bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Counterdiabatic control of biophysical processes

Efe Ilker,[1,2,3,*] Özenç Güngör,[4] Benjamin Kuznets-Speck,[4,5]
Joshua Chiel,[4,6] Sebastian Deffner,[7,8,†] and Michael Hinczewski[4,‡]

[1] *Laboratoire Physico-Chimie Curie, Institut Curie, PSL Research University, CNRS UMR 168, Paris, France*
[2] *Sorbonne Universités, UPMC Univ. Paris 06, Paris, France*
[3] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[4] *Department of Physics, Case Western Reserve University, Cleveland, OH, 44106, USA*
[5] *Biophysics Graduate Group, University of California, Berkeley, CA 94720, USA*
[6] *Department of Physics, University of Maryland, College Park, Maryland 20742, USA*
[7] *Department of Physics, University of Maryland, Baltimore County, Baltimore, MD 21250, USA*
[8] *Instituto de Física 'Gleb Wataghin', Universidade Estadual de Campinas, 13083-859, Campinas, São Paulo, Brazil*

The biochemical reaction networks that regulate living systems are all stochastic to varying degrees. The resulting randomness affects biological outcomes at multiple scales, from the functional states of single proteins in a cell to the evolutionary trajectory of whole populations. Controlling how the distribution of these outcomes changes over time—via external interventions like time-varying concentrations of chemical species—is a complex challenge. In this work, we show how counterdiabatic (CD) driving, first developed to control quantum systems, provides a versatile tool for steering biological processes. We develop a practical graph-theoretic framework for CD driving in discrete-state continuous-time Markov networks. We illustrate the formalism with examples from gene regulation and chaperone-assisted protein folding, demonstrating the possibility that nature can exploit CD driving to accelerate response to sudden environmental changes. We generalize the method to continuum Fokker-Planck models, and apply it to study AFM single-molecule pulling experiments in regimes where the typical assumption of adiabaticity breaks down, as well as an evolutionary model with competing genetic variants subject to time-varying selective pressures. The AFM analysis shows how CD driving can eliminate non-equilibrium artifacts due to large force ramps in such experiments, allowing accurate estimation of biomolecular properties.

A fundamental dichotomy for biological processes is that they are both intrinsically stochastic and tightly controlled. The stochasticity arises from the random nature of the underlying biochemical reactions, and has significant consequences in a variety of contexts: gene expression [1], motor proteins [2], protein folding [3], all the way up to the ecological interactions and evolution of entire populations of organisms [4, 5]. Theories for such systems often employ discrete state Markov models (or continuum analogues like Fokker-Planck equations) which describe how the probability distribution of system states evolves over time. On the other hand, biology utilizes a wide array of control knobs to regulate such distributions, most often through time-dependent changes in the concentration of chemical species that influence state transition rates. In many cases these changes occur due to environmental cues—either threatening or beneficial—and the system response must be sufficiently fast to avoid danger or gain advantage.

The interplay of randomness and regulation naturally leads us to ask about the limits of control: to what extent can a biological system be driven through a prescribed trajectory of probability distributions over a finite time interval? Beyond curiosity over whether nature actually tests these limits *in vivo*, this question also arises in experimental contexts. Certain biophysical methods like optical tweezers or atomic force microscopy (AFM) apply perturbations (e.g. mechanical force) to alter the state distribution of single biomolecules in order to extract their intrinsic properties [6]. Controlling the distribution can facilitate interpretation of the data. In synthetic biology [7] one may want to precisely specify the probabilistic behavior of genetic switches or other regulatory circuit components in response to a stimulus.

Control of a system is generally easiest to describe and quantify if the perturbation is applied slowly. For example, some tweezer or AFM experiments use an increasing force ramp to unfold single molecules or rupture molecular complexes [8]. Theoretical treatments of this process typically assume the force changes slowly enough (adiabatically) that the system remains in quasi-equilibrium [8–11]. The advantage of this assumption is that, at each moment of the experimental protocol, the approximate form of the state probability distribution is known from equilibrium thermodynamics. Deriving results for faster pulling rates is more challenging [12], but useful in order to compare experiments with molecular dynamics simulations. In natural settings, responses to rapid environmental changes may entail sharp changes in the concentrations of biochemical components. For instance, an ambient temperature increase of even a few degrees can significantly increase the probability that proteins misfold and aggregate. In response to such "heat shock", cells quickly upregulate the number of chaperones—specialized proteins that facilitate unfolding or disaggregating misfolded proteins [13–18].

There is no guarantee that the quasi-equilibrium as-

* ilker@pks.mpg.de
† deffner@umbc.edu
‡ michael.hinczewski@case.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

sumption holds throughout such a process, and thus the standard tools of equilibrium or near-equilibrium thermodynamics (i.e. linear response theory) do not necessarily apply. If we are driving a system over a finite-time interval, subject to fluctuations that take us far from equilibrium, can we still attain a degree of control? In particular, can we force the system to mimic quasi-equilibrium behavior, following a certain sequence of known target distributions, but at arbitrarily fast speeds?

Interestingly this situation strongly resembles questions from quantum control and quantum thermodynamics [19], where a new line of research has been dubbed "shortcuts to adiabaticity". In recent years a great deal of theoretical and experimental work has been dedicated to mathematical tools and practical schemes to suppress nonequilibrium excitations in finite-time, nonequilibrium processes. To this end, a variety of techniques have been developed: the use of dynamical invariants [20], the inversion of scaling laws [21], the fast-forward technique [22–29], optimal protocols from optimal control theory [30–33], optimal driving from properties of quantum work statistics [34], "environment" assisted methods [35], using the properties of Lie algebras [36], and approximate methods such as linear response theory [37–40], fast quasistatic dynamics [41], or time-rescaling [42, 43], to name just a few. See Refs. [44, 45] and references therein for comprehensive reviews of these techniques.

Among this plethora of different approaches, *counterdiabatic* (CD) or *transitionless quantum driving* stands out, since it is the only method that suppresses excitations away from the adiabatic manifold at all instants. In this paradigm [46–49] one considers a time-dependent Hamiltonian $H_0(t)$ with instantaneous eigenvalues $\{\epsilon_n(t)\}$ and eigenstates $\{|n(t)\rangle\}$. In the adiabatic limit no transitions between eigenstates occur [50], and each eigenstate acquires only a time-dependent phase that can be separated into a dynamical and a geometric contribution [51]. In other words, if we start in a particular eigenstate $|n(0)\rangle$ at $t = 0$, we remain in the corresponding instantaneous eigenstate $|n(t)\rangle$ at all later times, up to a phase. The goal of CD driving is to make the system follow the same target trajectory of eigenstates as in the adiabatic case, but over a finite time.

To accomplish this, a CD Hamiltonian $H(t)$ can be constructed, such that the adiabatic approximation associated with $H_0(t)$ is an exact solution of the dynamics generated by $H(t)$ under the time-dependent Schrödinger equation. It is reasonably easy to derive that time-evolution under [46–48],

$$\begin{aligned} H(t) &= H_0(t) + H_1(t) \\ &= H_0(t) + i\hbar \sum_n \left(|\partial_t n\rangle \langle n| - \langle n|\partial_t n\rangle |n\rangle \langle n| \right), \end{aligned} \tag{1}$$

maintains the system on the adiabatic manifold. Note that it is the *auxiliary Hamiltonian* $H_1(t)$ that enforces evolution along the adiabatic manifold of $H_0(t)$: if a system is prepared in an eigenstate $|n(0)\rangle$ of $H_0(0)$ and subsequently evolves under $H(t)$, then the term $H_1(t)$ effectively suppresses the non-adiabatic transitions out of $|n(t)\rangle$ that would arise in the absence of this term.

To date, a few dozen experiments have implemented and utilized such shortcuts to adiabaticity to, for instance, transport ions or load BECs into an optical trap without creating parasitic excitations [45]. However, due to the mathematical complexity of the auxiliary Hamiltonian (1), counterdiabatic driving has been restricted to "simple" quantum systems. Note that in order to compute $H_1(t)$ one requires the instantaneous eigenstates of the unperturbed Hamiltonian, which is practically, conceptually, and numerically a rather involved task.

On the other hand, the scope of CD driving is not limited to the quantum realm. Because of the close mathematical analogies between classical stochastic systems and quantum mechanics, it was recently recognized that the CD paradigm can also be formalized for classical scenarios [29, 49, 52–57]. The classical analogue of driving a system along a target trajectory of eigenstates is a trajectory of instantaneous stationary distributions. Last year, our group and collaborators developed the first biological application of CD driving: controlling the distribution of genotypes in an evolving cellular population via external drug protocols [58]. This type of "evolutionary steering" has various potential applications, most notably in designing strategies to combat drug resistance in bacterial diseases and tumors. The CD formalism in this case was built around a multi-dimensional Fokker-Planck model, generalizing the one-dimensional Fokker-Planck approach of Ref. [55].

To date however, there does not exist a universal framework for calculating CD strategies that covers the wide diversity of stochastic models used in biology, including both discrete state and continuum approaches. In the following, we develop such a framework, taking advantage of graph theory to construct a general CD algorithm that can be applied to systems of arbitrary complexity. The usefulness of this method is of course not confined to biology, but is relevant to other classical systems described by Markovian transitions between states. However biology provides a singularly fascinating context in which to explore CD driving, both because it sheds light on the possibility of control in complex stochastic systems with many interacting components, and provides an accessible platform for future experimental tests of these ideas.

Outline: In Sec. I we formulate a theory of CD driving for any discrete state Markov model. By looking at the properties of the probability current graph associated with the master equation of the model, we can express CD solutions in terms of spanning trees and fundamental cycles of the graph. Beyond its practical utility, the graphical approach highlights the degeneracy of CD driving: the potential existence of many distinct, physically realizable CD protocols that drive a system through the same target trajectory of probability distributions. The graphical approach is schematically summarized in Fig. 1, highlighting the components in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

most general form for CD solutions, Eq. (26).

In Sec. II we apply our formalism to two biological examples, a repressor-corepressor genetic regulatory switch, and a chaperone protein that catalyzes the unfolding of a misfolded protein in response to a heat shock. The examples allow us to investigate the physical constraints and thermodynamic costs associated with specific CD solutions, and the usefulness of the CD approach even in cases where an exact CD protocol cannot be physically implemented.

In Sec. III we show how CD driving in continuum systems (i.e. Fokker-Planck models with position-dependent diffusivity) is a special limiting case of our discrete state approach. We then apply the continuum theory to analysis of AFM pulling experiments on biomolecules, and show how CD driving can compensate for non-equilibrium artifacts, allowing us to extract molecular information even in non-adiabatic, fast pulling scenarios. This Fokker-Planck example is one-dimensional, but the Supplementary Information [SI] shows how our approach can be used on higher-dimensional continuum systems as well. We demonstrate a numerical CD solution for a two-dimensional Fokker-Planck equation describing an evolving cell population with three competing genetic variants, where the distribution of variants is driven along a target trajectory by time-varying selective pressures.

The examples in Secs. II, III, and the SI are self-contained, so after going over the general solution in Sec. I the reader is free to jump to any one that may be of particular interest. The diversity of the examples—from biochemical networks describing individual genes and proteins to the evolution of entire populations of cells—is meant to provide a practical guide on how to apply our theory to the kinds of models that regularly appear in biophysical contexts.

Sec. IV concludes with connections to other areas of nonequilibrium thermodynamics and questions for future work.

I. GENERAL THEORY OF COUNTERDIABATIC DRIVING IN DISCRETE STATE MARKOV MODELS

A. Setting up the counterdiabatic driving problem

1. *Master equation and the CD transition matrix*

Consider an N-state Markov system described by a vector $\boldsymbol{p}(t)$ whose component $p_i(t)$, $i = 1, \ldots, N$, is the probability of being in state i at time t. The distribution $\boldsymbol{p}(t)$ evolves under the master equation [59, 60],

$$\partial_t \boldsymbol{p}(t) = \Omega(\lambda_t)\boldsymbol{p}(t). \tag{2}$$

The off-diagonal element $\Omega_{ij}(\lambda_t)$, $i \neq j$, of the $N \times N$ matrix $\Omega(\lambda_t)$ represents the conditional probability per unit time to transition to state i, given that the system is currently in state j. The diagonal elements $\Omega_{ii}(\lambda_t) = -\sum_{j\neq i} \Omega_{ji}(\lambda_t)$ ensure each column of the matrix sums to zero [59]. The transition rates $\Omega_{ij}(\lambda_t)$ depend on a control protocol: a set of time-varying external parameters, denoted collectively by $\lambda(t) \equiv \lambda_t$. $\Omega(t)$ plays the role of the Hamiltonian $H_0(t)$ in the classical analogy.

The instantaneous stationary probability $\boldsymbol{\rho}(\lambda_t)$ associated with $\Omega(\lambda_t)$ is the right eigenvector with eigenvalue zero,

$$\Omega(\lambda_t)\boldsymbol{\rho}(\lambda_t) = 0. \tag{3}$$

When λ_t has a non-constant time dependence, $\boldsymbol{\rho}(\lambda_t)$ in general is not a solution to Eq. (2), except in the adiabatic limit when the control parameters are varied infinitesimally slowly, $\partial_t \lambda_t \to 0$. The sequence of distributions $\boldsymbol{\rho}(\lambda_t)$ as a function of λ_t defines a target trajectory for the system, analogous to the eigenstate trajectory $|n(t)\rangle$ in the quantum version of CD.

Given an instantaneous probability trajectory $\boldsymbol{\rho}(\lambda_t)$ defined by Eq. (3), we would like to find a counterdiabatic (CD) transition matrix $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ such that the new master equation,

$$\partial_t \boldsymbol{\rho}(\lambda_t) = \widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)\boldsymbol{\rho}(\lambda_t), \tag{4}$$

evolves in time with state probabilities described by $\boldsymbol{\rho}(\lambda_t)$. Here $\dot{\lambda}_t \equiv \partial_t \lambda_t$. We are thus forcing the system to mimic adiabatic time evolution, even when $\dot{\lambda}_t$ is nonzero. As we will see below, $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ will in general depend both on the instantaneous values of the control parameters λ_t and their rate of change $\dot{\lambda}_t$. In the limit of adiabatic driving we should recover the original transition matrix, $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t \to 0) = \Omega(\lambda_t)$. Solving for the CD protocol corresponds to determining the elements of the $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ matrix in Eq. (4) given a certain $\boldsymbol{\rho}(\lambda_t)$. This corresponds to finding the CD Hamiltonian $H(t)$ of Eq. (1) in the quantum case.

We can look at the counterdiabatic problem as a special case of a more general question: given a certain time-dependent probability distribution that is our target, what is the transition matrix of the master equation for which this distribution is a solution? In effect, this is the inverse of the typical approach for the master equation, where we know the transition matrix and solve for the distribution.

2. *Representing the system via an oriented current graph*

To facilitate finding CD solutions, we start by expressing the original master equation of Eq. (2) equivalently in terms of probability currents between states,

$$\partial_t p_i(t) = \sum_j J_{ij}(t), \quad i = 1, \ldots, N \tag{5}$$

where the current from state j to i is given by:

$$J_{ij}(t) \equiv \Omega_{ij}(\lambda_t)p_j(t) - \Omega_{ji}(\lambda_t)p_i(t). \tag{6}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We can interpret any pair of states (i, j) where either $\Omega_{ij}(\lambda_t) \neq 0$ or $\Omega_{ji}(\lambda_t) \neq 0$ at some point during the protocol as being connected via an edge on a graph whose vertices are the states $i = 1, \ldots, N$. Let E be the number of edges in the resulting graph, and define a numbering $\alpha = 1, \ldots, E$ and an arbitrary orientation for the edges such that each α corresponds to a specific edge and choice of current direction. For example if edge α was between states (i, j), and the choice of direction was from j to i, then we can define current $J_\alpha(t) \equiv J_{ij}(t)$ for that edge. Alternatively if the choice of direction was from i to j, then $J_\alpha(t) \equiv J_{ji}(t) = -J_{ij}(t)$. In this way we associate the master equation with a directed graph, a simple example of which is illustrated in Fig. 2. Eq. (5) can be rewritten in terms of the oriented currents $J_\alpha(t)$ as

$$\partial_t \boldsymbol{p}(t) = \nabla \boldsymbol{J}(t), \tag{7}$$

where $\boldsymbol{J}(t)$ is an E-dimensional vector with components $J_\alpha(t)$, and ∇ is an $N \times E$ dimensional matrix known as the incidence matrix of the directed graph [61] (closely related to the stoichiometric matrix of Ref. [62]). The components of ∇ are given by

$$\nabla_{i\alpha} = \begin{cases} 1 & \text{if the direction of edge } \alpha \text{ is toward } i \\ -1 & \text{if the direction of edge } \alpha \text{ is away from } i \\ 0 & \text{if edge } \alpha \text{ does not connect to } i \end{cases}. \tag{8}$$

The αth column of ∇ contains a single 1 and a single -1, since each edge must have an origin and a destination state. Conservation of probability is thus enforced by summing over rows in Eq. (7), since $\sum_i \nabla_{i\alpha} = 0$, and so $\sum_{i=1}^N \partial_t p_i(t) = 0$. Since any given row of Eq. (7) is thus linearly dependent on the other rows, it is convenient to work in the reduced representation of the equation where we leave out the row corresponding to a certain reference state (taken to be state N),

$$\partial_t \widehat{\boldsymbol{p}}(t) = \widehat{\nabla} \boldsymbol{J}(t). \tag{9}$$

Here $\widehat{\boldsymbol{p}}(t) = (p_1(t), \ldots, p_{N-1}(t))$ and the $(N-1) \times E$ dimensional reduced incidence matrix $\widehat{\nabla}$ is equal to ∇ with the Nth row removed. Our focus will be on systems where there is a unique instantaneous stationary probability vector $\boldsymbol{\rho}(\lambda_t)$ at every t. In this case the master equation necessarily corresponds to a connected graph in the oriented current picture [59]. By a well known result in graph theory, both the full and reduced incident matrices ∇ and $\widehat{\nabla}$ of a connected, directed graph with N vertices have rank $N-1$ [61]. This means that all $N-1$ rows of $\widehat{\nabla}$ are linearly independent for the systems we consider.

Having described the original master equation of Eq. (2) in terms of oriented currents, we can do the same for Eqs. (3) and (4). Let us define the oriented stationary current $\mathcal{J}_\alpha(t)$ for the distribution $\boldsymbol{\rho}(\lambda_t)$ as follows: if the αth edge is oriented from j to i then

$$\mathcal{J}_\alpha(t) \equiv \Omega_{ij}(\lambda_t)\rho_j(\lambda_t) - \Omega_{ji}(\lambda_t)\rho_i(\lambda_t). \tag{10}$$

The reduced representation of Eq. (3) corresponds to

$$\widehat{\nabla} \boldsymbol{\mathcal{J}}(t) = 0. \tag{11}$$

Analogously for the CD master equation, Eq. (4), we define the oriented current

$$\widetilde{\mathcal{J}}_\alpha(t) \equiv \widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)\rho_j(\lambda_t) - \widetilde{\Omega}_{ji}(\lambda_t, \dot{\lambda}_t)\rho_i(\lambda_t). \tag{12}$$

The time dependence of $\widetilde{\mathcal{J}}_\alpha$ is explicitly through λ_t and $\dot{\lambda}_t$, but we write it in more compact form as $\widetilde{\mathcal{J}}_\alpha(t)$ to avoid cumbersome notation. Then Eq. (4) can be expressed as

$$\partial_t \widehat{\boldsymbol{\rho}}(\lambda_t) = \widehat{\nabla} \widetilde{\boldsymbol{\mathcal{J}}}(t). \tag{13}$$

3. *Counterdiabatic current equation*

Subtracting Eq. (11) from Eq. (13) we find

$$\partial_t \widehat{\boldsymbol{\rho}}(\lambda_t) = \widehat{\nabla} \delta \boldsymbol{\mathcal{J}}(t), \tag{14}$$

where $\delta \boldsymbol{\mathcal{J}}(t) \equiv \widetilde{\boldsymbol{\mathcal{J}}}(t) - \boldsymbol{\mathcal{J}}(\lambda_t)$ is the difference between the CD and stationary current vectors. For the CD problem, we are given the original matrix elements $\Omega_{ij}(\lambda_t)$ and thus also have the corresponding stationary distribution values $\rho_i(\lambda_t)$ and stationary currents $\mathcal{J}_\alpha(\lambda_t)$. What we need to determine, via Eq. (14), are the CD currents $\widetilde{\boldsymbol{\mathcal{J}}}(t)$. We can then use Eq. (12) to solve for the CD matrix transition rates $\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)$. By construction, these satisfy Eq. (4), and hence define a CD protocol for the system.

As a first step, let us consider the invertibility of Eq. (14) to solve for $\delta \boldsymbol{\mathcal{J}}(t)$. The $(N-1) \times E$ dimensional matrix $\widehat{\nabla}$ is generally non-square: $N(N-1)/2 \geq E \geq N-1$ for a connected graph. Only in the special case of tree-like graphs (no loops) do we have $E = N-1$ and a square $(N-1) \times (N-1)$ matrix $\widehat{\nabla}$. Since the rank of $\widehat{\nabla}$ is $N-1$, as mentioned above, for tree-like graphs $\widehat{\nabla}$ is invertible and Eq. (14) can be solved without any additional complications:

$$\delta \boldsymbol{\mathcal{J}}(t) = \widehat{\nabla}^{-1} \partial_t \widehat{\boldsymbol{\rho}}(\lambda_t) \quad \text{iff } E = N-1. \tag{15}$$

As described in the next section, the elements of $\widehat{\nabla}^{-1}$ for a tree-like graph can be obtained directly through a graphical procedure, without the need to do any explicit matrix inversion.

In the case where $E > N-1$, the solution procedure is more involved, but the end result has a relatively straightforward form: the most general solution $\delta \boldsymbol{\mathcal{J}}(t)$ can always be expressed as a finite linear combination of a basis of CD solutions. How to obtain this basis, and its close relationship to the spanning trees and fundamental cycles of the graph, is the topic we turn to next.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

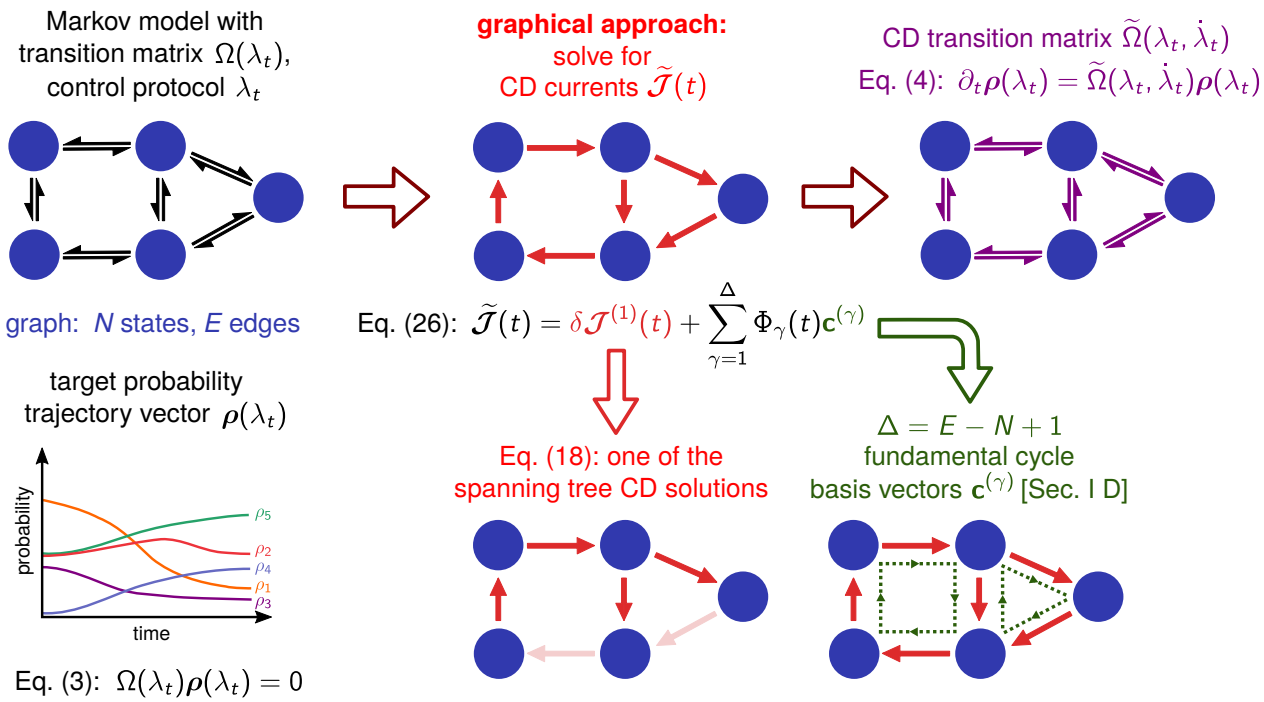

FIG. 1. Overview of the graphical approach for deriving CD solutions. We start with a Markov model defined by a transition matrix $\Omega(\lambda_t)$ dependent on the control protocol λ_t. Associated with this is a graph with N states, E edges, and a target trajectory $\boldsymbol{\rho}(\lambda_t)$ consisting of instantaneous stationary states of $\Omega(\lambda_t)$. The eventual goal is to find the CD transition matrix $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ where $\boldsymbol{\rho}(\lambda_t)$ is the solution to the associated master equation, Eq. (4). To facilitate this, we must first find the CD currents $\widetilde{\boldsymbol{\mathcal{J}}}(t)$, the main goal of the graphical approach. The most general form of the solution for $\widetilde{\boldsymbol{\mathcal{J}}}(t)$ is given by Eq. (26), and consists of two components: (i) a spanning tree CD solution $\delta\boldsymbol{\mathcal{J}}^{(1)}(t)$, given by Eq. (18) and derived via the procedure outlined in Sec. I B; (ii) a linear combination of the fundamental basis cycle vectors $\mathbf{c}^{(\gamma)}$, $\gamma = 1, \ldots, \Delta$, where $\Delta = E - N + 1$, as described in Sec. I D. The coefficient functions $\Phi_\gamma(t)$ are arbitrary.

B. General graphical solution for the counterdiabatic protocol

The graphical procedure described in this and the following two sections, culminating in the general solution of Eq. (26), is summarized in Fig. 1. To illustrate the procedure concretely, we will use the two-loop system shown in Fig. 2A as an example, where $N = 4$, $E = 5$. The solution for this case is relevant to the biophysical model for chaperone-assisted protein folding discussed later in the paper. Fig. 2A shows the rates $k_i(\lambda_t)$ and $r_i(\lambda_t)$ that determine the transition matrix $\Omega(\lambda_t)$, and Fig. 2B labels the oriented stationary currents $\mathcal{J}_\alpha(t)$, $\alpha = 1, \ldots, E$. Explicit expressions for $\rho_i(\lambda_t)$ and $\mathcal{J}_\alpha(t)$ in terms of the rates are given in Appendix A.

Every connected graph has a set of spanning trees: subgraphs formed by removing $\Delta \equiv E - N + 1$ edges such that the remaining $N - 1$ edges form a tree linking together all the N vertices. The number $\mathcal{T}$ of such spanning trees is related to the reduced incidence matrix through Kirchhoff's matrix tree theorem [61], $\mathcal{T} = \det\left(\widehat{\nabla}\widehat{\nabla}^T\right)$. For the current graph of Fig. 2B, this matrix is

$$\widehat{\nabla} = \begin{pmatrix} -1 & 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \end{pmatrix}, \tag{16}$$

and the number of trees is thus $\mathcal{T} = 8$.

Let us select one spanning tree to label as the reference tree. The choice is arbitrary, since any spanning tree can be a valid starting point for constructing the basis. The left side of Fig. 2C shows one such tree chosen for the two-loop example. Here $\Delta = 2$, so we have removed two edges: $\mathcal{J}_1$ and $\mathcal{J}_5$. From this reference tree we can derive Δ other distinct spanning trees using the following method: 1) Take one of the Δ edges that were removed to get the reference tree, and add it back to the graph. 2) This creates a loop in the graph, known as a fundamental cycle (highlighted in green in Fig. 2C) [61]. 3) Remove one of the other edges in that loop (not the one just added), such that the graph returns to being a spanning tree. This new tree is distinct from the reference because it contains one of the Δ edges not present in the reference tree. For example, in the top right of Fig. 2C, we added back edge 1, forming the fundamental cycle on the left loop. We then delete edge 2 from this loop, creating spanning tree 2. A similar procedure is used to construct

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

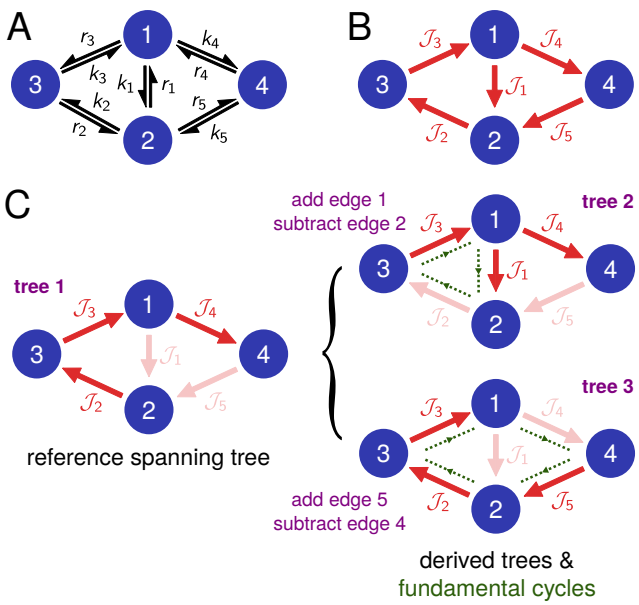

FIG. 2. A two-loop discrete state Markov model, with $N = 4$ states and $E = 5$ edges. A) The black arrows correspond to entries in the transition matrix $\Omega(\lambda_t)$: transition rates $k_i(\lambda_t)$ and $r_i(\lambda_t)$ between states that depend on an external protocol λ_t. B) The red arrows label the oriented stationary currents $\mathcal{J}_\alpha(\lambda_t)$, defined in Eq. (10). C) On the left, one of the spanning trees of the graph, chosen to be a reference for constructing the tree basis. Edges deleted to form the tree are shown in faint red. On the right, two trees in this set derived from the reference one. Each such derived tree has a one-to-one correspondence with a fundamental cycle of the graph (highlighted in green).

tree 3.

We denote the $\Delta+1$ trees (one reference + Δ derived trees) constructed in this manner as the tree basis. We will label the trees in the basis set with $\gamma = 1, \ldots, \Delta+1$, where $\gamma = 1$ corresponds to the reference. In general, this basis is a subset of all possible trees, since $\mathcal{T} \geq \Delta+1$. To every tree in the basis, we will associate a CD solution as follows. Let $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(\lambda_t, \dot{\lambda}_t)$ be a current difference vector that satisfies Eq. (14), but with the constraint that at every edge α that is not present in the γth tree, we have $\delta\mathcal{J}_\alpha^{(\gamma)}(t) = 0$. We call this a *fixed current* constraint, since it corresponds to not being able to perturb the current associated with that edge via external control parameters. For example imposing the restriction $\Omega_{ij} = \widetilde{\Omega}_{ij}$ and $\Omega_{ji} = \widetilde{\Omega}_{ji}$ for the pair (i, j) associated with edge α would make make $\delta\mathcal{J}_\alpha^{(\gamma)}(t) = 0$.

To find $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$, consider the $(N-1)\times E$-dimensional reduced incidence matrix $\widehat{\nabla}$ of the original graph; for example, Eq. (16) in the case of the two-loop graph of Fig. 2B. For a given spanning tree γ, we can construct an $(N-1)\times(N-1)$ submatrix $\widehat{\nabla}^{(\gamma)}$ from $\widehat{\nabla}$ by choosing the $N-1$ columns in $\widehat{\nabla}$ that correspond to edges present in γ. This submatrix $\widehat{\nabla}^{(\gamma)}$ is equal to the reduced incidence matrix of the spanning tree γ. Hence we know that it has rank $N-1$ and there exists an inverse $[\widehat{\nabla}^{(\gamma)}]^{-1}$. Let us now construct a "stretched inverse": an $E \times (N-1)$-dimensional matrix $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ where the rows are populated by the following rule. If the row corresponds to one of the Δ edges that was removed from the original graph to get the tree γ, it is filled with zeros; otherwise, it is filled with the corresponding row of $[\widehat{\nabla}^{(\gamma)}]^{-1}$. For the three trees in Fig. 2C, labeled $\gamma = 1, 2, 3$ clockwise from left, the matrices $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ have the following form:

$$[\widehat{\nabla}^{(1)}]_S^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & -1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$[\widehat{\nabla}^{(2)}]_S^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\widehat{\nabla}^{(3)}]_S^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}. \tag{17}$$

Moreover, it turns out one does not have to explicitly write down or invert $\widehat{\nabla}^{(\gamma)}$ in order to find the elements of $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$. We can take advantage of a known graphical procedure for constructing inverse reduced incidence matrices of connected tree-like graphs [63, 64]. To determine the ith column of the matrix $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$, start at the reference state (the state removed when constructing the reduced incidence matrix $\widehat{\nabla}$, which in our case is always state N). Among the edges of the spanning tree γ, there is a unique path that connects state N to state i. Following that path, if you encounter the current arrow $\mathcal{J}_\alpha$ oriented in the direction of the path, put a $+1$ in the row of $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ corresponding to $\mathcal{J}_\alpha$. Similarly if the current arrow is oriented opposite to the path, put a -1. All other entries in the ith column (current arrows not on the path, or not in the spanning tree) are set to zero. For example, consider the second column of $[\widehat{\nabla}^{(1)}]_S^{-1}$ in Eq. (17). The corresponds to the path from state $N = 4$ to state 2 in the tree on the left of Fig. 2C. This includes edges 4, 3, and 2, with the arrows along those edges all oriented opposite to the path. Hence the column has a -1 at the 4th, 3rd, and 2nd rows, and all other entries are set to zero.

By construction, each matrix $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ acts as a right pseudoinverse of $\widehat{\nabla}$, satisfying $\widehat{\nabla}[\widehat{\nabla}^{(\gamma)}]_S^{-1} = I_{N-1}$, where I_{N-1} is the $(N-1)\times(N-1)$ dimensional identity matrix. We can now write down a solution for $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$,

$$\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t) = [\widehat{\nabla}^{(\gamma)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\lambda_t). \tag{18}$$

If we act from the left on both sides by $\widehat{\nabla}$, we see that this form satisfies Eq. (14). The αth row of of $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ is zero if edge α corresponds to a fixed current constraint (edge not present in the tree γ). Thus $\delta\mathcal{J}_\alpha^{(\gamma)}(t) = 0$ for these α. Not only do the vectors $\delta\boldsymbol{\mathcal{J}}(t)$ associated with the tree basis constitute $\Delta + 1$ solutions to Eq. (14), they are

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

also linearly independent from one another. To see this, note that because of the procedure to construct derived trees (adding back a distinct edge that was removed in the reference tree), a tree with $\gamma \geq 2$ will have non-zero entry in $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$ in a position where every other tree (reference or derived) has a zero because of constraints. Hence the $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$ vector for each derived tree is linearly independent from all the other vectors in the basis.

We also know that any linear combination of solutions to Eq. (14) can be scaled by an overall normalization factor (to make the coefficients sum to one) so that it is also a solution to Eq. (14). Hence the following linear combination of basis vectors is a valid solution:

$$\delta\boldsymbol{\mathcal{J}}(t) = \sum_{\gamma=1}^{\Delta+1} w_\gamma(t)\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t), \tag{19}$$

Here $w_\gamma(t)$ are any real-valued functions where $\sum_{\gamma=1}^{\Delta+1} w_\gamma(t) = 1$ at each λ_t and $\dot{\lambda}_t$. As we argue in the next section, the tree basis is complete: any CD solution $\delta\boldsymbol{\mathcal{J}}(t)$ can be expressed in the form of Eq. (19). Note that Eq. (15) is a special case of Eq. (19). When the original graph is tree-like, $\Delta = 0$ and there is only one spanning tree ($\gamma = 1$), equivalent to the original graph. In this case $[\widehat{\nabla}^{(1)}]_S^{-1} = \widehat{\nabla}^{-1}$ and the sole coefficient function $w_1(t) = 1$ by normalization.

C. Completeness of the tree basis

To prove that any CD solution can be expressed as a linear combination of tree basis solutions $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$, let us first introduce Δ vectors of the following form:

$$\boldsymbol{\mathcal{V}}^{(\gamma)}(t) = \delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t) - \delta\boldsymbol{\mathcal{J}}^{(1)}(t), \tag{20}$$

for $\gamma = 2, \ldots, \Delta + 1$. Since both basis vectors on the right-hand side of Eq. (20) satisfy Eq. (14), we know that

$$\widehat{\nabla}\boldsymbol{\mathcal{V}}^{(\gamma)}(t) = \partial_t\widehat{\boldsymbol{\rho}}(\lambda_t) - \partial_t\widehat{\boldsymbol{\rho}}(\lambda_t) = 0. \tag{21}$$

Hence $\boldsymbol{\mathcal{V}}^{(\gamma)}(t)$ is a vector in the null space of $\widehat{\nabla}$. Moreover since the basis vectors $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$ are linearly independent, the set $\boldsymbol{\mathcal{V}}^{(\gamma)}(t)$ constitutes Δ linearly independent null vectors of $\widehat{\nabla}$. We can find the dimension of the null space, nullity($\widehat{\nabla}$), using the rank-nullity theorem: rank($\widehat{\nabla}$) + nullity($\widehat{\nabla}$) $= E$, where E is the number of columns in $\widehat{\nabla}$. Since rank($\widehat{\nabla}$) $= N - 1$ for a connected graph, as described earlier, we see that nullity($\widehat{\nabla}$) $= E - (N - 1) = \Delta$. Thus the Δ linearly independent vectors $\boldsymbol{\mathcal{V}}^{(\gamma)}(t)$ span the whole null space. If there existed a vector $\delta\boldsymbol{\mathcal{J}}(t)$ that satisfied Eq. (14) but could not be expressed as a linear combination of basis vectors, then the corresponding vector $\boldsymbol{\mathcal{V}}(t) = \delta\boldsymbol{\mathcal{J}}(t) - \delta\boldsymbol{\mathcal{J}}^{(1)}(t)$ would be a null vector that is linearly independent of all the $\boldsymbol{\mathcal{V}}^{(\gamma)}(t)$. But since the latter span the whole null space, this is impossible. Hence every CD solution $\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t)$ satisfying Eq. (14) must be expandable in the form of Eq. (19).

D. General solution in the cycle basis

The discussion in the previous section also allows us to rewrite the expansion in Eq. (19) in an alternative form that is convenient in practical applications. Using the fact that $\sum_{\gamma=1}^{\Delta+1} w_\gamma(t) = 1$, Eq. (19) can be equivalently expressed as

$$\begin{aligned}\delta\boldsymbol{\mathcal{J}}(t) &= \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \sum_{\gamma=2}^{\Delta+1} w_\gamma(t)\left(\delta\boldsymbol{\mathcal{J}}^{(\gamma)}(t) - \delta\boldsymbol{\mathcal{J}}^{(1)}(t)\right) \\ &= \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \sum_{\gamma=2}^{\Delta+1} w_\gamma(t)\boldsymbol{\mathcal{V}}^{(\gamma)}(t).\end{aligned} \tag{22}$$

Since the vectors $\boldsymbol{\mathcal{V}}^{(\gamma)}(t)$ form a basis for the null space of $\widehat{\nabla}$, the second term in the last line of Eq. (22), with its arbitrary coefficient functions $w_\gamma(t)$, is general enough to describe any vector function in the null space. With no loss of generality, we can rewrite this second term in another basis for the null space instead. A convenient choice is the fundamental cycle basis corresponding to some reference spanning tree (we need not choose the same reference as used to find $\delta\boldsymbol{\mathcal{J}}^{(1)}(t)$). The Δ fundamental cycles were identified in the procedure to construct derived trees. If we assign an arbitrary orientation to the cycles (clockwise or counterclockwise), then the E-dimensional cycle vector $\boldsymbol{c}^{(\gamma)}$, associated with the derived tree $\gamma + 1$, is defined as follows: a ± 1 at every row whose corresponding edge in the original graph belongs to the fundamental cycle, with a $+1$ if the edge direction is parallel to the cycle orientation, -1 if anti-parallel. All edges not belonging to the fundamental cycle are zero. For the reference tree in Fig. 2C the fundamental cycles are highlighted in green on the right of the panel. Here the two cycle vectors are:

$$\boldsymbol{c}^{(1)} = \begin{pmatrix}1\\1\\1\\0\\0\end{pmatrix}, \qquad \boldsymbol{c}^{(2)} = \begin{pmatrix}0\\1\\1\\1\\1\end{pmatrix}. \tag{23}$$

In general, the Δ fundamental cycle vectors form a basis for the null space of $\widehat{\nabla}$ [61].

In terms of the cycle vectors, Eq. (22) can be written as

$$\delta\boldsymbol{\mathcal{J}}(t) = \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \sum_{\gamma=1}^{\Delta} v_\gamma(t)\boldsymbol{c}^{(\gamma)}, \tag{24}$$

where $v_\gamma(t)$ for $\gamma = 1, \ldots, \Delta$ are another set of arbitrary coefficient functions. The convenience of Eq. (24) over Eq. (19) is that we only need to find one spanning tree solution $\delta\boldsymbol{\mathcal{J}}^{(1)}(t)$. Both have the same number of degrees of freedom: in the first case Δ coefficient functions $w_\gamma(t)$ for $\gamma = 2, \ldots, \Delta + 1$ (since $w_1(t)$ depends on the rest through the normalization constraint); in the second case Δ coefficient functions $v_\gamma(t)$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Finally we note that because of Eq. (11), the oriented stationary current vector $\boldsymbol{\mathcal{J}}(t)$ corresponding to the original protocol is in the null space of $\widehat{\nabla}$. Hence it can also be expanded in terms of the cycle vectors as

$$\boldsymbol{\mathcal{J}}(t) = \sum_{\gamma=1}^{\Delta} u_\gamma(t)\boldsymbol{c}^{(\gamma)}, \tag{25}$$

with some coefficient functions $u_\gamma(t)$. Since the CD currents $\widetilde{\boldsymbol{\mathcal{J}}}(t) = \boldsymbol{\mathcal{J}}(t) + \delta\boldsymbol{\mathcal{J}}(t)$, we can combine Eqs. (24) and (25) to get the most general expression for any set of currents that satisfies Eq. (13):

$$\widetilde{\boldsymbol{\mathcal{J}}}(t) = \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \sum_{\gamma=1}^{\Delta} \Phi_\gamma(t)\boldsymbol{c}^{(\gamma)}. \tag{26}$$

Here $\Phi_\gamma(t) \equiv u_\gamma(\lambda_t) + v_\gamma(t)$. Because the $v_\gamma(t)$ are arbitrary, the functions $\Phi_\gamma(t)$ are also arbitrary, and we still have the same Δ degrees of freedom to span the solution space.

E. Thermodynamic costs of CD driving

Consider a driving protocol where the control parameters λ_t are varied from λ_0 to λ_τ over a time interval τ. The system starts in the initial stationary state, $\boldsymbol{p}(0) = \boldsymbol{\rho}(\lambda_0)$. If we implement the CD protocol, the transition matrix is given by $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ and the dynamics are described by Eq. (4), with the probability distribution at time t equal to $\boldsymbol{\rho}(\lambda_t)$. The total entropy production rate $\dot{S}^{\rm tot}(t)$ at time t is given by [60],

$$\dot{S}^{\rm tot}(t) = k_B \sum_{ij} \widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)\rho_j(\lambda_t) \ln \frac{\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)\rho_j(\lambda_t)}{\widetilde{\Omega}_{ji}(\lambda_t, \dot{\lambda}_t)\rho_i(\lambda_t)}, \tag{27}$$

and is non-negative in accordance with the second law of thermodynamics, $\dot{S}^{\rm tot}(t) \geq 0$. For a system coupled to a single heat reservoir at temperature T, we can decompose $\dot{S}^{\rm tot}(t)$ into the rates of change of the system and reservoir entropies,

$$\dot{S}^{\rm tot}(t) = \dot{S}(t) + \dot{S}^{\rm r}(t), \tag{28}$$

where $S(t) = -k_B \sum_i \rho_i(\lambda_t) \ln \rho_i(\lambda_t)$ is the system entropy, and $S^r(t)$ is the reservoir entropy. The heat flow from the reservoir into the system is given by $-T\dot{S}^{\rm r}$, and there is some rate $\dot{W}$ at which external work is being done on the system during the driving. By energy conservation, the mean system energy $E(t) = \sum_i \rho_i(\lambda_t)E_i(\lambda_t)$, where $E_i(\lambda_t)$ is the energy of state i, must change according to:

$$\dot{E}(t) = -T\dot{S}^{\rm r} + \dot{W}. \tag{29}$$

Solving for $\dot{S}^r$ from Eq. (29) and plugging into Eq. (28), we can write

$$\begin{aligned} \dot{W}(t) &= \dot{E}(t) - T\dot{S}(t) + T\dot{S}^{\rm tot}(t) \\ &= \dot{F}(t) + T\dot{S}^{\rm tot}(t), \end{aligned} \tag{30}$$

where the free energy $F(t) = E(t) - TS(t)$.

In the special case where we begin and end in an equilibrium stationary state, $F(0) = F^{\rm eq}(\lambda_0)$ and $F(\tau) = F^{\rm eq}(\lambda_\tau)$. Here the equilibrium free energy $F^{\rm eq}(\lambda_t) = -k_B T \ln Z(\lambda_t)$, and $Z(\lambda_t) = \sum_i \exp(-\beta E_i(\lambda_t))$ is the corresponding partition function, with $\beta = (k_B T)^{-1}$. Integrating both sides of Eq. (30), we can relate the total work expended over the driving, ΔW, to the free energy difference of the system $\Delta F^{\rm eq} = F^{\rm eq}(\tau) - F^{\rm eq}(0)$:

$$\Delta W = \Delta F^{\rm eq} + T\Delta\dot{S}^{\rm tot}, \tag{31}$$

where for any observable $\dot{O}(t)$ we write $\Delta O \equiv \int_0^\tau dt\, \dot{O}(t)$. We can then define the dissipated (or irreversible) work $\Delta W^{\rm diss} \equiv \Delta W - \Delta F^{\rm eq} = T\Delta\dot{S}^{\rm tot}$ [65]. From the fact that $\dot{S}^{\rm tot}(t) \geq 0$ at all times t, we see that $\Delta\dot{S}^{\rm tot} \geq 0$ and hence the dissipated work during driving between two equilibrium states is always non-negative, $\Delta W^{\rm diss} \geq 0$. The dissipated work is a measure of the thermodynamic cost of the driving. From Eq. (31) we have to put in work ΔW at least equal $\Delta F^{\rm eq}$ to get between the starting and end states. $\Delta W^{\rm diss}$ tells us the additional work cost on top of this baseline.

From Eq. (28) we know $\Delta W^{\rm diss} = T\Delta S + T\Delta S^{\rm r}$. The first term just depends on the system entropy difference at the beginning and end of the protocol, and hence on the probability distributions $\boldsymbol{\rho}(\lambda_0)$ and $\boldsymbol{\rho}(\lambda_\tau)$. For any CD solution corresponding to the same target trajectory of distributions $\boldsymbol{\rho}(\lambda_t)$, the term $T\Delta S$ is always the same. However the excess work $T\Delta S^{\rm r}$ dissipated throughout the protocol can vary among CD solutions. To see this, note that from Eqs. (27)-(28) and the definition of $S(t)$ we can write the reservoir contribution to the total entropy change as [60]:

$$\Delta S^r = k_B \sum_{ij} \int_0^\tau dt\, \widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)\rho_j(\lambda_t) \ln \frac{\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)}{\widetilde{\Omega}_{ji}(\lambda_t, \dot{\lambda}_t)}. \tag{32}$$

This clearly depends not just on $\boldsymbol{\rho}(\lambda_t)$, but the values of the CD transition matrix elements $\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)$, which vary among CD solutions for the same target. A biological realization of this variation in $\Delta W^{\rm diss}$ among solutions will be explored in the repressor-corepressor protein example of the next section.

For a CD current solution $\widetilde{\boldsymbol{\mathcal{J}}}(t)$ of the form given in Eq. (26), each current component $\widetilde{\mathcal{J}}_\alpha(t)$ is related to the corresponding transition rates $\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)$ and $\widetilde{\Omega}_{ji}(\lambda_t, \dot{\lambda}_t)$ via Eq. (12). In many physical systems we may not have the ability to freely modify both transition rates—one may be fixed, and other tunable over some range by changing an external parameter (like concentration of a reactant in a biochemical system). But let us imagine a hypothetical scenario where we have the ability to implement any $\widetilde{\Omega}_{ij}(\lambda_t, \dot{\lambda}_t)$ and $\widetilde{\Omega}_{ji}(\lambda_t, \dot{\lambda}_t)$ compatible with Eqs. (12) and (26) for all the edges in the network. In this case we can show for any CD solution there always exists a choice of transition rates that makes $\Delta W^{\rm diss}$ arbitrarily small.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

To accomplish this, it is convenient to rewrite $\dot{S}^{\rm tot}(t)$ from Eq. (27) in terms of a sum over edges α in the oriented current graph [60],

$$\dot{S}^{\rm tot}(t) = k_B \sum_\alpha \widetilde{\mathcal{J}}_\alpha(t)\chi_\alpha(t), \tag{33}$$

where

$$\chi_\alpha(t) \equiv \ln \frac{\widetilde{\Omega}_{ij}(\lambda_t,\dot{\lambda}_t)\rho_j(\lambda_t)}{\widetilde{\Omega}_{ji}(\lambda_t,\dot{\lambda}_t)\rho_i(\lambda_t)}. \tag{34}$$

Here α is the label of the edge associated with (i,j). We take $\widetilde{\Omega}_{ij}$ to be the CD transition rate along the arrow direction of edge α, and $\widetilde{\Omega}_{ji}$ to be the reverse transition. The difference between the numerator and denominator inside the logarithm in Eq. (34) is just the current at that edge, which is determined by the CD solution:

$$\widetilde{\Omega}_{ij}(\lambda_t,\dot{\lambda}_t)\rho_j(\lambda_t) - \widetilde{\Omega}_{ji}(\lambda_t,\dot{\lambda}_t)\rho_i(\lambda_t) = \widetilde{\mathcal{J}}_\alpha(t). \tag{35}$$

To simultaneously satisfy Eqs. (34) and (35), the transition rates must be given by

$$\begin{aligned}\widetilde{\Omega}_{ij}(\lambda_t,\dot{\lambda}_t) &= \widetilde{\mathcal{J}}_\alpha(t)\frac{e^{\chi_\alpha(t)}}{\rho_j(\lambda_t)(e^{\chi_\alpha(t)}-1)},\\ \widetilde{\Omega}_{ji}(\lambda_t,\dot{\lambda}_t) &= \widetilde{\mathcal{J}}_\alpha(t)\frac{1}{\rho_i(\lambda_t)(e^{\chi_\alpha(t)}-1)}.\end{aligned} \tag{36}$$

The current $\widetilde{\mathcal{J}}_\alpha(t)$ is fixed by choosing a particular CD solution, but in this scenario we have the freedom to manipulate both $\widetilde{\Omega}_{ij}(\lambda_t,\dot{\lambda}_t)$ and $\widetilde{\Omega}_{ji}(\lambda_t,\dot{\lambda}_t)$ at each edge. We can thus choose the rates to make $|\chi_\alpha(t)| \ll 1$ for every α. Eq. (36) in this limit becomes

$$\begin{aligned}\widetilde{\Omega}_{ij}(\lambda_t,\dot{\lambda}_t) &\approx \frac{\widetilde{\mathcal{J}}_\alpha(t)}{\rho_j(\lambda_t)\chi_\alpha(t)},\\ \widetilde{\Omega}_{ji}(\lambda_t,\dot{\lambda}_t) &\approx \frac{\widetilde{\mathcal{J}}_\alpha(t)}{\rho_i(\lambda_t)\chi_\alpha(t)}.\end{aligned} \tag{37}$$

We are free to make $|\chi_\alpha(t)|$ arbitrarily small for fixed $\widetilde{\mathcal{J}}_\alpha(t)$, so long as $\chi_\alpha(t)$ and $\widetilde{\mathcal{J}}_\alpha(t)$ have the same sign. Approaching the limit $|\chi_\alpha(t)| \to 0$ makes $\dot{S}^{\rm tot}(t) \to 0$ from Eq. (33) and hence $\Delta W^{\rm diss} \to 0$. Of course to achieve this, both transition rates in Eq. (37) have to become arbitrarily large, while the current in Eq. (35) remains finite. It may seem counterintuitive that though this CD protocol can drive the system at arbitrary speed over a trajectory, the associated $\Delta W^{\rm diss}$ is small. However in this particular case the CD solution involves making the local "diffusivity" in the system large (if we imagine dynamics on the network as a discrete diffusion process). In other words we are reducing the effective friction to zero in order to eliminate dissipation.

In practice this extreme limit is not realistic. There are likely to be physical constraints that prevent us from simultaneously tuning each pair of rates in the network over an arbitrary range, so the CD implementation with $\Delta W^{\rm diss} \to 0$ is not realizable. However, among the family of realizable CD solutions for a given target trajectory, we can ask which one makes $\Delta W^{\rm diss}$ as small as possible. We illustrate this in the repressor-corepressor biological example described in the next section.

II. CD DRIVING IN BIOLOGICAL NETWORKS

To illustrate the general theory in specific biological contexts, we consider two examples of CD driving in biochemical networks. The first example is a simple genetic regulatory switch involving a repressor protein and corepressor ligand binding to an operator site on DNA, turning off the expression of a set of genes. Here it turns out there are enough control knobs—concentrations of repressors, corepressors, and repressor-corepressor complexes—to implement a whole family of exact CD solutions. Among this family we can then examine which ones satisfy certain physical constraints, or minimize thermodynamic costs. The second example involves a chaperone protein that binds to a misfolded substrate, catalyzing the unfolding of this misfolded protein and giving it another chance to fold into the correct ("native") state. The available control knobs—chaperone and ATP concentrations—are in fact insufficient to implement any exact CD solution. In this scenario, approximate CD driving by sharply increasing chaperone concentration can rapidly decrease the probability of a protein being in a misfolded state, a key factor in ameliorating the damage due to heat shock. We thus find that even when the CD control formalism cannot be applied in full, it can still provide a useful way to interpret time-dependent regulation in certain biological systems.

A. Repressor-corepressor model

The first system we consider is a common form of gene regulation in bacteria, illustrated schematically in Fig. 3A: a repressor protein has the ability to bind to an operator site on DNA. When bound, it interferes with the ability of RNA polymerase to attach to the nearby promoter site, preventing the transcription of the genes associated with the promoter. The system acts as a genetic switch, with the empty operator site the "on" state for gene expression, and the occupied operator site the "off" state. In many cases, additional regulatory molecules—inducers or corepressors— influence the binding affinity of repressor proteins [66]. In the present model, binding of the bare repressor to the operator site is weak (it unbinds easily), but the binding strength is enhanced in the presence of a particular small molecule—the corepressor. Hence the corepressor concentration acts like an input signal, with sufficiently high levels leading to the promoter site being occupied with high probability, and the associated genes being turned off.

As is generally the case with genetic regulation in bi-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

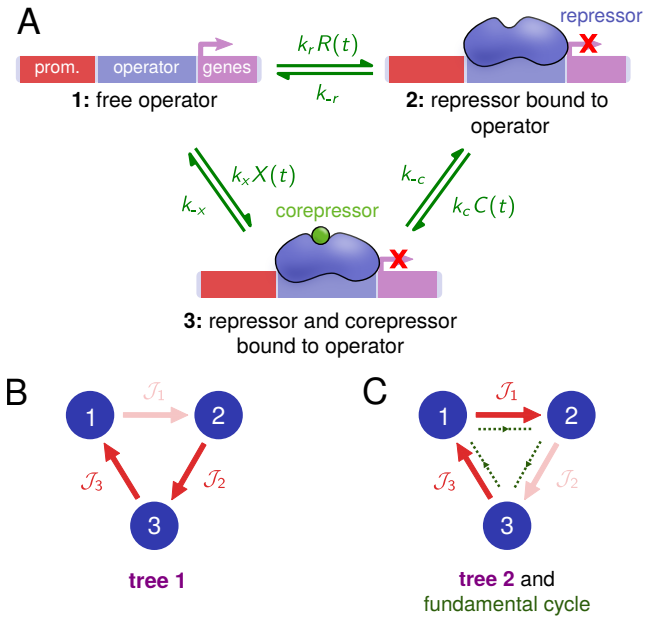

FIG. 3. A) Biochemical network of a repressor-corepressor model, showing an operator site on DNA in three different states: 1) free; 2) bound to a bare repressor protein; 3) bound to a represssor-corepressor complex. Transition rates between the states are shown in green. The binding reaction rates depend on three concentrations of molecules in solution: $R(t)$ for bare repressors, $C(t)$ for corepressors, and $X(t)$ for the complexes. B) One of the spanning trees for the associated network graph, with the edge deleted to form the tree shown in faint red. We take this to be the reference spanning tree for the tree basis. C) The other tree in the basis, with the corresponding fundamental cycle in green.

ology, the processes underlying repressor dynamics are stochastic [67]. Here we model the system through Markovian transitions among three discrete states: 1) free operator; 2) bare repressor bound to the operator; 3) repressor-corepressor complex bound to the operator. Transitions in both directions (clockwise and counterclockwise in Fig. 3A) are possible. In each pair of transition rates between neighboring states there is a binding reaction proportional to the concentration of a chemical species in solution. The relevant concentrations are those of bare repressors $R(t)$, corepressors $C(t)$, and repressor-corepressor complexes $X(t)$. The associated binding constants, which multiply the concentrations to give the binding rates, are k_r, k_c, and k_x respectively. The other transition rate in each pair is the corresponding unbinding reaction, k_{-r}, k_{-c}, or k_{-x}.

To be concrete, we base parameter values on the purine repressor (PurR) system of *E. coli* [66, 68]. The PurR protein turns off genes responsible for the de novo production of purines, a class of molecules including guanine and adenine that are essential ingredients in DNA/RNA and energy transducing molecules like ATP. If the cell has an excess of purines (for example from environmental sources), this is signaled by an abundance of the corepressors guanine or hypoxanthine (a purine derivative) that form complexes with PurR, enabling it to bind strongly with its operator site. This way, the cell can switch off the energetically expensive de novo production of purines when it is not needed. The entire biochemical network of Fig. 3A, including both clockwise and counterclockwise transitions, was experimentally measured for PurR, and the parameters are given by [68]: $k_r = 0.0191$ nM^{-1} min^{-1}, $k_c = 7.83 \times 10^{-4}$ nM^{-1} min^{-1}, $k_x = 0.9$ nM^{-1} min^{-1}, $k_{-r} = 1.68$ min^{-1}, $k_{-c} = 0.72$ min^{-1}, $k_{-x} = 0.072$ min^{-1}. Note that $k_{-x} \ll k_{-r}$ (the repressor-corepressor complex unbinds from the operator more slowly than bare repressor) and $k_x \gg k_r$ (it binds more easily), demonstrating the enhanced affinity of the complex to the operator relative to the bare repressor.

Three of the transition rates depend on the external concentration parameters $\lambda_t = (R(t), C(t), X(t))$, and time-dependent changes in these concentrations constitute a driving protocol that alters the probability landscape of the operator. Note that while our system description focuses on the state of the operator, the repressor and corepressor can also bind/unbind in solution away from the operator [68], and in some systems there are other molecules (like inducers) competing for the repressor in solution. The relative concentrations of the two reactants, $R(t)$, $C(t)$ and the complex $X(t)$ in solution can be expressed in terms of a chemical potential for the reaction:

$$\Delta\mu(t) = k_B T \ln\left(\frac{k_r R(t) k_c C(t) k_{-x}}{k_{-r} k_{-c} k_x X(t)}\right). \tag{38}$$

The fraction inside the logarithm is just the product of all the clockwise rates divided by the product of all the counterclockwise rates. When $\Delta\mu(t) = 0$ there is an equal chance of making clockwise versus counterclockwise transitions, and hence the reaction at that instant satisfies detailed balance. While $\Delta\mu(t) = 0$ may be possible in cellular conditions (for example if the binding/unbinding rates are very fast compared to the timescale of the protocol), it is not necessarily always the case [67]. Sudden environmental influxes of corepressor $C(t)$ might drive the concentrations out of detailed balance. This can also occur due to cellular processes that influence the availability of bare repressors—like changes in inducer abundance, in the case where inducers and corepressors compete for the same binding site on repressors. In general then we will take the solution concentrations $(R(t), C(t), X(t))$ to be arbitrary functions determined by processes outside of the system, and explore how these three control knobs can influence the state of the operator.

Imagine $\boldsymbol{\rho}(\lambda_t)$ is some target trajectory of distributions for the operator. For example (as we illustrate below) we might want to rapidly turn off gene expression, going from a situation where state 1 predominates to one where state 3 predominates. For any concentration protocol $\lambda_t = (R(t), C(t), X(t))$, the corresponding trajectory $\boldsymbol{\rho}(\lambda_t)$ is described explicitly by expressions in Appendix A. Given this target trajectory $\boldsymbol{\rho}(\lambda_t)$, we

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

can derive a family of counterdiabatic concentration protocols, in terms of alternative concentration functions $(\widetilde{R}(t), \widetilde{C}(t), \widetilde{X}(t))$ such that the operator is guaranteed have the distribution $\boldsymbol{\rho}(\lambda_t)$ at every t during the driving.

Following the graphical solution procedure of Sec. I B, we start with the fact that the oriented current graph corresponding to Fig. 3A has $N = 3$, $E = 3$. If $\mathcal{J}_i(t)$ is the current oriented clockwise starting from state i, the reduced incidence matrix for the one-loop graph is

$$\widehat{\nabla} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}. \tag{39}$$

Because $\Delta = E - N + 1 = 1$, we have $\Delta + 1 = 2$ trees in a tree basis. Taking the tree with edge 1 missing as the reference (tree 1 in Fig. 3B), we choose the other tree in the basis to be the one with edge 2 missing (tree 2 in Fig. 3C). Using the graphical algorithm, we can easily write down stretched inverse reduced incidence matrices for these trees:

$$[\widehat{\nabla}^{(1)}]_S^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad [\widehat{\nabla}^{(2)}]_S^{-1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}. \tag{40}$$

One can readily check that $\widehat{\nabla}[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ for $\gamma = 1, 2$ is the 2×2 identity matrix. There is a single fundamental cycle vector for the graph, shown as a dashed line in Fig. 3C, given by $\boldsymbol{c}^{(1)} = (1, 1, 1)$.

Using Eq. (18) for $\delta\boldsymbol{\mathcal{J}}^{(1)}(t)$ and Eq. (26) for the most general form of the CD currents $\widetilde{\boldsymbol{\mathcal{J}}}(t)$, we have

$$\begin{aligned} \widetilde{\boldsymbol{\mathcal{J}}}(t) &= \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \Phi_1(t)\boldsymbol{c}^{(1)} \\ &= [\widehat{\nabla}^{(1)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\lambda_t) + \Phi_1(t)\boldsymbol{c}^{(1)}, \end{aligned} \tag{41}$$

where $\Phi_1(t)$ is an arbitrary function. Given the definition of the oriented CD currents $\widetilde{\mathcal{J}}_\alpha(t)$ in Eq. (12), we can write Eq. (41) as a series of equations for each edge α:

$$\begin{aligned} \widetilde{\mathcal{J}}_1(t) &= k_r\widetilde{R}(t)\rho_1(\lambda_t) - k_{-r}\rho_2(\lambda_t) = \Phi_1(t), \\ \widetilde{\mathcal{J}}_2(t) &= k_c\widetilde{C}(t)\rho_2(\lambda_t) - k_{-c}\rho_3(\lambda_t) = -\partial_t\rho_2(\lambda_t) + \Phi_1(t), \\ \widetilde{\mathcal{J}}_3(t) &= k_{-x}\rho_3(\lambda_t) - k_x\widetilde{X}(t)\rho_1(\lambda_t) = \partial_t\rho_1(\lambda_t) + \Phi_1(t). \end{aligned} \tag{42}$$

Here we have assumed that our ability to influence the currents at each edge is via the CD concentration protocols $\widetilde{R}(t)$, $\widetilde{C}(t)$, and $\widetilde{X}(t)$, with all other parameters fixed. Solving Eq. (42) for these concentrations, we find

$$\begin{aligned} \widetilde{R}(t) &= \frac{\Phi_1(t) + k_{-r}\rho_2(\lambda_t)}{k_r\rho_1(\lambda_t)}, \\ \widetilde{C}(t) &= \frac{\Phi_1(t) - \partial_t\rho_2(\lambda_t) + k_{-c}\rho_3(\lambda_t)}{k_c\rho_2(\lambda_t)}, \\ \widetilde{X}(t) &= -\frac{\Phi_1(t) + \partial_t\rho_1(\lambda_t) - k_{-x}\rho_3(\lambda_t)}{k_x\rho_1(\lambda_t)}. \end{aligned} \tag{43}$$

Different choices of $\Phi_1(t)$ correspond to different CD solutions that drive the system through the same trajectory $\boldsymbol{\rho}(\lambda_t)$. The one additional constraint is that only $\Phi_1(t)$ functions that lead to non-negative concentrations in Eq. (43) at all t during driving are physically allowable.

To illustrate a family of CD solutions, let us choose a target trajectory $\boldsymbol{\rho}(\lambda_t)$ by specifying a concentration protocol. To mimic a rapid switch in gene expression, we select $C(t)$ to sharply increase in a sigmoidal fashion, with $R(t)$ kept at a constant level and $X(t)$ in detailed balance with $C(t)$ and $R(t)$:

$$\begin{aligned} &R(t) = R_0, \quad C(t) = C_0 + (C_f - C_0)\frac{e^{k(t-t_0)}}{1 + e^{k(t-t_0)}}, \\ &X(t) = \frac{k_r R(t) k_c C(t) k_{-x}}{k_{-r} k_{-c} k_x}, \end{aligned} \tag{44}$$

where $R_0 = 20$ nM, $C_0 = 0.2$ μM, $C_f = 20$ μM, $k = 3$ min^{-1}, $t_0 = 5$ min, and the remaining parameters described above. The form of $X(t)$ means that $\Delta\mu(t) = 0$ at all t for this protocol. As a result the stationary current for this protocol is zero at all edges, $\mathcal{J}_1(t) = \mathcal{J}_2(t) = \mathcal{J}_3(t) = 0$. The concentrations in Eq. (44) are shown as dotted curves in Fig. 4D. The elements of the corresponding stationary distribution $\boldsymbol{\rho}(\lambda_t)$, calculated using the expressions in Appendix A, are depicted as solid curves in Fig. 4A. They represent a transition from a system dominated by state 1 at the beginning of the protocol to one dominated by state 3 at the end (the gene turning mostly "off"). As expected, if we actually drove the system using the protocol of Eq. (44) the resulting probability distribution $\boldsymbol{p}(t)$ (dashed curves in Fig. 4A) would lag behind the target distribution $\boldsymbol{\rho}(\lambda_t)$.

To keep the system on target, we need to use one of the CD concentration protocols given by Eq. (43). For example, consider the one associated with tree 1, which means setting $\Phi_1(t) = 0$ at all t, so that from Eq. (41) we have $\widetilde{\boldsymbol{\mathcal{J}}}(t) = [\widehat{\nabla}^{(1)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\lambda_t)$. The corresponding CD concentrations are shown in violet in Fig. 4D. Tree 1 has edge 1 missing, so the CD current there is constrained to be equal to the original stationary current, $\widetilde{\mathcal{J}}_1(t) = \mathcal{J}_1(t) = 0$. Since the edge 1 current determines $\widetilde{R}(t)$, we see that $\widetilde{R}(t) = R(t)$ (the violet curve overlaps with the dotted one), while both $\widetilde{C}(t)$ and $\widetilde{X}(t)$ differ from the original protocol. In a similar way, we can look at the CD protocol associated with tree 2. To implement this, making Eq. (41) take the form $\widetilde{\boldsymbol{\mathcal{J}}}(t) = [\widehat{\nabla}^{(2)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\lambda_t)$, with $[\widehat{\nabla}^{(2)}]_S^{-1}$ given by Eq. (40), we set $\Phi_1(t) = \partial_t\rho_2(\lambda_t)$. The corresponding CD concentrations are colored teal in Fig. 4D. Note that since edge 2 is missing in tree 2, here $\widetilde{C}(t) = C(t)$.

Even though the original protocol satisfies detailed balance, the tree 1 and tree 2 CD solutions that force the system to stay along the target trajectory $\boldsymbol{\rho}(\lambda_t)$ violate detailed balance. As can be seen from Fig. 4C, $\Delta\mu(t)$ for both these solutions veers significantly negative during the driving (where $\Delta\mu(t)$ for any CD solution is calculated via Eq. (38) with the CD concen-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

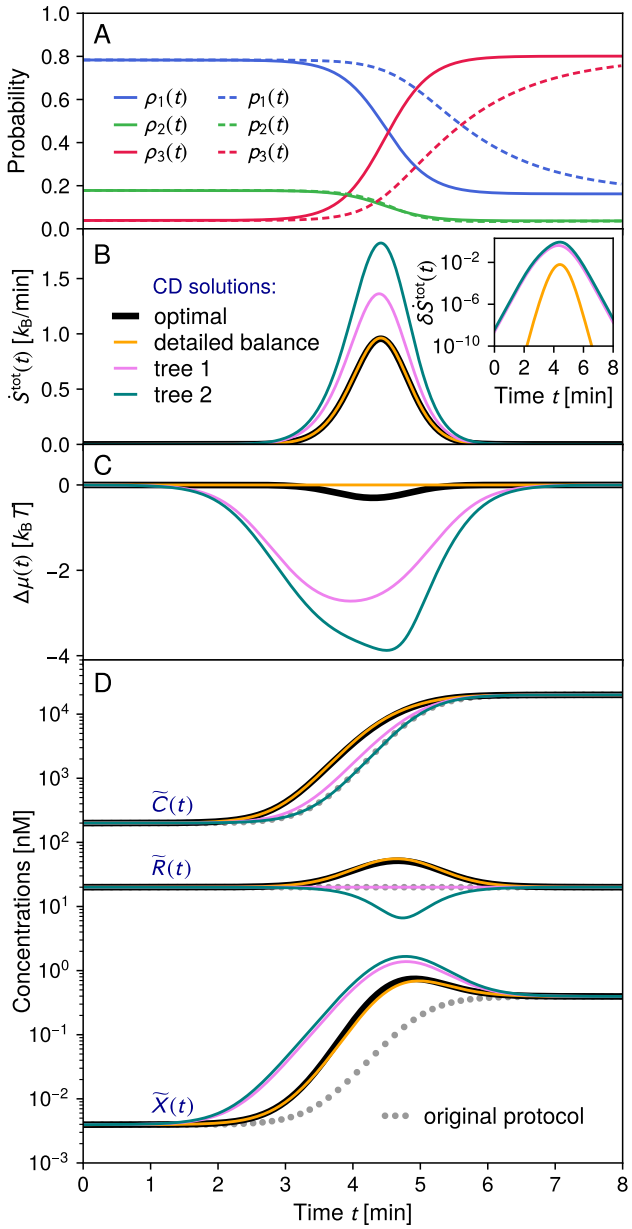

FIG. 4. A) Components of the target stationary distribution trajectory $\boldsymbol{\rho}(\lambda_t)$ (solid curves) for the repressor-corepressor system, defined by the protocol of Eq. (44). In the absence of CD driving, the actual distribution $\boldsymbol{p}(t)$ (dashed curves) lags behinds the target. B-D) Characteristics of four different CD solutions that all drive the system along the target trajectory $\boldsymbol{\rho}(\lambda_t)$. The four solutions are: spanning tree 1 (violet, corresponding to Fig. 3B); spanning tree 2 (teal, corresponding to Fig. 3C); the solution satisfying detailed balance, $\Delta\mu(t) = 0$ at all t (yellow); and the optimal solution that minimizes $\dot{S}^{\rm tot}(t)$ at all t (thick black). For each solution we depict: B) the total entropy production rate $\dot{S}^{\rm tot}(t)$, with the inset showing the difference $\delta\dot{S}^{\rm tot}(t) \equiv \dot{S}^{\rm tot}(t) - \dot{S}^{\rm tot,opt}(t)$ between non-optimal and optimal rates on a log scale [units: k_B/min]; C) the instantaneous chemical potential $\Delta\mu(t)$, calculated using Eq. (38) with the CD concentrations; D) the CD concentrations $\widetilde{C}(t)$, $\widetilde{R}(t)$, $\widetilde{X}(t)$.

trations from Eq. (43) substituted for the original concentrations). Interestingly, one can find a CD solution where the concentrations satisfy $\Delta\mu(t) = 0$, by solving $\Phi_1(t)$ at each t that enforces this condition. The resulting CD protocol is shown in yellow in Fig. 4D. Note that despite the fact that $\Delta\mu(t) = 0$ for this solution, the target trajectory $\boldsymbol{\rho}(\lambda_t)$ is not an instantaneous stationary distribution of the CD transition matrix $\widetilde{\Omega}(\lambda_t, \dot{\lambda}_t)$ during the driving. Hence there is still nonzero entropy production $\dot{S}^{\rm tot}(t)$, as seen in Fig. 4B. All three CD solutions discussed so far have different $\dot{S}^{\rm tot}(t)$ profiles, and hence involve different amount of dissipated work $\Delta W^{\rm diss}$ (the integrated area under the $\dot{S}^{\rm tot}(t)$ curve). We can also determine the CD solution for the given target trajectory that has the smallest $\Delta W^{\rm diss}$, by numerically finding $\Phi_1(t)$ at each t that minimizes $\dot{S}^{\rm tot}(t)$. This yields the CD protocol shown as a thick black curve in Fig. 4D. It is close, but not exactly equal to, the $\Delta\mu(t) = 0$ protocol, exhibiting slightly negative $\Delta\mu(t)$ at intermediate times (Fig. 4C). The inset of Fig. 4B shows the difference $\delta\dot{S}^{\rm tot}(t) \equiv \dot{S}^{\rm tot}(t) - \dot{S}^{\rm tot,opt}(t)$ between each non-optimal solution and the optimal one. The $\Delta\mu(t) = 0$ solution is significantly closer to optimal entropy production than the two tree solutions.

This simple three state biological example illustrates the variety of physically realizable CD solutions that exist in certain systems. The CD solutions can have quite distinct physical characteristics, i.e. concentrations that vary by up to an order of magnitude among different protocols in Fig. 4D, or different dissipation profiles in Fig. 4B. Yet all of them drive the system through the same target trajectory $\boldsymbol{\rho}(\lambda_t)$. Thus there can be a rich array of options available to nature (or to an experimentalist engineering a synthetic system) to achieve a specific probabilistic target. Indeed in this scenario an interesting question for future studies would be to ask whether certain driving options would be evolutionarily favored over others because of selection pressures due to energetic costs [69].

In contrast to this abundance of exact solutions, the next example explores the other extreme: what if the available control knobs are insufficient to exactly implement CD driving? Can the CD theoretical framework still provide insights?

B. Chaperone model

Many newly synthesized proteins, susceptible to misfolding, become trapped in long-lived metastable states that are prone to aggregation. Since aggregates present a danger to the survival of the cell, there exists an elaborate rescue machinery of molecular chaperone proteins that facilitate unfolding or disaggregating misfolded proteins [13–16]. In the case of *E.coli*, which has the most extensively studied chaperone network, certain components like the GroEL-GroES system are obligatory for survival [70]. Environmental stresses further exacerbate the problem, and an increase of ambient temperature by

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

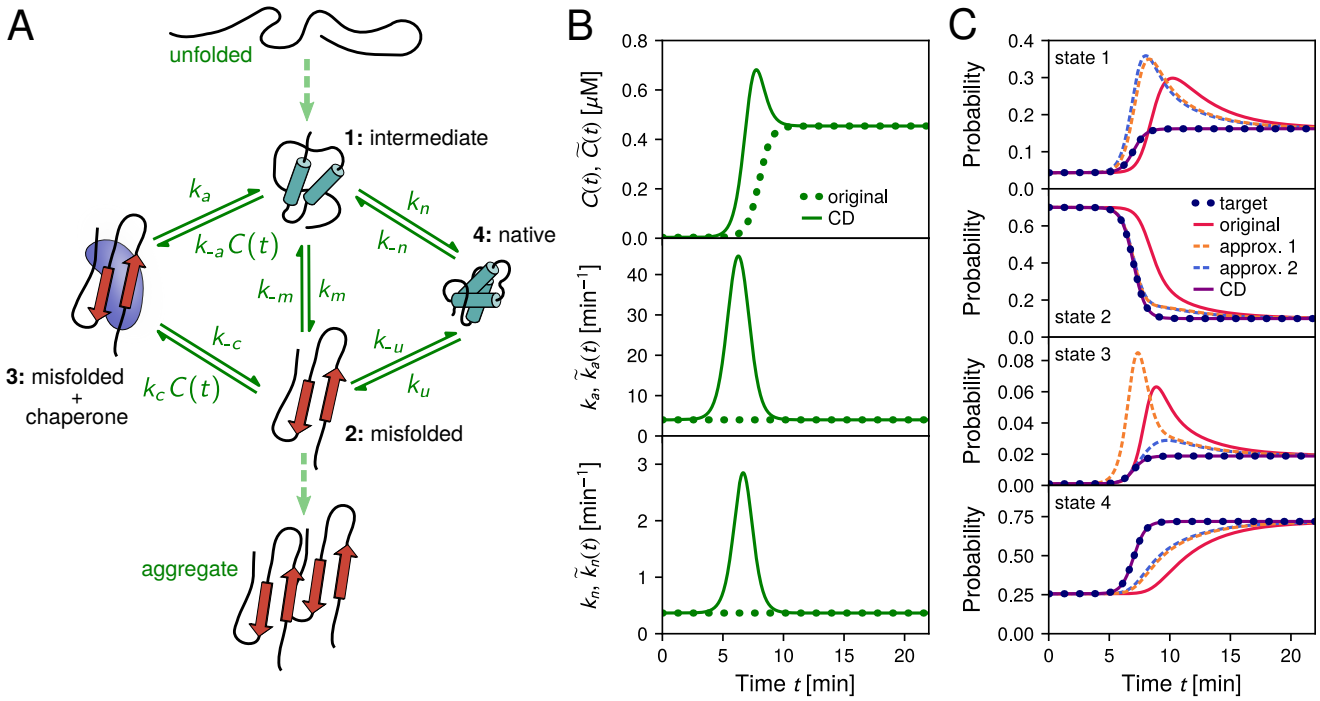

FIG. 5. A) Conformational states of a protein interacting with a chaperone. Transition rates in our kinetic network model are indicated by solid green arrows. Related transitions outside the scope of the model are shown as dashed arrows. B) A chosen control protocol, labeled "original" (dotted curves), and the corresponding exact CD solution (solid curves) based on the reference spanning tree of Fig. 2C. Top: chaperone concentrations $C(t)$ and $\widetilde{C}(t)$ for the original and CD protocols respectively; middle: chaperone enzymatic reaction rates k_a and $\widetilde{k}_a(t)$; bottom: the intermediate-to-native transition rates k_n and $\widetilde{k}_n(t)$. C) Probabilities versus time for the four system states. The dotted curve shows that stationary distribution target trajectory defined by the original protocol, which exhibits a rapid decrease of the misfolded (state 2) probability given by Eq. (48). The red solid curve shows the actual state probabilities under the original protocol, which diverge from the target. In contrast, the exact CD solution (purple solid curve) perfectly follows the target. The dashed curves show two approximations to the CD solution, when we lack the ability to vary all three quantities shown in panel B. Approximation 1 implements the CD protocol for $\widetilde{C}(t)$, but lets k_a and k_n stay fixed. Approximation 2 implements the CD protocol for $\widetilde{C}(t)$ and $\widetilde{k}_a(t)$, while keeping k_n fixed.

just a few degrees can significantly enhance protein misfolding and consequently aggregation [17]. Responding to a heat shock requires creating extra capacity, since even under normal conditions the majority of chaperones are occupied by misfolded proteins [15] (i.e. occupancy for GroEL can approach 100% for fast-growing *E. coli* [16]). This is accomplished by rapidly upregulating the number of chaperones to cope with additional misfolded proteins [17, 18].

Most chaperones require constant power input in the form of ATP hydrolysis. As a result the stationary probability distribution of conformational states for a protein interacting with a chaperone will generally be out of equilibrium (non-Boltzmannian) [71, 72]. When the chaperone concentration increases after a heat shock (for example following a sudden rise to a new temperature [73]), the protein is driven away from the previous stationary distribution, and eventually relaxes to a new stationary distribution once the chaperone concentrations reach steady-state values at the new temperature. Chaperone upregulation during heat shock therefore serves as a natural example of nonequilibrium driving in a biological system.

In this section we consider a four-state Markov model for chaperone-assisted protein unfolding, inspired by earlier models like those of Refs. [71, 72]. We focus on a network of four states for a particular substrate ("client") protein, and one type of chaperone, depicted in Fig. 5A: 1) an intermediate conformational state of the protein, along the folding pathway between the unfolded and native states; 2) a misfolded protein state, prone to aggregation; 3) the misfolded protein bound to chaperone; 4) the correctly folded "native" state. These four states can interconvert with transition rates denoted in the figure (further details below). The model is a small biochemical module within a broader set of processes, some of which are depicted schematically with dashed arrows in the figure: protein synthesis and the initial folding to the intermediate state, and aggregation of the misfolded proteins. Our focus will be on a single protein once it enters the intermediate state, and then transitions among the four states. Similarly we ignore the aggregation process, occurring over much larger timescales than the transitions in the network. We model the dynamics in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

aftermath of a heat shock [17, 74]: a sudden jump to some high temperature T, which then remains fixed as the system adapts. The conditions favor misfolding over the native folding pathway. In the absence of chaperones, state 2 (misfolded) would be most likely, and over longer timescales this would eventually result in a build-up of aggregates. To prevent such a fate, which can be lethal for the cell, the cell engages in a complex heat shock defense mechanism. A key part of the defense is rapidly increasing the concentration of chaperones, lowering the probability of the protein being in state 2. The heat shock response is an example of nature regulating the state distribution of proteins via at least one time-varying parameter (chaperone concentration), in a situation where a fast shift in the distribution is crucial for survival. It is thus an interesting scenario to investigate the possibility of CD-like control.

As a first step, let us detail the various transitions in the network. The protein can interconvert between states 1 and 2 with rates k_m and k_{-m}, which satisfy a local detailed balance relationship,

$$\frac{k_{-m}}{k_m} = e^{-\beta\epsilon_m}, \tag{45}$$

where $\epsilon_m > 0$ is the free energy difference between the intermediate and misfolded states. A chaperone can bind to the misfolded protein at rate $k_c C(t)$, where $C(t)$ is the concentration of unoccupied chaperones and k_c is the binding constant. Once bound, the chaperone catalyzes the partial unfolding of the misfolded state to the intermediate state at rate k_a. This conversion may involve several substeps and hydrolysis of multiple ATP molecules, but we simplify the process to a single reaction step hydrolyzing one ATP molecule, with some rate function k_a.

Though typically negligible compared to the forward rate k_a, the reverse rate $k_{-a}C(t)$, proportional to chaperone concentration, must be formally defined in order to have a thermodynamically complete description of the system. Since a full traversal of the left loop clockwsise (states $1 \to 2 \to 3 \to 1$) involves hydrolysis of an ATP molecule, the product of clockwise/counterclockwise rate ratios over the entire cycle is related to the chemical potential difference $\Delta\mu$ of ATP hydrolysis:

$$\frac{k_m k_c C(t) k_a}{k_{-m} k_{-c} k_{-a} C(t)} = \frac{k_m k_c k_a}{k_{-m} k_{-c} k_{-a}} = e^{\beta\Delta\mu}. \tag{46}$$

Transitions from the intermediate to native state occur with rate k_n, and from the native to misfolded state with rate k_u. The corresponding reverse rates k_{-n} and k_{-u} are related to the forward ones through the local detailed balance relations

$$\frac{k_{-n}}{k_n} = e^{-\beta\epsilon_n}, \qquad \frac{k_{-u}}{k_u} = e^{-\beta\epsilon_u}. \tag{47}$$

Here ϵ_n and ϵ_u are the free energy differences between the intermediate and native, and between the native and misfolded states respectively. Since going from states $2 \to 4 \to 1$ should yield the same cumulative free energy difference as going directly from $2 \to 1$, we know that $\epsilon_m = \epsilon_u + \epsilon_n$.

We base the parameter values in our model on those associated with the chaperone GroEL assisting the folding of the substrate protein MDH, estimated from fitting to experimental data [71]: $k_m = 0.37$ min^{-1}, $k_n = 0.366$ min^{-1}, $k_u = 0.025$ min^{-1}, $k_{-u} = 7.78 \times 10^{-3}$ min^{-1}, $k_c = 1.7 \times 10^6$ M^{-1}min^{-1}, $k_a = 4$ min^{-1}. In cases where only upper or lower bounds on the parameters could be determined, we used the values at the bound. Using Eq. (47) and the values of k_u and k_{-u} yield an estimate of $\epsilon_u = 1.17\ k_B T$. We do not know the precise value of ϵ_m from the experimental fitting, but we assume a typical value of $\epsilon_m = 3\ k_B T$, which then gives $\epsilon_n = \epsilon_m - \epsilon_u = 1.83\ k_B T$. Similarly, we set $k_{-c} = 0.1$ min^{-1} as the unbinding rate of the chaperone, a typical scale assuming strong binding affinity between the chaperone and substrate. The remaining unknown parameters can now be determined using Eqs. (45)-(47) (setting the ATP hydrolysis potential difference $\Delta\mu = 22\ k_B T$ [75]): $k_{-m} = 0.0184$ min^{-1}, $k_{-n} = 0.0585$ min^{-1}.

The chaperone network of Fig. 5A has exactly the same form as the two-loop network of Fig. 2A, with the mapping $k_1 = k_m$, $k_2 = k_c C(t)$, $k_3 = k_a$, $k_4 = k_n$, $k_5 = k_u$, $r_1 = k_{-m}$, $r_2 = k_{-c}$, $r_3 = k_{-a}C(t)$, $r_4 = k_{-n}$, $r_5 = k_{-u}$. There is a unique stationary state for every distinct value of chaperone concentration $C(t)$. The time-varying concentration protocol plays the role of the external control parameter, $\lambda_t = C(t)$. Explicit expressions for the stationary probabilities are given in Appendix A.

In the initial aftermath of the step rise in temperature, we imagine a scenario where the available concentration of free chaperones is $C(0)$, and the system rapidly has relaxed to a stationary state corresponding to $C(0)$. This $t = 0$ time point is the start of the driving. We assume $C(0)$ is insufficient to cope with the conditions favoring misfolded proteins, so chaperones are upregulated to some final concentration $C(\tau)$ by the end of the driving ($t = \tau$). Given the parameters described above, and a small starting chaperone concentration $C(0) = 3.55$ nM, the stationary probability values are shown as the $t = 0$ initial points on the dotted curve in Fig. 5C. The misfolded probability $\rho_2(0) = 0.7$ and the native probability $\rho_4(0) = 0.256$. At the end of the protocol we choose a final concentration $C(\tau) = 0.454\ \mu$M. Once the system relaxes to the stationary state at this concentration, the corresponding misfolded and native probabilities are $\rho_2(\tau) = 0.1$ and $\rho_4(\tau) = 0.719$. Driving the system between these two stationary states thus significantly reduces the probability of misfolding and increases the probability of being in the native state.

Can this driving be accomplished in a finite time, following a particular target trajectory of stationary states? Let us consider an example target trajectory $\boldsymbol{\rho}(\lambda_t)$ where the misfolded probability shifts from 0.7 to 0.1 over several minutes,

$$\rho_2(\lambda_t) = A \tanh(\kappa(t - t_0)) + B \tag{48}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

with $A = -0.3$, $B = 0.4$, $t_0 = 2$ min, $\kappa = 1$ min^{-1}. Using the stationary probability expressions from Appendix A, one can solve Eq. (48) for the $\lambda_t = C(t)$ curve corresponding to this trajectory, which we label as the "original" protocol (dotted curve in the top panel of Fig. 5B). As expected, if one were to directly implement this original protocol the actual state probabilities $\boldsymbol{p}(t)$ (red curves in Fig. 5C) diverge from the target (dotted curves in Fig. 5C).

The complete set of possible CD solutions for the two-loop system was derived earlier. Let us take for example the solution corresponding to the reference spanning tree in Fig. 2C. The associated current perturbations are $\delta\boldsymbol{\mathcal{J}}^{(1)}(t) = [\widehat{\nabla}^{(1)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\lambda_t)$, with $[\widehat{\nabla}^{(1)}]_S^{-1}$ given by Eq. (17). To solve for the CD transition rates, we need to combine these perturbations with the stationary currents for the original protocol, $\widetilde{\boldsymbol{\mathcal{J}}}(t) = \boldsymbol{\mathcal{J}}(t) + \delta\boldsymbol{\mathcal{J}}^{(1)}(t)$. The stationary currents for the various edges (as labeled according to Fig. 2B) are $\mathcal{J}_1(t) = \mathcal{J}_L(t) - \mathcal{J}_R(t)$, $\mathcal{J}_2(t) = \mathcal{J}_3(t) = \mathcal{J}_L(t)$, $\mathcal{J}_4(t) = \mathcal{J}_5(t) = \mathcal{J}_R(t)$. Expressions for the left-loop stationary current $\mathcal{J}_L(t)$ and the right-loop stationary current $\mathcal{J}_R(t)$, which depend on the control parameter $C(t)$, are shown in Appendix A (see also Eq. (51) below). Putting everything together, we have three edges α where $\delta\mathcal{J}_\alpha^{(1)} \neq 0$, and hence we need to modify the currents there to implement CD driving:

$$\begin{aligned}
\widetilde{\mathcal{J}}_2(t) &= k_c\widetilde{C}(t)\rho_2(\lambda_t) - k_{-c}\rho_3(\lambda_t) \\
&= \mathcal{J}_L(t) - \partial_t\rho_2(\lambda_t), \\
\widetilde{\mathcal{J}}_3(t) &= \widetilde{k}_a(t)\rho_3(\lambda_t) - k_{-a}\widetilde{C}(t)\rho_1(\lambda_t) \\
&= \mathcal{J}_L(t) - \partial_t\rho_2(\lambda_t) - \partial_t\rho_3(\lambda_t), \\
\widetilde{\mathcal{J}}_4(t) &= \widetilde{k}_n(t)\rho_1(\lambda_t) - k_{-n}\rho_4(\lambda_t) \\
&= \mathcal{J}_R(t) - \partial_t\rho_1(\lambda_t) - \partial_t\rho_2(\lambda_t) - \partial_t\rho_3(\lambda_t).
\end{aligned} \tag{49}$$

This particular CD solution could be implemented if we employ a modified concentration protocol $\widetilde{C}(t)$ and time-dependent transition rates $\widetilde{k}_a(t)$, $\widetilde{k}_n(t)$. Solving for these from Eq. (49) we find:

$$\begin{aligned}
\widetilde{C}(t) &= C(t) - \frac{1}{k_c}\frac{\partial_t\rho_2(\lambda_t)}{\rho_2(\lambda_t)}, \\
\widetilde{k}_a(t) &= k_a - \frac{\partial_t\rho_2(\lambda_t) + \partial_t\rho_3(\lambda_t)}{\rho_3(\lambda_t)} - \frac{k_{-a}}{k_c}\frac{\rho_1(\lambda_t)\partial_t\rho_2(\lambda_t)}{\rho_2(\lambda_t)}, \\
\widetilde{k}_n(t) &= k_n - \frac{\partial_t\rho_1(\lambda_t) + \partial_t\rho_2(\lambda_t) + \partial_t\rho_3(\lambda_t)}{\rho_1(\lambda_t)}.
\end{aligned} \tag{50}$$

To get the final form for Eq. (50) we have used the fact that

$$\begin{aligned}
\mathcal{J}_L(t) &= k_cC(t)\rho_2(\lambda_t) - k_{-c}\rho_3(\lambda_t) \\
&= k_a\rho_3(\lambda_t) - k_{-a}\rho_1(\lambda_t), \\
\mathcal{J}_R(t) &= k_n\rho_1(\lambda_t) - k_{-n}\rho_4(\lambda_t).
\end{aligned} \tag{51}$$

Eq. (50) implements the reference spanning tree CD solution for any desired target trajectory $\boldsymbol{\rho}(\lambda_t)$. For the specific trajectory corresponding to Eq. (48) we show the time-dependent CD protocol as solid curves in Fig. 5B. Qualitatively, the CD protocol takes the form of a transient peak or overshoot in the three parameters $\widetilde{C}(t)$, $\widetilde{k}_a(t)$, and $\widetilde{k}_n(t)$ relative to the original protocol. This transient perturbation eliminates the discrepancy between $\boldsymbol{p}(t)$ and $\boldsymbol{\rho}(\lambda_t)$, as we see in the purple CD curves in Fig. 5C, where the system now exactly follows the target (dotted) trajectory. From the form of Eq. (50), it is clear that in the limit of adiabatically slow driving, where $\partial_t\boldsymbol{\rho}(\lambda_t) \to 0$, the perturbation vanishes and the CD solution approaches the original protocol, as expected.

However in an actual chaperone system the full CD solution in Eq. (50) cannot be implemented exactly, because we cannot simultaneously vary chaperone concentration and both rates via external parameters. Carrying out the concentration protocol $\widetilde{C}(t)$ alone is in principal realizable, and in Fig. 5C we show the behavior of the system under this approximate CD approach (labeled "approximation 1"), where $\widetilde{C}(t)$ is given by Eq. (50), but $\widetilde{k}_a(t) = k_a$, $\widetilde{k}_n(t) = k_n$ are fixed. As expected, the state probabilities no longer exactly follow the target trajectory, but the driving is partially effective. In particular the state 2 (misfolded) probability still follows the target closely, dropping steeply over the course of several minutes. The reason for this is that the transition rates on all edges in the network that connect to state 2 agree with the exact CD solution: k_m, k_{-m}, k_u, k_{-u} are fixed in the reference spanning tree solution, and $\widetilde{C}(t)$ follows Eq. (50). On the other hand, the edges involving k_a and k_n diverge from the exact CD solution, and so the probabilities of states connected by these edges (1, 3, and 4) veer further from the target than in the case of state 2.

Even partially implemented in this manner, the biological benefit of CD driving is clear: rapidly reducing the probability of a misfolded protein helps protect the cell against the threat of aggregation. In fact the $\widetilde{C}(t)$ protocol shown in Fig. 5B—its quick rise and overshoot, eventually leveling off at a higher concentration—has qualitative analogues in the actual heat shock response of organisms. In the immediate aftermath of a temperature increase in yeast, heat shock proteins needed for rapid response (like molecular chaperones) typically peak in concentration within the first ten minutes [17, 74], before reducing to lower levels. These later levels are still elevated relative to the pre-shock state, and remain so for hours while long-term recovery processes kick in. A similar overshoot occurs in the first few minutes after heat shock for the sigma factor σ^{32} of *E. coli*, a regulatory protein responsible for promoting the synthesis of a variety of heat shock proteins [76]. We then see downstream echoes of this overshoot in levels of mRNA transcribed from chaperone genes like *dnaK* and *ibp* [73].

Can the approximation to the CD solution be improved by adding another control knob? Since the enzymatic action of the chaperone, reverting the misfolded protein back to an intermediate state, involves the hydrolysis of ATP, the rate k_a depends on ATP concen-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

tration. If that concentration could vary significantly over time, then a time-dependent protocol like $\widetilde{k}_a(t)$ in Eq. (50) is imaginable. The curves labeled "approximation 2" in Fig. 5C show the results of this modified approximation, where $\widetilde{C}(t)$ and $\widetilde{k}_a(t)$ follow Eq. (50), and $\widetilde{k}_n(t) = k_n$ is fixed. Since now all edges connected to state 3 and state 2 have rates obeying the exact CD solution, the probabilities of both these states follow the target closely. For state 2 there is limited improvement relative to the first approach, with the state 2 probabilities nearly the same for both approximations. At least for this parameter set, the biological implications of the added control knob are small, since control of state 3 is not linked to a direct threat in the same way as state 2: misfolded proteins bound to chaperones are protected from aggregation. And thus it is not surprising that unlike the rise in chaperone concentration, increasing ATP concentrations is not a universal feature of the heat shock response. There is one documented case however where it does occur: in *E. coli* the concentration of ATP increases sharply by about a factor of two in the first minutes after heat shock, and then gradually decreases [73], an observation additionally supported by metabolic evidence [77]. This has rough similarities to the transient increase seen in our $\widetilde{k}_a(t)$ protocol of Fig. 5B. So while manipulating $\widetilde{k}_a(t)$ does seem feasible (to an extent) in certain real systems, its precise significance in chaperone-mediated driving is debateable: the transient ATP increase in *E. coli* might be related to other aspects of the heat shock response.

In summary, some biological systems may not have enough controllable degrees of freedom to implement a full CD solution. However, a partial implementation, focused on edges connected to certain states of interest, allows us to approximately follow a target trajectory for that subset of states. This significantly broadens the scope of applicability of the CD approach, beyond just scenarios where exact CD is possible.

III. CD DRIVING IN CONTINUOUS SYSTEMS

Let us consider discrete-state Markov models on lattice graphs (also known as grid graphs). In these cases the states can be visualized as points on some d-dimensional lattice, with transitions occurring between neighboring lattice points. If we imagine the states as actual positions in a d-dimensional space, and allow the lattice spacing to become infinitesimal as the number of states $N \to \infty$, then the behavior of such models should approach continuum diffusive dynamics described by Fokker-Planck equations. Thus, taking appropriate limits, we should be able to use our formalism to derive CD solutions for Fokker-Planck systems. In Sec. IIIA, we describe how to do this for a $d = 1$ lattice, while the SI describes a $d = 2$ example. The $d = 1$ derivation recovers the Fokker-Planck CD driving results of Refs. [55, 78], while for $d = 2$ we show a generalized version of the CD driving used for the evolutionary Fokker-Planck system in Ref. [58]. The higher-dimensional case also shows how the graphical algorithm enables a discretized approach for numerically solving continuum CD problems, capable of handling fine meshes with large N. Beyond this validation, we demonstrate how CD driving works for systems exhibiting position-dependent diffusivity (not considered in Refs. [55, 78]) and in Sec. IIIB illustrate the usefulness of this feature with a concrete example from AFM force spectroscopy on biomolecules.

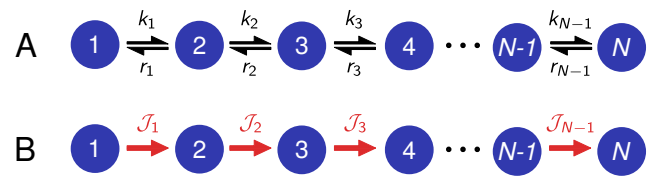

FIG. 6. A) N-state Markov model on a one-dimensional lattice graph, with $E = N - 1$ edges. Black arrows correspond to transitions between neighboring states, $k_i(\lambda_t)$ and $r_i(\lambda_t)$, $i = 1, \ldots, N-1$, which depend on the control protocol λ_t. B) Oriented stationary currents $\mathcal{J}_\alpha(t)$, $\alpha = 1, \ldots, N-1$. These currents form the only spanning tree for the graph.

A. Generalization to Fokker-Planck systems

To connect our earlier formalism to Fokker-Planck dynamics, let us first describe a one-dimensional Fokker-Planck equation for the time evolution of a probability density $p(x,t)$,

$$\partial_t p(x,t) = -\frac{\partial}{\partial x}\left[A(x)p(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[D(x)p(x,t)\right], \quad (52)$$

where x is our position variable, $A(x,t)$ is the drift function, and $D(x)$ is the position-dependent local diffusivity. Though $D(x)$ is often taken to be a constant, $D(x) = D$, here we allow it to be position-dependent for generality (and since it will prove useful for our later force spectroscopy example). We focus on the case where the drift $A(x,t) = -D(x)\partial_x E(x,\lambda_t)$, and hence arises from forces due to a potential energy $E(x,\lambda_t)$ that may be dependent on time-varying control parameters λ_t. Eq. (52) can then be rewritten as

$$\begin{aligned}\partial_t p(x,t) &= -\frac{\partial}{\partial x}\left[-D(x)\rho(x,\lambda_t)\frac{\partial}{\partial x}\frac{p(x,t)}{\rho(x,\lambda_t)}\right], \\ &\equiv -\frac{\partial}{\partial x}J(x,t),\end{aligned} \quad (53)$$

where

$$\rho(x,\lambda_t) = \frac{e^{-\beta E(x,\lambda_t)}}{Z(\lambda_t)}. \quad (54)$$

From the structure of Eq. (53) it is clear that $\rho(x,\lambda_t)$ is the instantaneous stationary distribution that makes the right-hand side vanish. We assume the energy function $E(x,\lambda_t) \to \infty$ as $x \to x_L$ and $x \to x_R$, defining a domain

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of x of width $\Delta x = x_R - x_L$. Thus the partition function $Z(\lambda_t) = \int_{x_L}^{x_R} dx\, \exp(-\beta E(x, \lambda_t))$ is well-defined. An infinite domain would correspond to the special case where $\Delta x \to \infty$. The second line of Eq. (53) defines a probability current density $J(x,t)$, in terms of which the Fokker-Planck equation takes the form of a continuity equation for probability.

To apply our general solution approach for discrete Markov systems, let us construct a one-dimensional lattice graph Markov model with N states, shown in Fig. 6A, that approximates the Fokker-Planck equation as $N \to \infty$. State i corresponds to position $x_i = x_L + ia$, where $a = \Delta x/N$ is the lattice spacing, which becomes infinitesimal for large N. In this limit the probability $p_i(t)$ of being in state i is related to the probability density $p(x,t)$ through $a^{-1}p_i(t) \to p(x_i, t)$.

In the discrete model the nonzero transition matrix elements correspond to the forward (right) arrows, $\Omega_{i+1,i}(\lambda_t) = k_i(\lambda_t)$, and the backward (left) arrows, $\Omega_{i,i+1}(\lambda_t) = r_i(\lambda_t)$, for $i = 1, \ldots, N-1$. We choose the following forms for the transition rates [79]:

$$\begin{aligned} k_i(\lambda_t) &= \frac{D_i}{a^2} e^{-\frac{1}{2}\beta(E_{i+1}(\lambda_t) - E_i(\lambda_t))}, \\ r_i(\lambda_t) &= \frac{D_i}{a^2} e^{\frac{1}{2}\beta(E_{i+1}(\lambda_t) - E_i(\lambda_t))}. \end{aligned} \tag{55}$$

Here $D_i \equiv D(x_i)$ and $E_i(\lambda_t) \equiv E(x_i, \lambda_t)$ are the discrete versions of the local diffusivity and potential energy. The ratio of the forward and backward transitions satisfies the local detailed balance relationship

$$\frac{k_i(\lambda_t)}{r_i(\lambda_t)} = e^{-\beta(E_{i+1}(\lambda_t) - E_i(\lambda_t))}. \tag{56}$$

As a result, the instantaneous stationary distribution for this system assumes a form analogous to Eq. (54),

$$\rho_i(\lambda_t) = \frac{e^{-\beta E_i(\lambda_t)}}{\mathcal{Z}(\lambda_t)}, \tag{57}$$

where $\mathcal{Z}(\lambda_t) = \sum_{i=1}^{N} \exp(-\beta E_i(\lambda_t))$. To check whether the transition rates of Eq. (55) give the correct Fokker-Planck equation in the continuum limit, we note that the master equation for the discrete system can be written as:

$$\partial_t p_i(t) = \sum_j \Omega_{ij}(\lambda_t) p_j(t) = -J_{i+1}(t) + J_i(t), \tag{58}$$

where the current from state i to $i+1$ is given by

$$J_i(t) = k_i(\lambda_t) p_i(t) - r_i(\lambda_t) p_{i+1}(t). \tag{59}$$

Eq. (58) is the discrete analogue of the second line in Eq. (53), with the conversion $J_i(t) \to J(x_i, t)$, $a^{-1}p_i(t) \to p(x_i, t)$. Plugging Eq. (55) into Eq. (59), we can rewrite the current $J_i(t)$ as

$$J_i(t) = -\frac{1}{a} D_i \sqrt{\rho_{i+1}(\lambda_t)\rho_i(\lambda_t)} \frac{1}{a} \left[\frac{p_{i+1}(t)}{\rho_{i+1}(\lambda_t)} - \frac{p_i(t)}{\rho_i(\lambda_t)} \right]. \tag{60}$$

Eq. (60) goes to the correct limit in the continuum case, becoming the current density in the square brackets in Eq. (53). To see this, note that

$$\begin{aligned} & a^{-1}\sqrt{\rho_{i+1}(\lambda_t)\rho_i(\lambda_t)} \\ & \quad \to [\rho(x_i + a, \lambda_t)\rho(x_i, \lambda_t)]^{1/2} \\ & \quad \approx [(\rho(x_i, \lambda_t) + a\partial_x \rho(x_i, \lambda_t))\, \rho(x_i, \lambda_t)]^{1/2} \\ & \quad = \rho(x_i, \lambda_t) + \mathcal{O}(a), \end{aligned} \tag{61}$$

where $\mathcal{O}(a)$ denotes corrections of order a. Thus Eq. (55) is a valid discretization of the Fokker-Planck system. It is not unique (other choices are possible, as shown in the SI) but any valid discretization should lead to the same CD results in the continuum limit.

With the discretization validated, we can now proceed to applying the general solution procedure. The oriented current graph (N states, $E = N-1$ edges) is tree-like, so the graph itself is the only spanning tree. Using the graphical algorithm we can write down the $(N-1) \times (N-1)$ dimensional stretched inverse reduced incidence matrix for this tree,

$$[\widehat{\nabla}^{(1)}]_S^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & -1 & 0 & 0 & \cdots & 0 \\ -1 & -1 & -1 & 0 & \cdots & 0 \\ -1 & -1 & -1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & -1 & \cdots & -1 \end{pmatrix}. \tag{62}$$

Because the graph is tree-like, the stretched inverse is also the ordinary inverse of the reduced incidence matrix, $[\widehat{\nabla}^{(1)}]_S^{-1} = \widehat{\nabla}^{-1}$. From Eqs. (55) and (57) we can deduce that the stationary currents have zero magnitude:

$$\mathcal{J}_i(t) = k_i(\lambda_t)\rho_i(\lambda_t) - r_i(\lambda_t)\rho_{i+1}(\lambda_t) = 0. \tag{63}$$

Hence we know that $\widetilde{\boldsymbol{\mathcal{J}}}(t) = \delta\boldsymbol{\mathcal{J}}(t)$. Moreover, since there are no cycles in the graph, Eq. (18) gives us the full CD current solution:

$$\widetilde{\boldsymbol{\mathcal{J}}}(t) = [\widehat{\nabla}^{(1)}]_S^{-1} \partial_t \widehat{\boldsymbol{\rho}}(\lambda_t). \tag{64}$$

Let us assume CD rates $\widetilde{k}_i(t)$ and $\widetilde{r}_i(t)$ of a form analogous to Eq. (55),

$$\begin{aligned} \widetilde{k}_i(t) &= \frac{\widetilde{D}_i(t)}{a^2} e^{-\frac{1}{2}\beta(\widetilde{E}_{i+1}(t) - \widetilde{E}_i(t))}, \\ \widetilde{r}_i(t) &= \frac{\widetilde{D}_i(t)}{a^2} e^{\frac{1}{2}\beta(\widetilde{E}_{i+1}(t) - \widetilde{E}_i(t))}, \end{aligned} \tag{65}$$

where $\widetilde{D}_i(t)$ represents a modified, potentially time-dependent, local diffusivity which we allow for generality, and $\widetilde{E}_i(t)$ is the energy associated with state i in the CD protocol. In many cases it may not be possible to control the local diffusivity via external parameters, and hence it remains unchanged, $\widetilde{D}_i(t) = D_i$. However as will be seen from the structure of the CD solution described below, we have in principle the freedom to choose $\widetilde{D}_i(t)$ to be

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

any non-negative function. The energy perturbation at each site due to the CD protocol is $U_i(t) = \widetilde{E}_i(t) - E_i(\lambda_t)$. To solve for these CD perturbations $U_i(t)$, the first step is to rewrite Eq. (64) using Eq. (62) and the expression for $\widetilde{\mathcal{J}}(t)$ in terms of the CD transition rates:

$$\widetilde{k}_i(t)\rho_i(\lambda_t) - \widetilde{r}_i(t)\rho_{i+1}(\lambda_t) = -\sum_{j=1}^{i} \partial_t \widehat{\rho}_j(\lambda_t). \tag{66}$$

After plugging in Eq. (65) for the CD rates, and Eq. (57) for the stationary distribution, Eq. (66) can be written as:

$$\begin{aligned} &-2a^{-2}\widetilde{D}_i(t)\sqrt{\widehat{\rho}_i(\lambda_t)\widehat{\rho}_{i+1}(\lambda_t)}\sinh\left[\frac{\beta(U_{i+1}(t)-U_i(t))}{2}\right] \\ &\qquad = -\sum_{j=1}^{i}\partial_t\widehat{\rho}_j(\lambda_t). \end{aligned} \tag{67}$$

We can invert this to find a recursion relation for the $U_i(t)$,

$$U_{i+1}(t) - U_i(t) = \frac{2}{\beta}\sinh^{-1}\left[\frac{a^2\sum_{j=1}^{i}\partial_t\widehat{\rho}_j(\lambda_t)}{2\widetilde{D}_i(t)\sqrt{\widehat{\rho}_i(\lambda_t)\widehat{\rho}_{i+1}(\lambda_t)}}\right]. \tag{68}$$

Given an arbitrary choice of function $U_1(t)$ (which corresponds to the freedom of redefining the zero level for energies), we can use consecutive applications of Eq. (68) to solve for $U_i(t)$, $i = 2, \ldots, N$.

The final step is to transform the CD results back to the continuum, where the CD energies can be expressed as $\widetilde{E}(x,t) = E(x,\lambda_t) + U(x,t)$. The perturbations $U(x,t)$ can be found from the continuum analogue of Eq. (68),

$$\frac{\partial U(x,t)}{\partial x} = \frac{1}{\beta\widetilde{D}(x,t)\rho(x,t)}\int_{x_L}^{x} dx'\partial_t\rho(x',\lambda_t). \tag{69}$$

To derive this we have expanded in small a and used the fact that $\sinh^{-1}(\epsilon) \approx \epsilon$ to lowest order in ϵ. In the continuum limit $\sum_{j=1}^{i} a\,\partial_t\widehat{\rho}_j(\lambda_t) \to \int_{x_L}^{x} dx'\partial_t\rho(x',\lambda_t)$ and $a^{-1}\sqrt{\widehat{\rho}_i(\lambda_t)\widehat{\rho}_{i+1}(\lambda_t)} \to \rho(x,\lambda_t)$, to leading order. This follows from the same argument as Eq. (61), setting $x = x_i$.

From Eq. (69) we can directly solve for $U(x,t)$,

$$\begin{aligned} U(x,t) &= U_0(t) \\ &\quad + \int_{x_0}^{x} dx' \frac{1}{\beta\widetilde{D}(x',t)\rho(x',\lambda_t)}\int_{x_L}^{x'} dx''\partial_t\rho(x'',\lambda_t), \end{aligned} \tag{70}$$

where x_0 is an arbitrary reference position and $U_0(t)$ is an arbitrary energy offset function (which does not affect the driving).

In practice, a particular CD protocol means simultaneously implementing the diffusivity $\widetilde{D}(x,t)$ and perturbing the energy landscape by $U(x,t)$. As mentioned earlier, in many experimental scenarios control of diffusivity will not be possible, so the only available CD protocols will involve keeping the diffusivity equal to the value in the original system, $\widetilde{D}(x,t) = D(x)$. One special case of this is a position-independent diffusivity $D(x) = D$ that is not varied during the CD protocol. This was solved by Li *et al.* [78] and Patra & Jarzynski [55] using alternative approaches, and their expressions for the CD perturbation are equivalent to our Eq. (69) with the substitution $\widetilde{D}(x,t) = D$.

From the perspective of thermodynamic costs, Eq. (33) for our discrete-state system takes the form

$$\begin{aligned} &\dot{S}^{\text{tot}}(t) \\ &\quad = k_B\sum_{i=1}^{N-1}\widetilde{\mathcal{J}}_i(t)\ln\frac{\widetilde{k}_i(t)\rho_i(\lambda_t)}{\widetilde{r}_i(t)\rho_{i+1}(\lambda_t)} \\ &\quad = -\frac{1}{T}\sum_{i=1}^{N-1}\widetilde{\mathcal{J}}_i(t)(U_{i+1}(t)-U_i(t)) \\ &\quad = 2k_B\sum_{i=1}^{N-1}\widetilde{\mathcal{J}}_i(t)\sinh^{-1}\left[\frac{a^2\widetilde{\mathcal{J}}_i(t)}{2\widetilde{D}_i(t)\sqrt{\widehat{\rho}_i(\lambda_t)\widehat{\rho}_{i+1}(\lambda_t)}}\right], \end{aligned} \tag{71}$$

where we have used the CD rates from Eq. (65) and $\widetilde{\mathcal{J}}_i(t) = -\sum_{j=1}^{i}\partial_t\widehat{\rho}_j(\lambda_t)$ from Eqs. (62)-(64). The functional form for $\dot{S}^{\text{tot}}(t)$ is always non-negative, since $y\sinh^{-1}(cy) \geq 0$ for any y when $c \geq 0$. In the limit of adiabatically slow driving, $\partial_t\widehat{\rho}_j(\lambda_t) \to 0$, we see that $\widetilde{\mathcal{J}}_i(t) \to 0$ and hence the entropy production rate $\dot{S}^{\text{tot}}(t) \to 0$. As noted in Sec. IE, under the (unlikely) scenario that one can control the local diffusivity $\widetilde{D}_i(t)$ and make it large during the CD protocol, then $\dot{S}^{\text{tot}}(t)$ can be made small even for fast driving.

In the continuum limit, Eq. (71) becomes

$$\dot{S}^{\text{tot}}(t) = k_B\int_{x_L}^{x_R} dx\,\frac{\widetilde{\mathcal{J}}^2(x,t)}{\widetilde{D}(x,t)\rho(x,t)}, \tag{72}$$

where $\widetilde{\mathcal{J}}_i(t) \to \widetilde{\mathcal{J}}(x,t)$ is the continuum CD current. This expression has the same form as the standard Fokker-Planck result for $\dot{S}^{\text{tot}}(t)$ [80], with the CD current and CD local diffusivity substituted for the original ones.

B. AFM force spectroscopy on biomolecules

As an example application of the Fokker-Planck CD theory described in the previous section, let us consider the analysis of biomolecular force spectroscopy experiments [6]. We will focus on a single-molecule atomic force microscopy (AFM) setup, illustrated schematically in Fig. 7A, though our approach readily generalizes to other force spectroscopy setups, like those involving optical tweezers [81, 82]. In our case a complex of two biomolecules bound to each other (or alternatively a single biomolecule) is attached to two polymeric handles

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(i.e. DNA or polyproteins) in solution. The bottom handle is bound to the surface of a stage whose vertical position is controlled by the experimentalist. The top handle is attached to the tip of the AFM cantilever, and the other end of the cantilever is fixed to a stationary support. Let us denote the vertical distance between the unloaded position of the cantilever tip and the current stage position ℓ_t. In a single experimental run, the experimentalist moves the stage downward, increasing ℓ_t according to some pre-defined control protocol, so we will take our control parameter $\lambda_t \equiv \ell_t$. A simple protocol involves starting at some initial separation ℓ_0, and then increasing ℓ_t at some constant velocity v. As the handles become taut, the force on the handle / biomolecule system increases, and the cantilever begins to deflect downward. Eventually the tension becomes large enough that the complex ruptures (the two biomolecules unbind), or if we are studying a single biomolecule, it unfolds. In either case the rupture/unfolding event is apparent to the experimentalist as a sudden decrease in the amount of force registered in the AFM cantilever (a decrease in the deflection). The time t at which the event occurs is recorded, and the experimental run is done. Many iterations of the same protocol on the same biomolecular system will typically yield different values of t for each iteration, because the rupture/unfolding process is inherently stochastic. The resulting distribution $p_{\rm rup}(t)$ of rupture times t is the main experimental quantity of interest.

What can we learn about the biomolecular complex from knowing $p_{\rm rup}(t)$ for a given protocol ℓ_t? Often we are interested in extracting a more basic property of the system, like the mean lifetime $\tau(F)$ of the bond holding the complex together, at a certain constant applied force F. Though the experiment is done with time-varying force (because a force ramp accelerates rupture and allows more events to be recorded), $p_{\rm rup}(t)$ in principle contains information to deduce $\tau(F)$ under hypothetical constant force conditions [8, 9]. Knowledge of $\tau(F)$ can in turn be used (under certain assumptions) to characterize properties of the underlying free energy landscape of the bond, like the transition-state distance or energy barrier to rupture [10, 83]. Alternatively, $p_{\rm rup}(t)$ can be used to infer the presence or absence of long-lived heterogeneous bound states of the complex [11]. However most such approaches to analyze $p_{\rm rup}(t)$ [8–11] assume that the experiment is being done quasi-adiabatically: in other words, the pulling speed v must be slow enough that at each value of ℓ_t the entire system is approximately in instantaneous equilibrium. Some analysis is also possible in the asymptotic limit of extremely fast pulling [12], but the quasi-adiabatic limit remains the most common starting point for building models to interpret single-molecule force spectroscopy data.

What if we could mimic quasi-adiabatic behavior even when pulling at arbitrarily fast speeds? To see how this can be done through the CD theory above, let us first construct a Fokker-Planck description of the system dynamics.

1. *Constructing a model of an AFM pulling experiment*

Let x_t be the vertical distance between stage surface and the cantilever tip (Fig. 7A) at time t. This is the vertical extension of what we will refer to as the molecular portion of the system (the two handles + the biomolecular complex). The amount by which the cantilever is deflected vertically is given by $\ell_t - x_t$, and hence the force exerted by the cantilever on the molecular portion is:

$$F_{\rm cant}(x_t, \ell_t) = k_{\rm cant}(\ell_t - x_t), \tag{73}$$

where $k_{\rm cant}$ is the cantilever spring constant. We take $k_{\rm cant} = 10$ pN/nm as typical for AFM cantilevers. Our stage position protocol has the following form:

$$\ell_t = \ell_0 + \frac{1}{\zeta}\log\left(1 + e^{\zeta(t-t_r)v}\right), \tag{74}$$

with parameters $\ell_0 = 200$ nm, $\zeta = 0.5$ nm^{-1}, $v = 10^5$ nm/s, and $t_r = 0.2$ ms. As illustrated by the red curve in Fig. 7B, the protocol starts at ℓ_0 for $t \ll t_r$, and then smoothly transitions to a constant velocity ramp, with $\ell_t \approx v(t - t_r)$ for $t \gg t_r$. The smoothing parameter ζ controls the abruptness of the transition, which we have made continuously differentiable for mathematical convenience. As we will see, the velocity $v = 10^5$ nm/s is rapid enough that the system is out of instantaneous equilibrium in the absence of CD driving.

We estimate the force $F_{\rm mol}(x_t)$ exerted by the molecular portion of the system on the cantilever as follows. The complex will typically have a much stiffer effective spring constant than the handles, and since the two handles and complex are in series, the primary contribution to $F_{\rm mol}(x_t)$ is from the elastic properties of the polymeric handles. We can approximate this force using an extensible semiflexible polymer model, appropriate for example for DNA handles. In this case $F_{\rm mol}(x_t)$ satisfies the following approximate relation [84, 85]:

$$F_{\rm mol}(x_t) \approx \frac{k_BT}{l_p}\left[\frac{1}{4\left(1 - \frac{x_t}{L} + \frac{F_{\rm mol}(x_t)}{K}\right)^2} - \frac{1}{4} + \frac{x_t}{L} - \frac{F_{\rm mol}(x_t)}{K}\right]. \tag{75}$$

We assume parameters typical for double-stranded DNA: persistence length $l_p = 50$ nm, stretch modulus $K = 1100$ pN [85]. We take the combined contour length of both handles to be $L = 300$ nm. Note $F_{\rm mol}(x_t)$ appears on both the left and right sides of Eq. (75). For a given molecular extension x_t, Eq. (75) can be numerically solved for the corresponding value of $F_{\rm mol}(x_t)$, and hence we can construct a numerical interpolation of the function $F_{\rm mol}(x_t)$.

Thus the dynamics of the cantilever tip, described by vertical position x_t relative to the stage surface, are subject to the cantilever and molecular forces pulling in opposite directions, yielding a total force

$$F(x_t, \ell_t) = F_{\rm cant}(x_t, \ell_t) - F_{\rm mol}(x_t). \tag{76}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

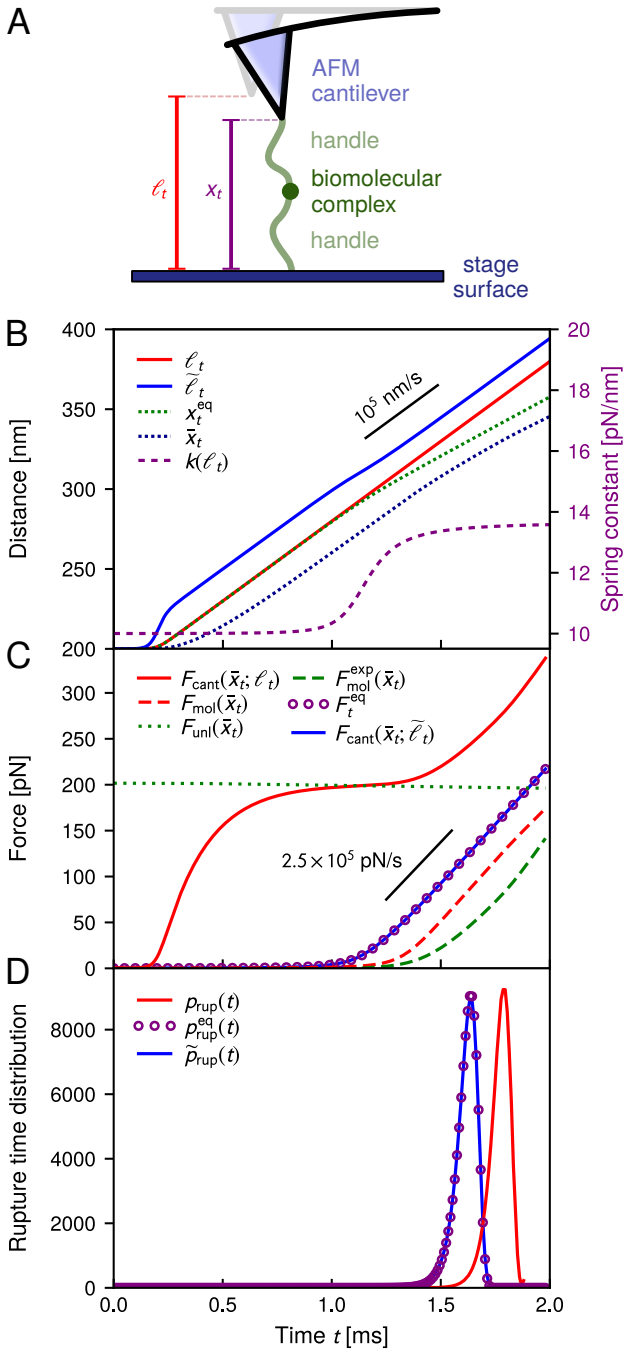

FIG. 7. A) Schematic of an AFM pulling experiment designed to probe rupture properties of a biomolecular complex. B) Original control protocol ℓ_t and the corresponding CD protocol $\widetilde{\ell}_t$. Under the original protocol the mean extension $\bar{x}_t$ lags behind the instantaneous equilibrium extension $x_t^{\rm eq}$. The effective spring constant $k(\ell_t)$ during the original protocol is also shown for comparison. C) Forces in the system under both original and CD protocols (details in the text). D) The rupture time distribution $p_{\rm rup}(t)$ under the original protocol, which disagrees with the quasi-equilibrium expectation $p_{\rm rup}^{\rm eq}(t)$. In contrast, the distribution $\widetilde{p}_{\rm rup}(t)$ under CD driving closely mimics the quasi-equilirium result.

We can interpret $F(x,\ell_t)$ to be due to an effective potential energy defined through

$$E(x,\ell_t) = -\int_0^x dx' F(x',\ell_t). \tag{77}$$

Note the dependence on the protocol ℓ_t, which means the energy landscape changes continuously in time. Since the region $x<0$ is forbidden, we take $E(x,\ell_t)=\infty$ for $x<0$. Let us denote the position at which the potential energy is minimized at time t to be $x_t^{\rm eq}$. This is the equilibrium position where the cantilever and molecular forces cancel, $F_{\rm cant}(x_t^{\rm eq},\ell_t) = F_{\rm mol}(x_t^{\rm eq}) \equiv F_t^{\rm eq}$. In the vicinity of $x_t^{\rm eq}$ the potential energy can be approximated as a harmonic well,

$$E(x,\ell_t) \approx E(x_t^{\rm eq},\ell_t) + \frac{1}{2}k(\ell_t)(x-x_t^{\rm eq})^2, \tag{78}$$

with a protocol-dependent effective spring constant

$$k(\ell_t) = k_{\rm cant} + \frac{\partial F_{\rm mol}}{\partial x}(x_t^{\rm eq}). \tag{79}$$

Fig. 7B shows numerical results for the equilibrium position $x_t^{\rm eq}$ versus t (dotted curve) for the protocol parameters described above, as well as the effective spring constant $k(\ell_t)$ (dashed curve). Initially the molecular portion of the system is slack, and the spring constant is dominated by the cantilever contribution, $k(\ell_t)\approx k_{\rm cant}$. In this regime $x_t^{\rm eq}$ remains close to ℓ_t and the equilibrium force $F_t^{\rm eq}$ (shown as circles in Fig. 7C) is small. For $t \gtrsim 1$ ms the molecular portion starts becoming taut, giving an added contribution to the spring constant $k(\ell_t)$. The equilibrium position $x_t^{\rm eq}$ begins to diverge from ℓ_t, as the cantilever gets deflected downwards, and the corresponding force $F_t^{\rm eq}$ starts to increase roughly as a linear ramp.

To complete the model, we need to write down a Fokker-Planck equation for x_t. We can treat x_t as a position variable that fluctuates in a time-varying potential $E(x,\ell_t)$. Though x_t represents the junction of the AFM cantilever and the molecular portion, the associated diffusivity of x_t is dominated by the viscous drag of the cantilever, whose dimensions are typically much larger than the molecular portion. Thus to good approximation we can set the diffusivity equal to that of the cantilever, using a phenomenological model of the drag on a cantilever in the vicinity of a surface that has proven to fit well in earlier experimental analyses of AFM pulling experiments [86, 87]:

$$D(x_t) = \frac{k_B T(x_t + h_{\rm eff})}{6\pi\eta a_{\rm eff}^2}, \tag{80}$$

where temperature $T=298$ K, $\eta = 0.89$ mPa·s is the viscosity of water, and $a_{\rm eff}$, $h_{\rm eff}$ are parameters with dimensions of length. We set $a_{\rm eff}=25$ μm, $h_{\rm eff}=5$ μm, values in the typical experimental range [86, 87]. Note the decrease of diffusivity $D(x_t)$ with decreasing x_t, which approximately captures wall drag effects due

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

to hydrodynamic interactions as the tip approaches the surface [88, 89].

If the biomolecular complex were not allowed to rupture, the resulting dynamics would be described by a probability density $p(x,t)$ to find the system with tip-to-surface distance x at time t, obeying a Fokker-Planck equation of the form of Eq. (53),

$$\partial_t p(x,t) = \frac{\partial}{\partial x} D(x) e^{-\beta E(x,\ell_t)} \frac{\partial}{\partial x} e^{\beta E(x,\ell_t)} p(x,t), \quad (81)$$

To account for rupture, we interpret $p(x,t)$ as referring to the intact system, and add a sink term to Eq. (81) to model the loss of probability over time as the system undergoes stochastic rupture events [79]. There are a variety of possible functional forms to describe the rupture rate at a given force (or equivalently the inverse mean bond lifetime $\tau^{-1}(F)$), but the simplest one is the Bell model [83], where the probability per unit time for the biomolecular complex to rupture under force F is given by $\tau^{-1}(F) = k_0 \exp(\beta F d)$. Here k_0 is the base rupture rate when the handles are slack (negligible force on the complex), and the exponential term accounts for the increase in rupture rate as the tension F in the molecular portion of the system increases. We choose experimentally typical parameter values of $k_0 = 0.1$ s^{-1}, $d = 0.4$ nm, the latter parameter often interpreted as the distance between the bound and transition state along the reaction coordinate.

Then the final form for the Fokker-Planck equation, incorporating the sink term due to rupture, is:

$$\begin{aligned}\partial_t p(x,t) = &\frac{\partial}{\partial x} D(x) e^{-\beta E(x,\ell_t)} \frac{\partial}{\partial x} e^{\beta E(x,\ell_t)} p(x,t) \\ &- k_0 e^{\beta F_{\rm mol}(x) d} p(x,t),\end{aligned} \quad (82)$$

where we use $F_{\rm mol}(x)$ in the sink term, since this is the tension felt at the molecular complex. To connect $p(x,t)$ to the experimental observable $p_{\rm rup}(t)$, we note that the probability $\Sigma(t)$ that the system is intact (not ruptured) at time t is given by

$$\Sigma(t) = \int_{-\infty}^{\infty} dx\, p(x,t). \quad (83)$$

where we have taken the domain bounds to be $x_L = -\infty$, $x_R = \infty$. Then the probability $p_{\rm rup}(t)dt$ that the system ruptured between times t and $t + dt$ is just $\Sigma(t) - \Sigma(t + dt) \approx -dt\, \partial_t \Sigma(t)$. Hence

$$p_{\rm rup}(t) = -\partial_t \Sigma(t). \quad (84)$$

The other observable of interest is the mean extension $\bar{x}_t$ at time t (conditioned on the system being intact), defined by

$$\bar{x}_t = \frac{1}{\Sigma(t)} \int_{-\infty}^{\infty} dx\, x p(x,t). \quad (85)$$

2. *Pulling in the absence of counterdiabatic driving*

First, we numerically solve the Fokker-Planck equation of Eq. (82) with the protocol ℓ_t described by Eq. (74). At the given velocity, $v = 10^5$ nm/s, the system before rupture is clearly not in the quasi-adiabatic regime. The mean extension $\bar{x}_t$ (green dotted curve in Fig. 7B) lags behind the instantaneous equilibrium extension $x_t^{\rm eq}$ (blue dotted curve). This means that at the mean extension the cantilever and molecular forces do not balance each other out: $F_{\rm cant}(\bar{x}_t, \ell_t)$ (red solid curve in Fig. 7C) does not equal $F_{\rm mol}(\bar{x}_t)$ (red dashed curve), or the instantaneous equilibrium force $F_t^{\rm eq}$ (purple circles) one would expect for quasi-adiabatic pulling.

In an actual experimental setting, the force on the molecule at the moment of rupture would be estimated in the following way. Since the force on the cantilever can be determined from measurements of its deflection, the mean force on the cantilever $F_{\rm cant}(\bar{x}_t, \ell_t)$ right before the rupture would be recorded, and then compared to the mean force on the unloaded cantilever $F_{\rm unl}(\bar{x}_t)$ right after the rupture. The latter is nonzero because the cantilever still feels fluid drag even after the complex has ruptured, $F_{\rm unl}(\bar{x}_t) = k_B T v / D(\bar{x}_t)$ (green dotted curve in Fig. 7C). The difference $F_{\rm mol}^{\rm exp}(\bar{x}_t) = F_{\rm cant}(\bar{x}_t, \ell_t) - F_{\rm unl}(\bar{x}_t)$ (green dashed curve) is the experimentally estimated force on the molecule at rupture. $F_{\rm mol}^{\rm exp}(\bar{x}_t)$ is closer to the actual mean force on the molecule, $F_{\rm mol}(\bar{x}_t)$, than the uncorrected $F_{\rm cant}(\bar{x}_t, \ell_t)$ before factoring out hydrodynamic drag. However because of non-equilibrium effects $F_{\rm mol}^{\rm exp}(\bar{x}_t)$ is still an underestimate of $F_{\rm mol}(\bar{x}_t)$ [87].

The rupture time distribution $p_{\rm rup}(t)$ is shown in Fig. 7D. If one were to naively apply the assumption of instantaneous equilibrium in analyzing this data, as is done in the most common approaches to analyze pulling experiments [8, 9], we end up with significant errors. The quasi-adiabatic assumption rests on believing that at any given time t in the experiment, the applied force is changing sufficiently slowly that the system behaves as if it equilibrates at the instantaneous force $F_t^{\rm eq}$. For the Bell model the instantaneous rate of rupture at this force would be $\tau^{-1}(F_t^{\rm eq}) = k_0 \exp(\beta F_t^{\rm eq} d)$, and the intact probability would then obey the simple equation

$$\partial_t \Sigma_t = -k_0 e^{\beta F_t^{\rm eq} d} \Sigma(t). \quad (86)$$

The solution to this equation is

$$\Sigma(t) = \exp\left(-k_0 \int_{t_0}^{t} dt'\, e^{\beta F_{t'}^{\rm eq} d}\right). \quad (87)$$

We only focus on times $t \geq t_0$, where t_0 is the effective start time of recording rupture events in the experiment: the time when the force on the molecule begins to ramp up. From Eq. (84) the corresponding rupture time distribution is given by

$$p_{\rm rup}^{\rm eq}(t) = k_0 \exp\left(\beta F_t^{\rm eq} d - k_0 \int_{t_0}^{t} dt'\, e^{\beta F_{t'}^{\rm eq} d}\right). \quad (88)$$

Here the superscript eq denotes that the distribution is derived assumed quasi-equilibrium at all times before rupture. Typically $F_t^{\rm eq}$ is approximated by fitting a linear ramp to the experimentally estimated force on the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

	$r_{\rm exp}$ [pN/s]	t_0 [ms]	k_0 [s^{-1}]	d [nm]
best-fit to data (no CD)	2.3×10^5	1.4	3.1	0.45
expected (ground truth)	2.5×10^5	1.1	0.10	0.40
best-fit to data (with CD)	2.6×10^5	1.1	0.12	0.39

TABLE I. Analysis of an AFM pulling numerical simulation. $r_{\rm exp}$ is the force ramp rate determined from the simulated data, t_0 the approximate starting time of the ramp, k_0 the zero-force rupture rate of the biomolecular complex, and d the transition-state distance of the complex. The first row shows the best-fit parameter values based on fitting to the rupture time distribution $p_{\rm rup}(t)$ from the original protocol using the typical quasi-adiabatic assumption. The second row shows the expected values if the experiment actually satisfied the quasi-adiabatic assumption (which it does not). The final row shows the best-fit values for the distribution $\widetilde{p}_{\rm rup}(t)$ using the CD protocol. The parameters extracted from the CD protocol closely match the true values for the complex.

molecule, $F^{\rm exp}_{\rm mol}(t) \approx r_{\rm exp}(t-t_0)$ with fitting parameters $r_{\rm exp}$ and t_0. Substituting $r_{\rm exp}(t-t_0)$ for $F^{\rm eq}_t$ in Eq. (88), we can evaluate the integral in the expression to get

$$p^{\rm eq}_{\rm rup}(t) = k_0 \exp\Bigg[\beta r_{\rm exp}(t-t_0)d + \frac{k_0}{\beta r_{\rm exp} d}\left(1 - e^{\beta r_{\rm exp}(t-t_0)d}\right)\Bigg]. \tag{89}$$

The experimental analysis procedure would then consist of two steps: i) use the estimate $F^{\rm exp}_{\rm mol}(t)$ to extract $r_{\rm exp}$ and t_0, ii) plug these values into Eq. (89), and then fit to the experimental $p_{\rm rup}(t)$ to extract the Bell model parameters k_0 and d. If we apply this quasi-adiabatic approach to our numerical experimental results for $F^{\rm exp}_{\rm mol}(t)$ and $p_{\rm rup}(t)$ from Fig. 7C-D, the best-fit results yield the values shown in the first row of Table I. The expected results, if the quasi-adiabatic assumption were true, using the actual equilibrium force $F^{\rm eq}_t$ instead of $F^{\rm exp}_{\rm mol}$ and the true values for the Bell model, are shown in the second row. The quasi-adiabatic analysis leads to a significant discrepancy in the extracted value for k_0, which is off by an order of magnitude, and smaller errors in the other parameters. We can also see this by comparing $p^{\rm eq}_{\rm rup}(t)$ from Eq. (89) with the expected parameter values (purple circles in Fig. 7D) with the numerical results for $p_{\rm rup}(t)$ (red curve). The two distributions are non-trivially displaced from another, with the fast, non-adiabatic pulling leading to longer than expected rupture times.

3. *Pulling with counterdiabatic driving*

To see how CD driving can alleviate the non-adiabatic discrepancies described above, we first need to determine the CD protocol. Given the potential energy of Eq. (78) we evaluate the corresponding instantaneous equilibrium distribution from Eq. (54),

$$\rho(x,\ell_t) = \sqrt{\frac{\beta k(\ell_t)}{2\pi}} \exp\left(-\frac{1}{2}\beta k(\ell_t)(x - x^{\rm eq}_t)^2\right). \tag{90}$$

Since we do not have the experimental means to easily control diffusivity, we choose the diffusivity for the protocol to be the same as in the original system from Eq. (80), $\widetilde{D}(x,\ell_t) = D(x)$. Using Eq. (70), the energy perturbation needed to achieve CD driving is

$$U(x,t) = U_0(t) + \int_{x_0}^{x} dx' \frac{k_B T}{D(x')\rho(x',\ell_t)} \int_{-\infty}^{x'} dx'' \partial_t \rho(x'',\ell_t), \tag{91}$$

with some arbitrary x_0 and $U_0(t)$. The inner integral in Eq. (91) can be evaluated exactly, but it turns out an approximate form is sufficient in our case: instead of integrating over the full expression for $\partial_t\rho(x'',\ell_t)$, we ignore the terms that involve $\partial_t k(\ell_t)$. This is because the time derivative of the effective spring constant is nearly zero over the time range of interest except for a narrow time window where $k(\ell_t)$ shifts upwards, as seen from Fig. 7B. Using this approximation, the inner integral evaluates to

$$\int_{-\infty}^{x'} dx'' \dot{\rho}(x'',\ell_t) \approx \sqrt{\frac{\beta k(\ell_t)}{2\pi}} \dot{x}^{\rm eq}_t \exp\left(-\frac{1}{2}\beta k(\ell_t)(x' - x^{\rm eq}_t)^2\right), \tag{92}$$

where $\dot{x}^{\rm eq}_t \equiv (d/dt)x^{\rm eq}_t$, and Eq. (91) simplifies to

$$U(x,t) = U_0(t) + k_B T \dot{x}^{\rm eq}_t \int_{x_0}^{x} \frac{dx'}{D(x')}. \tag{93}$$

This energy perturbation corresponds to modifying the total force on the cantilever tip to a new value

$$\widetilde{F}(x,\ell_t) = F(x,\ell_t) + \frac{\partial U(x,t)}{\partial x} = F(x,\ell_t) + \frac{k_B T \dot{x}^{\rm eq}_t}{D(x)}. \tag{94}$$

Since $F_{\rm cant}(x_t,\ell_t) = k_{\rm cant}(\ell_t - x_t)$, this can be accomplished by replacing ℓ_t with an alternative protocol $\widetilde{\ell}_t$,

$$\widetilde{\ell}_t = \ell_t + \frac{k_B T \dot{x}^{\rm eq}_t}{k_{\rm cant} D(x)}. \tag{95}$$

Under the new protocol, the cantilever force $F_{\rm cant}(\bar{x}_t,\widetilde{\ell}_t)$ (blue curve in Fig. 7C) now overlaps with the the expected equilibrium force $F^{\rm eq}_t$. Crucially, the protocol requires only knowledge about components of the experimental setup, and is independent of the intrinsic properties of the biomolecular complex (k_0, d) that we are trying to estimate from the experiment. $k_{\rm cant}$ is a known parameter of the cantilever, and the parameters that determine $D(x)$ through Eq. (80) can be deduced from calibration experiments in the absence

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of any attached molecule. The equilibrium extension curve $x_t^{\rm eq}$ that is used to calculate $\dot{x}_t^{\rm eq}$ is the solution to $F_{\rm cant}(x_t^{\rm eq}, \ell_t) = F_{\rm mol}(x_t^{\rm eq})$. Under the assumption that the biomolecular complex is small and stiff relative to the polymeric handles, the function $F_{\rm mol}(x_t)$ given by Eq. (75) is determined by the handle properties (contour length, persistence length, and stretch modulus) which all can be fit via calibration experiments. Thus once the apparatus is parametrized in this manner, one can carry out experiments on a biomolecular complex with unknown k_0 and d. The stage can be programmed to execute the CD protocol $\widetilde{\ell}_t$ in Eq. (95), and with repeated trials we can construct an experimental estimate for the rupture time distribution.

We do this for the numerical model system described above, and the rupture time distribution $\widetilde{p}_{\rm rup}(t)$ under the CD protocol is shown in blue in Fig. 7D. As expected, it closely overlaps with the distribution $p_{\rm rup}^{\rm eq}(t)$ predicted by the quasi-equilibrium assumption. As a result, the best-fit parameters (third row of Table I) all agree well with the ground truth values, and in particular we have eliminated the large discrepancy in the estimated k_0. Thus CD driving allows us to mimic quasi-equilibrium conditions and apply quasi-equilibrium analysis to accurately extract biomolecular properties even under fast, non-adiabatic pulling.

IV. CONCLUDING REMARKS

Our theory of classical CD driving and its biological applications open up a variety of questions for future work. The fact that there can exist many CD protocols for the same target trajectory, with distinct thermodynamic properties, means that one can search among these protocols for those that optimize certain quantities—like minimizing dissipated work under given physical constraints. Optimal control of nonequilibrium and finite-time processes is an active research area [90–93], with connections to techniques like Monge-Kantorovich transport theory [94] and trajectory-observable biasing within the framework of large deviations [95, 96]. Situating CD driving within the broader context of these earlier optimal control approaches is an interesting topic for further study, both generically and in specific biological implementations in areas like ecology and evolution [97, 98].

Driving a system between long-lived states is also subject to universal bounds or "speed limits" [99–101] that constrain the speed of driving in terms of dissipated work. Does CD driving saturate these bounds in certain circumstances? If so, are there biological implications, for example cases where natural selection has pushed a control process close to the theoretical limit? Finally, we can relax the goal of the control, and target only the initial and final distributions, rather than a specific trajectory between them. In quantum systems there are so-called fast-forward techniques to tackle this problem [22–29], but what are the classical stochastic analogues?

In summary, stochastic processes and their biological realizations are an ideal laboratory for investigating nonequilibrium control ideas. The CD driving framework we have developed is a particularly useful starting point, because the control protocols can be expressed analytically in terms of easy-to-calculate graph properties of the underlying Markov model. We can thus in principle explore a wide swath of CD solutions, and identify generic features of control in diverse biological systems sharing similar graph topologies. The practicality of our formulation makes it well suited for deriving CD driving prescriptions in specific experimental contexts, like evolving cell populations [58] or the example systems in the current work. Because the CD protocols involve accessible control knobs—varying drug/protein concentrations, or the stage position in the AFM case— we believe near-term experimental validation is within reach. CD driving is thus poised to become an important tool both for implementing control of biological systems in the lab, and understanding how that control operates in nature.

Appendix A: Stationary probabilities and currents in one-loop and two-loop examples

Explicit expressions for stationary probability components $\rho_i(\lambda_t)$ in the one-loop system of Fig. 3A and the two-loop system of Fig. 2A can be calculated by solving Eq. (3) for the system, or equivalently employing the graphical method of Ref. [102]. The result is that each component can be expressed as $\rho_i(\lambda_t) = D_i(\lambda_t)/D(\lambda_t)$, where $D(\lambda_t) = \sum_{i=1}^N D_i(\lambda_t)$ for functions $D_i(\lambda_t)$, $i = 1, \ldots, N$ defined below.

For the one-loop example, let $k_i(\lambda_t)$ be the clockwise (forward) transition starting from state i, and $r_i(\lambda_t)$ the corresponding counterclockwise (reverse) rate. These can be expressed in terms of binding/unbinding constants and concentrations as shown in Fig. 3A. The functions $D_i(\lambda_t)$ are given by:

$$\begin{aligned} D_1 &= k_2 k_3 + r_1(k_3 + r_2) \\ D_2 &= k_1(k_3 + r_2) + r_2 r_3 \\ D_3 &= k_1 k_2 + (k_2 + r_1) r_3 \end{aligned} \tag{96}$$

For conciseness, in the equation above we have not shown the λ_t dependence of all the rate variables. The stationary currents in the system are all equal, and given by $\mathcal{J}_1 = \mathcal{J}_2 = \mathcal{J}_3 = (k_1 k_2 k_3 - r_1 r_2 r_3)/D$.

For the two-loop system, the labeling of the rates $k_i(\lambda_t)$ and $r_i(\lambda_t)$ is shown in Fig. 2A. The functions

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$D_i(\lambda_t)$ are given by:

$$\begin{aligned} D_1 =& k_2k_3\left(k_5+r_4\right)+\left(k_3+r_2\right)\left(k_5r_1+r_4\left(r_1+r_5\right)\right), \\ D_2 =& k_5r_2\left(k_4+r_3\right)+k_1\left(k_3+r_2\right)\left(k_5+r_4\right)+k_3k_4k_5 \\ &+r_2r_3r_4, \\ D_3 =& k_2k_5\left(k_4+r_3\right)+k_2r_3r_4+k_1k_2\left(k_5+r_4\right) \\ &+r_1r_3\left(k_5+r_4\right)+r_3r_4r_5 \\ D_4 =& k_4r_1\left(k_3+r_2\right)+\left(k_1+k_4\right)r_5\left(k_3+r_2\right)+k_2k_3k_4 \\ &+r_2r_3r_5. \end{aligned} \tag{97}$$

In this case there is one stationary current in the left loop, $\mathcal{J}_L \equiv J_2 = J_3$, and one in the right loop, $\mathcal{J}_R \equiv J_4 = J_5$. Edge 1, shared between the loops, has stationary current $\mathcal{J}_1 = \mathcal{J}_L - \mathcal{J}_R$. The left and right currents can be expressed as $\mathcal{J}_L = D_L/D$ and $\mathcal{J}_R = D_R/D$, where the two numerator functions are given by:

$$\begin{aligned} D_L =& k_1k_2k_3\left(k_5+r_4\right)-r_2r_3\left(r_1\left(k_5+r_4\right)+r_4r_5\right) \\ &+k_2k_3k_4k_5, \\ D_R =& k_4k_5r_1\left(k_3+r_2\right)-r_4r_5\left(k_1\left(k_3+r_2\right)+r_2r_3\right) \\ &+k_2k_3k_4k_5. \end{aligned} \tag{98}$$

ACKNOWLEDGMENTS

The authors would like to thank the stimulating environment provided by the Telluride Science Research Center, where this project was conceived. M.H. acknowledges support from the U.S. National Science Foundation (NSF) under Grant No. MCB-1651650. E.I. acknowledges support from the Labex CelTisPhyBio (ANR-11-LABX-0038, ANR-10-IDEX-0001-02).

[1] Johan Paulsson, "Models of stochastic gene expression," Phys. Life Rev. **2**, 157–175 (2005).
[2] Mauro L Mugnai, Changbong Hyeon, Michael Hinczewski, and D Thirumalai, "Theoretical perspectives on biological machines," Rev. Mod. Phys. **92**, 025001 (2020).
[3] Kyle A Beauchamp, Robert McGibbon, Yu-Shan Lin, and Vijay S Pande, "Simple few-state models reveal hidden complexity in protein folding," Proc. Natl. Acad. Sci. **109**, 17807–17813 (2012).
[4] Guy Sella and Aaron E Hirsh, "The application of statistical physics to evolutionary biology," Proc. Natl. Acad. Sci. **102**, 9541–9546 (2005).
[5] Daniel Nichol, Peter Jeavons, Alexander G Fletcher, Robert A Bonomo, Philip K Maini, Jerome L Paul, Robert A Gatenby, Alexander RA Anderson, and Jacob G Scott, "Steering evolution with sequential therapy to prevent the emergence of bacterial antibiotic resistance," PLoS Comput. Biol. **11**, e1004493 (2015).
[6] Jordanka Zlatanova, Stuart M Lindsay, and Sanford H Leuba, "Single molecule force spectroscopy in biology using the atomic force microscope," Prog. Biophys. Mol. Biol. **74**, 37–61 (2000).
[7] Ahmad S Khalil and James J Collins, "Synthetic biology: applications come of age," Nat. Rev. Genet. **11**, 367–379 (2010).
[8] Evan Evans and Ken Ritchie, "Dynamic strength of molecular adhesion bonds," Biophys. J. **72**, 1541–1555 (1997).
[9] Olga K Dudko, Gerhard Hummer, and Attila Szabo, "Theory, analysis, and interpretation of single-molecule force spectroscopy experiments," Proc. Natl. Acad. Sci. **105**, 15755–15760 (2008).
[10] Olga K Dudko, Gerhard Hummer, and Attila Szabo, "Intrinsic rates and activation free energies from single-molecule pulling experiments," Phys. Rev. Lett. **96**, 108101 (2006).
[11] Michael Hinczewski, Changbong Hyeon, and Devarajan Thirumalai, "Directly measuring single-molecule heterogeneity using force spectroscopy," Proc. Natl. Acad. Sci. **113**, E3852–E3861 (2016).
[12] Jakob T Bullerjahn, Sebastian Sturm, and Klaus Kroy, "Theory of rapid force spectroscopy," Nat. Commun. **5**, 4463 (2014).
[13] George H Lorimer, "A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo." FASEB J. **10**, 5–9 (1996).
[14] Devarajan Thirumalai and George H Lorimer, "Chaperonin-mediated protein folding," Annu. Rev. Biophys. Biomol. Struct. **30**, 245–269 (2001).
[15] Michael J Kerner, Dean J Naylor, Yasushi Ishihama, Tobias Maier, Hung-Chun Chang, Anna P Stines, Costa Georgopoulos, Dmitrij Frishman, Manajit Hayer-Hartl, Matthias Mann, *et al.*, "Proteome-wide analysis of chaperonin-dependent protein folding in escherichia coli," Cell **122**, 209–220 (2005).
[16] Mantu Santra, Daniel W Farrell, and Ken A Dill, "Bacterial proteostasis balances energy and chaperone utilization efficiently," Proc. Natl. Acad. Sci. **114**, E2654–E2661 (2017).
[17] Klaus Richter, Martin Haslbeck, and Johannes Buchner, "The heat shock response: life on the verge of death," Mol. Cell **40**, 253–266 (2010).
[18] Davide Roncarati and Vincenzo Scarlato, "Regulation of heat-shock genes in bacteria: from signal sensing to gene expression output," FEMS Microbiol. Rev. **41**, 549–574 (2017).
[19] S. Deffner and S. Campbell, *Quantum Thermodynamics* (Morgan & Claypool Publishers, 2019).
[20] X. Chen, A. Ruschhaupt, S. Schmidt, A. del Campo, D. Guéry-Odelin, and J. G. Muga, "Fast optimal frictionless atom cooling in harmonic traps: Shortcut to adiabaticity," Phys. Rev. Lett. **104**, 063002 (2010).
[21] A. del Campo and M. G. Boshier, "Shortcuts to adiabaticity in a time-dependent box," Sci. Rep. **2**, 648 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[22] S. Masuda and K. Nakamura, "Fast-forward of adiabatic dynamics in quantum mechanics," Proc. R. Soc. A **466**, 1135 (2010).

[23] S. Masuda and K. Nakamura, "Acceleration of adiabatic quantum dynamics in electromagnetic fields," Phys. Rev. A **84**, 043434 (2011).

[24] E. Torrontegui, S. Martínez-Garaot, A. Ruschhaupt, and J. G. Muga, "Shortcuts to adiabaticity: Fast-forward approach," Phys. Rev. A **86**, 013601 (2012).

[25] E. Torrontegui, X. Chen, M. Modugno, S. Schmidt, A. Ruschhaupt, and J. G. Muga, "Fast transport of Bose—Einstein condensates," New J. Phys. **14**, 013031 (2012).

[26] S. Masuda, K. Nakamura, and A. del Campo, "High-fidelity rapid ground-state loading of an ultracold gas into an optical lattice," Phys. Rev. Lett. **113**, 063003 (2014).

[27] A. Kiely, J. P. L. McGuinness, J. G. Muga, and A. Ruschhaupt, "Fast and stable manipulation of a charged particle in a penning trap," J. Phys. B: At. Mol. Opt. Phys. **48**, 075503 (2015).

[28] S. Deffner, "Shortcuts to adiabaticity: suppression of pair production in driven dirac dynamics," New Journal of Physics **18**, 012001 (2015).

[29] C. Jarzynski, S. Deffner, A. Patra, and Y. Subaş ı, "Fast forward to the classical adiabatic invariant," Phys. Rev. E **95**, 032122 (2017).

[30] X. Chen, E. Torrontegui, Di. Stefanatos, J. Li, and J. G. Muga, "Optimal trajectories for efficient atomic transport without final excitation," Phys. Rev. A **84**, 043415 (2011).

[31] D. Stefanatos, "Optimal shortcuts to adiabaticity for a quantum piston," Automatica **49**, 3079 (2013).

[32] S. Campbell, G. De Chiara, M. Paternostro, G. M. Palma, and R. Fazio, "Shortcut to Adiabaticity in the Lipkin-Meshkov-Glick Model," Phys. Rev. Lett. **114**, 177206 (2015).

[33] S. Deffner, "Optimal control of a qubit in an optical cavity," J. Phys. B: At. Mol. Opt. Phys. **47**, 145502 (2014).

[34] G. Xiao and J. Gong, "Suppression of work fluctuations by optimal control: An approach based on Jarzynski's equality," Phys. Rev. E **90**, 052132 (2014).

[35] S. Masuda and S. A. Rice, "Fast-Forward Assisted STIRAP," J. Phys. Chem. A **119**, 3479 (2015), arXiv:1410.1883 [cond-mat.mes-hall].

[36] E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, "Hamiltonian engineering via invariants and dynamical algebra," Phys. Rev. A **89**, 043408 (2014).

[37] M. V. S. Bonança and S. Deffner, "Optimal driving of isothermal processes close to equilibrium," J. Chem. Phys. **140**, 244119 (2014).

[38] T. V. Acconcia, M. V. S. Bonança, and S. Deffner, "Shortcuts to adiabaticity from linear response theory," Phys. Rev. E **92**, 042148 (2015).

[39] M. V. S. Bonança and S. Deffner, "Minimal dissipation in processes far from equilibrium," Phys. Rev. E **98**, 042103 (2018).

[40] Sebastian Deffner and Marcus V. S. Bonança, "Thermodynamic control —an old paradigm with new applications," EPL (Europhysics Letters) **131**, 20001 (2020).

[41] S. Martínez-Garaot, A. Ruschhaupt, J. Gillet, Th. Busch, and J. G. Muga, "Fast quasiadiabatic dynamics," Phys. Rev. A **92**, 043406 (2015).

[42] Bertúlio de Lima Bernardo, "Time-rescaled quantum dynamics as a shortcut to adiabaticity," Phys. Rev. Research **2**, 013133 (2020).

[43] Agniva Roychowdhury and Sebastian Deffner, "Time-rescaling of dirac dynamics: Shortcuts to adiabaticity in ion traps and weyl semimetals," Entropy **23** (2021), 10.3390/e23010081.

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

[45] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, "Shortcuts to adiabaticity: Concepts, methods, and applications," Rev. Mod. Phys. **91**, 045001 (2019).

[46] M. Demirplak and S. A. Rice, "Adiabatic population transfer with control fields," J. Chem. Phys. A **107**, 9937 (2003).

[47] M. Demirplak and S. A Rice, "Assisted adiabatic passage revisited," J. Phys. Chem. B **109**, 6838 (2005).

[48] M. V. Berry, "Transitionless quantum driving," J. Phys. A: Math. Theor. **42**, 365303 (2009).

[49] S. Deffner, C. Jarzynski, and A. del Campo, "Classical and quantum shortcuts to adiabaticity for scale-invariant driving," Phys. Rev. X **4**, 021013 (2014).

[50] A. Messiah, *Quantum Mechanics*, Vol. II (John Wiley & Sons, Amsterdam, The Netherlands, 1966).

[51] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," Proc. R. Soc. A **392**, 45 (1984).

[52] C. Jarzynski, "Generating shortcuts to adiabaticity in quantum and classical dynamics," Phys. Rev. A **88**, 040101 (2013).

[53] S. Vaikuntanathan and C. Jarzynski, "Escorted free energy simulations: Improving convergence by reducing dissipation," Phys. Rev. Lett. **100**, 190601 (2008).

[54] I. A. Martínez, A. Petrosyan, D. Guéry-Odelin, E. Trizac, and S. Ciliberto, "Engineered swift equilibration of a brownian particle," Nature Physics **12**, 843 (2016).

[55] Ayoti Patra and Christopher Jarzynski, "Shortcuts to adiabaticity using flow fields," New J. Phys. **19**, 125009 (2017).

[56] A. Patra and C. Jarzynski, "Classical and quantum shortcuts to adiabaticity in a tilted piston," The Journal of Physical Chemistry B **121**, 3403–3411 (2017), pMID: 27700088, https://doi.org/10.1021/acs.jpcb.6b08769.

[57] A. G. Frim, A. Zhong, S.-F. Chen, D. Mandal, and M. R. DeWeese, "Engineered swift equilibration for arbitrary geometries," Phys. Rev. E **103**, L030102 (2021).

[58] Shamreen Iram, Emily Dolson, Joshua Chiel, Julia Pelesko, Nikhil Krishnan, Özenç Güngör, Benjamin Kuznets-Speck, Sebastian Deffner, Efe Ilker, Jacob G Scott, and Michael Hinczewski, "Controlling the speed and trajectory of evolution with counterdiabatic driving," Nat. Phys. **17**, 135–142 (2021).

[59] Nicolaas Godfried Van Kampen, *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).

[60] Massimiliano Esposito and Christian Van den Broeck, "Three faces of the second law. I. master equation formulation," Phys. Rev. E **82**, 011143 (2010).

[61] Narsingh Deo, *Graph theory with applications to engineering and computer science* (Courier Dover Publications, 2017).

[62] Riccardo Rao and Massimiliano Esposito, "Nonequilib-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

rium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics," Phys. Rev. X **6**, 041064 (2016).

[63] J Resh, "The inverse of a nonsingular submatrix of an incidence matrix," IEEE Trans. Circuit Theory **10**, 131–132 (1963).

[64] Jean H Bevis, Frank J Hall, and Irving J Katz, "Integer generalized inverses of incidence matrices," Linear Algebra Appl. **39**, 247–258 (1981).

[65] Massimiliano Esposito and Christian Van den Broeck, "Second law and landauer principle far from equilibrium," EPL (Europhysics Letters) **95**, 40004 (2011).

[66] Maria A Schumacher, Kang Yell Choi, Fu Lu, Howard Zalkin, and Richard G Brennan, "Mechanism of corepressor-mediated specific dna binding by the purine repressor," Cell **83**, 147–155 (1995).

[67] Otto G Berg, Johan Paulsson, and Måns Ehrenberg, "Fluctuations in repressor control: thermodynamic constraints on stochastic focusing," Biophys. J. **79**, 2944–2953 (2000).

[68] Han Xu, Markos Moraitis, Ross J Reedstrom, and Kathleen S Matthews, "Kinetic and thermodynamic studies of purine repressor binding to corepressor and operator dna," J. Biol. Chem. **273**, 8958–8964 (1998).

[69] E Ilker and M Hinczewski, "Modeling the growth of organisms validates a general relation between metabolic costs and natural selection," Phys. Rev. Lett. **122**, 238101 (2019).

[70] Olivier Fayet, T Ziegelhoffer, and C Georgopoulos, "The GroES and GroEL heat shock gene products of escherichia coli are essential for bacterial growth at all temperatures." J. Bacteriol. **171**, 1379–1385 (1989).

[71] Shaon Chakrabarti, Changbong Hyeon, Xiang Ye, George H Lorimer, and D Thirumalai, "Molecular chaperones maximize the native state yield on biological times by driving substrates out of equilibrium," Proc. Natl. Acad. Sci. , 201712962 (2017).

[72] Pierre Goloubinoff, Alberto S Sassi, Bruno Fauvet, Alessandro Barducci, and Paolo Los Rios, "Chaperones convert the energy from atp into the nonequilibrium stabilization of native proteins," Nat. Chem. Biol. **14**, 388 (2018).

[73] Jaakko Soini, Christina Falschlehner, Christina Mayer, Daniela Böhm, Stefan Weinel, Johanna Panula, Antti Vasala, and Peter Neubauer, "Transient increase of atp as a response to temperature up-shift in escherichia coli," Microb. Cell. Fact. **4**, 9 (2005).

[74] Michael B Eisen, Paul T Spellman, Patrick O Brown, and David Botstein, "Cluster analysis and display of genome-wide expression patterns," Proceedings of the National Academy of Sciences **95**, 14863–14868 (1998).

[75] Ron Milo and Rob Phillips, *Cell biology by the numbers* (Garland Science, 2015).

[76] David B Straus, William A Walter, and Carol A Gross, "The heat shock response of e. coli is regulated by changes in the concentration of $\sigma 32$," Nature **329**, 348 (1987).

[77] Fréderic Anglès, Marie-Pierre Castanié-Cornet, Nawel Slama, Mickael Dinclaux, Anne-Marie Cirinesi, Jean-Charles Portais, Fabien Létisse, and Pierre Genevaux, "Multilevel interaction of the dnak/dnaj (hsp70/hsp40) stress-responsive chaperone machine with the central metabolism," Sci. Rep. **7**, 41341 (2017).

[78] Geng Li, HT Quan, and ZC Tu, "Shortcuts to isothermality and nonequilibrium work relations," Phys. Rev. E **96**, 012144 (2017).

[79] DJ Bicout and Attila Szabo, "Electron transfer reaction dynamics in non-debye solvents," J. Chem. Phys. **109**, 2325–2338 (1998).

[80] Udo Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," Rep. Prog. Phys. **75**, 126001 (2012).

[81] Michael Hinczewski, Yann von Hansen, and Roland R Netz, "Deconvolution of dynamic mechanical networks," Proc. Natl. Acad. Sci. **107**, 21493–21498 (2010).

[82] Michael Hinczewski, J Christof M Gebhardt, Matthias Rief, and D Thirumalai, "From mechanical folding trajectories to intrinsic energy landscapes of biopolymers," Proc. Natl. Acad. Sci. **110**, 4500–4505 (2013).

[83] George I Bell, "Models for the specific adhesion of cells to cells," Science **200**, 618–627 (1978).

[84] John F Marko and Eric D Siggia, "Stretching dna," Macromolecules **28**, 8759–8770 (1995).

[85] Michelle D Wang, Hong Yin, Robert Landick, Jeff Gelles, and Steven M Block, "Stretching dna with optical tweezers," Biophys. J. **72**, 1335–1346 (1997).

[86] J Alcaraz, L Buscemi, M Puig-de Morales, J Colchero, A Baro, and D Navajas, "Correction of microrheological measurements of soft samples with atomic force microscopy for the hydrodynamic drag on the cantilever," Langmuir **18**, 716–721 (2002).

[87] Harald Janovjak, Jens Struckmeier, and Daniel J Müller, "Hydrodynamic effects in fast afm single-molecule force measurements," Eur. Biophys. J. **34**, 91–96 (2005).

[88] JR Blake, "A note on the image system for a Stokeslet in a no-slip boundary," Proc. Camb. Philos. Soc. **70**, 303–310 (1971).

[89] Yann von Hansen, Michael Hinczewski, and Roland R Netz, "Hydrodynamic screening near planar boundaries: Effects on semiflexible polymer dynamics," J. Chem. Phys. **134**, 06B609 (2011).

[90] Tim Schmiedl and Udo Seifert, "Optimal finite-time processes in stochastic thermodynamics," Phys. Rev. Lett. **98**, 108301 (2007).

[91] David A Sivak and Gavin E Crooks, "Thermodynamic metrics and optimal paths," Phys. Rev. Lett. **108**, 190602 (2012).

[92] T. R. Gringrich, G. M. Rotskoff, G. E. Crooks, and P. L. Geissler, "Near-optimal protocols in complex nonequilibrium transformations," Proc. Natl. Acad. Sci. **113**, 10263–10268 (2016).

[93] A. Das and D. T. Limmer, "Variational design principles for nonequilibrium colloidal assembly," J. Chem. Phys. **154**, 014107 (2021).

[94] Erik Aurell, Carlos Mejía-Monasterio, and Paolo Muratore-Ginanneschi, "Optimal protocols and optimal transport in stochastic thermodynamics," Phys. Rev. Lett. **106**, 250601 (2011).

[95] R. Chetrite and H. Touchette, "Variational and optimal control representations of conditioned and driven processes," J. Stat. Mech. **2015**, P12001 (2015).

[96] A. Das, D. C. Rose, J. P. Garrahan, and D. T. Limmer, "Reinforcement learning of rare diffusive dynamics," arXiv:2105.04321 (2021).

[97] Armita Nourmohammad and Ceyhun Eksin, "Optimal evolutionary control for artificial selection on molecular phenotypes," Phys. Rev. X **11**, 011044 (2021).

[98] Michael Lässig and Ville Mustonen, "Eco-evolutionary

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

control of pathogens," Proc. Natl. Acad. Sci. **117**, 19694–19704 (2020).

[99] Benjamin Kuznets-Speck and David T. Limmer, "Dissipation bounds the amplification of transition rates far from equilibrium," Proc. Natl. Acad. Sci. **118** (2021), https://doi.org/10.1073/pnas.2020863118.

[100] Naoto Shiraishi, Ken Funo, and Keiji Saito, "Speed limit for classical stochastic processes," Phys. Rev. Lett. **121**, 070601 (2018).

[101] Sebastian Deffner and Steve Campbell, "Quantum speed limits: from heisenberg's uncertainty principle to optimal quantum control," J. Phys. A. **50**, 453001 (2017).

[102] Jürgen Schnakenberg, "Network theory of microscopic and macroscopic behavior of master equation systems," Rev. Mod. Phys. **48**, 571 (1976).

[103] Gareth J Baxter, Richard A Blythe, and Alan J McKane, "Exact solution of the multi-allelic diffusion model," Math. Biosci. **209**, 124–170 (2007).

[104] Ville Mustonen and Michael Lässig, "Fitness flux and ubiquity of adaptive evolution," Proc. Natl. Acad. Sci. **107**, 4248–4253 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplementary Information: Counterdiabatic control of biophysical processes

CD DRIVING OF A 2D FOKKER-PLANCK SYSTEM: APPLICATION TO AN EVOLUTIONARY MODEL

As an illustration of the versatility of the graphical CD algorithm described in the main text, we apply it here to a Markovian model on a two-dimensional lattice graph. For small lattice spacing and a large number of states, this provides a discretized numerical approach for implementing CD driving in a two-dimensional Fokker-Planck system. Though we focus on the 2D case for simplicity, the method readily generalizes to higher-dimensional systems.

Our example is based on an evolutionary Fokker-Planck model [103, 104] describing an evolving population of $\mathcal{N}$ organisms, each of which belongs to one of $\mathcal{M}$ genetic variants with potentially different fitnesses. The population is assumed to be at carrying capacity, with births exactly compensating for deaths, and hence $\mathcal{N}$ is fixed. Mutations at birth allow for different variants to arise, and these compete via natural selection. The relative magnitudes of the selective pressures can be tuned via external parameters, as is the case for example in bacterial populations evolving under externally controlled concentrations of antibiotic drugs. As we demonstrated in a recent work [58], such tunability enables the possibility of CD driving in evolutionary systems: guiding the probability distribution of variants in the population through a target sequence of distributions over time. This level of control could facilitate novel drug protocols to combat the evolution of drug resistance. Our main focus in the earlier paper was an approximate CD solution that had certain practical advantages for experimental implementation in microbial contexts. Here we look explicitly at the exact CD solution, which requires the ability to implement selective pressures that depend on the relative proportion of the different variants. One way to accomplish this would be through time-varying culling strategies: killing different proportions of each variant depending on their current distribution in the population.

1. 2D Fokker-Planck model description

We consider a system with $\mathcal{M} = 3$ variants, where the fraction of the population in variant i is x_i, $i = 1, \ldots, 3$. For large population size $\mathcal{N}$, we can approximately take x_i to be a continuous variable between 0 and 1, with $\sum_{i=1}^{3} x_i = 1$. Since x_3 is linearly dependent on x_1 and x_2, we can specify the state of the system by the 2D vector $\boldsymbol{x} = (x_1, x_2)$. The probability distribution $p(\boldsymbol{x}, t)$ of observing state $\boldsymbol{x}$ obeys the Fokker-Planck equation [103, 104]:

$$\frac{\partial p(\boldsymbol{x}, t)}{\partial t} = -\sum_{i=1}^{2} \frac{\partial}{\partial x_i} \left[A_i(\boldsymbol{x}, t) p(\boldsymbol{x}, t)\right] + \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial^2}{\partial x_i \partial x_j} \left[D_{ij}(\boldsymbol{x}) p(\boldsymbol{x}, t)\right]. \tag{S1}$$

The first term on the right represents the deterministic contribution to the dynamics, with the mean state trajectory $\bar{\boldsymbol{x}}(t)$ governed by the set of equations $\partial_t \bar{x}_i(t) = A_i(\boldsymbol{x}, t)$. The function $A_i(\boldsymbol{x}, t)$ has the form

$$A_i(\boldsymbol{x}, t) = \mu(1 - x_i) - 2\mu x_i + \sum_{j=1}^{2} g_{ij}(\boldsymbol{x}) s_j(t). \tag{S2}$$

Here we have assumed that all three variants can mutate into each other with the same rate μ, so $\mu(1 - x_i)$ is the contribution of the other variants mutating into type i and $-2\mu x_i$ is the loss of type i due to mutations into the other two variants. The final term in Eq. (S2) encapsulates the effect of natural selection. Let us take $f_i(t)$ to be fitness of type i, which varies due to some externally controlled protocol. Then we define the selection coefficient $s_i(t) = f_i(t)/f_3(t) - 1$ for $i = 1, 2$ as a measure of the relative fitness of type i compared to type 3. We choose the latter to be our reference, or "wild" type. A positive value for $s_i(t)$ means type i is fitter than the wild type at time t, while a negative value means it is less fit. We typically assume fitness differences between competitive variants are small, with $|s_i(t)| \ll 1$. The selection coefficient vector $\boldsymbol{s}(t) = (s_1(t), s_2(t))$ plays the role of the control parameter λ_t. In practice it could reflect the effects of time-varying drug concentrations or other controllable environmental factors [58]. The $g_{ij}(\boldsymbol{x})$ element of the 2×2 matrix $g(\boldsymbol{x})$ represents the contribution of the jth selection coefficient to the mean behavior of type i. The matrix elements are given by [104]:

$$g_{ij}(\boldsymbol{x}) = \begin{cases} x_i(1 - x_i) & i = j \\ -x_i x_j & i \neq j \end{cases}. \tag{S3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Finally, the second term on the right hand side of Eq. (S1) represents diffusive spreading of the distribution due to the stochastic nature of reproduction in a population of finite size $\mathcal{N}$. The 2×2 diffusivity matrix $D_{ij}(\boldsymbol{x})$ has elements of the form [103, 104],

$$D_{ij}(\boldsymbol{x}) = \frac{g_{ij}(\boldsymbol{x})}{2\mathcal{N}}. \tag{S4}$$

The proportionality between $D(\boldsymbol{x})$ and the matrix $g(\boldsymbol{x})$, which controls the response of the system to relative fitness differences, is an evolutionary analogue to the fluctuation-dissipation theorem, with population size $\mathcal{N}$ playing the role of inverse temperature.

Because of the constraints that $x_1 \geq 0$, $x_2 \geq 0$ and $x_1 + x_2 = 1 - x_3 \leq 1$, the distribution $p(\boldsymbol{x}, t)$ is defined over a 2D probability simplex. To simplify the geometry, we introduce new variables $u_1 = x_1$, $u_2 = x_2/(1 - x_1)$ [103]. The domain of the state vector $\boldsymbol{u} = (u_1, u_2)$ is now the unit square: $0 \leq u_1 \leq 1$, $0 \leq u_2 \leq 1$. Using this mapping the corresponding distribution $p(\boldsymbol{u}, t)$ obeys a Fokker-Planck equation of the form,

$$\begin{aligned} \frac{\partial p(\boldsymbol{u}, t)}{\partial t} &= -\sum_{i=1}^{2} \frac{\partial}{\partial u_i} \left[\mathcal{A}_i(\boldsymbol{u}, t) p(\boldsymbol{u}, t)\right] + \sum_{i=1}^{2} \frac{\partial^2}{\partial u_i^2} \left[\mathcal{D}_i(\boldsymbol{u}) p(\boldsymbol{u}, t)\right] \\ &\equiv -\sum_{i=1}^{2} \frac{\partial}{\partial u_i} J_i(\boldsymbol{u}, t). \end{aligned} \tag{S5}$$

Here $\boldsymbol{\mathcal{A}}(\boldsymbol{u}, t)$ and $\boldsymbol{\mathcal{D}}(\boldsymbol{u}, t)$ are vector functions with components given by

$$\begin{aligned} \mathcal{A}_1(\boldsymbol{u}, t) &= \mu - 3\mu u_1 + u_1(1 - u_1)(s_1(t) - u_2 s_2(t)), \\ \mathcal{A}_2(\boldsymbol{u}, t) &= \frac{\mu - 2\mu u_2}{1 - u_1} + u_2(1 - u_2) s_2(t), \\ \mathcal{D}_1(\boldsymbol{u}) &= \frac{u_1(1 - u_1)}{2\mathcal{N}}, \\ \mathcal{D}_2(\boldsymbol{u}) &= \frac{u_2(1 - u_2)}{2\mathcal{N}(1 - u_1)}. \end{aligned} \tag{S6}$$

and $\boldsymbol{J}(\boldsymbol{u}, t)$ is the probability current density with components given by

$$J_i(\boldsymbol{u}, t) = \mathcal{A}_i(\boldsymbol{u}, t) p(\boldsymbol{u}, t) - \frac{\partial}{\partial u_i} \left[\mathcal{D}_i(\boldsymbol{u}) p(\boldsymbol{u}, t)\right], \qquad i = 1, 2. \tag{S7}$$

To derive Eq. (S5) we have taken into account the Jacobian for the change of variables which relates the probability densities in the old and new variables as: $p(\boldsymbol{u}, t) = (1 - u_1) p(\boldsymbol{x}, t)$ [103].

Eq. (S5) has an instantaneous stationary distribution $\rho(\boldsymbol{u}, \boldsymbol{s}(t))$ of the form

$$\rho(\boldsymbol{u}, \boldsymbol{s}(t)) = \frac{e^{-E(\boldsymbol{u}, \boldsymbol{s}(t))}}{Z(\lambda_t)}, \tag{S8}$$

where

$$E(\boldsymbol{u}, \boldsymbol{s}(t)) = 2\mathcal{N}(1 - u_1)(s_1(t) - s_2(t) u_2) - \log\left[4\left(u_1(1 - u_1) u_2(1 - u_2)\right)^{2\mu\mathcal{N} - 1} (1 - u_1)^{2\mu\mathcal{N}}\right], \tag{S9}$$

and the normalization constant $Z(\lambda_t)$ can be determined numerically from the integral $\int_0^1 du_1 \int_0^1 du_2\, \rho(\boldsymbol{u}, \lambda_t) = 1$. The current densities in Eq. (S7) can be rewritten in terms of the stationary distribution as

$$J_i(\boldsymbol{u}, t) = -\mathcal{D}_i(\boldsymbol{u}) \rho(\boldsymbol{u}, \boldsymbol{s}(t)) \frac{\partial}{\partial u_i} \left[\frac{p(\boldsymbol{u}, t)}{\rho(\boldsymbol{u}, \boldsymbol{s}(t))}\right], \qquad i = 1, 2. \tag{S10}$$

It is then easy to see that the corresponding stationary probability current densities $\boldsymbol{\mathcal{J}}(\boldsymbol{u}, t)$ vanish:

$$\mathcal{J}_i(\boldsymbol{u}, t) = J_i(\boldsymbol{u}, t)|_{p(\boldsymbol{u}, t) = \rho(\boldsymbol{u}, \boldsymbol{s}(t))} = 0, \qquad i = 1, 2. \tag{S11}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

2. Formulating the CD driving problem

To define an example target sequence of distributions for our CD driving, let us adopt the following simple linear forms for the time-varying selection coefficient protocol,

$$s_1(t) = s_1^0 + (s_1^\tau - s_1^0)t/\tau, \qquad s_2(t) = s_2^0 + (s_2^\tau - s_2^0)t/\tau. \tag{S12}$$

The protocol runs from initial time $t = 0$ to final time $t = \tau$, with time in the evolutionary Fokker-Planck system measured in units of generations. We choose parameter values as follows: $\mathcal{N} = 5000$, $\mu = 0.001$, $s_1^0 = -0.003$, $s_1^\tau = 0.005$, $s_2^0 = 0.005$, $s_2^\tau = -0.01$, $\tau = 1000$. Initially $s_2(0) > 0 > s_1(0)$, and hence the initial stationary distribution $\rho(\boldsymbol{u}, \boldsymbol{s}(0))$ is dominated by type 2 organisms, corresponding to large u_2 and small u_1, as shown in the left panel of Fig. S1A. As time increases, the $\boldsymbol{s}(t)$ protocol from Eq. (S12) leads to a continuous sequence of distributions $\rho(\boldsymbol{u}, \boldsymbol{s}(t))$ that shift towards large u_1 and small u_2, as seen in the center and right panels of Fig. S1A. At the final time, $s_1(\tau) > 0 > s_2(\tau)$, and hence the distribution is dominated by type 1 organisms. The green dashed line in Fig. S1A shows the positions of the stationary mean value $\bar{\boldsymbol{u}}(t) = \int_0^1 du_1 \int_0^1 du_2\, \boldsymbol{u}\rho(\boldsymbol{u}, \boldsymbol{s}(t))$ for this target trajectory of distributions.

The goal of CD driving in this system is to find modified selection coefficient functions $\widetilde{\boldsymbol{s}}(\boldsymbol{u}, t)$ for which the target distribution trajectory $\rho(\boldsymbol{u}, \boldsymbol{s}(t))$ is a solution to the corresponding Fokker-Planck equation. In other words we want a modified version of Eq. (S5),

$$\frac{\partial \rho(\boldsymbol{u}, \boldsymbol{s}(t))}{\partial t} = -\sum_{i=1}^{2} \frac{\partial}{\partial u_i} \widetilde{\mathcal{J}}_i(\boldsymbol{u}, t), \tag{S13}$$

where

$$\begin{aligned}
\widetilde{\mathcal{J}}_i(\boldsymbol{u}, t) &= \widetilde{\mathcal{A}}_i(\boldsymbol{u}, t)\rho(\boldsymbol{u}, \boldsymbol{s}(t)) - \frac{\partial}{\partial u_i}\left[\mathcal{D}_i(\boldsymbol{u})\rho(\boldsymbol{u}, \boldsymbol{s}(t))\right], \\
\widetilde{\mathcal{A}}_1(\boldsymbol{u}, t) &= \mu - 3\mu u_1 + u_1(1-u_1)(\widetilde{s}_1(\boldsymbol{u}, t) - u_2\widetilde{s}_2(\boldsymbol{u}, t)), \\
\widetilde{\mathcal{A}}_2(\boldsymbol{u}, t) &= \frac{\mu - 2\mu u_2}{1-u_1} + u_2(1-u_2)\widetilde{s}_2(\boldsymbol{u}, t).
\end{aligned} \tag{S14}$$

Note that because the original stationary current densities vanish, as shown in Eq. (S11), we can express the CD current densities as

$$\widetilde{\mathcal{J}}_i(\boldsymbol{u}, t) = \delta\mathcal{J}_i(\boldsymbol{u}, t) \equiv \delta\mathcal{A}_i(\boldsymbol{u}, s(t))\rho(\boldsymbol{u}, s(t)) \tag{S15}$$

where $\delta\mathcal{A}_i(\boldsymbol{u}, t) \equiv \widetilde{\mathcal{A}}_i(\boldsymbol{u}, t) - \mathcal{A}_i(\boldsymbol{u}, t)$ is the perturbation to the $\mathcal{A}_i(\boldsymbol{u}, t)$ function due to the modified CD selection coefficient. From Eqs. (S6) and (S14) these perturbations take the form,

$$\begin{aligned}
\delta\mathcal{A}_1(\boldsymbol{u}, t) &= u_1(1-u_1)(\delta s_1(\boldsymbol{u}, t) - u_2\delta s_2(\boldsymbol{u}, t)), \\
\delta\mathcal{A}_2(\boldsymbol{u}, t) &= u_2(1-u_2)\delta s_2(\boldsymbol{u}, t),
\end{aligned} \tag{S16}$$

where $\delta\boldsymbol{s}(\boldsymbol{u}, t) = \widetilde{\boldsymbol{s}}(\boldsymbol{u}, t) - \boldsymbol{s}(t)$ is the CD perturbation to the original selection coefficient protocol. A CD solution means finding $\widetilde{\mathcal{J}}_i(\boldsymbol{u}, t)$ that satisfy Eq. (S13). If those are known, then the perturbations to the selection protocol that implement the currents can be found by inverting Eqs. (S15)-(S16):

$$\delta s_1(\boldsymbol{u}, t) = \frac{1}{\rho(\boldsymbol{u}, \boldsymbol{s}(t))}\left[\frac{\widetilde{\mathcal{J}}_1(\boldsymbol{u}, t)}{u_1(1-u_1)} + \frac{\widetilde{\mathcal{J}}_2(\boldsymbol{u}, t)}{1-u_2}\right], \qquad \delta s_2(\boldsymbol{u}, t) = \frac{\widetilde{\mathcal{J}}_2(\boldsymbol{u}, t)}{\rho(\boldsymbol{u}, \boldsymbol{s}(t))u_2(1-u_2)}. \tag{S17}$$

Note that the perturbation $\delta\boldsymbol{s}(\boldsymbol{u}, t)$ will in general depend on $\boldsymbol{u}$. Translated back to the original variables $\boldsymbol{x}$, this means applying selection pressures in a way that depends on the instantaneous proportion of different variants. In cases where this is not practically achievable, we can also do an approximate form of CD by replacing $\delta\boldsymbol{s}(\boldsymbol{u}, t)$ with an averaged version $\delta\bar{\boldsymbol{s}}(t)$ independent of $\boldsymbol{u}$, as described below.

3. Calculating CD solutions using the graphical algorithm

To apply the graphical CD algorithm described in the main text, we will discretize the $\boldsymbol{u}$ unit box by dividing each axis into L equal segments of length $a = 1/L$, creating an $L \times L$ lattice graph shown in Fig. S1. Each state, $n = 1, \ldots, N$, in the corresponding Markov model will be associated with a square of area a^2 centered at position

$$\boldsymbol{u}^{(n)} = a\left(\frac{1}{2} + (n-1) \bmod L, -\frac{1}{2} + \lceil n/L \rceil\right), \tag{S18}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

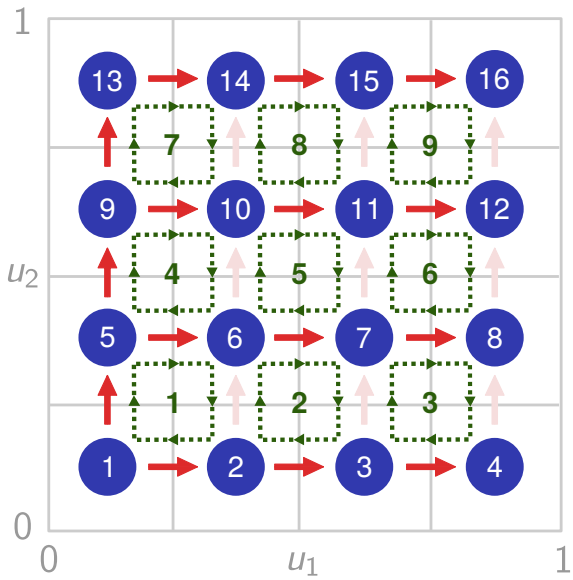

FIG. S1. An example of discretizing the $\boldsymbol{u}$ unit box for the evolutionary system using an $L \times L$ lattice graph, with $L = 4$. The numbered blue circles represent states $n = 1, \ldots, N$, located at positions $\boldsymbol{u}^{(n)}$ given by Eq. (S18), with $N = 16$. There are $E = 24$ edges in the graph, depicted as light or dark red arrows representing oriented currents. The subset of $N - 1$ dark red arrows is a spanning tree. A set of fundamental cycles, numbered $\gamma = 1, \ldots, \Delta$, with $\Delta = 9$, is shown as green dashed squares.

where $\lceil x \rceil$ denotes the smallest integer greater or equal to x. Note, $N = L^2$. There are $E = 2L(L-1)$ edges in the lattice graph, and hence $\Delta = E - N + 1 = (L-1)^2$. Fig. S1 shows the case of $L = 4$, with $N = 16$ states at positions $\boldsymbol{u}^{(n)}$ denoted by blue circles, $E = 24$ oriented currents denoted by dark or light red arrows, and a set of $\Delta = 9$ fundamental cycles denoted by green dashed lines. The $N - 1 = 15$ dark red arrows constitute one example of a spanning tree for the oriented current graph.

To ensure that the discrete Markov model corresponds to the Fokker-Planck equation of Eq. (S5) in the continuum ($L \to \infty$) limit, we choose the following scheme for the transition rate $\Omega_{mn}(\boldsymbol{s}(t))$ from state n to m:

$$\Omega_{mn}(\boldsymbol{s}(t)) = \begin{cases} a^{-2}\mathcal{D}_1(\boldsymbol{u}^{(n)})e^{-\left(E(\boldsymbol{u}^{(m)},\boldsymbol{s}(t))-E(\boldsymbol{u}^{(n)},\boldsymbol{s}(t))\right)} & \text{if } m = n+1 \text{ and } n \bmod L \neq 0 \\ a^{-2}\mathcal{D}_1(\boldsymbol{u}^{(m)}) & \text{if } m = n-1 \text{ and } m \bmod L \neq 0 \\ a^{-2}\mathcal{D}_2(\boldsymbol{u}^{(n)})e^{-\left(E(\boldsymbol{u}^{(m)},\boldsymbol{s}(t))-E(\boldsymbol{u}^{(n)},\boldsymbol{s}(t))\right)} & \text{if } m = n+L \text{ and } n \leq L(L-1) \\ a^{-2}\mathcal{D}_2(\boldsymbol{u}^{(m)}) & \text{if } m = n-L \text{ and } n > L \\ 0 & \text{otherwise} \end{cases} . \tag{S19}$$

The top two rows on the right-hand side above are the horizontal right and left transitions in the lattice graph, and the next two rows are the vertical up and down transitions. With this choice of matrix $\Omega(\boldsymbol{s}(t))$, any pair of non-zero forward and reverse rates satisfy the local detailed balance relation,

$$\frac{\Omega_{mn}(\boldsymbol{s}(t))}{\Omega_{nm}(\boldsymbol{s}(t))} = e^{-\left(E(\boldsymbol{u}^{(m)},\boldsymbol{s}(t))-E(\boldsymbol{u}^{(n)},\boldsymbol{s}(t))\right)}, \tag{S20}$$

which guarantees that the instantaneous stationary distribution is given by the discrete analogue of Eq. (S8),

$$\rho_n(\boldsymbol{s}(t)) = \frac{e^{-E(\boldsymbol{u}^{(n)},\boldsymbol{s}(t))}}{\mathcal{Z}(\boldsymbol{s}(t))}, \tag{S21}$$

where $\mathcal{Z}(\boldsymbol{s}(t)) = \sum_{n=1}^{N} \exp(-E(\boldsymbol{u}^{(n)}, \boldsymbol{s}(t)))$. We can also check the form of the current for edge α, corresponding to an arrow from state n to a state m directly to the right or above. Using Eq. (S19) this can be rewritten as:

$$J_\alpha(t) = \Omega_{mn}(\boldsymbol{s}(t))p_n(t) - \Omega_{nm}(\boldsymbol{s}(t))p_m(t) = -\frac{1}{a}\mathcal{D}_i(\boldsymbol{u}^{(n)})\rho_m(\boldsymbol{s}(t))\frac{1}{a}\left[\frac{p_m(t)}{\rho_m(\boldsymbol{s}(t))} - \frac{p_n(t)}{\rho_n(\boldsymbol{s}(t))}\right]. \tag{S22}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

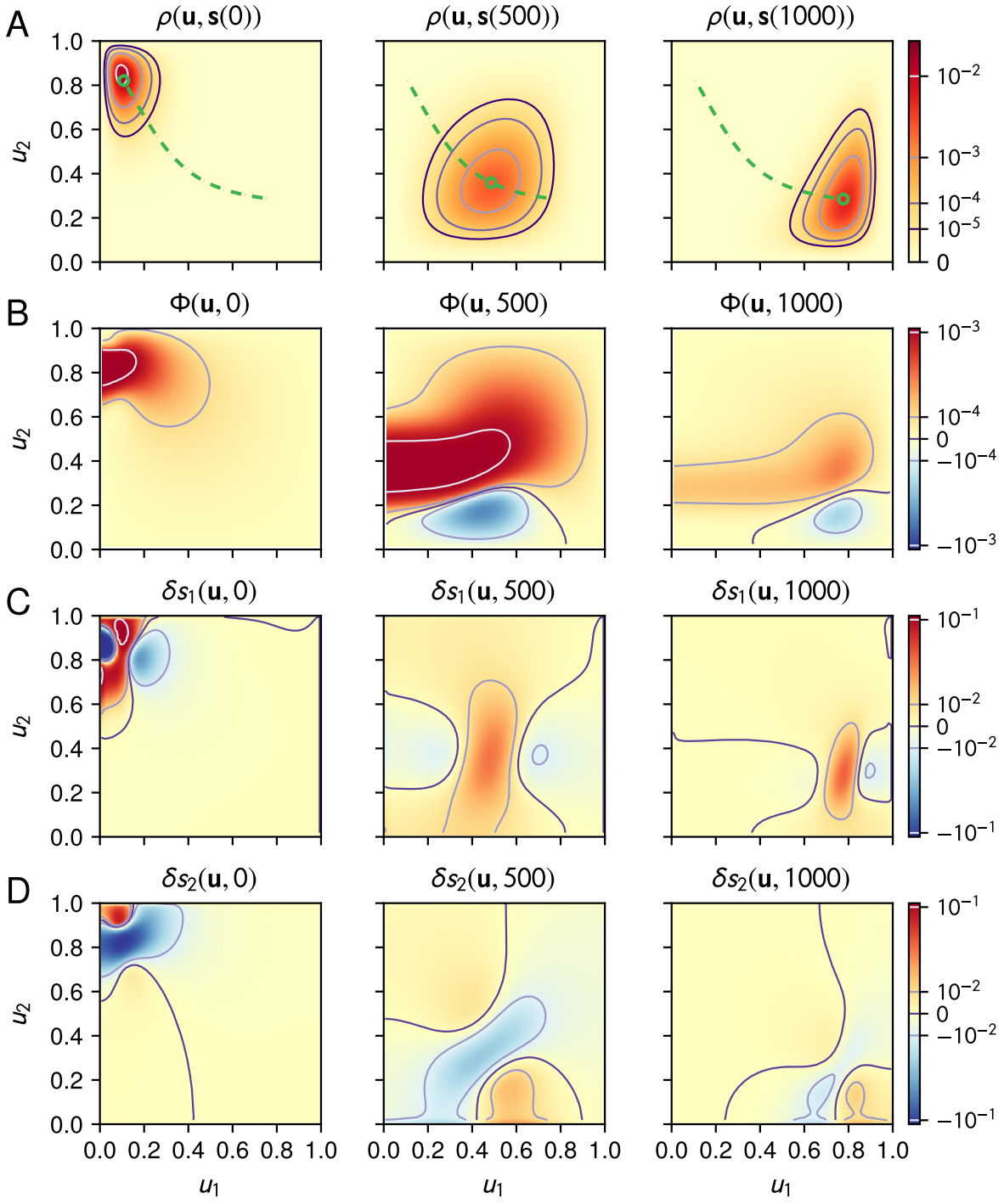

FIG. S2. Numerically calculated CD driving results for the evolutionary system, using a discretization with $L = 80$. All panels are shown in terms of u_1 and u_2 on the axes, and the three columns represent snapshots at three times $t = 0$, 500, and 1000 generations. A) The target distributions $\rho(\boldsymbol{u}, \boldsymbol{s}(t))$, defined by the selection coefficient protocol given in Eq. (S12). The goal of CD is to steer the evolving population through this trajectory of distributions. The green dashed curves represent the instantaneous stationary mean $\bar{\boldsymbol{u}}(t)$, and the green circle is the value of the mean in that snapshot. B) The cycle coefficient functions $\Phi(\boldsymbol{u}, t)$. This specific choice corresponds to the CD solution that minimizes the overall CD current magnitude ($\widetilde{\mathcal{J}}^2(t)$ in the discrete model) at each time t. C-D) The components of the selection coefficient perturbations $\delta s_i(\boldsymbol{u}, t)$, $i = 1, 2$ that implement this particular CD solution.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Here $i = 1$ if the edge is horizontal and $i = 2$ if vertical. In the limit as $L \to \infty$ and $a = 1/L \to 0$ we see that $a^{-1}J_\alpha(t) \to J_i(\boldsymbol{u}, t)$ from Eq. (S10). The factor a^{-1} converts the discrete current to a current density in the continuum. In taking this limit, note that discrete state probabilities are related to 2D probability densities as $a^{-2}\rho_n(\boldsymbol{s}(t)) \to \rho(\boldsymbol{u}, \boldsymbol{s}(t))$. Thus we see Eq. (S19) leads to the correct continuum formulation. This choice for discretized transition rates is not unique, but it will be convenient for our purposes.

Once the discretization is defined, we can directly apply the general solution method. Any possible CD solution for this system can be expressed in the form of main text Eq. (26),

$$\widetilde{\boldsymbol{\mathcal{J}}}(t) = \delta\boldsymbol{\mathcal{J}}^{(1)}(t) + \sum_{\gamma=1}^{\Delta} \Phi_\gamma(t)\boldsymbol{c}^{(\gamma)}. \tag{S23}$$

Here $\delta\boldsymbol{\mathcal{J}}^{(1)}(t)$ is a spanning tree solution (i.e. associated with the tree shown as dark red arrows in Fig. S1),

$$\delta\boldsymbol{\mathcal{J}}^{(1)}(t) = [\widehat{\nabla}^{(\gamma)}]_S^{-1}\partial_t\widehat{\boldsymbol{\rho}}(\boldsymbol{s}(t)), \tag{S24}$$

and we allow arbitrary coefficient functions $\Phi_\gamma(t)$ multiplying the cycle vectors $\boldsymbol{c}^{(\gamma)}$. All components of this solution are relatively straightforward to implement numerically, even for large graphs. The results in Fig. S2, for example, correspond to $L = 80$, where $N = 6400$, $E = 12640$, and $\Delta = 6241$. The vector $\partial_t\widehat{\boldsymbol{\rho}}(\boldsymbol{s}(t))$ is calculated by differentiating the components of the stationary distribution in Eq. (S21). The $E \times (N-1)$ dimensional matrix $[\widehat{\nabla}^{(\gamma)}]_S^{-1}$ is constructed by the graphical procedure described in main text Sec. IIB. Finally, the simplest cycle vector basis for the lattice graph is shown in green in Fig. S1. In this case each cycle vector $\boldsymbol{c}^{(\gamma)}$ involves the edges making up a square plaquette of neighboring states. We can associate the γth cycle vector with the position $\boldsymbol{u}^{(\gamma)}$ around which the loop is centered:

$$\boldsymbol{u}^{(\gamma)} = a\left(1 + (\gamma - 1) \bmod (L-1), \lceil \gamma/(L-1) \rceil\right). \tag{S25}$$

This allows us to plot a given choice of $\Phi_\gamma(t)$ as a heat map $\Phi(\boldsymbol{u}, t)$ for large L, as shown for example in Fig. S2B. Similarly any function of state number n, such as $\rho_n(\boldsymbol{s}(t))$, can be plotted as a heat map $\rho(\boldsymbol{u}, \boldsymbol{s}(t))$ using the positions $\boldsymbol{u}^{(n)}$ defined in Eq. (S18), as seen in Fig. S2A.

Any choice of $\Phi_\gamma(t)$ at each t leads to a valid CD solution, giving us an infinite family of solutions where we can seek out specific ones with desirable properties. A natural choice is the solution that leads to the overall smallest magnitude of CD currents, minimizing $\widetilde{\boldsymbol{\mathcal{J}}}^2(t)$ at each t. This solution is in a sense the smallest amount of "driving" that is necessary to still get the system to follow the target trajectory, and can be readily found using any standard numerical optimization algorithm. For the target trajectory defined by the control protocol of Eq. (S12), snapshots of the current-minimizing $\Phi(\boldsymbol{u}, t)$ at three different times are shown in Fig. S2B.

Once $\Phi_\gamma(t)$ is chosen, giving us numerical values for the CD current vector $\widetilde{\boldsymbol{\mathcal{J}}}(t)$, we can convert those to estimates of the current densities $\widetilde{\mathcal{J}}_i(\boldsymbol{u}, t)$, $i = 1, 2$, in the continuum. The horizontal densities $\widetilde{\mathcal{J}}_1(\boldsymbol{u}, t)$ correspond to $a^{-1}\widetilde{\mathcal{J}}_\alpha(t)$ for horizontal edges α, with each edge assigned the position $\boldsymbol{u}^{(n)}$ at the origin of the current arrow. We find $\widetilde{\mathcal{J}}_2(\boldsymbol{u}, t)$ analogously, except using the vertical edges in the graph. Knowledge of $\widetilde{\mathcal{J}}_i(\boldsymbol{u}, t)$ then allows us to use Eq. (S17) to calculate the selection coefficient perturbations necessary to implement the driving. For our choice of the current-minimizing $\Phi(\boldsymbol{u}, t)$ in Fig. S2B, the heatmaps of $\delta s_1(\boldsymbol{u}, t)$ and $\delta s_2(\boldsymbol{u}, t)$ are shown in Fig. S2C and D respectively.

4. Approximate CD solution

Since the perturbation $\delta\boldsymbol{s}(\boldsymbol{u}, t)$ to achieve exact CD driving is in general dependent on $\boldsymbol{u}$ and hence the fraction $\boldsymbol{x}$ of different variants in the population, it requires a way of monitoring the distribution of variants throughout the driving process. When this monitoring is not possible, a simpler approximate strategy can be used, where we apply an averaged perturbation $\delta\bar{\boldsymbol{s}}(t)$ independent of $\boldsymbol{u}$:

$$\delta\bar{\boldsymbol{s}}(t) = \int_0^1 du_1 \int_0^1 du_2\, \rho(\boldsymbol{u}, \boldsymbol{s}(t))\delta\boldsymbol{s}(\boldsymbol{u}, t). \tag{S26}$$

Since this is weighted by the instaneous stationary distribution, the approximation should provide an overall driving that works well for typical realizations of the system, particularly in the large $\mathcal{N}$ case where the distribution $\rho(\boldsymbol{u}, s(t))$ is narrow.

Plugging Eq. (S17) into Eq. (S26) we find

$$\delta\bar{s}_1(t) = \int_0^1 du_1 \int_0^1 du_2 \left[\frac{\widetilde{\mathcal{J}}_1(\boldsymbol{u}, t)}{u_1(1-u_1)} + \frac{\widetilde{\mathcal{J}}_2(\boldsymbol{u}, t)}{1-u_2}\right], \qquad \delta\bar{s}_2(t) = \int_0^1 du_1 \int_0^1 du_2\, \frac{\widetilde{\mathcal{J}}_2(\boldsymbol{u}, t)}{u_2(1-u_2)}. \tag{S27}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.13.448255; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

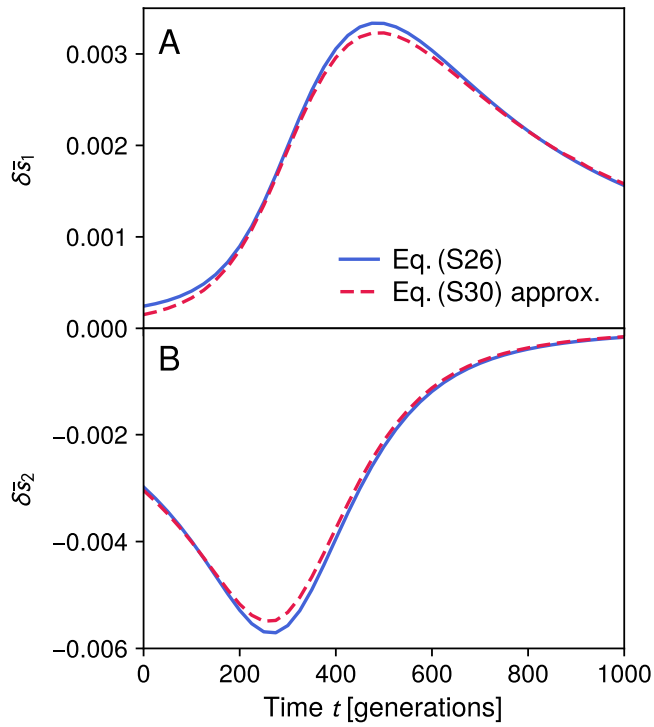

FIG. S3. Averaged CD perturbation $\delta\bar{\boldsymbol{s}}(t)$, calculated using two methods: directly via Eq. (S26) [blue solid curves], and using the approximation of Eq. (S30) [red dashed curves]. Panels A and B show the components $\delta\bar{s}_1(t)$ and $\delta\bar{s}_2(t)$ respectively.

Similar to the discussion in Ref. [58], for the large $\mathcal{N}$ case we can focus on the region of the $\boldsymbol{u}$ space in the vicinity of the instantaneous stationary mean $\bar{\boldsymbol{u}}(t)$. Replacing the components of $\boldsymbol{u}$ with $\bar{\boldsymbol{u}}(t)$ in the denominators of Eq. (S27), and pulling the resulting constants out of the integrals, we get the approximate expressions

$$\begin{aligned}\delta\bar{s}_1(t) &\approx \frac{1}{\bar{u}_1(t)(1-\bar{u}_1(t))}\int_0^1 du_1\int_0^1 du_2\,\widetilde{\mathcal{J}}_1(\boldsymbol{u},t)+\frac{1}{1-\bar{u}_2(t)}\int_0^1 du_1\int_0^1 du_2\,\widetilde{\mathcal{J}}_2(\boldsymbol{u},t),\\ \delta\bar{s}_2(t) &\approx \frac{1}{\bar{u}_2(t)(1-\bar{u}_2(t))}\int_0^1 du_1\int_0^1 du_2\,\widetilde{\mathcal{J}}_2(\boldsymbol{u},t).\end{aligned} \tag{S28}$$

We can now take advantage of the identity

$$\int_0^1 du_1\int_0^1 du_2\,\widetilde{\mathcal{J}}_j(\boldsymbol{u},t)=\partial_t\bar{u}_j(t),\qquad j=1,2. \tag{S29}$$

This can be derived from Eq. (S13) by multiplying both sides of that equation by u_j and integrating over the entire $\boldsymbol{u}$ unit box, then using integration by parts and the fact that the currents normal to the boundaries of the unit box vanish. Plugging Eq. (S29) into Eq. (S28) we find:

$$\delta\bar{s}_1(t)\approx\frac{\partial_t\bar{u}_1(t)}{\bar{u}_1(t)(1-\bar{u}_1(t))}+\frac{\partial_t\bar{u}_2(t)}{1-\bar{u}_2(t)},\qquad \delta\bar{s}_2(t)\approx\frac{\partial_t\bar{u}_2(t)}{\bar{u}_2(t)(1-\bar{u}_2(t))}. \tag{S30}$$

This approximation to Eq. (S26) is useful since it only depends on knowledge of the mean trajectory $\bar{\boldsymbol{u}}(t)$ and its derivative $\partial_t\bar{\boldsymbol{u}}(t)$. Note that any of the infinite family of CD solutions $\delta\boldsymbol{s}(\boldsymbol{u},t)$, integrated via Eq. (S26), will yield approximately the same result, Eq. (S30). This is true so long as the dominant contributions to $\rho(\boldsymbol{u},s(t))$ are in the vicinity of $\bar{\boldsymbol{u}}(t)$, justifying the approximation in Eq. (S28). For the $L=80$ CD driving results in Fig. S2, the corresponding averaged perturbation from Eq. (S26) is shown as a blue solid curve in Fig. S3. The Eq. (S30) approximation is plotted for comparison as a red dashed curve. The two curves match quite well, as expected given the narrowness of the $\rho(\boldsymbol{u},s(t))$ distributions along our target trajectory, seen in Fig. S2A.