Optimal finite-time erasure of a classical bit

Patrick R. Zulkowski*

Department of Physics, University of California, Berkeley, CA 94720 and
Redwood Center for Theoretical Neuroscience, University of California, Berkeley CA 94720

Michael R. DeWeese†

Department of Physics, University of California, Berkeley, CA 94720
Redwood Center for Theoretical Neuroscience, University of California, Berkeley CA 94720 and
Helen Wills Neuroscience Institute, University of California, Berkeley CA 94720
(Dated: October 12, 2013)

Information erasure inevitably leads to heat dissipation. Minimizing this dissipation will be crucial for developing small-scale information processing systems, but little is known about the optimal procedures required. We have obtained closed-form expressions for maximally efficient erasure cycles for deletion of a classical bit of information stored by the position of a particle diffusing in a double-well potential. We find that the extra dissipation beyond the Landauer bound is proportional to the square of the Hellinger distance between the initial and final states divided by the cycle duration, which quantifies how far out of equilibrium the system is driven. Finally, we demonstrate close agreement between the exact optimal cycle and the protocol found using a linear response framework.

Introduction. Optimization schemes for thermodynamic processes occurring in finite time will be needed for applications in which energetic or entropic costs are undesirable [1, 2]. An important class of such processes consists of mesoscopic information processing systems operating out of equilibrium. Optimization will aid technological development in the decades to come as computational demands approach limits imposed by physical law [3].

Moreover, understanding these systems will provide insight into the foundations of nonequilibrium statistical mechanics. Investigations into the interplay between information and thermodynamics seem to have originated with Maxwell's hypothetical demon and its implications for the second law of thermodynamics [4]. Much groundbreaking work followed from the Maxwell demon paradox including Szilard's engine revealing a quantitative link between thermodynamic work and information [5], Landauer's observation of the physical nature of information [6] and Bennett's interpretation of the paradox in terms of the relation between logical and thermodynamic reversibility [7].

In recent times, research into nonequilibrium statistical mechanics of small-scale systems has shed more light on the thermodynamic role of information [8]. Most notable is experimental verification [9] of the theoretical prediction of microscopic violations of Landauer's principle with the preservation of the principle on average [10], analogous to experimental and theoretical work on fluctuations theorems demonstrating that entropy-reducing processes can occur microscopically whereas the second law holds on average [11]. Research into feedback and measurement of mesoscopic nonequilibrium systems has improved our understanding of the role information plays in the second law [8, 12]. Other work has focused on developing techniques to optimize thermodynamic quantities arising in small-scale systems designed to store and erase classical information [13–15], including the derivation of a refined second law [15]. Recent work has also focused on the general problem of predicting optimal protocols to drive systems between stationary states with minimal dissipation [16–23].

Here we obtain closed-form expressions for the dissipation of maximally efficient cyclical protocols for a simple system designed to store and delete a classical bit of information. The system storing this bit consists of an overdamped Brownian colloidal particle diffusing in a one-dimensional double-square-well potential separated by a potential barrier stabilizing the memory. We take as control parameters the height of the potential barrier and the difference in minima of the two wells.

When our two simultaneously adjustable parameters are optimally controlled, we find that the dissipation falls off as the inverse of the cycle duration, asymptoting to the Landauer bound in the long duration limit. This is consistent with pioneering studies of erasure for a similar model system [15] as well as a single-level quantum dot [13, 14], but unlike those studies, we have derived an explicit formula for the minimal dissipation, valid for arbitrary temperature, providing specific testable predictions for existing experimental setups [9, 24]. Our solutions are non-trivial in that both control parameters are continuously varying in time, but they are easily described, which complements previous results for a non-parametric model [15].

For durations that are long compared to a characteristic timescale, we obtain a simple expression for the dissipation that depends on the difference between the initial
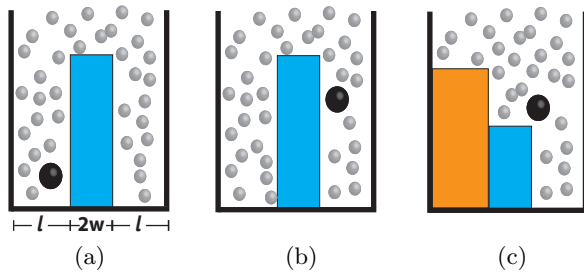
FIG. 1. Double-well potential for storage of a single classical bit. (a,b) The system begins in thermal equilibrium with equipotential wells and a potential barrier of height much larger than the thermal fluctuation scale. Observing the particle (black dot) to the left (right) of the potential barrier corresponds to memory value 1 (0). The width $2w$ of the central barrier and the width l of each well satisfy $2w/l \ll 1$. (c) Optimally-efficient erasure protocols are sought in which the "tilt" V_l (orange) and the barrier height V_b (blue) are control parameters. After the erasure step, the particle is much more likely to be in the right well, regardless of where it originated.

and final spatial distributions of the particle. Interestingly, the extra dissipation beyond the Landauer bound for the optimal finite-time protocol is proportional to the square of the Hellinger distance, which is always greater than zero for any nonzero change in the probabilities of finding the particle in the left or right potential well, unlike the Landauer bound itself, which can be zero or even negative depending on the change in entropy of the particle's spatial distribution.

Finally, we demonstrate that a recently developed geometrical framework [16, 17] for finding optimal protocols based on the inverse diffusion tensor predicts nearly identical solutions to our exact optimal protocols in this parameter regime, which is an encouraging sign for finding optimal protocols in other model systems.

Model of Classical Information Erasure. We consider the following model to represent a single classical bit of information: an overdamped Brownian colloidal particle diffusing in a one-dimensional double-well potential in contact with a thermal bath of temperature T [9, 10] (Fig. (1)). The wells are initially separated by a potential barrier whose height is much larger than the energy scale $\beta^{-1} \equiv k_B T$ set by thermal fluctuations, ensuring stability of memory. The system is prepared so that the particle has equal probability of being found in either well. This may be achieved, for example, by selecting the initial position of the particle to be at the midpoint of the potential barrier and waiting a sufficiently long relaxation period [10]. If the particle is found in the left-hand (right-hand) well, the memory value is defined to be 1 (0).

The time evolution of the particle's position $x(t)$ is

governed by Brownian dynamics

$$\dot{x} = -\frac{1}{\gamma} \partial_x U(x(t), t) + F(t) \qquad (1)$$

for Gaussian white noise $F(t)$ satisfying

$$\langle F(t) \rangle = 0 \ , \ \langle F(t)F(t') \rangle = \frac{2}{\beta\gamma}\delta(t - t'). \qquad (2)$$

Here, γ is the Cartesian friction coefficient and $U(x, t)$ is a generic double-well potential satisfying $U(x, t) \to \infty$ as $|x| \to \infty$. We will find the equivalent statistical description in terms of the Fokker-Planck equation

$$\partial_t \rho = D\left[\partial_x \left(\beta U'(x, t)\rho\right) + \partial_x^2 \rho\right] \equiv -\partial_x G \qquad (3)$$

convenient, where $\rho(x, t)$ is the position probability density, $G(x, t)$ is the probability current, and D is the diffusion coefficient.

Out of equilibrium, the system's probability distribution over microstates fundamentally depends on the history of the control parameters $\boldsymbol{\lambda}$, which we denote by the control parameter protocol $\boldsymbol{\Lambda}$; $\langle \cdot \rangle_{\boldsymbol{\Lambda}}$ denotes the average over the nonequilibrium probability distribution arising from the parameter protocol $\boldsymbol{\Lambda}$.

We are primarily interested in optimizing finite-time erasure efficiency over cyclic protocols for the classical single bit model described above. When classical information is being erased, the difference in Shannon entropies of the final and initial probability distributions must satisfy $\triangle S \equiv S_f - S_i < 0$, which would allow us to define the erasure efficiency $\epsilon \equiv -\triangle S/ \left(k_B \langle \beta Q \rangle_{\boldsymbol{\Lambda}}\right)$ as the ratio of this decrease in Shannon entropy to the average heat $\langle Q \rangle_{\boldsymbol{\Lambda}}$ released into the thermal bath [13, 25]. However, in addition to erasure, we will also consider arbitrary initial and final spatial distributions for the particle, so we will state our results in terms of dissipation rather than efficiency.

Our goal will be to minimize the dissipated heat subject to constraints on the initial and final probability distributions. Since we are constraining the initial and final probability distributions, we expect our optimal protocols to have jump discontinuities at the endpoints based on experience with optimization in the context of stochastic thermodynamics in general [21–23, 26, 27] and erasure efficiency in particular [14, 15]. These jump discontinuities warrant caution when defining thermodynamic quantities such as the average dissipated heat [27] (see Supplementary Material).

To simplify the mathematics, we consider a piecewise constant potential as illustrated in Fig. 1 and similar to the model considered in [28]. This model admits a reasonable "discretization" of the system, providing a means of calculating optimal protocols exactly. We use the discrete approximation of [29] to obtain transition rates for the master equations [30]

$$\frac{dp_i}{dt} = \sum_{j \neq i} r_{j \to i} \ p_j - \sum_{j \neq i} r_{i \to j} \ p_i \qquad (4)$$

governing the time evolution of p_i (see Supplementary Material). Here and throughout, p_l (p_r) is the probability of the particle being on the left (right) of the barrier, corresponding to memory value 1 (0).

From the perspective of our optimization problem, minimizing the dissipation is equivalent to minimizing the average work done on the system, so in addition to yielding an exact solution, this problem is amenable to a recently developed geometric framework for calculating optimal protocols. This framework, originally developed in the linear response regime [16], utilizes the equivalence of optimal protocols and geodesics of the inverse diffusion tensor on the space of control parameters [16–18]. The inverse diffusion tensor can be calculated directly from the Fokker-Planck equation Eq. (3), allowing us to compare the exact answer with this approximate solution. Formulating the optimization problem in terms of a geodesic problem on a manifold gives us the opportunity to use powerful and elegant methods from Riemannian geometry [17, 18], and it can provide approximate solutions in other cases that cannot be solved exactly (see Supplementary Material).

We take as control parameters the "tilt" V_l and the potential barrier height V_b (see Fig. 1(c)), and we initially focus on the class of protocols resulting in (partial) erasure of the classical bit. For example, increasing V_l and decreasing V_b appropriately as in Fig. 1(c) ensures near unity probability of finding the particle in the right well (*i.e.*, memory value of 0) regardless of its initial state.

We consider protocols consisting of two stages. During the first (erasure) stage, the initial equilibrium distribution transitions to a final nonequilibrium distribution in which the system is overwhelmingly likely to have memory value 0. In the second (reset) stage, the control parameters are brought instantaneously back to their original values while keeping the particle probability distribution constant. We allow these protocols to have jump discontinuities at the endpoints of each stage, and the optimal protocols will indeed exhibit them.

Exact Optimizer. For the discrete dynamics, the average heat produced during the cycle is

$$\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} = Q_b + \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{p}_l \ln\left(\dot{p}_l + p_l\right) + \dot{p}_r \ln\left(\dot{p}_r + p_r\right) \right], \tag{5}$$

with a boundary term defined as

$$Q_b \equiv p_b(t_f) \left[\ln\left(p_b(t_f)\right) - 1 \right] - p_b(0) \left[\ln\left(p_b(0)\right) - 1 \right], \tag{6}$$

where $\bar{t} \equiv \frac{2D}{l^2} t$ and $f(t^{\pm}) \equiv \lim_{\delta \to 0^+} f(t \pm \delta)$ (see Supplementary Material). The boundary term Q_b depends only on the probability distributions at the endpoints. Moreover, we see that the "bulk" term of the average heat

functional is a sum $I[p_l] + I[p_r]$, where

$$I[z] \equiv \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{z} \ln\left(\dot{z} + z\right) \right]. \tag{7}$$

Therefore, to extremize the average heat functional, we can solve the Euler-Lagrange equations for I.

Suppose $z(t)$ satisfies the Euler-Lagrange equation for the Lagrangian $L[z, \dot{z}] \equiv \dot{z} \ln\left(\dot{z} + z\right)$. Then it must be true that

$$\dot{z} \frac{\partial L}{\partial \dot{z}} - L = \frac{(\dot{z})^2}{\dot{z} + z} \tag{8}$$

is a constant.

Therefore, probability distributions extremizing $\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}}$ satisfy $\dot{p}_i = K_i (\dot{p}_i + p_i)$ for positive constants K_i and $i = l, r$. Over the course of the erasure stage, p_l (p_r) decreases (increases). The constants K_i may in turn be numerically fixed by imposing the constraints $p_l(0) = \frac{1}{2(1+\gamma)} = p_r(0)$, $p_l(t_f) = \delta$, $p_r(t_f) = 1 - 2\delta$, where δ and γ are small and positive. The uniqueness of our solution combined with the Second Law guarantee that this is the minimum we sought.

In the long duration limit, we obtain a simple result (see Supplementary Material)

$$\langle \beta Q \rangle_{\mathbf{\Lambda}_{opt}} \approx \frac{-\triangle S}{k_B} + \frac{4K}{\overline{t_f}}, \tag{9}$$

where

$$K \equiv \left(\sqrt{p_r(t_f)} - \sqrt{p_r(0)} \right)^2 + \left(\sqrt{p_l(t_f)} - \sqrt{p_l(0)} \right)^2 \tag{10}$$

is twice the square of the Hellinger distance [31], a measure of similarity between pairs of probability distributions. Note that K contains no terms for $p_b(0)$ and $p_b(t_f)$, which are both small for the cases we consider.

Consistent with previous studies [13–15], we find that the total dissipation for our optimized protocols consists of the sum of two terms: one given by the Landauer bound, which is proportional to the decrease in Shannon entropy resulting from the erasure of information, and a second term that falls as $1/t_f$ (Eq. (9)). Fortunately, we have arrived at a simple closed-form expression for the total dissipation, valid for arbitrary temperature, that can be experimentally tested using existing setups [9, 24] by comparing our optimal path through the two dimensional parameter space with alternate protocols. Fig. 2(a) depicts an optimally efficient finite-time erasure cycle constructed based on the calculations in this section. For the parameter values selected to generate Fig. 2, the erasure efficiency ϵ is about 94%.

Inverse Diffusion Tensor-Based Approximation. As originally constructed [16], the formulation of the inverse diffusion tensor assumes smooth protocols on the entire domain of definition. However, we anticipated that the

optimal solution would have discontinuities at the end-points. Fortunately, we were able to search over all protocols with end point discontinuities by using the derivative truncation method [17, 18] to obtain a numerical solution of the approximate optimizer (see Supplementary Material). Note the strong agreement between the exact solution and the approximate solution (Fig. 2(b)) based on the inverse diffusion tensor.
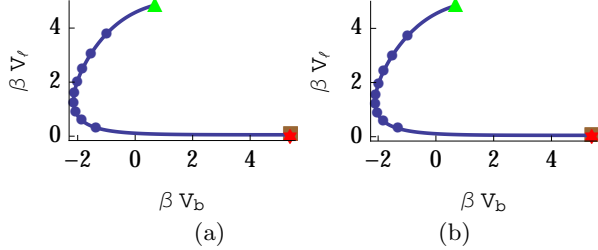

FIG. 2. Optimally efficient finite-time erasure cycles. (a) The optimal cycle consists of two parts: the erasure stage and the reset stage. The erasure stage begins at $(\ln\{w/(l\gamma)\}, 0)$ (red star) then jumps to the initial point of the erasure protocol (brown square). The erasure protocol (blue) proceeds for time t_f until reaching its terminus (green triangle). The reset stage consists of the jump from this terminus to the parameter values defining the original equilibrium state (red star). Blue dots indicate points separated by equal times along the erasure stage. For these parameters $(\overline{t_f} \equiv (2D/l^2) t_f = 50$, $\delta = 0.01$, $\gamma = \exp(-10)$, $w/l = 0.01$) the efficiency of the optimal cycle is 94.01%. (b) An approximate optimal efficiency erasure cycle determined by the inverse diffusion tensor framework is nearly identical to the exact solution shown in panel (a).

Beyond Erasure.   So far we have focused on the case of complete erasure of one bit, for which both the Landauer bound and the added dissipation necessary to achieve erasure in finite time are always positive. However, our formalism is completely general, valid for any change in the spatial distribution of the particle. Importantly, even in this broader setting, the second term is always nonnegative, as the Hellinger distance (and its square) is positive for any nonzero difference between the initial and final probability distributions, unlike the Landauer term, which can be zero or even negative.

Discussion.   We have obtained a simple, closed-form expression for the dissipated heat of optimally efficient, finite-time erasure cycles, providing falsifiable predictions for currently achievable experiments. The solutions we have found are nontrivial, in that both of our control parameters are continuously varying throughout the optimal protocol, yet our parametric solutions can be easily described. In addition to erasure, our solutions are valid for any initial and final particle distributions.

We find that the total dissipation for the optimal protocol consists of the Landauer bound plus a nonnegative second term proportional to the square of the Hellinger
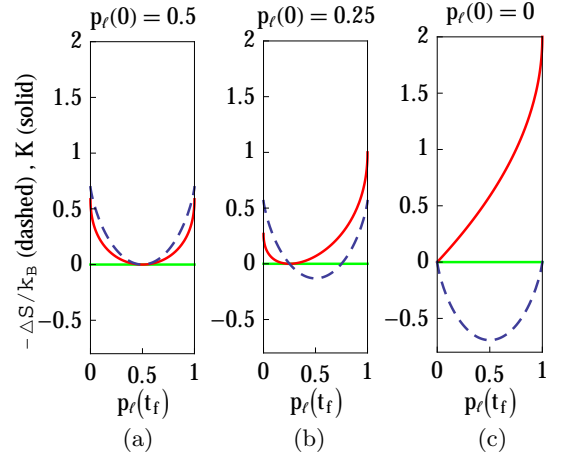

FIG. 3. The Landauer bound and the added dissipation necessary for finite-time cycles exhibit different dependences on the spatial distribution of the particle. (a) If the particle is initially distributed equally between the two wells, $p_l(0) = 0.5$, where $p_l(0)$ is the probability of being found in the left well at $t = 0$, then the Landauer bound, proportional to $-\Delta S$ (dashed blue curve), is zero for no change in the likelihood of finding the particle on the left ($p_l(t_f) = 0.5$) and positive for any other final distribution. This is always true for the second term in the full dissipation (solid red curve), which is proportional to K (Eq. (10)). In all panels, $p_b(0) = p_b(t_f) = 0$, where $p_b(t)$ is the probability of finding the particle at the central barrier at time t. (b,c) For other initial conditions, the Landauer bound can be positive, zero, or negative depending on how the particle's spatial distribution changes.

distance between the initial and final particle distributions (Eqs. (9,10)). Fittingly, one can think of this second term as a measure of how far out of equilibrium the system must be during the driving protocol, as it is the ratio of the "distance" between the initial and final probability distributions and the time allowed to make the transition. Indeed, one can show [31] that K (Eq. 10) is a lower bound on the relative entropy $D[\mathbf{p}(t_f) \; \| \; \mathbf{p}(0)] \equiv \Sigma_i p_i(t_f) \log[p_i(t_f)/p_i(0)]$ between the distributions, which is precisely the dissipation that would result from allowing the system to relax from the final distribution back to the initial equilibrium distribution with the control parameters held fixed to their initial values [32–34]. This is a tight bound in many cases; for example, for perfect erasure of one bit, $K \approx 0.59$ and $D[\mathbf{p}(t_f) \| \mathbf{p}(0)] \approx 0.69$ are comparable.

The exact and approximate optimal erasure cycles we found are nearly identical and both achieved high efficiencies for the finite cycle duration selected here, suggesting that the inverse diffusion tensor formalism could be an indispensable tool for predicting optimally efficient finite-time erasure cycles for complex model systems more relevant for biology or engineering.

---

[*] pzulkowski@berkeley.edu
[†] deweese@berkeley.edu

[1] B. Andresen, Angew. Chem., Int. Ed. **50**, 2690 (2011).
[2] L. Chen and F. Sun, eds., *Advances in finite time thermodynamics: analysis and optimization* (Nova Science Publishers, New York, 2004).
[3] M. Frank, Comput. Sci. Eng. **4**, 16 (2002).
[4] J. Maxwell, *Theory of Heat* (Appleton, London, 1871).
[5] L. Szilard, Z.Phys. **53**, 840 (1929).
[6] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[7] C. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[8] T. Sagawa, *Thermodynamics of information processing in small systms*, Ph.D. thesis, The University of Tokyo, Japan.
[9] A. Beacuterut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Nature **483** (2012).
[10] R. Dillenschneider and E. Lutz, Phys. Rev. Lett. **102** (2009).
[11] R. Klages, W. Just, and C. Jarzynski, eds., *Nonequilibrium statistical physics of small systems* (Wiley-VCH Verlag and Co. KGaA, Weinheim, Germany, 2013).
[12] T. Sagawa and M. Ueda, Phys. Rev. E **85** (2012).
[13] G. Diana, G. B. Bagci, and M. Esposito, Phys. Rev. E **87** (2013).
[14] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, Europhys. Lett. **89** (2010).
[15] E. Aurell, K. Gawedzki, C. Mejia-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, J. Stat. Phys. **147** (2012).
[16] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 190602 (2012).
[17] P. R. Zulkowski, D. A. Sivak, G. E. Crooks, and M. R. DeWeese, Phys. Rev. E **86** (2012).
[18] P. R. Zulkowski, D. A. Sivak, and M. R. DeWeese, "Optimal control of transitions between nonequilibrium steady states," (2013), arXiv:1303.6596.
[19] D. K. Shenfeld, H. Xu, M. P. Eastwood, R. O. Dror, and D. E. Shaw, Phys. Rev. E **80**, 046705 (2009).
[20] D. C. Brody and D. W. Hook, J. Phys. A **42**, 023001 (2009).
[21] A. Gomez-Marin, T. Schmiedl, and U. Seifert, J. Chem. Phys. **129**, 024114 (8) (2008).
[22] T. Schmiedl and U. Seifert, Phys. Rev. Lett. **98**, 108301 (2007).
[23] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Phys. Rev. Lett. **106**, 250601 (4) (2011).
[24] Y. Jun and J. Bechhoefer, Phys Rev E **86**, 061106 (2012).
[25] M. Plenio and V. Vitelli, Contemp. Phys. **42** (2001).
[26] U. Seifert, Reports on Progress in Physics **75**, 126001 (2012).
[27] H. Then and A. Engel, Phys. Rev. E **77** (2008).
[28] M. M. Barkeshli, "Dissipationless information erasure and landauer's principle," (2006), arXiv:0504323v3.
[29] J. Xing, H. Wang, and G. Oster, Biophys. J. **89**, 1551 (2005).
[30] R. Zwanzig, *Nonequilibrium statistical mechanics* (Oxford University Press, New York, 2001).
[31] R. D. Reiss, *Approximate Distributions of Order Statistics with Applications to Non-Parametric Statistics*, Springer Series in Statistics (Springer-Verlag, 1989).
[32] R. Shaw, *The Dripping Faucet as a Model Chaotic System* (Aerial, Santa Cruz, CA, 1984).
[33] B. Gaveau and L. S. Schulman, Phys. Lett. A **229**, 347 (1997).
[34] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 150601 (2012).

Optimal finite-time erasure of a classical bit: supplementary material

Patrick R. Zulkowski*
Department of Physics, University of California, Berkeley, CA 94720 and
Redwood Center for Theoretical Neuroscience, University of California, Berkeley CA 94720

Michael R. DeWeese†
Department of Physics, University of California, Berkeley, CA 94720
Redwood Center for Theoretical Neuroscience, University of California, Berkeley CA 94720 and
Helen Wills Neuroscience Institute, University of California, Berkeley CA 94720

(Dated: October 13, 2013)

DEFINITION OF AVERAGE DISSIPATED HEAT

Suppose $\boldsymbol{\lambda}$ is smooth for $t \in (0, t_f)$ but possesses jump discontinuities at $t = 0, t_f$. Then $U(x, t) \equiv U(x, \boldsymbol{\lambda}(t))$ is smooth on the interval $(\delta, t_f - \delta)$ and it is possible to write

$$U(x(t_f - \delta), t_f - \delta) - U(x(\delta), \delta) =$$
$$\int_{\delta}^{t_f - \delta} dt \left[\frac{d\boldsymbol{\lambda}}{dt}\right]^T \cdot \frac{\partial U}{\partial \boldsymbol{\lambda}}(x(t), \boldsymbol{\lambda}(t)) +$$
$$\int_{\delta}^{t_f - \delta} \frac{\partial U}{\partial x}(x(t), \boldsymbol{\lambda}(t)) \circ dx(t), \quad (1)$$

where the integral over the fluctuating quantity x is computed using Stratonovich integral calculus [1–3] and δ is small and positive. We may also write this expression as

$$U(x(t_f), t_f) - U(x(0), 0) =$$
$$\left[U(x(t_f), t_f) - U(x(t_f - \delta), t_f - \delta)\right] -$$
$$\left[U(x(0), 0) - U(x(\delta), \delta)\right] +$$
$$\int_{\delta}^{t_f - \delta} dt \left[\frac{d\boldsymbol{\lambda}}{dt}\right]^T \cdot \frac{\partial U}{\partial \boldsymbol{\lambda}}(x(t), \boldsymbol{\lambda}(t)) +$$
$$\int_{\delta}^{t_f - \delta} \frac{\partial U}{\partial x}(x(t), \boldsymbol{\lambda}(t)) \circ dx(t). \quad (2)$$

The nonequilibrium ensemble average of the left-hand side is given by

$$\left\langle U(x(t_f), t_f) - U(x(0), 0)\right\rangle_{\boldsymbol{\Lambda}} =$$
$$\int_{\mathbb{R}} dx \left[U(x, t_f)\rho(x, t_f) - U(x, 0)\rho(x, 0)\right]. \quad (3)$$

The last equality follows from the fact that $\langle \delta(x(t) - x)\rangle_{\boldsymbol{\Lambda}} = \rho(x, t)$ which is assumed continuously differentiable on $(0, t_f)$ and continuous on $[0, t_f]$. Note that $\left\langle U(x(t_f), t_f) - U(x(0), 0)\right\rangle_{\boldsymbol{\Lambda}} \equiv \triangle U$ depends only on $\boldsymbol{\lambda}(0)$, $\boldsymbol{\lambda}(t_f)$ and the initial and final probability distributions.

Taking the nonequilibrium ensemble average of both sides of Eq. (2) and then taking the limit $\delta \to 0$, we find

the average work done on the system [2, 3]

$$\langle W\rangle_{\boldsymbol{\Lambda}} \equiv \triangle U + \langle Q\rangle_{\boldsymbol{\Lambda}} = \int_{0^+}^{t_f^-} dt \left[\frac{d\boldsymbol{\lambda}}{dt}\right]^T \cdot \left\langle \frac{\partial U}{\partial \boldsymbol{\lambda}}(\boldsymbol{\lambda}(t))\right\rangle_{\boldsymbol{\Lambda}}$$
$$+ \int_{\mathbb{R}} \rho(x, t_f) \, U(x, t)|_{t=t_f^-}^{t=t_f} - \int_{\mathbb{R}} \rho(x, 0) \, U(x, t)|_{t=0^+}^{t=0}$$
$$(4)$$

where we abuse notation by defining

$$\int_{\mathbb{R}} dx \, \frac{\partial U}{\partial \boldsymbol{\lambda}}(x, \boldsymbol{\lambda}(t)) \, \rho(x, t) \equiv \left\langle \frac{\partial U}{\partial \boldsymbol{\lambda}}(\boldsymbol{\lambda}(t))\right\rangle_{\boldsymbol{\Lambda}}. \quad (5)$$

The last two terms of the average work explicitly take into account the cost of jump discontinuities at the beginning and end of the protocol [4].

By definition of average dissipation for nonequilibrium transitions from stationary states [5, 6],

$$\langle \beta W\rangle_{\boldsymbol{\Lambda}} = \int_{0^+}^{t_f^-} dt \left[\frac{d\boldsymbol{\lambda}}{dt}\right]^T \cdot \left\langle \frac{\partial \phi}{\partial \boldsymbol{\lambda}}(\boldsymbol{\lambda}(t))\right\rangle_{\boldsymbol{\Lambda}} + \beta \triangle F$$
$$+ \int_{\mathbb{R}} \rho(x, t_f) \, \beta U(x, t)|_{t=t_f^-}^{t=t_f} - \int_{\mathbb{R}} \rho(x, 0) \, \beta U(x, t)|_{t=0^+}^{t=0}$$
$$(6)$$

where $\triangle F \equiv F(\boldsymbol{\lambda}(t_f^-)) - F(\boldsymbol{\lambda}(0^+))$, $F(\boldsymbol{\lambda}) \equiv -\beta^{-1} \ln\left(\int_{\mathbb{R}} \exp\{-\beta U(x; \boldsymbol{\lambda})\}\right)$, $\phi(x, \boldsymbol{\lambda}) \equiv -\ln \rho_{eq}(x; \boldsymbol{\lambda})$ and $\rho_{eq}(x; \boldsymbol{\lambda}) = \exp\{\beta (F(\boldsymbol{\lambda}) - U(x, \boldsymbol{\lambda}))\}$ is the Boltzmann distribution.

DISCRETE DYNAMICS

We use the discrete approximation of [7] to obtain transition rates for the master equations governing the time evolution of p_i. Here and throughout, p_i is the probability of the colloidal particle occupying the i-th compartment. Alternatively, we may interpret p_l (p_r) as the probability of the memory value being 1 (0).

The double-well potential has a natural decomposition into "compartments": we define the interval $(-l-w, -w)$ as compartment 1, $(-w, w)$ as compartment 2 and $(w, l+w)$ as compartment 3, which we will denote as l (left),

b (barrier), and r (right) respectively. We may discretize the continuum dynamics a la [7] to obtain master equations governing the probability of finding the particle in compartment i at any given time.

If we define

$$q_i(x) = e^{-\beta U(x)} \int_{x_{i-1}}^{x} dx' \, e^{\beta U(x')} \qquad (7)$$

for $x_0 = -l - w$, $x_1 = -w$, $x_2 = w$, $x_3 = l + w$ and

$$e^{-\beta G_i} = \int_{x_{i-1}}^{x_i} dx' \, e^{-\beta U(x')}, \qquad (8)$$

then the transition rates are given by

$$r_{i \to i+1} = D h_i^+[U] \ , \ r_{i+1 \to i} = D h_i^-[U] \qquad (9)$$

where

$$h_i^+[U] = \frac{e^{-\beta G_{i+1}}}{e^{-\beta G_i} \int_{x_i}^{x_{i+1}} dx \, q_i(x) - e^{-\beta G_{i+1}} \int_{x_{i-1}}^{x_i} dx \, q_i(x)} \qquad (10)$$

and

$$h_i^-[U] = \frac{e^{-\beta G_i}}{e^{-\beta G_i} \int_{x_i}^{x_{i+1}} dx \, q_i(x) - e^{-\beta G_{i+1}} \int_{x_{i-1}}^{x_i} dx \, q_i(x)}. \qquad (11)$$

A straightforward calculation determines the nontrivial

transition rates for our model system:

$$r_{l \to b} = \frac{2D}{l^2} \frac{1}{1 + 4 \left(\frac{w}{l}\right)^2 \frac{\eta}{\xi}}$$

$$r_{b \to l} = \frac{2D}{l^2} \frac{\eta}{\xi} \frac{1}{1 + 4 \left(\frac{w}{l}\right)^2 \frac{\eta}{\xi}}$$

$$r_{b \to r} = \frac{2D}{l^2} \frac{1}{4 \left(\frac{w}{l}\right)^2 + \frac{\xi}{1-\eta-\xi}}$$

$$r_{r \to b} = \frac{2D}{l^2} \frac{1}{1 + 4 \left(\frac{w}{l}\right)^2 \left(\frac{1-\eta-\xi}{\xi}\right)}. \qquad (12)$$

The quantities ξ and η serve as an intermediate coordinate system in the geometric construction of optimal protocols. They are explicitly written in terms of physical coordinates (V_b, V_l) in Eq. (42). The transition rates simplify considerably if we assume terms proportional to $(w/l)^2$ are negligible. Physically, this means that the width of the barrier is negligible compared to the width of the wells. In that case,

$$r_{l \to b} \approx \frac{2D}{l^2} \ , \ r_{b \to l} \approx \frac{2D}{l^2} \frac{\eta}{\xi}$$

$$r_{b \to r} \approx \frac{2D}{l^2} \frac{1 - \eta - \xi}{\xi} \ , \ r_{r \to b} \approx \frac{2D}{l^2}. \qquad (13)$$

Explicitly, the master equations [8] are

$$\frac{dp_i}{dt} = \sum_{j \neq i} r_{j \to i} \, p_j - \sum_{j \neq i} r_{i \to j} \, p_i. \qquad (14)$$

THE INVERSE DIFFUSION TENSOR

The components of the inverse diffusion matrix [6] are

$$\zeta_{ij}(\boldsymbol{\lambda}) \equiv \int_0^\infty dt' \left\langle \frac{\partial \phi}{\partial \lambda^i}(t') \, \frac{\partial \phi}{\partial \lambda^j}(0) \right\rangle_{\boldsymbol{\lambda}}. \qquad (15)$$

This matrix is symmetric and positive semi-definite in general for systems relaxing to an equilibrium state [9] and it defines a Riemannian geometry on the space of parameters.

We calculate the components of the inverse diffusion tensor using both the continuum and discrete dynamics. In the continuum case, the components are given by

$$\zeta_{ij}(\boldsymbol{\lambda}) = -\frac{1}{DZ(\boldsymbol{\lambda})} \int_{-(l+w)}^{l+w} dx \int_{-(l+w)}^{x} dx' \int_{-(l+w)}^{x'} dx'' \, \partial_{\lambda^i} \phi(x; \boldsymbol{\lambda}) \, \partial_{\lambda^j} \phi(x''; \boldsymbol{\lambda}) \, e^{\beta \left(U(x'; \boldsymbol{\lambda}) - U(x; \boldsymbol{\lambda}) - U(x''; \boldsymbol{\lambda})\right)}, \qquad (16)$$

which is a specific case of Eq. (32) constructed below. Here, $Z(\boldsymbol{\lambda}) \equiv \int_{-(l+w)}^{l+w} dx \, e^{-\beta U(x; \boldsymbol{\lambda})}$ is the classical partition function.

The procedure developed in [6, 10] for calculating components of the inverse diffusion tensor is specific to harmonic potentials. For more general potentials, the

Laplace transform may be used to extract the components of the tensor directly from the Fokker-Planck equation:

$$\partial_t \rho = D\big[\partial_x\left(\beta U'(x,t)\rho\right) + \partial_x^2\rho\big] \equiv -\partial_x G. \qquad (17)$$

Here, we focus on the most relevant situation in which

the potential is finite on an interval $[a,b]$ and equals $+\infty$ otherwise. We further impose reflecting-wall boundary conditions on the corresponding probability current; *i.e.* $G(a,t) = G(b,t) = 0$. The argument used below was adapted from [11].

We begin by rewriting Eq. (15) as

$$\zeta_{ij}(\boldsymbol{\lambda}) = \int_0^\infty dt' \left[\int_a^b dx_0 \; \rho_{eq}(x_0;\boldsymbol{\lambda}) \; \partial_{\lambda^j}\phi(x_0;\boldsymbol{\lambda}) \left(\int_a^b dx \; \rho(x,t';x_0) \; \partial_{\lambda^i}\phi(x;\boldsymbol{\lambda}) \right) \right], \qquad (18)$$

where $\rho(x,t;x_0)$ satisfies Eq. (17) with initial condition $\rho(x,t=0;x_0) = \delta(x-x_0)$ in addition to the reflecting-wall boundary conditions. For simplicity, define $m(t;x_0,\boldsymbol{\lambda}) \equiv \int_a^b dx \; \rho(x,t;x_0) \; \partial_{\lambda^i}\phi(x;\boldsymbol{\lambda})$ so that

$$\zeta_{ij}(\boldsymbol{\lambda}) = \int_a^b dx_0 \, \rho_{eq}(x_0;\boldsymbol{\lambda}) \, \partial_{\lambda^j}\phi(x_0;\boldsymbol{\lambda}) \int_0^\infty dt' \, m(t';x_0,\boldsymbol{\lambda}). \qquad (19)$$

Evaluating $\int_0^\infty dt' \; m(t';x_0,\boldsymbol{\lambda})$ is the first step towards calculating the inverse diffusion tensor components. We compute the Laplace transform of $m(t;x_0,\boldsymbol{\lambda})$

$$\hat{m}(s;x_0,\boldsymbol{\lambda}) \equiv \int_0^\infty dt' \; m(t';x_0,\boldsymbol{\lambda}) \; e^{-st'} \qquad (20)$$

and take the limit as $s \to 0^+$. For sake of clarity, we temporarily suppress the dependence upon x_0 and $\boldsymbol{\lambda}$.

Integrating by parts,

$$\int_0^\infty dt' \; \frac{dm}{dt'}(t') \; e^{-st'} = s \; \hat{m}(s) - m(0). \qquad (21)$$

Note that $m(\infty)$ vanishes since $\lim_{t\to\infty} \rho(x,t;x_0) = \rho_{eq}(x;\boldsymbol{\lambda})$; *i.e.* the system equilibrates after a sufficiently long time has elapsed. By definition of m,

$$m'(t) = \int_a^b dx \; \partial_t\rho(x,t;x_0) \; \partial_{\lambda^i}\phi(x;\boldsymbol{\lambda}). \qquad (22)$$

In terms of the probability current $G(x,t)$,

$$\hat{m}(s) = \frac{m(0) - \int_a^b dx \; \partial_x\hat{G}(x,s) \; \partial_{\lambda^i}\phi(x;\boldsymbol{\lambda})}{s}. \qquad (23)$$

Therefore, to compute $\hat{m}(s)$, we need the Laplace transform of the probability current.

The Fokker-Planck equation may be used to derive an equation for the probability current:

$$\partial_t G(x,t) = D\big[\beta U'(x;\boldsymbol{\lambda})\partial_x G(x,t) + \partial_x^2 G(x,t)\big]. \qquad (24)$$

Taking the Laplace transform of both sides, we have

$$s \, \hat{G}(x,s) - G(x,0) = D\big[\beta U'(x;\boldsymbol{\lambda}) \, \partial_x\hat{G}(x,s) + \partial_x^2\hat{G}(x,s)\big], \qquad (25)$$

which follows from $\lim_{t\to\infty} G(x,t) = 0$. Multiplying both sides by s and defining $H(x,s) \equiv s \, \hat{G}(x,s)$,

$$s \, H(x,s) - s \, G(x,0) = D\big[\beta U'(x;\boldsymbol{\lambda}) \, \partial_x H(x,s) + \partial_x^2 H(x,s)\big]. \qquad (26)$$

According to [11], we may obtain a solution to Eq. (26) by expanding $H(x,s)$ as a series in s. If we define $H(x,s) \equiv H_0(x) + s \, H_1(x) + \dots$, then

$$0 = \beta U'(x;\boldsymbol{\lambda}) \, H_0'(x) + H_0''(x) \qquad (27)$$

follows from substituting the expansion into Eq. (26) and comparing the coefficients of s^0 on both sides. The reflecting-wall boundary conditions $G(a,t) = G(b,t) = 0$ imply $H_i(a) = H_i(b) = 0$ for all i. Therefore, we conclude that $H_0(x) = 0$ identically.

Similarly,

$$-G(x,0) = D\big[\beta U'(x;\boldsymbol{\lambda}) \, \partial_x H_1(x) + \partial_x^2 H_1(x)\big]. \qquad (28)$$

The boundary conditions $H_1(a) = H_1(b) = 0$ and $\rho(x,0;x_0) = \delta(x-x_0)$ uniquely specify $H_1(x)$ as

$$H_1(x) = \theta(x - x_0) - \Pi(x;\boldsymbol{\lambda}) \qquad (29)$$

where $\Pi(x;\boldsymbol{\lambda}) = \int_a^x dx' \; \rho_{eq}(x;\boldsymbol{\lambda})$ and θ is the Heaviside step function.

Continuing this iterative process for finding H_i, we see that

$$H_2(x) = -\frac{1}{D} \int_a^x dx' \, \rho_{eq}(x';\boldsymbol{\lambda}) \left(\int_a^b dx'' \, e^{-\beta U(x'';\boldsymbol{\lambda})} \int_a^{x''} dx''' \, e^{\beta U(x''';\boldsymbol{\lambda})} \left[\theta(x''' - x_0) - \Pi(x''';\boldsymbol{\lambda}) \right] \right)$$

$$+ \frac{1}{D} \int_a^x dx' \, e^{-\beta U(x';\boldsymbol{\lambda})} \int_a^{x'} dx'' \, e^{\beta U(x'';\boldsymbol{\lambda})} \left[\theta(x'' - x_0) - \Pi(x'';\boldsymbol{\lambda}) \right]. \tag{30}$$

Fortunately, this is all that is needed as a short calculation using Eq. (23) shows that

$$\lim_{s \to 0^+} \hat{m}(s) = \int_0^\infty dt' \, m(t') = -\int_a^b dx \, \partial_x H_2(x) \, \partial_{\lambda^i} \phi(x;\boldsymbol{\lambda}). \tag{31}$$

Substituting this result into Eq. (19) and defining $Z(\boldsymbol{\lambda}) \equiv \int_a^b dx \, e^{-\beta U(x;\boldsymbol{\lambda})}$, we find

$$\zeta_{ij}(\boldsymbol{\lambda}) = -\frac{1}{DZ(\boldsymbol{\lambda})} \int_a^b dx \int_a^x dx' \int_a^{x'} dx'' \, \partial_{\lambda^i} \phi(x;\boldsymbol{\lambda}) \, \partial_{\lambda^j} \phi(x'';\boldsymbol{\lambda}) \, e^{\beta \left(U(x';\boldsymbol{\lambda}) - U(x;\boldsymbol{\lambda}) - U(x'';\boldsymbol{\lambda}) \right)}. \tag{32}$$

Since the potential is piecewise constant, it is possible to evaluate the iterated integrals in Eq. (16) explicitly. However, the resulting expressions are quite complicated when left in terms of the "physical" parameters V_b and V_l. It is mathematically advantageous at this stage to make a coordinate transformation in parameter space so that the metric tensor components are compact. Define

$$\eta \equiv \frac{le^{-\beta V_l}}{Z}, \quad \xi \equiv \frac{2we^{-\beta V_b}}{Z}, \tag{33}$$

where $Z = l + le^{-\beta V_l} + 2we^{-\beta V_b}$. The inverse diffusion tensor in this coordinate system is

$$\zeta^{cont}(\boldsymbol{\lambda}) = \frac{1}{D} \begin{pmatrix} \frac{l^2}{3} \frac{1-\xi}{\eta(1-\eta-\xi)} + \frac{4w^2}{\xi} & \frac{2w^2}{\xi} + \frac{l^2}{3} \frac{1}{1-\eta-\xi} \\ \frac{2w^2}{\xi} + \frac{l^2}{3} \frac{1}{1-\eta-\xi} & \frac{4w^2}{3\xi} + \frac{l^2}{3(1-\eta-\xi)} \end{pmatrix}. \tag{34}$$

Since we are ignoring $(w/l)^2$ terms,

$$\zeta^{cont}(\boldsymbol{\lambda}) \approx \frac{l^2}{3D} \frac{1}{1-\eta-\xi} \begin{pmatrix} \frac{1-\xi}{\eta} & 1 \\ 1 & 1 \end{pmatrix}. \tag{35}$$

For the discrete dynamics, Eq. (15) may be written as

$$\zeta_{ij}(\boldsymbol{\lambda}) = \int_0^\infty dt \sum_{\sigma,\sigma'} p_{\sigma'}(t|\sigma) \, p_\sigma^* \, \partial_{\lambda^j} \phi_\sigma \, \partial_{\lambda^i} \phi_{\sigma'}. \tag{36}$$

Here, p_σ^* denotes the equilibrium probability distribution

$$p_l^* = \eta, \quad p_b^* = \xi, \quad p_r^* = 1 - \eta - \xi, \tag{37}$$

$\phi_\sigma \equiv -\ln p_\sigma^*$, and $p_{\sigma'}(t|\sigma)$ represents the solution to the master equations Eq. (14) satisfying the initial condition $p_{\sigma'}(0|\sigma) = \delta_{\sigma,\sigma'}$ for *fixed* η, ξ.

To obtain $p_{\sigma'}(t|\sigma)$, we recognize that $p_b(t|\sigma) = 1 - p_l(t|\sigma) - p_r(t|\sigma)$ and write the master equations Eq. (14) with rates Eq. (13) as a linear system of equations:

$$\frac{d}{d\bar{t}} \begin{pmatrix} p_l \\ p_r \end{pmatrix} = -\begin{pmatrix} 1 + \frac{\eta}{\xi} & \frac{\eta}{\xi} \\ \frac{1-\eta-\xi}{\xi} & \frac{1-\eta}{\xi} \end{pmatrix} \begin{pmatrix} p_l \\ p_r \end{pmatrix} + \begin{pmatrix} \frac{\eta}{\xi} \\ \frac{1-\eta-\xi}{\xi} \end{pmatrix}, \tag{38}$$

where $\bar{t} \equiv \frac{2D}{l^2} t$. This system may be solved by standard methods [12]. We find

$$\begin{pmatrix} p_l(t|1) \\ p_r(t|1) \end{pmatrix} = \begin{pmatrix} \eta \left(\frac{\xi}{1-\xi} e^{-\frac{\bar{t}}{\xi}} + 1 \right) + \frac{1-\eta-\xi}{1-\xi} e^{-\bar{t}} \\ (1-\eta-\xi) \left(\frac{\xi}{1-\xi} e^{-\frac{\bar{t}}{\xi}} + 1 \right) - \frac{1-\eta-\xi}{1-\xi} e^{-\bar{t}} \end{pmatrix},$$

$$\begin{pmatrix} p_l(t|2) \\ p_r(t|2) \end{pmatrix} = \left(1 - e^{-\frac{\bar{t}}{\xi}} \right) \begin{pmatrix} \eta \\ 1-\eta-\xi \end{pmatrix},$$

$$\begin{pmatrix} p_l(t|3) \\ p_r(t|3) \end{pmatrix} = \begin{pmatrix} \eta \left(\frac{\xi}{1-\xi} e^{-\frac{\bar{t}}{\xi}} + 1 \right) - \frac{\eta}{1-\xi} e^{-\bar{t}} \\ (1-\eta-\xi) \left(\frac{\xi}{1-\xi} e^{-\frac{\bar{t}}{\xi}} + 1 \right) + \frac{\eta}{1-\xi} e^{-\bar{t}} \end{pmatrix}. \tag{39}$$

Evaluation of Eq. (36) using Eq. (39) yields

$$\zeta^{disc}(\boldsymbol{\lambda}) = \frac{l^2}{2D} \frac{1}{1-\eta-\xi} \begin{pmatrix} \frac{1-\xi}{\eta} & 1 \\ 1 & 1 \end{pmatrix}. \tag{40}$$

It follows that $\zeta^{cont}(\boldsymbol{\lambda}) = 2/3 \, \zeta^{disc}(\boldsymbol{\lambda})$ when $(w/l)^2$ is negligible. Both dynamics then yield precisely the same geodesics since Christoffel symbols are invariant under constant scalings of the metric tensor [13].

Ignoring constant prefactors, (approximate) optimal protocols are precisely the geodesics of the line element

$$d\Sigma^2 = \frac{1}{1-\eta-\xi} \left[\frac{(1-\xi)}{\eta} (d\eta)^2 + 2 \, d\eta d\xi + (d\xi)^2 \right]. \tag{41}$$

The Ricci scalar vanishes identically; therefore, there must exist a coordinate transformation in which the line element is Euclidean [13]. Indeed, if we define $x = 2\sqrt{1 - \eta - \xi}$, $y = 2\sqrt{\eta}$, then $d\Sigma^2 = dx^2 + dy^2$.

Geodesics are most conveniently calculated in (x, y)-coordinates. Physically, we should express quantities in terms of the (V_b, V_l)-coordinate system. For convenience, we list here the explicit formulae allowing us to transform between the two coordinate systems via the intermediate coordinates (η, ξ):

$$\eta = \frac{le^{-\beta V_l}}{l + le^{-\beta V_l} + 2we^{-\beta V_b}} \ , \ \xi = \frac{2we^{-\beta V_b}}{l + le^{-\beta V_l} + 2we^{-\beta V_b}},$$
$$x = 2\sqrt{1 - \eta - \xi} \ , \ y = 2\sqrt{\eta}, \qquad (42)$$

$$\xi = 1 - \frac{1}{4}\left(x^2 + y^2\right) \ , \ \eta = \frac{y^2}{4},$$
$$\beta V_b = \ln\left[\frac{2w}{l}\frac{1 - \eta - \xi}{\xi}\right] \ , \ \beta V_l = \ln\left[\frac{1 - \eta - \xi}{\eta}\right]. \quad (43)$$

As originally constructed [9], the formulation of the inverse diffusion tensor assumes smooth protocols on the entire domain of definition. Here we constructed an approximation for the optimizer in the interior of the erasure stage.

The question of what endpoints should be selected for the approximate optimizer is answered by enforcing the probability constraints at the endpoints. This is achieved by using the derivative truncation method [6, 10] to obtain the approximate time evolution of the probability distributions when the parameters are adjusted according to the geodesic protocol. This in turn gives us approximate formulae for the probability distributions at the beginning and end of the erasure stage in terms of the endpoints of the geodesic. A numerical solution of these constraints is easily obtained.

EXACT OPTIMIZER

For the discrete dynamics, Eq. (6) simplifies to

$$\langle \beta W \rangle_{\mathbf{\Lambda}} = \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{\boldsymbol{\lambda}} \cdot \sum_i p_i(\bar{t}) \frac{\partial \phi_i}{\partial \boldsymbol{\lambda}} \right] + \beta \triangle F$$
$$+ \sum_i p_i(t_f) \, \beta U_i(\bar{t})|_{\bar{t}=t_f^-}^{\bar{t}=\overline{t_f}} - \sum_i p_i(0) \, \beta U_i(\bar{t})|_{\bar{t}=0^+}^{\bar{t}=0} \quad (44)$$

where

$$\beta \triangle F \equiv \ln\left[\frac{l + le^{-\beta V_l(0^+)} + 2we^{-\beta V_b(0^+)}}{l + le^{-\beta V_l(t_f^-)} + 2we^{-\beta V_b(t_f^-)}}\right]. \quad (45)$$

Integrating by parts, $\langle \beta W \rangle_{\mathbf{\Lambda}}$ equals

$$\sum_i p_i(t_f)\beta U_i(\overline{t_f}) - \sum_i p_i(0)\beta U_i(0) - \int_{0^+}^{\overline{t_f}^-} d\bar{t} \sum_i \phi_i \dot{p}_i. \quad (46)$$

It follows immediately that the average work performed during the reset stage of the cycle is given by

$$\langle \beta W \rangle_{\mathbf{\Lambda}_{reset}} = \sum_i p_i(t_f) \left[\beta U_i(0) - \beta U_i(t_f) \right] \quad (47)$$

and so

$$\langle \beta W \rangle_{\mathbf{\Lambda}_{cycle}} = \sum_i \left[p_i(t_f) - p_i(0) \right] \beta U_i(0) - \int_{0^+}^{\overline{t_f}^-} d\bar{t} \sum_i \phi_i \dot{p}_i. \quad (48)$$

By the first law of stochastic thermodynamics,

$$\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} = - \int_{0^+}^{\overline{t_f}^-} d\bar{t} \sum_i \phi_i \dot{p}_i. \quad (49)$$

The average heat over the cycle is given explicitly by

$$\int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{p}_l \ln\left(\eta\right) + \dot{p}_b \ln\left(\xi\right) + \dot{p}_r \ln\left(1 - \xi - \eta\right) \right], \quad (50)$$

or, since $p_b = 1 - p_l - p_r$,

$$\int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{p}_l \ln\left(\frac{\eta}{\xi}\right) + \dot{p}_r \ln\left(\frac{1 - \xi - \eta}{\xi}\right) \right]. \quad (51)$$

We may write this integral and hence the total average heat lost over the cycle explicitly as a functional of the probabilities and then optimize over these variables. We begin by rewriting Eq. (38) as

$$\frac{1}{1 - p_l - p_r} \begin{pmatrix} \dot{p}_l + p_l \\ \dot{p}_r + 1 - p_l \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{\eta}{\xi} \\ \frac{1}{\xi} \end{pmatrix}, \quad (52)$$

which allows us to solve for η and ξ explicitly in terms of the probabilities and their time derivatives:

$$\xi = \frac{1 - p_l - p_r}{1 + \dot{p}_l + \dot{p}_r} \ , \ \eta = \frac{\dot{p}_l + p_l}{1 + \dot{p}_l + \dot{p}_r}. \quad (53)$$

Using these expressions and performing another integration by parts,

$$\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} = Q_b + \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{p}_l \ln\left(\dot{p}_l + p_l\right) + \dot{p}_r \ln\left(\dot{p}_r + p_r\right) \right] \quad (54)$$

for boundary term

$$Q_b \equiv p_b(t_f) \left[\ln\left(p_b(t_f)\right) - 1 \right] - p_b(0) \left[\ln\left(p_b(0)\right) - 1 \right]. \quad (55)$$

The boundary term Q_b depends only on the probability distributions at the endpoints. Moreover, we see that the "bulk" term of the average heat functional is a sum $I[p_l] + I[p_r]$ for

$$I[z] \equiv \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{z} \ln\left(\dot{z} + z\right) \right]. \quad (56)$$

Therefore, for all variations δp_i (for a fixed i) vanishing at the endpoints,

$$\delta \langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} = \delta I[p_i] \qquad (57)$$

implies that each component of a stationary probability distribution of the average heat must satisfy the Euler-Lagrange equation for S.

Suppose $z(t)$ satisfies the Euler-Lagrange equation for the Lagrangian $L[z, \dot{z}] \equiv \dot{z} \ln (\dot{z} + z)$. Then it must be true that

$$\dot{z} \frac{\partial L}{\partial \dot{z}} - L = \frac{(\dot{z})^2}{\dot{z} + z} \qquad (58)$$

is a constant.

Therefore, probability distributions extremizing $\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}}$ satisfy $\dot{p}_i = K_i (\dot{p}_i + p_i)$ for positive constants K_i and $i = l, r$. Over the course of the erasure stage, p_l (p_r) decreases (increases). Hence,

$$\dot{p}_l = \frac{K_l}{2} \left(1 - \sqrt{1 + \frac{4}{K_l} p_l} \right),$$
$$\dot{p}_r = \frac{K_r}{2} \left(1 + \sqrt{1 + \frac{4}{K_r} p_r} \right). \qquad (59)$$

These equations may be integrated to obtain implicit expressions for the optimal probability time course:

$$\sqrt{1 + \frac{4p_l(0)}{K_l}} - \sqrt{1 + \frac{4p_l(\bar{t})}{K_l}} + \ln \left[\frac{\sqrt{1 + \frac{4p_l(0)}{K_l}} - 1}{\sqrt{1 + \frac{4p_l(\bar{t})}{K_l}} - 1} \right] = \bar{t} \qquad (60)$$

and

$$\sqrt{1 + \frac{4p_r(\bar{t})}{K_r}} - \sqrt{1 + \frac{4p_r(0)}{K_r}} - \frac{1}{2} \ln \left[\frac{1 + \sqrt{1 + \frac{4p_r(\bar{t})}{K_r}}}{\sqrt{1 + \frac{4p_r(\bar{t})}{K_r}} - 1} \right]$$
$$+ \frac{1}{2} \ln \left[\frac{1 + \sqrt{1 + \frac{4p_r(0)}{K_r}}}{\sqrt{1 + \frac{4p_r(0)}{K_r}} - 1} \right] - \frac{1}{2} \ln \left(\frac{p_r(\bar{t})}{p_r(0)} \right) = \bar{t}. \qquad (61)$$

The constants K_i may in turn be numerically fixed by imposing the constraints $p_l(0) = \frac{1}{2(1+\gamma)} = p_r(0)$, $p_l(t_f) = \delta$, $p_r(t_f) = 1 - 2\delta$. The optimal probability time course is completely determined once the K_i are set. As a consequence, Q_b and also $\eta(t)$ and $\xi(t)$ for $t \in (0, t_f)$ (see Eq. (53)) are completely determined.

OPTIMAL EFFICIENCY

Recall that the optimal time course for the probability distribution is given by $\dot{p}_i = K_i (\dot{p}_i + p_i)$, where $i = l, r$ and $K_i > 0$. It is possible to use this relation to obtain an upper bound on the K_i in terms of the protocol duration assuming that the probabilities do not vanish

during the time course. This assumption is valid given the monotonicity of the probability time course and the non-vanishing initial and final values for the p_i.

Rearranging the equation for the optimal time course, we have

$$\left(\frac{1 - K_i}{K_i} \right) \dot{p}_i = p_i. \qquad (62)$$

Integrating both sides over the duration of the protocol, we have

$$\left(\frac{1 - K_i}{K_i} \right) \triangle p_i = \int_0^{\overline{t_f}} d\bar{t} \, p_i, \qquad (63)$$

where the right-hand side may be bounded by $\overline{t_f} \, p_i^{min} \leq \int_0^{\overline{t_f}} dt \, p_i \leq \overline{t_f} \, p_i^{max}$.

It follows immediately that

$$\frac{1}{1 + \overline{t_f} \left(\frac{p_i^{max}}{\triangle p_i} \right)} \leq K_i \leq \frac{1}{1 + \overline{t_f} \left(\frac{p_i^{min}}{\triangle p_i} \right)} \qquad (64)$$

if $\triangle p_i > 0$; a similar set of inequalities holds if $\triangle p_i < 0$. Therefore, in the long duration regime, K_i is a very small quantity. This observation is important in determining the optimal efficiency in the long time regime. Recall that

$$\langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} = Q_b + \int_{0^+}^{\overline{t_f}^-} d\bar{t} \left[\dot{p}_l \ln (\dot{p}_l + p_l) + \dot{p}_r \ln (\dot{p}_r + p_r) \right] \qquad (65)$$

for boundary term

$$Q_b \equiv p_b(t_f) \left[\ln (p_b(t_f)) - 1 \right] - p_b(0) \left[\ln (p_b(0)) - 1 \right] \qquad (66)$$

and $\epsilon \equiv (-\triangle S) / \left(k_B \langle \beta Q \rangle_{\mathbf{\Lambda}_{cycle}} \right)$ for change in Shannon entropy

$$\triangle S \equiv S_f - S_i \equiv - \sum_i p_i^f \ln p_i^f + \sum_i p_i^o \ln p_i^o. \qquad (67)$$

To find the approximate efficiency in the long duration limit, we must approximate the integrals $\int_{0^+}^{\overline{t_f}^-} d\bar{t} \, \dot{p}_i \ln (\dot{p}_i + p_i)$. We first consider $i = l$: since

$$\dot{p}_l = \frac{K_l}{2} \left(1 - \sqrt{1 + \frac{4}{K_l} p_l} \right), \qquad (68)$$

we may rewrite the integral as

$$\int_{p_l^o}^{p_l^f} dz \, \ln \left[\frac{K_l}{2} \left(1 - \sqrt{1 + \frac{4}{K_l} z} \right) + z \right] \qquad (69)$$

after a change of variables. We may factorize the logarithm as

$$\ln \left[\frac{K_l}{2z} \left(1 - \sqrt{1 + \frac{4}{K_l} z} \right) + 1 \right] + \ln(z). \qquad (70)$$

In the long duration limit, we have

$$\ln\left[\frac{K_l}{2z}\left(1-\sqrt{1+\frac{4}{K_l}z}\right)+1\right]\approx-\sqrt{\frac{K_l}{z}}. \qquad (71)$$

Therefore,

$$\int_{p_l^o}^{p_l^f}dz\ \ln\left[\frac{K_l}{2}\left(1-\sqrt{1+\frac{4}{K_l}z}\right)+z\right]\approx$$
$$p_l^f\ln p_l^f-p_l^f-p_l^o\ln p_l^o+p_l^o-2\sqrt{K_l}\left(\sqrt{p_l^f}-\sqrt{p_l^o}\right). \qquad (72)$$

A similar calculation demonstrates

$$\int_{p_r^o}^{p_r^f}dz\ \ln\left[\frac{K_r}{2}\left(1+\sqrt{1+\frac{4}{K_r}z}\right)+z\right]\approx$$
$$p_r^f\ln p_r^f-p_r^f-p_r^o\ln p_r^o+p_r^o+2\sqrt{K_r}\left(\sqrt{p_r^f}-\sqrt{p_r^o}\right). \qquad (73)$$

The total average heat dissipated during the cycle is equal to the sum of these two integrals plus the boundary term Eq. (66). The first two terms of each integral approximation plus the boundary term simply yield the Landauer term $-\triangle S$. This follows from the observation that the sum over p_i for $i=l,b,r$ must be 1 both at the beginning and end of the time course.

Therefore,

$$\langle\beta Q\rangle_{\mathbf{\Lambda}_{cycle}}\approx-\triangle S/k_B+2\sqrt{K_r}\left(\sqrt{p_r^f}-\sqrt{p_r^o}\right)$$
$$-2\sqrt{K_l}\left(\sqrt{p_l^f}-\sqrt{p_l^o}\right). \qquad (74)$$

Furthermore, since

$$\frac{\dot{p}_l}{p_l}=\frac{K_l}{2p_l}\left(1-\sqrt{1+\frac{4}{K_l}p_l}\right)\approx-\sqrt{\frac{K_l}{p_l}} \qquad (75)$$

and

$$\frac{\dot{p}_r}{p_r}=\frac{K_r}{2p_r}\left(1+\sqrt{1+\frac{4}{K_r}p_r}\right)\approx\sqrt{\frac{K_r}{p_r}}, \qquad (76)$$

we find that

$$-\sqrt{K_l}\ \overline{t_f}\approx2\left(\sqrt{p_l^f}-\sqrt{p_l^o}\right),$$
$$\sqrt{K_r}\ \overline{t_f}\approx2\left(\sqrt{p_r^f}-\sqrt{p_r^o}\right). \qquad (77)$$

It follows that in the long duration limit,

$$\langle\beta Q\rangle_{\mathbf{\Lambda}_{cycle}}\approx\frac{-\triangle S}{k_B}+\frac{4}{\overline{t_f}}\left[\left(\triangle\sqrt{p_r}\right)^2+\left(\triangle\sqrt{p_l}\right)^2\right]. \qquad (78)$$

---

* pzulkowski@berkeley.edu

† deweese@berkeley.edu

[1] E. Aurell, K. Gawedzki, C. Mejia-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, J. Stat. Phys. **147** (2012).

[2] U. Seifert, Reports on Progress in Physics **75**, 126001 (2012).

[3] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).

[4] H. Then and A. Engel, Phys. Rev. E **77** (2008).

[5] T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).

[6] P. R. Zulkowski, D. A. Sivak, and M. R. DeWeese, "Optimal control of transitions between nonequilibrium steady states," (2013), arXiv:1303.6596.

[7] J. Xing, H. Wang, and G. Oster, Biophys. J. **89**, 1551 (2005).

[8] R. Zwanzig, *Nonequilibrium statistical mechanics* (Oxford University Press, New York, 2001).

[9] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 190602 (2012).

[10] P. R. Zulkowski, D. A. Sivak, G. E. Crooks, and M. R. DeWeese, Phys. Rev. E **86** (2012).

[11] A. Pankratov, Phys. Lett. A **255**, 17 (1999).

[12] W. Boyce and R. DiPrima, *Elementary differential equations and boundary value problems*, seventh ed. (Wiley, New York, 2000).

[13] M. P. do Carmo, *Riemannian geometry* (Birkhäuser, Boston, 1992).