Unified, Geometric Framework for Nonequilibrium Protocol Optimization

Shriram Chennakesavalu* and Grant M. Rotskoff†
Department of Chemistry, Stanford University, Stanford, California 94305, USA

(Received 2 May 2022; accepted 7 February 2023; published 6 March 2023)

Controlling thermodynamic cycles to minimize the dissipated heat is a long-standing goal in thermodynamics, and more recently, a central challenge in stochastic thermodynamics for nanoscale systems. Here, we introduce a theoretical and computational framework for optimizing nonequilibrium control protocols that can transform a system between two distributions in a minimally dissipative fashion. These protocols optimally transport a system along paths through the space of probability distributions that minimize the dissipative cost of a transformation. Furthermore, we show that the thermodynamic metric—determined via a linear response approach—can be directly derived from the same objective function that is optimized in the optimal transport problem, thus providing a unified perspective on thermodynamic geometries. We investigate this unified geometric framework in two model systems and observe that our procedure for optimizing control protocols is robust beyond linear response.

DOI: 10.1103/PhysRevLett.130.107101

Understanding how to efficiently control thermodynamic cycles is a truly foundational problem in thermodynamics. Our modern mathematical framework for macroscopic thermodynamics emerged from efforts to describe the transfer of heat into work and to quantify the wasted or excess heat dissipated to the environment [1]. As it has become possible to probe the thermodynamics of nanoscale systems, both experimentally and in computer simulations, the role of thermal fluctuations has reoriented our interpretation of the fundamental constraints imposed by the second law; at small scales, fluctuation theorems precisely quantify the relationship between entropy production and irreversible dynamics [2–4]. Exploiting nonequilibrium dynamics to understand equilibrium properties like free energy differences has been realized both experimentally and computationally via the Jarzynski equality [5,6]. However, the statistical accuracy of such calculations requires minimizing dissipation over the nonequilibrium ensemble of trajectories by choosing an appropriate external driving protocol [7,8]. What is more, fundamental questions about the design and properties of nanoscale machines from biology to engineering require theoretical tools to carefully interrogate dissipation in stochastic systems.

Despite the importance of measuring dissipation, it has proved challenging to do so accurately in nanoscale systems, with only indirect proxies available. While bounds like the thermodynamic uncertainty relations [9,10] can be used to aid inference, these relations do not necessarily tightly constrain the dissipation and cannot be directly correlated with it in general [11]. Nevertheless, recent experimental and computational advances have reinvigorated efforts to design optimal controllers for nanoscale systems. Optimizing a protocol through the use of "thermodynamic geometry" [12–16], an approach in which the dissipation is quantified through a Riemannian path length in the space of protocols, has proved among the most productive strategies for this problem [8,17–19]. The metric itself is derived via a perturbative expansion [8] and hence applies only in the limit of driving that is sufficiently slow or when the magnitude of the perturbation is sufficiently small.

Separately, initially spurred by developments in the study of variational solutions to certain partial differential equations [20–22], a distinct geometry, based on optimal transport theory, has been connected to nonequilibrium dissipation. In this formulation, distances are measured not with a Riemannian metric but directly between probability distributions by determining a minimum cost transport plan that moves the probability mass from an initial distribution ρ_A to a given target ρ_B. The cost defines the Wasserstein metric, which, in the Monge formulation, is formally defined through an optimization problem,

$$\mathcal{W}_2^2(\rho_A, \rho_B) = \inf_T \int_\Omega |\boldsymbol{x} - T(\boldsymbol{x})|^2 \rho_A(\boldsymbol{x}) d\boldsymbol{x}, \qquad (1)$$

where T ranges over all valid maps or transportation plans that send ρ_A to ρ_B. The Wasserstein metric is a lower bound on the dissipative cost to transform ρ_A to ρ_B in a finite time, and, importantly, provides an alternate geometric framework for minimizing dissipation [23]. Unlike the perturbative formulation that leads to the thermodynamic Riemannian metric, this approach makes no approximation to quantify the total change entropy along a dissipative transformation. However, the constrained minimization problem (1) that one must solve to compute the Wasserstein distance is notoriously challenging, both analytically and numerically [24].

Here, we introduce a theoretical and computational framework for optimizing control protocols that realize

geodesics in the Wasserstein metric, which we refer to as displacement interpolations [25]. These are paths through the space of probability distributions that minimize the dissipative cost of the finite-time transformation. Furthermore, we show that the thermodynamic metric can be derived directly from the same objective function that we optimize in the optimal transport problem, emphasizing that the two coincide in the limit of slow driving. Our result provides a unified geometric framework for minimizing the dissipative costs of nonequilibrium transformations. Crucially, the approach we propose is numerically tractable without globally computing the thermodynamic metric, a numerically costly procedure, especially when the dimensionality of the protocol is large. We investigate this approach to minimum dissipation control in two simple models of nanoscale engines, and we compare control protocols determined via both the thermodynamic metric and the geometry of optimal transport. Remarkably, we observe that our procedure for optimizing control protocols is robust outside the linear response regime, which leads to a significant improvement in protocol design over the thermodynamic metric when the driving is fast.

Connection between the thermodynamic metric and the optimal transport problem.—We consider the problem of transforming an initial equilibrium distribution $\rho_A = \rho(\cdot, 0)$ into a target distribution $\rho_B = \rho(\cdot, t_f)$ with a nonequilibrium driving protocol $\boldsymbol{\lambda}$ of duration t_f. At sufficiently long times, we assume that the system relaxes to an equilibrium distribution $\rho_0(\boldsymbol{x}) = e^{-\beta U(\boldsymbol{x},\boldsymbol{\lambda})}/Z(\boldsymbol{\lambda})$, where U denotes the potential energy of the system. The results we derive below apply to systems that evolve according to overdamped Langevin dynamics, though we see that the methods we develop also apply to open quantum systems [26]. While we do not consider underdamped Langevin dynamics here, minimizing dissipation in these systems is an active area of inquiry [27,28].

We consider first a system with coordinates $\boldsymbol{x} \in \Omega \subset \mathbb{R}^d$ subject to the following overdamped Langevin equation,

$$\dot{\boldsymbol{x}} = -\nabla U(\boldsymbol{x}, \boldsymbol{\lambda}(t)) + \sqrt{2\beta^{-1}}\boldsymbol{\eta}(t), \tag{2}$$

where $\boldsymbol{\eta}$ is a Gaussian random variable with $\langle \boldsymbol{\eta}(t) \rangle = 0$ and $\langle \boldsymbol{\eta}_i(t)\boldsymbol{\eta}_j(t') \rangle = \delta(t - t')\delta_{ij}$. The external protocol $\boldsymbol{\lambda}$ changes as a function of time, which drives the system away from equilibrium. As a result, the time-dependent distribution $\rho(\boldsymbol{x}, t)$ of states may be non-Boltzmann, but it does satisfy a Fokker-Planck equation,

$$\partial_t \rho + \nabla \cdot (\boldsymbol{v}\rho) = 0, \tag{3}$$

where the local velocity is

$$\boldsymbol{v}(\boldsymbol{x}, t) = -\nabla U(\boldsymbol{x}(t), \boldsymbol{\lambda}(t)) - \beta^{-1}\nabla \log \rho(\boldsymbol{x}, t). \tag{4}$$

Nonconservative forces can be incorporated into this framework and only change the expression for the local velocity [23]. The total entropy production can be written in terms of the heat flow to the environment using the stochastic thermodynamics convention for the heat flow [29,30],

$$\beta \mathcal{Q} = -\int \beta(t) \nabla U(\boldsymbol{x}(t), \boldsymbol{\lambda}(t)) \circ d\boldsymbol{x}(t), \tag{5}$$

and the Gibbs entropy of the system. Combining these terms, we obtain a quadratic form for the total entropy production along a nonequilibrium transformation [23,31,32],

$$\Delta\Sigma_{\mathrm{tot}} = \int_0^{t_f} \beta(t) \int_\Omega \boldsymbol{v}^T(\boldsymbol{x}, t)\boldsymbol{v}(\boldsymbol{x}, t)\rho(\boldsymbol{x}, t) d\boldsymbol{x} dt. \tag{6}$$

Hence, to minimize the dissipation associated with a transformation from the thermodynamic state specified by $\boldsymbol{\lambda}(0)$ to a state in which the control parameters are fixed at $\boldsymbol{\lambda}(t_f)$, we must identify a minimizer $\boldsymbol{\lambda}_*(t)$ of Eq. (6), noting that local velocity and the density both depend on $\boldsymbol{\lambda}$. Formally, we solve

$$\boldsymbol{\lambda}_* = \underset{\boldsymbol{\lambda}:[0,t_f]\to\mathbb{R}^k}{\operatorname{argmin}} \quad \underset{\rho(\cdot,0)=\rho_A,\quad \rho(\cdot,t_f)=\rho_B}{} \Delta\Sigma_{\mathrm{tot}}[\boldsymbol{\lambda}]. \tag{7}$$

This minimization problem has a geometric interpretation: the minimum of Eq. (6) over all $(\boldsymbol{v}, \rho)$ satisfying Eqs. (3) and (4) with the boundary conditions that $\rho(\cdot, 0) = \rho_A$ and $\rho(\cdot, t_f) = \rho_B$ are exactly the Benamou-Brenier formulation of the Wasserstein optimal transport distance [22,33]. In the present work, we optimize not over velocity fields and densities but rather protocol $\boldsymbol{\lambda}$ which, when sufficiently flexible, provides substantial control over the distribution. While this constraint means that we may not saturate the optimal transport distance in general, the examples we consider here are ones in which the protocol provides complete control over the distribution. Furthermore, we believe that a protocol optimization framework is more physically practical than a setting in which the requisite velocity fields to minimize Eq. (6) are inaccessible to any external controller. Minimizing Eq. (6) provides an alternative formulation of the Wasserstein optimal transport problem defined in Eq. (1), as explained in the Supplemental Material [32]. The connection between optimal transport and dissipation is well known in the partial differential equations literature [34] and was subsequently connected to stochastic thermodynamics by Aurell and co-workers [23,31]; recently this connection was discussed in Ref. [35] to provide a framework for general thermodynamic speed limits (cf. Refs. [12,14,23]).

Computing the minimizer $\boldsymbol{\lambda}_*$ analytically is generically challenging because of the nonlinear dependence of the distribution on the protocol. Within a linear response approximation, the form of the distribution simplifies considerably, and the quadratic functional can be written

as an explicit function of $\boldsymbol{\lambda}$; we now demonstrate that carrying out the minimization of Eq. (6) with respect to $\boldsymbol{\lambda}$ recovers the Riemannian thermodynamic metric. Following the dynamical linear response framework of Zwanzig [36] (see also Ref. [37]), we assume that $\boldsymbol{\lambda}$ changes slowly relative to the rate of relaxation of the system. We expand the instantaneous density around the equilibrium density with the control parameters fixed:

$$\rho(\boldsymbol{x},t)=\rho_0(\boldsymbol{x},\boldsymbol{\lambda}(t))+\epsilon\rho_1(\boldsymbol{x},\boldsymbol{\lambda}(t))+\mathcal{O}(\epsilon^2). \tag{8}$$

At first order in ϵ we obtain

$$\rho_1(\boldsymbol{x},t)=\beta\rho_0(\boldsymbol{x},t)\int_0^\infty \delta\boldsymbol{\Theta}(\boldsymbol{x}^{\boldsymbol{\lambda}}(s),t)\cdot\dot{\boldsymbol{\lambda}}(t)ds, \tag{9}$$

where $\delta\boldsymbol{\Theta}(\boldsymbol{x},s)$ denotes the deviation of the generalized forces from their average, $\partial_{\boldsymbol{\lambda}}U(\boldsymbol{x},\boldsymbol{\lambda}(t))-\partial_{\boldsymbol{\lambda}}F(\boldsymbol{\lambda}(t))$. The only part of the quadratic functional that depends on $\boldsymbol{\lambda}$ is the work; to leading order in ϵ an explicit computation [32] shows that the protocol-dependent dissipation is

$$L[\boldsymbol{\lambda}]=\int_0^{t_f}\dot{\boldsymbol{\lambda}}^T(t)\zeta(\boldsymbol{\lambda}(t))\dot{\boldsymbol{\lambda}}(t)dt, \tag{10}$$

with

$$\zeta(\boldsymbol{\lambda}(t))=\beta\int_0^\infty\langle\delta\boldsymbol{\Theta}(s)\delta\boldsymbol{\Theta}^T(0)\rangle_{\boldsymbol{\lambda}(t)}ds. \tag{11}$$

Minimizing the expression (10) yields a geodesic with respect to the positive definite symmetric form ζ; these geodesics in protocol space are minimum dissipation protocols within the linear response approximation [12,17]. The $\mathcal{W}_2$-optimal local velocity associated with the thermodynamic geodesic is explicitly given by the negative, temperature scaled spatial gradient of the perturbative correction $\epsilon\rho_1$, simply plugging Eq. (8) into Eq. (4).

Computational approach.—We consider two general paradigms for determining protocols that minimize Eq. (6). The first involves specifying a set of intermediate distributions and learning a protocol $\boldsymbol{\lambda}_*(t)$ that drives the system along a Wasserstein geodesic between these intermediate distributions. Alternatively, we specify a protocol $\boldsymbol{\lambda}(t)$ and determine an optimal speed function $\phi_*(t)$ such that the system driven under $\boldsymbol{\lambda}(\phi_*(t))$ minimizes Eq. (6). The first paradigm ensures that an engine reaches the desired intermediate distribution even when the driving is fast relative to the relaxation time of the system. The latter approach constrains the protocol but is less computationally and experimentally demanding.

To drive the system through prespecified intermediate distributions (see Fig. 1), we use gradient-based optimization to learn protocols $\boldsymbol{\lambda}_*(t)$. Given an initial distribution ρ_A and a final distribution ρ_B, we compute a Wasserstein geodesic that interpolates these two distributions. This path

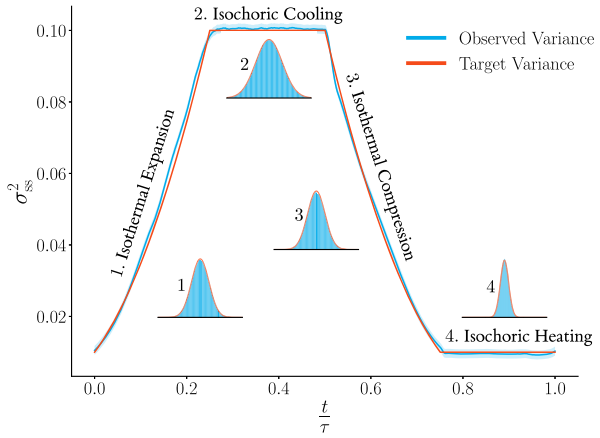

FIG. 1. Brownian engine driven according to a Stirling cycle through target intermediate Gaussian distributions with mean $\mu=0$ and variance σ^2_{ss}. Variance of the target distributions on the Wasserstein geodesics (in red) along each stage are computed exactly between corresponding end points. In the fast-driving regime ($\tau=10$), observed variance of trajectories under $\boldsymbol{\lambda}_*$ closely approximate target variance.

through the space of probability distributions is known as a displacement interpolation [33,34]. Computing this geodesic exactly is computationally demanding for arbitrary ρ_A and ρ_B, so, for a general system, one must instead estimate a displacement interpolation using computational optimal transport algorithms, such as the Sinkhorn algorithm [24]. We anticipate that this approximate calculation of the Wasserstein geodesic will be tractable even for relatively high-dimensional systems [38]. In the present work, we do not directly contend with this approximation; for the two minimal models considered here, we can compute the Wasserstein geodesics exactly.

Determining a displacement interpolation $\rho_*(t)$ does not directly yield a protocol that drives the system through the set of probability distributions that constitute the geodesic. While it would be possible to identify $\boldsymbol{\lambda}_*$ from ρ_* in the quasistatic limit, when the system is driven far from equilibrium, we must optimize the protocol so that the empirical distribution remains close to $\rho_*(t)$. We represent $\boldsymbol{\lambda}_*$ using a neural network and carry out our optimization using automatic differentiation. The gradient information is stored throughout the dynamical evolution of the system, and gradients are explicitly backpropagated through the trajectory. Automatic differentiation is a natural way to optimize $\boldsymbol{\lambda}_*$ for the nonequilibrium systems considered here, as it enables dynamical information to be implicitly incorporated into a learned protocol. We note that for large protocol durations, it is impractical to differentiate throughout the entire trajectory; instead, we carry out our optimization over short-time intervals.

Optimizing a nanoscale Brownian engine.—We consider a minimal model for a nanoscale Brownian engine, consisting of a Brownian particle in a harmonic potential.

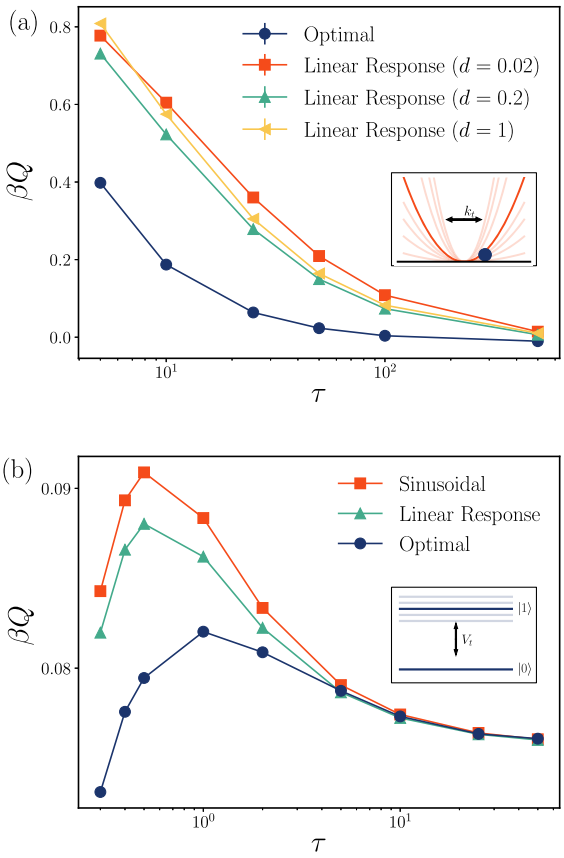

FIG. 2. The dissipation profiles for two nanoscale engines plotted across different protocol durations τ. (a) For a nanoscale Brownian engine (see inset), the dissipation for the optimal protocol $\boldsymbol{\lambda}_*$ (dark blue circle) is lower than the dissipation for protocols computed according to a linear response approximation $\boldsymbol{\lambda}_{\mathrm{LR}}$ for different smoothness parameters d [40]. (b) For a superconducting qubit (see inset), the optimal protocol $\boldsymbol{\lambda}(\phi_*)$ (dark blue circle) is less dissipative than a protocol computed using a linear response approximation $\boldsymbol{\lambda}(\phi_{\mathrm{LR}})$ (green triangle) [19] or a base sinusoidal protocol $\boldsymbol{\lambda}(t)$ (red square).

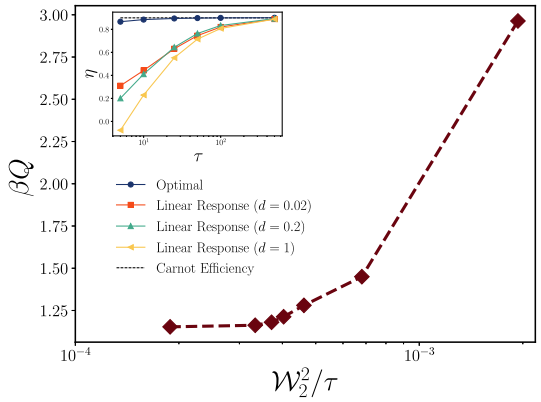

FIG. 3. The dissipation plotted against $\mathcal{W}_2^2/\tau$ (the squared Wasserstein distance normalized by protocol duration) for the isothermal compression of a Brownian engine across different τ. An increasing dissipative cost is incurred as $\mathcal{W}_2^2/\tau$ increases, consistent with Eq. (6). Inset: the efficiencies of the Brownian engine plotted against the protocol duration τ for a learned optimal protocol $\boldsymbol{\lambda}_*$ and protocols determined from a linear approximation [40]. Even for fast drivings, the Brownian engine driven by $\boldsymbol{\lambda}_*$ (dark blue circle) can achieve near Carnot efficiency and has a higher efficiency in comparison to the Brownian engine driven by $\boldsymbol{\lambda}_{\mathrm{LR}}$.

The temperature T of the heat bath and the strength of the harmonic potential k are controlled to mimic a Stirling cycle [39]. The model is depicted schematically in the inset of Fig. 2(a). In our implementation, each stage of the engine cycle has a fixed duration of $\tau/4$. Given T_h and T_c, the maximum and minimum allowed temperature of the bath, and k_h and k_l, the maximum and minimum allowed strength of the harmonic potential, we can exactly specify the equilibrium distributions at the end points of each step of the Stirling cycle, which are Gaussian with $\mu = 0$ and $\sigma^2 = T/k$. Because the distributions are Gaussian, we can exactly determine a displacement interpolation ρ_* between the end points of each stage.

We carry out this optimization for a range of protocol durations τ, the shortest of which are far from the linear response regime and the longest of which are essentially quasistatic. For each τ we consider, we optimize the protocol $\boldsymbol{\lambda}_*$ separately. We compare our results with optimal protocols determined using the linear response approximation [40], which we denote $\boldsymbol{\lambda}_{\mathrm{LR}}$. Here, $\boldsymbol{\lambda}_{\mathrm{LR}}$ also drives the engine along a Stirling cycle and has the same T_h, T_c, k_h, and k_l as $\boldsymbol{\lambda}_*$. Importantly, we observe that, in the fast-driving regime, the Brownian engine driven under $\boldsymbol{\lambda}_*$ is significantly less dissipative than the Brownian engine driven under $\boldsymbol{\lambda}_{\mathrm{LR}}$, as shown in Fig. 2. As the engine undergoes isothermal compression—the most dissipative step—the Brownian engine under $\boldsymbol{\lambda}_{\mathrm{LR}}$ significantly deviates from the displacement interpolation. However, under $\boldsymbol{\lambda}_*$, the Brownian engine is able to closely realize this geodesic, resulting in a consistently lower dissipation profile. Finally, for $\boldsymbol{\lambda}_*$ we further analyze the relationship between $\mathcal{W}_2^2/\tau$ and the dissipation for this step and observe that as $\mathcal{W}_2^2/\tau$ increases, the dissipation increases (Fig. 3), as expected from Eq. (6).

Remarkably, we see that the Brownian engine driven under $\boldsymbol{\lambda}_*$ enables us to realize protocols that have a higher efficiency η compared to protocols driven by the linear response protocol $\boldsymbol{\lambda}_{\mathrm{LR}}$ (see Fig 3, inset). In fact, the optimal transport protocols approach the Carnot efficiency even for fast driving. In this regime, the linear response protocol $\boldsymbol{\lambda}_{\mathrm{LR}}$ does not maximally expand the engine, thus limiting the work done by it. Because we learn a distinct $\boldsymbol{\lambda}_*$ for each τ, the Brownian engine *does* expand nearly maximally, resulting in greater work extraction. Crucially, the system also remains near the displacement interpolants, ensuring minimal dissipation and thus a higher efficiency.

Minimum dissipation control of a model superconducting qubit.—To show that our computational framework

also applies to Markovian quantum dynamics, we investigate a model of a superconducting qubit engine. We note that for discrete state Markov processes, extending the Benamou-Brenier formulation exactly requires a distinct metric [41]; however, we optimize $\mathcal{W}_1$ and demonstrate numerically that this lower bound to heat (cf. Ref. [26], Theorem 2) also yields a powerful variational principle for optimal control problems. In this model [Fig. 2(b), inset], we modulate the temperature T of the environment and the level splitting of the qubit V as a function of time. We optimize protocols (i) using a prespecified protocol and varying the speed at which the protocol is traversed and (ii) by setting target intermediate distributions and optimizing $V(t)$ and $T(t)$.

For (i), we used a sinusoidal protocol $\boldsymbol{\lambda}(t)$, as specified in the Supplemental Material [32]. We compare three different protocols $\boldsymbol{\lambda}(t)$, $\boldsymbol{\lambda}(\phi_{\mathrm{LR}}(t))$, and $\boldsymbol{\lambda}(\phi_*(t))$, where ϕ_* was determined by minimizing the length $\mathcal{W}_2$ along the protocol and ϕ_{LR} was determined in Ref. [19]. We observe that for the qubit engine driven under $\boldsymbol{\lambda}(\phi_*(t))$, the dissipation is lower compared to the engine driven by both $\boldsymbol{\lambda}(t)$ and $\boldsymbol{\lambda}(\phi_{\mathrm{LR}}(t))$ for fast driving, as shown in Fig. 2(b). We find that the work obtained using $\boldsymbol{\lambda}(\phi_{\mathrm{LR}}(t))$ is higher than that of $\boldsymbol{\lambda}(\phi_*(t))$, resulting in the efficiency under $\boldsymbol{\lambda}(\phi_{\mathrm{LR}}(t))$ being marginally better than that of $\boldsymbol{\lambda}(\phi_*(t))$. Of course, the expression that we optimize [Eq. (6)] only includes dissipation as an objective, so there is no guarantee of higher efficiency.

For (ii), we optimize a protocol $\boldsymbol{\lambda}_*$ to drive the qubit engine through prespecified intermediate distributions $\rho_*(t)$. We considered the distributions that the system relaxed to at $t = 0$, $\tau/4$, $\tau/2$, and $3\tau/4$ in the quasistatic limit under the sinusoidal protocol, and specified ρ_* to be the displacement interpolation of these distributions. We computed these geodesics exactly using the density matrix. As with the Brownian heat engine, we used automatic differentiation to learn a protocol $\boldsymbol{\lambda}_*(t)$. The optimal protocol $\boldsymbol{\lambda}_*(t)$ is less dissipative than all other protocols, and the corresponding steady state distribution $\hat{\rho}(t)$ closely realizes ρ_*. Ultimately, the results observed in the two systems considered demonstrate the utility of the unified geometric framework and the computational approach introduced here in learning minimally dissipative protocols for nonequilibrium control.

This work was supported by the Department of Energy, Office of Basic Energy Sciences, Division of Materials Science and Engineering, under Grant No. DE-SC0022917.

*shriramc@stanford.edu
†rotskoff@stanford.edu

[1] S. Carnot, Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance, Ann. Sci. Éc. Norm. Supér. **2e série, 1**, 393 (1872).
[2] G. E. Crooks, Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, J. Stat. Phys. **90**, 1481 (1998).
[3] J. L. Lebowitz and H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).
[4] J. Kurchan, Fluctuation theorem for stochastic dynamics, J. Phys. A **31**, 3719 (1998).
[5] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).
[6] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, Science **296**, 1832 (2002).
[7] P. Maragakis, F. Ritort, C. Bustamante, M. Karplus, and G. E. Crooks, Bayesian estimates of free energies from nonequilibrium work data in the presence of instrument noise, J. Chem. Phys. **129**, 024102 (2008).
[8] G. M. Rotskoff, G. E. Crooks, and E. Vanden-Eijnden, Geometric approach to optimal nonequilibrium control: Minimizing dissipation in nanomagnetic spin systems, Phys. Rev. E **95**, 012148 (2017).
[9] A. C. Barato and U. Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, Phys. Rev. Lett. **114**, 158101 (2015).
[10] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation Bounds All Steady-State Current Fluctuations, Phys. Rev. Lett. **116**, 120601 (2016).
[11] T. R. Gingrich, G. M. Rotskoff, and J. M. Horowitz, Inferring dissipation from current fluctuations, J. Phys. A **50**, 184004 (2017).
[12] G. E. Crooks, Measuring Thermodynamic Length, Phys. Rev. Lett. **99**, 100602 (2007).
[13] P. Salamon and R. S. Berry, Thermodynamic Length and Dissipated Availability, Phys. Rev. Lett. **51**, 1127 (1983).
[14] P. Salamon, J. D. Nulton, and R. S. Berry, Length in statistical thermodynamics, J. Chem. Phys. **82**, 2433 (1985).
[15] G. Ruppeiner, Thermodynamics: A Riemannian geometric model, Phys. Rev. A **20**, 1608 (1979).
[16] F. Schlögl, Thermodynamic metric and stochastic measures, Z. Phys. B **59**, 449 (1985).
[17] D. A. Sivak and G. E. Crooks, Thermodynamic Metrics and Optimal Paths, Phys. Rev. Lett. **108**, 190602 (2012).
[18] G. M. Rotskoff and G. E. Crooks, Optimal control in nonequilibrium systems: Dynamic Riemannian geometry of the Ising model, Phys. Rev. E **92**, 060102(R) (2015).
[19] K. Brandner and K. Saito, Thermodynamic Geometry of Microscopic Heat Engines, Phys. Rev. Lett. **124**, 040602 (2020).
[20] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics No. 58 (American Mathematical Society, Providence, 2003).
[21] R. Jordan, D. Kinderlehrer, and F. Otto, The variational formulation of the Fokker–Planck equation, SIAM J. Math. Anal. **29**, 1 (1998).
[22] J.-D. Benamou, Y. Brenier, and K. Guittet, The Monge–Kantorovitch mass transfer and its computational fluid

mechanics formulation, Int. J. Numer. Methods Fluids **40**, 21 (2002).
[23] E. Aurell, K. Gawędzki, C. Mejía-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, Refined second law of thermodynamics for fast random processes, J. Stat. Phys. **147**, 487 (2012).
[24] M. Cuturi, Sinkhorn distances: Lightspeed computation of optimal transport, in *Advances in Neural Information Processing Systems 26*, edited by C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Curran Associates, Inc., Lake Tahoe, Nevada, 2013), pp. 2292–2300.
[25] R. J. McCann, A convexity principle for interacting gases, Adv. Math. **128**, 153 (1997).
[26] T. Van Vu and K. Saito, Thermodynamic Unification of Optimal Transport: Thermodynamic Uncertainty Relation, Minimum Dissipation, and Thermodynamic Speed Limits, Phys. Rev. X **13**, 011013 (2023).
[27] S. Dago and L. Bellon, Dynamics of Information Erasure and Extension of Landauer's Bound to Fast Processes, Phys. Rev. Lett. **128**, 070604 (2022).
[28] P. Muratore-Ginanneschi, On extremals of the entropy production by 'Langevin–Kramers' dynamics, J. Stat. Mech. (2014) P05013.
[29] K. Sekimoto, Langevin equation and thermodynamics, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[30] U. Seifert, Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).
[31] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Optimal Protocols and Optimal Transport in Stochastic Thermodynamics, Phys. Rev. Lett. **106**, 250601 (2011).
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.107101 for additional computational and theoretical details.
[33] G. Peyré and M. Cuturi, Computational optimal transport: With applications to data science, Found. Trends Mach. Learn. **11**, 355 (2019).
[34] C. Villani, *Optimal Transport: Old and New*, Grundlehren Der Mathematischen Wissenschaften No. 338 (Springer, Berlin, 2009).
[35] M. Nakazato and S. Ito, Geometrical aspects of entropy production in stochastic thermodynamics based on Wasserstein distance, Phys. Rev. Res. **3**, 043093 (2021).
[36] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).
[37] G. A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, Texts in Applied Mathematics Vol. 60 (Springer, New York, 2014).
[38] J. Altschuler, J. Niles-Weed, and P. Rigollet, Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration, in *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., 2017), Vol. 30.
[39] V. Blickle and C. Bechinger, Realization of a micrometre-sized stochastic heat engine, Nat. Phys. **8**, 143 (2012).
[40] K. Brandner, K. Saito, and U. Seifert, Thermodynamics of Micro- and Nano-Systems Driven by Periodic Temperature Variations, Phys. Rev. X **5**, 031019 (2015).
[41] J. Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. **261**, 2250 (2011).