b_i, b_f) \equiv \left(\sqrt{b_i} - \sqrt{b_f}\right)^2 + \left(\sqrt{1-b_i} - \sqrt{1-b_f}\right)^2 .$$

Though our chosen bit bias trajectory is not optimal, as discussed in Appendix B, numerical integration shows that the contribution to the dissipated work $F_2[b(\cdot)] \propto K(b_i, b_f)$. While we do not yet have an explanation of the correspondence between these two different estimates of dissipated work in finite time, the results' similarity is suggestive. We should point out, though, that other bias trajectories could be chosen that are not proportional to the Hellinger distance and may be less dissipative. It may be a coincidence that our chosen bit bias trajectory yielded results similar to the linear response analysis of Ref. [7].

E. Robust Information Storage Requires Work

With a potential $V(x, 0) = \alpha x^4 - \beta x^2$ that stores information in metastable distributional states, there is a finite lifetime of that information. In this symmetric double well, with one well corresponding to informational state 0 and the other to 1, the lifetime can be quantified in terms of the average time $\langle \tau_{0 \to 1} \rangle$ it takes for a particle to switch between these states. In the overdamped regime this *information lifetime* is given by [30, 32]:

$$\langle \tau_{0 \to 1} \rangle = \frac{2\pi}{\mu \sqrt{|\ddot{V}(x_0, 0) \ddot{V}(x_B, 0)|}} e^{\Delta E_B / k_{\rm B} T} , \quad (24)$$

where by $\ddot{V}(x, 0) = \partial_x^2 V(x, 0)$ we denote the curvature of the default potential energy landscape, $x_0 = -\sqrt{\beta/2\alpha}$ is the location of the minimum in the metastable 0 well, $x_B = 0$ is the location of the barrier maximum, and $\Delta E_B = V(x_B, 0) - V(x_0, 0)$ is the height of the barrier above the minimum. The latter is a useful measure of the barrier's energy scale and, thus, how robustly the potential stores information. By explicit calculation we obtain $\ddot{V}(x_B, 0) = -2\beta$, $\ddot{V}(x_0, 0) = 4\beta$, and:

$$\Delta E_B = \frac{\beta^2}{4\alpha} . \quad (25)$$

Hence, the information lifetime is:

$$\langle \tau_{0 \to 1} \rangle = \frac{\pi}{\mu \beta \sqrt{2}} e^{\beta^2 / 4\alpha k_{\rm B} T} . \quad (26)$$

Note that $\langle \tau_{0 \to 1} \rangle$ scales as the system length L squared, due to the $\ddot{V}$ terms in Eq. (24)'s denominator. Beyond this, the information lifetime is more strongly controlled by the energy scale of the energy landscape, which can be characterized by ΔE_B. It has nearly exponential dependence on this energy scale:

$$\frac{\langle \tau_{0 \to 1} \rangle}{L^2} \propto \frac{e^{\Delta E_B / k_{\rm B} T}}{\Delta E_B} .$$

Thus, we can capture this dependence by evaluating the information lifetime and scaling by the length. Comparing $f_1[p(\cdot)]$ to $\langle \tau_{0 \to 1} \rangle$ (i.e. $F_1[p(\cdot)] = f_1[p(\cdot)]/L^2$ to $\langle \tau_{0 \to 1} \rangle/L^2$) reveals an interesting correspondence between dissipation and information lifetime, as well as identifying a term that depends on the default potential's well depth.

As illustrated in Fig. 7, with increasing well depth ΔE_B the bistable distribution becomes increasingly peaked at the local minima, and the information lifetime increases nearly exponentially; as predicted by Eq. (24). Interestingly, $f_1[p(\cdot)]$, which is proportional to the ex-
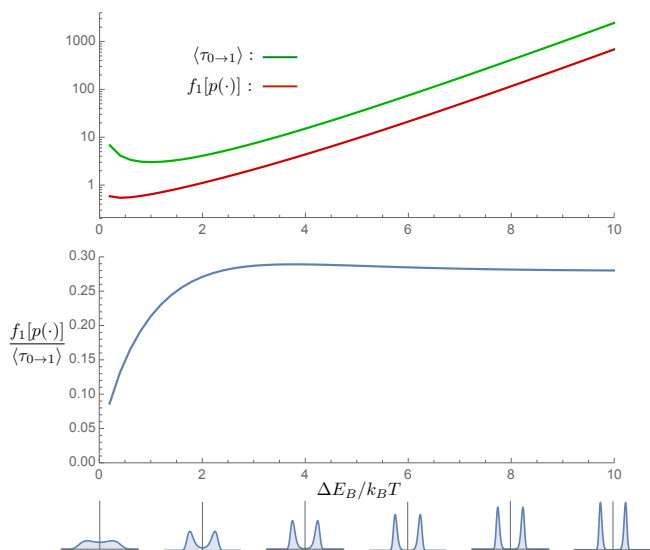
FIG. 7. Energy barrier dependence: (Top) Changing energy barrier height ΔE_B relative to the thermal energy scale $k_B T$, $f_1[p(\cdot)]$ and so the required dissipated work increase nearly exponentially. This corresponds to an increase in the separation between the distribution in the 0 and 1 states, as shown by the six distributions along the horizontal axis (Bottom). Increasing ΔE_B, greater well separation, leads to more robust information storage, as shown by the information lifetime $\langle \tau_{0 \to 1} \rangle$ (Top). (Center) Moreover, the information lifetime, which scales just below exponentially, appears to scale at the same rate as the dissipated energy when the barrier is at least twice $k_B T$. $k_B T = 1$, $\alpha = 1$, and $\mu = 1$ for these calculations, while β is used to change the energy barrier ΔE_B as in Eq. (25).

cess work production required during erasure, scales at roughly the same rate as the information lifetime, if the length scale is held fixed. As shown in Fig. 7, the work required to erase increases nearly exponentially with the height of the energy barrier between the wells and stays nearly proportional to the information lifetime. The exception to this occurs for very small barrier heights, where the potential's equilibrium distribution is not clearly bimodal and there is nearly unobstructed flow between the information states.

Thus, for finite-time erasure there is a clear energy cost to robust information storage, which is multiplicatively separable from both the bias difference, as well as protocol total duration τ. Though we described how the excess work scales as the square of the length scale L, this dependency is directly contained in the functional $f_1[p(\cdot)]$. This reinforces the relationship to the information lifetime $\langle \tau_{0 \to 1} \rangle$, which also scales as the inverse length scale, due to the term $\sqrt{|\ddot{V}(x_0, 0)\ddot{V}(x_B, 0)|}$. However, we also see the direct effect of the energy-barrier height ΔE_B through $F_1[p(\cdot)]$ and its near proportionality to $\langle \tau_{0 \to 1} \rangle / L^2$.

V. CONCLUSION

Counterdiabatic control is a new tool for thermodynamic computing that executes precisely-controlled information processing in finite time at finite cost. It is implemented via an additional term in the potential energy—the counterdiabatic potential—which guides the microstate distribution along a path that results in the desired computation. We analyzed the work required for counterdiabatic information processing, developing a full suite of resource trade-offs. Since, as we showed, counterdiabatic protocols are the unique control that guides the system distribution along a desired trajectory, these trade-offs apply broadly to any Hamiltonian control in overdamped Fokker-Planck dynamics in one dimension. Other than the expected technical complications, the overall control strategy will generalize to higher-dimensional state spaces.

We described how to deploy counterdiabatic protocols in combination with quenching as a general strategy for finite-time metastable information processing. Since counterdiabatic control exactly specifies the system's final distribution, it is distinct from previous treatments of finite-time information processing using geometric control, which focused on driving an external (thermodynamic) parameter to a desired value with minimal work.

We showed that the work performed during a counterdiabatic protocol separates into the change in equilibrium free energy ΔF^{eq} and the counterdiabatic work $\langle W^{CD} \rangle$, which is also the dissipated work and, thus, proportional to the protocol's entropy production. We showed that $\langle W^{CD} \rangle$ scales as the inverse of the protocol duration τ—reinforcing previous analyses of finite-time thermodynamic processes that showed the work required for information processing increases with computation rate [6, 7, 26]. We also showed that dissipated work scales as the square of the system length scale L. That is, it is more difficult to move distributions long distances in the same finite time. The time and distance dependence together imply that going twice as far at the same speed takes twice the energy, rather similar to locomotive machines.

We then showed that counterdiabatic protocols can process information by adding quenching at a protocol's beginning and end. Quenching allows rapidly evolving between nonequilibrium metastable states, which store information. Applying the approach, we considered a symmetric double-well system and calculated the work production for various types of finite-time bit manipulation. This analysis demonstrated that, in addition to the dependence on length scale and duration, counterdiabatic work depends on erasure fidelity and information storage robustness.

Evaluating the multiplicative component $F_2[b(\cdot)]$ of the counterdiabatic work, we found that the dissipation increases with the bit bias difference between the initial and final distributions. More specifically, it is proportional to the Hellinger distance for our chosen class of control protocols. Given an initial equilibrium and unbiased metastable distribution, the closer the final metastable distribution is to giving all-0 informational states—increased erasure fidelity—the more the operation costs. However, there is an upper bound on the dissipated work. Thus, it is possible to perform perfect erasure in finite time at finite cost. It is also possible to flip a bit in finite time with finite work, as shown in Fig. 7's leftmost plot. Perfect fidelity, though, does not mean results are held indefinitely.

The factor $f_1[p(\cdot)]$ in the expression for the counterdiabatic work depends only on the default equilibrium distribution and so it captures the dependence on information storage robustness. That is, with increased well depth—and so metastable state robustness—the dissipated work increases nearly exponentially. Numerical calculations demonstrate that work scales at the same rate as the information lifetime, which is the Kramers estimate [30] of the average time it takes to jump between wells.

A much richer and more detailed picture of resource tradeoffs in thermodynamic computing emerges. Most concisely, the required work decomposes as follows:

$$\langle W \rangle = k_{\mathrm{B}} T \ln 2 (\mathrm{H}[Y_0] - \mathrm{H}[Y_\tau]) + \frac{L^2}{\tau} F_1[p(\cdot)] F_2[b(\cdot)] \ .$$

Landauer's Principle for thermodynamic computing, the first term on the right, is the work required to reversibly implement a change in metastably-stored information; it is equal to the change in the physical processor's nonequilibrium free energy. Counterdiabatic protocols complement and extend this principle. They reveal, in the second term on the right, an additional cost in the form of dissipated work that depends on duration τ, length scale L, bias difference through $F[b(\cdot)]$, energy scale of information storage ΔE_B through $F_1[p(\cdot)]$, and information lifetime through the product $L^2 F_1[p(\cdot)]$. The result is a rather more complete picture of finite-time, accurate thermodynamic computing.

Appendix A: Uniqueness of Counterdiabatic Protocols

Typically, via the Perron-Frobenius operator, the equations of motion over a space $\mathcal{X}$ are used to evolve the distribution $\Pr(X_t)$ over states $x \in \mathcal{X}$ for a time interval $t \in (0, \tau)$ from an initial distribution $\Pr(X_0)$. The inverse problem, of determining the equations of motion from the evolution of states, is more challenging. For overdamped Fokker-Planck dynamics, Ref. [11] shows how to determine the counterdiabatic control protocol $V(x,t) = V^Q(x,t) + V^{CD}(x,t)$ directly from the desired evolution of $\Pr(X_t^d)$ and, hence, determine the equations of motion. The equations of motion are specified by a changing potential landscape $V(x,t)$. However, while the counterdiabatic potential is a solution to the inverse problem, given distribution trajectory $\{\Pr(X_t^d)\}$, such solutions a priori need not be unique. Here, we show that the counterdiabatic protocol is the unique protocol that generates the distribution trajectory.

We start by assuming that $V(x,t)$ induces the evolution of $\Pr(X_t)$ over the time interval $(0, \tau)$. This means that it solves the Fokker-Planck equation:

$$\frac{\partial \Pr(X_t^d = x)}{\partial t} = \mu \frac{\partial}{\partial x}\left(\Pr(X_t^d = x)\frac{\partial V(x,t)}{\partial x}\right) + \mu k_{\mathrm{B}} T \frac{\partial^2 \Pr(X_t^d = x)}{\partial x^2} \ . \quad \text{(A1)}$$

If the potential is not the unique dynamic solving this equation, then there exists potential energy landscape:

$$V'(x,t) = V(x,t) + \Delta V(x,t) \ ,$$

that also solves this equation with nonzero $\Delta V(x,t)$. That is:

$$\frac{\partial \Pr(X_t = x)}{\partial t} = \mu \frac{\partial}{\partial x}\left(\Pr(X_t^d = x)\frac{\partial V(x,t) + \Delta V(x,t)}{\partial x}\right) + \mu k_{\mathrm{B}} T \frac{\partial^2 \Pr(X_t^d = x)}{\partial x^2} \ .$$

Subtracting Eq. (A1) gives:

$$0 = \mu \frac{\partial}{\partial x}\left(\Pr(X_t^d = x)\frac{\partial \Delta V(x,t)}{\partial x}\right) \ .$$

Solving for the difference between the two possible solu-

tions leads to the conclusion that all possible solutions for the difference have the form:

$$\Delta V(x,t) = C(t) + K(t) \int_0^x \frac{dx'}{\Pr(X_t^d = x')} ,$$

where $C(t)$ and $K(t)$ can vary with time, but are constant in the positional variable x. $C(t)$ is an expected and trivial additional component: one can add an additional flat potential to any protocol without physical consequence beyond the change in total potential between the start and end: $C(\tau) - C(0)$. However, $K(t)$ corresponds to an additional force, which can have meaningful affect on the work invested during a control protocol. Thus, it appears that there are multiple ways to solve for potential energy underlying the state dynamics. However, if the state space $\mathcal{X}$ is truly an unbounded spatial degree of freedom, topologically equivalent to the real numbers $\mathbb{R}$, then the additional solutions corresponding to nonzero $K(t)$ are unphysical.

Framed another way, these additional components in possible alternative solutions correspond to addition to the force field $F'(x,t) = F(x,t) + \Delta F(x,t)$, where:

$$\Delta F(x,t) = -\frac{\partial \Delta V(x,t)}{\partial x}$$
$$= -\frac{K(t)}{\Pr(X_t^d = x)} ,$$

and the force is defined $F(x,t) \equiv -\partial_x V(x,t)$. While the strength of this force field varies spatially, its sign is the same for all x at a given time, meaning that the forces at every point are aligned in the same direction. This additional force corresponds to an addition to the drift velocity $v'_{\text{drift}}(x,t) = v_{\text{drift}}(x,t) + \Delta v_{\text{drift}}(x,t)$, given by:

$$\Delta v_{\text{drift}}(x,t) = \mu \Delta F(x,t) ,$$

where the drift velocity is $v_{\text{drift}} \equiv \mu F(x,t)$. Finally, this adds to the probability current $J'(x,t) = J(x,t) + \Delta J(x,t)$. This turns out to be constant over position:

$$\Delta J(x,t) = \Pr(X_t^d = x)\Delta v_{\text{drift}}(x,t)$$
$$= -\mu K(t) . \tag{A2}$$

This constant probability current cannot be realized in an infinite positional variable since, despite locally preserving the probability distribution, probability flows out at one extreme end of the spatial degree of freedom.

To explicitly prove that an additional constant probability current is impossible in positional space $\mathcal{X}$ topologically conjugate to the real line $\mathbb{R}$, note that the Fokker-Planck equation Eq. (A1) is the continuity equation $\partial_t \Pr(X_t = x) = -\partial_x J(x,t)$. There is an integral form of

this equation, which relates the change in probability in a region $[x_0, x_1]$ to the probability current through the boundary of the region:

$$\partial_t \int_{x_0}^{x_1} dx \Pr(X_t^d = x) = J(x_0,t) - J(x_1,t) .$$

In order for $J'(x,t)$ to satisfy the Fokker-Planck equation, it must also satisfy $\partial_t \int_{x_0}^{x_1} dx \Pr(X_t^d = x) = J'(x_0,t) - J'(x_1,t)$. So far, there is no contradiction, since:

$$J'(x_0,t) - J'(x_1,t) = J(x_0,t) - \mu K(t) - J(x_1,t) + \mu K(t)$$
$$= J(x_0,t) - J(x_1,t).$$

However, in the special case with $x_0 = -\infty$—the region of interest is all $x \leq x_1$—then the only boundary of the region is at x_1, such that:

$$\partial_t \int_{-\infty}^{x_1} dx \Pr(X_t^d = x) = -J(x_1,t)$$
$$= -J'(x,t) .$$

For this to be true, $K(t)$ must vanish, and so there cannot be any additional drift term. That is, up to an additional flat potential $C(t)$, the counterdiabatic control protocols are the unique way to guide the system along a desired distribution trajectory $\{\Pr(X_t^d)\}$.

This proof does not preclude additional solutions with nonzero $K(t)$ if the position variable has circular topology on a finite range $[x_0, x_1]$. This would mean that x_0 and x_1 are effectively adjacent such that there can be probability current at both points. In this case, there are always at least two boundary surfaces for any region, so it is impossible to use the integral continuity equation as above. The additional probability current $K(t)$ is possible, due to probability flow between x_0 and x_1, which was not possible between ∞ and $-\infty$ in the previous case. However, this additional probability current corresponds to a force that points in the same direction along the loop, meaning that system is being driven cyclically. And so, the dynamics cannot be implemented with Hamiltonian control and must rely on some free energy resource to be sustained.

Appendix B: Symmetric Metastable Erasure

In metastable erasure, we assume the system is in a metastable distribution of the initial symmetric equilibrium potential $V(x,0) = V(-x,0)$ during the entire protocol. If the two metastable informational states are $Y = 0$, corresponding to $x \in (-\infty, 0]$, and $Y = 1$, corresponding to $x \in (0, \infty)$, then we can describe a probabil-

ity distribution trajectory as:

$$\Pr(X_t = x) = \begin{cases} \Pr(X^{\text{eq}} = x) 2 \Pr(Y_t = 0) & \text{if } x \leq 0 \\ \Pr(X^{\text{eq}} = x) 2 \Pr(Y_t = 1) & \text{if } x > 0 \end{cases}.$$

We can then reparametrize in terms of the bit bias, that is the probability of the $Y = 0$ informational state:

$$b(t) = \Pr(Y_t = 0) .$$

As in Sec. IV A, let $p(x) = \Pr(X^{\text{eq}} = x)$ denote the equilibrium distribution, which inherits the symmetry of the double well potential:

$$p(x) = p(-x) .$$

We then express the evolving metastable distribution as a function of control parameter:

$$\Pr(X_t = x) = \begin{cases} 2p(x)b(t) & \text{if } x \leq 0 \\ 2p(x)(1 - b(t)) & \text{if } x > 0 \end{cases}.$$

This expression allows us to simplify the counterdiabatic potential and counterdiabatic work, as follows.

$$V^{CD}(x,t) = \frac{1}{\mu} \int_0^x dx' \frac{1}{\Pr(X_t = x')} \int_{-\infty}^{x'} dx'' \partial_t \Pr(X_t = x'')$$

$$= \begin{cases} \frac{1}{\mu} \int_0^x dx' \frac{1}{2bp(x')} \int_{-\infty}^{x'} dx'' 2p(x'') \dot{b} & \text{if } x \leq 0 \\ \frac{1}{\mu} \int_0^x dx' \frac{1}{2(1-b)p(x')} (\int_{-\infty}^{0} dx'' 2p(x'') \dot{b} \\ \quad + \int_0^{x'} dx'' 2p(x'') \partial_t(1-b)) & \text{if } x > 0 \end{cases}$$

$$= \begin{cases} \frac{1}{\mu} \int_0^x dx' \frac{1}{bp(x')} \int_{-\infty}^{x'} dx'' p(x'') \dot{b} & \text{if } x \leq 0 \\ \frac{1}{\mu} \int_0^x dx' \frac{1}{(1-b)p(x')} (\int_{-\infty}^{0} dx'' p(x'') \dot{b} \\ \quad - \int_0^{x'} dx'' p(-x'') \dot{b}) & \text{if } x > 0 \end{cases}$$

$$= \begin{cases} \frac{1}{\mu} \int_0^x dx' \frac{1}{bp(x')} \int_{-\infty}^{x'} dx'' p(x'') \dot{b} & \text{if } x \leq 0 \\ \frac{1}{\mu} \int_0^x dx' \frac{1}{(1-b)p(x')} (\int_{-\infty}^{0} dx'' p(x'') \dot{b} \\ \quad - \int_{-x'}^{0} dx'' p(x'') \dot{b}) & \text{if } x > 0 \end{cases}$$

$$= \begin{cases} \frac{1}{\mu} \int_0^x dx' \frac{1}{bp(x')} \int_{-\infty}^{x'} dx'' p(x'') \dot{b} & \text{if } x \leq 0 \\ \frac{1}{\mu} \int_0^x dx' \frac{1}{(1-b)p(x')} \int_{-\infty}^{-x'} dx'' p(x'') \dot{b} & \text{if } x > 0 \end{cases}$$

$$= \begin{cases} \frac{1}{\mu} \int_0^x dx' \frac{1}{bp(x')} \int_{-\infty}^{-|x'|} dx'' p(x'') \dot{b} & \text{if } x \leq 0 \\ \frac{1}{\mu} \int_0^x dx' \frac{1}{(1-b)p(x')} \int_{-\infty}^{-|x'|} dx'' p(x'') \dot{b} & \text{if } x > 0 \end{cases}$$

$$= \begin{cases} \frac{\dot{b}}{b} \frac{1}{\mu} \int_0^x dx' \frac{1}{p(x')} \int_{-\infty}^{-|x'|} dx'' p(x'') & \text{if } x \leq 0 \\ \frac{\dot{b}}{1-b} \frac{1}{\mu} \int_0^x dx' \frac{1}{p(x')} \int_{-\infty}^{-|x'|} dx'' p(x'') & \text{if } x > 0 \end{cases},$$

where $b = b(t)$ and $\dot{b} = \partial_t b(t)$. The second line follows from the first, since $p(x'') = p(-x'')$.

We can substitute $u = -x'$ again since $\int_0^x \frac{1}{p(x')} dx' =$

$-\int_0^{-x} du \frac{1}{p(u)}$. And so, if we define:

$$h(x) = \frac{1}{\mu} \int_0^{|x|} dx' \frac{1}{p(x')} \int_{-\infty}^{-|x'|} dx'' p(x'') ,$$

then:

$$V^{CD}(x,t) = h(x) \times \begin{cases} -\frac{\dot{b}}{b} & \text{if } x \leq 0 \\ \frac{\dot{b}}{1-b} & \text{if } x > 0 \end{cases}. \tag{B1}$$

The resulting counterdiabatic work is:

$$\langle W^{CD} \rangle = \int_0^\tau dt \int_{-\infty}^{\infty} dx \Pr(X_t = x) \partial_t U(x,t)$$

$$= \int_0^\tau dt \int_{-\infty}^{0} dx\, p(x) b\, h(x) \partial_t \left(\frac{-\dot{b}}{b} \right)$$

$$\quad + \int_0^\tau dt \int_0^{\infty} dx\, p(x)(1-b) h(x) \partial_t \left(\frac{\dot{b}}{1-b} \right)$$

$$= \int_0^{\infty} dx\, p(x) h(x) \times$$

$$\int_0^\tau dt \left(b\partial_t \left(\frac{-\dot{b}}{b} \right) + (1-b)\partial_t \left(\frac{\dot{b}}{1-b} \right) \right)$$

$$= \int_0^{\infty} dx\, p(x) h(x) \times$$

$$\int_0^\tau dt \left(-b \left(\frac{\partial_t^2 b}{b} - \frac{\dot{b}^2}{b^2} \right) + (1-b) \left(\frac{\partial_t^2 b}{1-b} + \frac{(\dot{b})^2}{(1-b)^2} \right) \right)$$

$$= \int_0^{\infty} dx\, p(x) h(x) \int_0^\tau dt \left(-\partial_t^2 b + \frac{(\dot{b})^2}{b} + \partial_t^2 b + \frac{(\dot{b})^2}{(1-b)} \right)$$

$$= \int_0^{\infty} dx\, p(x) h(x) \int_0^\tau dt \left(\frac{(\dot{b})^2}{b} + \frac{(\dot{b})^2}{(1-b)} \right)$$

$$= \int_0^{\infty} dx\, p(x) h(x) \int_0^\tau dt \frac{(\dot{b})^2}{b - b^2}$$

$$= f_1[p(\cdot)] \times f_2[b(\cdot)] .$$

The second line follows from the first since $p(x)$ and $h(x)$ are symmetric. The functions appearing on the last line are given by:

$$f_2[b(\cdot)] = \int_0^\tau dt \frac{(\dot{b})^2}{b - b^2}$$

and:

$$
\begin{aligned}
f_1[p(\cdot)] &= \int_0^\infty dx\, p(x) h(x) \\
&= \int_0^\infty dx\, p(x) \frac{1}{\mu} \int_0^{|x|} dx' \frac{1}{p(x')} \int_{-\infty}^{-|x'|} dx''\, p(x'') \\
&= \frac{1}{\mu} \int_0^\infty dx\, p(x) \int_0^{x} dx' \frac{1}{p(x')} \int_{-\infty}^{-x'} dx''\, p(x'') \ .
\end{aligned}
$$

Thus, the counterdiabatic work is the product of two factors: one dependent on the bias trajectory $b(t)$, containing all dependence on erasure fidelity, and the other dependent on the equilibrium potential $p(x)$, containing all dependence on information storage robustness—the information lifetime.

One is tempted to use the functional f_2 to find the bias trajectory that minimizes dissipation. f_2 can be expressed as the integral of a Lagrangian:

$$
f_2[b(\cdot)] = \int_0^\tau \mathcal{L}(b(t), b'(t)) dt \ ,
$$

where:

$$
\mathcal{L}(b(t), b'(t)) = \frac{b'(t)^2}{b - b^2} \ .
$$

This implies that, for the optimal path satisfying the equation of motion:

$$
\frac{\partial \mathcal{L}}{\partial b} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial b'} \ . \tag{B2}
$$

Integrating these equations of motion, given the constraint of starting at initial bias $b(0) = b_i$ and ending at final bias $b(\tau) = b_f$ would determine the most thermodynamically efficient path $b(t)$ for transiting between different metastable distributions. However, this is challenging and remains unsolved. So, instead, consider a simpler protocol.

We choose a bias trajectory:

$$
b(t) = b(0) \cos(t\pi/2\tau)^2 + b(\tau) \sin(t\pi/2\tau)^2 \ ,
$$

which has vanishing derivative at the protocol's beginning and end, such that the desired distribution $\Pr(X_t^d)$ has zero time derivative at the initial and final times. This means that the counterdiabatic potential energy is zero at the protocol's beginning and end. As a result, the system is in equilibrium at the beginning and end of the counterdiabatic step in the protocol. Substituting this into the expression for $f_2[b(\cdot)]$ above, we evaluate numerically, and see that the counterdiabatic work is proportional to twice Hellinger distance:

$$
K(b_i, b_f) = \left(\sqrt{b_i} - \sqrt{b_f}\right)^2 + \left(\sqrt{1-b_i} - \sqrt{1-b_f}\right)^2 \ .
$$

Despite not knowing the proportionality, we see that the maximum of $f_2[b(\cdot)]$ occurs when a bit is perfectly flipped, yielding a multiplicative contribution π^2/τ. Substituting our chosen bias trajectory into the expression for optimality in Eq. (B2), we see that it does not satisfy the equality, and so is not optimal.

---

[1] R. Landauer. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.*, 5(3):183–191, 1961.

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

[3] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948.

[4] 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.

[5] A. J. P. Garner, J. Thompson, V. Vedral, and M. Gu. Thermodynamics of complexity and pattern manipulation. *Phys. Rev. E*, 95:042140, 2017.

[6] P. R. Zulkowski, D. A. Sivak, G. E. Crooks, and M. R. DeWeese. Geometry of thermodynamic control. *Physical Review E*, 86(041148), 2012.

[7] P. R. Zulkowski and M. R. DeWeese. Optimal finite-time erasure of a classical bit. *Phys. Rev. E*, 89:052140, May 2014.

[8] P. R. Zulkowski and M. R. DeWeese. Optimal control of

overdamped systems. *Phys. Rev. E*, 92(032117), 2015.

[9] T. R. Gingrich, G. M. Rotskoff, G. E. Crooks, and P. L. Geissler. Near-optimal protocols in complex nonequilibrium transformations. *Proc Natl. Acad. Sci. USA*, 113(37), 2016.

[10] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga. Chapter 2—Shortcuts to Adiabaticity. *Adv. At. Mol. Opt. Phys.*, 62:117, 2013.

[11] A. Patra and C. Jarzynski. Shortcuts to adiabaticity using flow fields. *New J. Physics*, 19:125009, 2017.

[12] I. A. Martinez, A. Petrosyan, D. Guery-Odelin, E. Trizac, and S. Ciliberto. Engineered swift equilibration of a Brownian particle. *Nature Physics*, 12, 2016.

[13] C. H. Bennett. Notes on landauer's principle, reversible computation, and maxwell's demon. *Stud. Hist. Phil. Modl. Phys.*, 34(3):501–510, 2003.

[14] S. Deffner and C. Jarzynski. Information processing and the second law of thermodynamics: An inclusive, Hamiltonian approach. *Phys. Rev. X*, 3:041003, 2013.

[15] Symbolic dynamics is a long-lived subfield of dynam-

ics systems [16]. Our use of it here is relatively simple, highlighting (i) how mesoscopic symbols capture (or not) collections of microscopic states and (ii) the concern of proper coarse-graining to locate information storage and processing. Fully deploying the symbolic dynamics for thermodynamic computing must wait for a different venue. Early results, however, do develop the symbolic dynamics of thermally-activated (noisy) systems [33].

[16] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding.* Cambridge University Press, New York, 1995.

[17] M. Gavrilov and J. Bechhoefer. Arbitrarily slow, non-quasistatic, isothermal transformations. *Europhys. Lett.*, 114(5), 2016.

[18] K. Proesmans, Y. Dreher, M. Gavrilov, J. Bechhoefer, and C. Van den Broeck. Brownian duet: A novel tale of thermodynamic efficiency. *Phys. Rev. X*, 6(4), 2016.

[19] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Let.*, 45:712, 1980.

[20] J. P. Crutchfield and B. S. McNamara. Equations of motion from a data series. *Complex Systems*, 1:417–452, 1987.

[21] J. P. Crutchfield. Between order and chaos. *Nature Physics*, 8(January):17–24, 2012.

[22] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *J. Stat. Phys.*, 136(6):1005–1034, 2009.

[23] C. Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Cond. Matt. Phys.*, 2:329–351, 2011.

[24] M. Esposito and C. van den Broeck. Second law and Landauer principle far from eqilibrium. *Europhys. Lett*, 95:40004, 2011.

[25] K. Takara, H.-H. Hasegawa, and D. J. Driebe. Generalization of the second law for a transition between nonequilibrium states. *Phys. Lett. A*, 375(88-92), 2010.

[26] S. Lahiri, J. Sohl-Dickstein, and S. Ganguli. A universal tradeoff between power, precision and speed in physical communication. *arXiv:1603.07758*, 2016.

[27] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz. Experimental verification of landauer's principle linking information and thermodynamics. *Nature*, 483:187–190, 2012.

[28] Y. Jun, M. Gavrilov, and J. Bechhoefer. High-precision test of Landauer's principle in a feedback trap. *Phys. Rev. Lett.*, 113(19), 2014.

[29] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, New York, second edition, 2006.

[30] P. Hanggi, P. Talkner, and M. Borkovec. Reaction-rate theory: Fifty years after Kramers. *Rev. Mod. Phys.*, 62(2), 1990.

[31] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck. Dissipation: The phase-space perspective. *Phys. Rev. Lett.*, 98:080602, 2007.

[32] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[33] J. P. Crutchfield and N. H. Packard. Symbolic dynamics of noisy chaos. *Physica*, 7D(1-3):201 – 223, 1983.