Deterministic particle flows for constraining stochastic nonlinear systems

Dimitra Maoutsa[1,2,*] and Manfred Opper[1,3,†]

[1]*Department of Theoretical Computer Science, Technical University of Berlin, Marchstraße 23, 10587 Berlin, Germany*
[2]*Institute of Mathematics, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany*
[3]*Centre for Systems Modelling and Quantitative Biomedicine, University of Birmingham, B15 2TT, United Kingdom*

(Received 26 February 2022; accepted 7 July 2022; published 17 October 2022)

Devising optimal interventions for constraining stochastic systems is a challenging endeavor that has to confront the interplay between randomness and dynamical nonlinearity. Existing intervention methods that employ stochastic path sampling scale poorly with increasing system dimension and are slow to converge. Here we propose a generally applicable and practically feasible methodology that computes the optimal interventions in a *noniterative* scheme. We formulate the optimal dynamical adjustments in terms of deterministically sampled probability flows approximated by an interacting particle system. Applied to several biologically inspired models, we demonstrate that our method provides the necessary optimal controls in settings with terminal, transient, or generalized collective state constraints and arbitrary system dynamics.

DOI: 10.1103/PhysRevResearch.4.043035

I. INTRODUCTION

Most biological and physical systems are continuously subjected to noise arising either from intrinsic fluctuations of their constituents, or from external environmental variations at multiple timescales [1–4]. The stochastic nature of these influences confers on these systems complex behavior [5–7], but also renders them remarkably unpredictable—by generating noise induced transitions [8,9], intervening in intracellular communication [10], and compromising the precision of biological functions.

Yet, concrete understanding of characteristics, properties, and functions of biological processes often requires external interventions either by precise steering of state trajectories, or by enforcing design constraints that limit their evolution. Characteristic examples of such interventions in systems biology include modulating transcription pathways to decrease response time [11], improving stability of epigenetic states in gene regulatory networks [12], or modifying cell differentiation in multicellular organisms [13]. One then may be interested in statistical properties of constrained trajectories (e.g., for computing averages of macroscopic observables) or in obtaining precise control protocols that implement the imposed limitations.

In most settings, the optimality of the imposed interventions plays critical role. Performing unreasonably strong perturbations may damage the underlying biological tissue, or result in dynamical changes that considerably deviate from physiological biological functions. Translated to mathematics this implies the requirement for the interventions to induce the minimum possible deviation from the typical evolution of the unconstrained system.

Such problems can usually be posed as stochastic optimal control problems. This research area has recently attracted the interest in the context of stochastic thermodynamics [14–17] and quantum control [18,19], i.e., for estimating the free energy differences between two equilibrium states [20,21], or for identifying optimal protocols that drive a system from one equilibrium to another in finite time [17]. Similar problems appear also often in chemistry, biology, finance, and engineering, required for computation of rare event probabilities [22,23], state estimation of partially observed systems [24–26], or for precise manipulation of stochastic systems to target states [27,28] with applications in artificial selection [29,30], motor control [31], epidemiology [32], and more [33–38]. Albeit the prior developments, the problem of controlling nonlinear systems in the presence of random fluctuations remains still considerably challenging.

Central role in (stochastic) optimal control theory plays the Hamilton-Jacobi-Bellman (HJB) equation [39], a nonlinear second-order partial differential equation (PDE), characterizing the value function of the control problem required for computing the optimal controls. Existing approaches for devising optimal interventions can be broadly divided into two classes: the first class treats the HJB equation directly, while a second class optimizes the interventions iteratively by employing stochastic path sampling. Directly treating the HJB often involves space discretizing PDE solvers, that in general, scale poorly with system dimension [40,41]. By introducing certain structural assumptions for the control problem in Ref. [27], Kappen proposed the path integral (PI) control formalism that linearizes the HJB, and via the Feynman-Kac formula reduces the solution of the stochastic control

*dimitra.maoutsa@tu-berlin.de
†manfred.opper@tu-berlin.de

problem to the computation of a path integral. Thereafter, several methods have either treated the linearized HJB with function approximations [42], or employed path integral approximation methods [43–45]. The second class of methods optimizes the interventions directly in iterative schemes. A subset of those methods is inspired by the PI-control literature, but instead of focusing on approximating the (exponentiated) value function, they directly optimize the controls by employing information theoretic metrics [22,46–49]. In particular, the path integral cross-entropy (PICE) method [46,47], employs importance sampling to generate paths from a stochastic system with a provisional control and applies appropriate reweighting to iteratevely converge to the optimal interventions.

In this paper, we borrow ideas from the inference formulation of optimal control [50–53], and take a new look at the problem by providing a sample based solution that nevertheless avoids stochastic path sampling. More precisely, we reformulate the optimal controls in terms of the solutions of two forward (filtering) equations, and employ recent deterministic particle methods for propagating probability flows [54], properly adapted to fit our purposes. Building on the theory of time-reversed SDEs [55], we obtain an exact representation of the optimally adjusted drift of the underlying stochastic system in terms of the logarithmic gradient of two forward probability flows. The latter are estimated from an interacting particle approximation to the logarithmic gradient of sampled densities using a variational formulation developed in the field of machine learning.

We show that we can successfully intervene in a series of biologically inspired systems in time constrained settings, subjected to terminal, path, and generalized collective state constraints. We further demonstrate how various problem parameters influence the estimated interventions, and compare our framework to the already established path integral cross entropy method [46].

II. DETERMINISTIC PARTICLE FLOW CONTROL: THEORETICAL BACKGROUND

A. Constraining stochastic systems with deterministic forcing

Biological and physical systems are often subjected to intrinsic or extrinsic noise sources that influence their dynamics. Characteristic examples include molecular reactions and chemical kinetics [56], populations of animal species, biological neurons [57], and evolutionary dynamics [58,59]. Stochastic differential equations (SDEs) effectively capture the phenomenology of the dynamics of such systems by both considering deterministic and stochastic forces affecting their state variables $X_t \in \mathcal{R}^d$ following:

$$dX_t = f(X_t, t)dt + \sigma dW_t. \tag{1}$$

In Eq. (1), the drift $f(\cdot,\cdot) : \mathcal{R}^d \times \mathcal{R} \to \mathcal{R}^d$ is a smooth typically nonlinear function that captures the deterministic part of the driving forces, while W stands for a d-dimensional vector of independent Wiener processes acting as white noise sources, representing contributions from unaccounted degrees of freedom, thermal fluctuations, or external perturbations. We denote the noise strength by $\sigma \in \mathcal{R}$. For the sake of brevity, we consider here additive noise, but the formalism easily generalizes for multiplicative and nonisotropic noise, i.e., for a state dependent diffusion matrix $\sigma(x, t)$ (See Ref. [60], Sec. IV). In the following, we refer to this system as the *uncontrolled* system.

Under multiple independent realizations, the stochastic nature of Eq. (1) gives rise to an ensemble of trajectories starting from an initial state $X_0 = x_0$. This ensemble captures the likely evolution of the considered system at later time points. We may characterize the unfolding of this trajectory ensemble in terms of a probability density $p_t(x)$ for the system state X_t, whose evolution is governed by the Fokker-Planck equation

$$\frac{\partial p_t(x)}{\partial t} = \nabla \cdot \left[-f(x, t)p_t(x) + \frac{\sigma^2}{2}\nabla p_t(x) \right] = \mathcal{L}_f p_t(x), \tag{2}$$

with initial condition $p_0(x) = \delta(x - x_0)$, and $\mathcal{L}_f$ denoting the Fokker-Planck operator.

Due to the stochastic nature of the system of Eq. (1), exact pinpointing of its state at some later time point T is in general not possible. Yet, often, we desire to drive stochastic systems to predefined target states within a specified time interval. Characteristic examples include designing artificial selection strategies for population dynamics [29], or triggering phenotype switches during cell fate determination [28]. Similar needs for manipulation are also relevant for nonbiological, but rather technical systems, e.g., for control of robotic or artificial limbs [61,62]. In all these settings, external system interventions become essential.

Here, we are interested in introducing constraints $\mathcal{C}$ to the system of Eq. (1) acting within a predefined time interval $[0,\ T]$. The set of possible constraints $\mathcal{C}$ comprises terminal $\chi(X_T)$, and path constraints $U(x, t)$, for $t \leqslant T$, depending on whether the desired limiting conditions apply for the entire interval, and/or only at the terminal time point. The path constraints $U(x, t) : \mathcal{R}^d \times \mathcal{R} \to \mathcal{R}$ penalize trajectories (*paths*) to render specific regions of the state space more (un)likely to be visited, while the terminal constraint $\chi(x) : \mathcal{R}^d \to \mathcal{R}$ influences the system state X_T at the final time T.

To incorporate the constraints $\mathcal{C}$ into the system, we define a modified dynamics, the *controlled* dynamics, through a change of probability measure of the path ensemble $\mathbb{P}_f$ induced by its uncontrolled counterpart. More precisely, we define the path measure $\mathbb{Q}^*$, induced by the controlled system, by a *reweighting* of paths $X_{0:T}$ generated from the uncontrolled one [Eq. (1)] over the time interval $[0,\ T]$ (Ref. [60], Sec. I Eq. (2)). Path weights are given by the likelihood ratio (*Radon-Nikodym derivative*)

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}_f}(X_{0:T}) = \frac{\chi(X_T)}{Z} \exp\left[-\int_0^T U(X_t, t)dt \right], \tag{3}$$

where Z is the normalizing constant

$$Z = \left\langle \chi(X_T) \exp\left(-\int_0^T U(X_t, t)dt \right) \right\rangle_{\mathbb{P}_f}, \tag{4}$$

and $\langle\cdot\rangle_{\mathbb{P}_f}$ denotes the expectation over paths of the uncontrolled system.

By a direct calculation (see Ref. [60], Sec. I) it can be shown that the infinite dimensional path measure $\mathbb{Q}^*$ is the

solution of the variational problem

$$\mathbb{Q}^* = \arg\min_{\mathbb{Q}} \left\{ KL(\mathbb{Q}||\mathbb{P}_f) + \int_0^T \langle U(X_t,t)\rangle_{\mathbb{Q}}\, dt - \ln \chi(X_T) \right\}, \tag{5}$$

where $KL(\mathbb{Q}||\mathbb{P}_f)$ stands for the relative entropy [*Kullback-Leibler (KL) divergence*] between the controlled and the uncontrolled path measures.

It can be shown that the optimal path measure $\mathbb{Q}^*$ is induced by a time- and state-dependent perturbation $u(x,t) : \mathcal{R}^d \times \mathcal{R} \to \mathcal{R}^d$ of the deterministic forces $f(x,t)$ acting on the uncontrolled system [46]. Thus we express the controlled dynamics as a space- and time-dependent perturbation of the uncontrolled system

$$\begin{aligned} dX_t &= (f(X_t,t) + u(X_t,t))\, dt + \sigma dW_t \\ &= g(X_t,t)\, dt + \sigma dW_t. \end{aligned} \tag{6}$$

To identify the optimal interventions, using the explicit result for the KL divergence between the control and uncontrolled path measures, we recast the optimization problem of Eq. (5) to the minimization of the cost functional $\mathcal{J}$

$$\mathcal{J} \doteq \min_u \left\langle \int_0^T \left(\frac{1}{2\sigma^2} \|u(X_t,t)\|^2 + U(X_t,t) \right) dt - \ln \chi(X_T) \right\rangle_{\mathbb{Q}}. \tag{7}$$

The first part of the cost functional penalizes large interventions $u(x,t)$, and results from minimizing the relative entropy between the path measures induced by the controlled and uncontrolled dynamics, $KL(\mathbb{Q}||\mathbb{P}_f)$. The second term constrains the transient behavior of the system through the path costs $U(x,t)$, while $\chi(X_T)$, influences only the terminal system state.

Finding exact optimal controls for general stochastic control problems amounts to solving the Hamilton-Jacobi-Bellman (HJB) equation (see Ref. [39]), a *nonlinear*, partial differential equation (PDE) that is in general computationally demanding to treat directly.

The control cost formulation of Eq. (7) gives rise to a cluster of stochastic control problems known as *Kullback-Leibler (KL) control* [63] or *path integral (PI) control* [27,51] in the literature (see Ref. [60], Sec. II for details). For this class of problems the logarithmic Hopf-Cole transformation [64], i.e., setting $\mathcal{J}(x,t) = -\ln(\varphi_t(x))$, linearizes the Hamilton-Jacobi Bellman equation [27], and the optimally perturbed drift simplifies into (see Ref. [60], Sec. II)

$$g(x,t) = f(x,t) + \sigma^2 \nabla \ln \varphi_t(x), \tag{8}$$

where the function $\varphi_t(x)$ is a solution to the backward PDE

$$\frac{\partial \varphi_t(x)}{\partial t} + \mathcal{L}_f^\dagger \varphi_t(x) - U(x,t)\varphi_t(x) = 0, \tag{9}$$

with terminal condition $\varphi_T(x) = \chi(x)$, and $\mathcal{L}_f^\dagger$ denoting the adjoint Fokker-Planck operator. For $U(x,t) \equiv 0$, Eq. (9) reduces to the Kolmogorov backward equation for the uncontrolled system. This formulation of the controlled drift has been derived also in the field of statistical mechanics by applying the Doob's h transform [23,65–68]. Yet, due to the intractability of the logarithmic gradient of $\varphi_t(x)$ for most practical problems, the drift of Eq. (8) has been applied only on simple systems.

Although the controlled drift admits a well defined expression in the terms of the solution of the backward partial differential equation of Eq. (9), direct solutions with space discretizing schemes [40,41] often suffer from high computational complexity with increasing dimensionality and become inefficient for most practical settings. On the other hand, stochastic path sampling frameworks building on the equivalence between path reweighting and optimal control, like the *path integral cross entropy method* [46,47], follow iterative procedures that progressively converge to the optimal controls. (Note also recent neural network advances towards this direction [69,70].)

B. Constrained flows from time-reversed SDEs

Here, to circumvent the need for backward-in-time integration of the backward PDE, we express the optimal interventions $u(x,t)$ in terms of two *forward* probability flows. To that end, we consider a factorization for the path probability density $q_t(x)$ arising from the *controlled* system into two terms that account for past and future constraints separately [68],

$$q_t(x) \propto \varrho_t(x)\varphi_t(x). \tag{10}$$

In Eq. (10), $\varphi_t(x)$ fulfills the backward PDE [Eq. (9)], and embodies *prospective (future)* constraints to the time t, while $\varrho_t(x)$ denotes a (non-normalized) forward probability flow that accounts for *concurrent* and *retrospective (past)* constraints, and is the solution of the forward PDE

$$\begin{aligned} \frac{\partial \varrho_t(x)}{\partial t} &= \mathcal{L}_f \varrho_t(x) - U(x,t)\varrho_t(x) \\ &= -\nabla \cdot (f(x,t)\varrho_t(x)) + \frac{\sigma^2}{2}\nabla^2 \varrho_t(x) - U(x,t)\varrho_t(x). \end{aligned} \tag{11}$$

In the absence of path constraints [$U(x,t) \equiv 0$], Eq. (11) reduces to the Fokker-Planck equation for the uncontrolled system. On the other hand, in the presence of path constraints [$U(x,t) \neq 0$] the resulting evolution equation becomes more complicated. By integrating Eq. (11) at time t over a small time interval δt, we obtain

$$\varrho_{t+\delta t}(x) = e^{\delta t(\mathcal{L}_f - U(x,t))}\varrho_t(x) \tag{12}$$

$$= e^{-\delta t U(x,t)} e^{\delta t \mathcal{L}_f} \varrho_t(x) + O((\delta t)^2), \tag{13}$$

in terms of operator exponentials [71]. This formulation admits an interpretation as the concatenation of two processes: the propagation of the density described by the uncontrolled Fokker-Planck equation [Eq. (2)], followed by a multiplication of the resulting density by a factor $e^{-\delta t U(x,t)}$. This second process, is known in *filtering* problems for stochastic dynamics [72], where the current estimate of the system state X_t results from a multiplication of the likelihood of the noisy observations $e^{-\delta t U(x,t)}$ with the density capturing the prior belief of the state X_t. Hence, we call equation Eq. (11) the

forward filtering equation. In turn, the factorization of Eq. (10) is reminiscent to the representation of smoothing densities for hidden Markov models as products of forward and backward messages.

A direct calculation shows that the factorized probability flow $q_t(x)$ [Eq. (10)], i.e., the probability flow characterizing the evolution of the constrained system state, fulfills the Fokker-Planck equation

$$\frac{\partial q_t(x)}{\partial t} = \mathcal{L}_g q_t(x) = -\nabla \cdot (g(x,t) q_t(x)) + \frac{\sigma^2}{2} \nabla^2 q_t(x), \tag{14}$$

with initial condition $q_0(x) = \varrho_0(x)$, and $\mathcal{L}_g$ denoting the Fokker-Planck operator for the optimally adjusted drift $g(x,t)$.

The factorization of Eq. (10) allows for a new representation of the optimal drift $g(x,t)$ by eliminating the backward flow $\varphi_t(x)$ in favor of $\varrho_t(x)$

$$g(x,t) = f(x,t) + \sigma^2 (\nabla \ln q_t(x) - \nabla \ln \varrho_t(x)). \tag{15}$$

The new formulation of the optimal drift still requires the logarithmic gradient of the constrained flow $q_t(x)$, and therefore does not allow for direct simulation of controlled paths. Yet, this formulation of the optimal drift turns Eq. (14) into an equation resembling a Fokker-Planck equation, but with a negative diffusion term

$$\frac{\partial q_t(x)}{\partial t} = \nabla \cdot [(\sigma^2 \nabla \ln \varrho_t(x) - f(x,t)) q_t(x)] - \frac{\sigma^2}{2} \nabla^2 q_t(x). \tag{16}$$

By introducing the backward time variable $\tau = T - t$, and setting $\tilde{q}_\tau(x) = q_{T-\tau}(x)$ we obtain a new Fokker–Planck equation with properly signed diffusion

$$\frac{\partial \tilde{q}_\tau(x)}{\partial \tau} = -\nabla \cdot [(\sigma^2 \nabla \ln \varrho_{T-\tau}(x) - f(x, T-\tau)) \tilde{q}_\tau(x)] + \frac{\sigma^2}{2} \nabla^2 \tilde{q}_\tau(x), \tag{17}$$

with initial condition $\tilde{q}_0(x) \propto \varrho_T(x) \chi(x)$.

Hence, we have represented the optimal control

$$u^*(x,t) = \sigma^2 (\nabla \ln \tilde{q}_{T-t}(x) - \nabla \ln \varrho_t(x)), \tag{18}$$

as the difference of the logarithmic gradients of two probability densities (*score functions*).

This result suggests a possible numerical strategy for obtaining the optimal interventions: First, solve the forward filtering equation for $\varrho_t(x)$ [Eq. (11)] using, e.g., sequential Monte Carlo methods [73], and estimate the logarithmic gradients $\nabla \ln \varrho_{T-\tau}(x)$. Subsequently, sample stochastic paths of the SDE associated with the Fokker–Planck equation [Eq. (17)] and use the trajectories to approximate the logarithmic gradient $\nabla \ln \tilde{q}_{T-t}(x)$ (Fig. 1). In fact, a similar approach was recently used to solve so-called *Schrödinger bridge* problems [35,74]. The latter can be also be understood as a specific control problem for SDEs with only control energy costs [i.e., $U(x) \equiv 0$] in the cost functional [Eq. (7)], but in addition the probability densities $q_0(x)$ and $q_T(x)$ at initial and final times are specified as extra constraints. In contrast to these approaches, we apply and generalize a recent *deterministic*, particle framework for solving Fokker-Planck equations introduced by the authors in Ref. [54] to solve generic path integral control problems of the type of Eq. (7). This new technique avoids stochastic path sampling and reduces significantly temporal fluctuations, delivering thereby accurate Fokker-Planck equation solutions for relatively low number of employed particles [54]. In addition, the computation of the logarithmic gradients is already an integral part of the method for computing the deterministic particle dynamics.

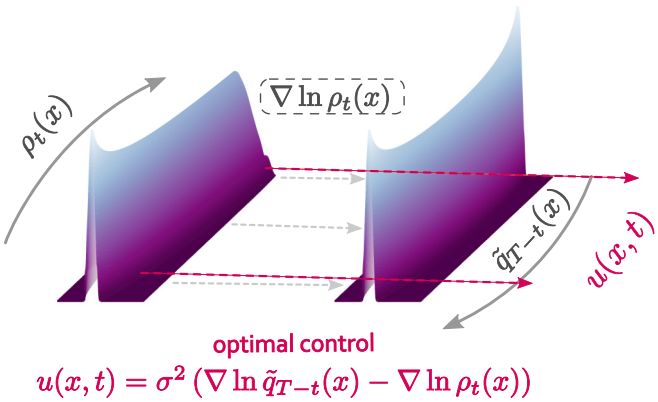

FIG. 1. Schematic of forward and time-reversed probability flows for deriving state- and time-dependent dynamical interventions $u(x,t)$. We initially sample the flow $\varrho_t(x)$ for the time interval $[0, T]$. By employing the logarithmic gradient (*score*) of $\varrho_t(x)$, we evolve the time-reversed constrained probability flow $\tilde{q}_\tau(x)$. The optimal state- and time-dependent dynamical interventions $u(x,t)$ result from the difference of the logarithmic gradients of the two probability flows.

C. Deterministic particle flow (DPF) control

To sample the forward densities $\varrho_t(x)$ and $\tilde{q}_\tau(x)$, we build on the idea that a Fokker-Planck equation can be rewritten as a *Liouville equation* [75] for an ensemble of deterministic dynamical systems where the logarithmic gradient of the ensemble density acts as an additional force. In particular, for an SDE with drift $f(x,t)$ and diffusion σ, we rewrite the Fokker-Planck equation [Eq. (1)] for the probability density $p_t(x)$ of the system state in the form (see Ref. [54] for details)

$$\frac{\partial p_t(x)}{\partial t} = -\nabla \cdot \left[\left(f(x,t) - \frac{\sigma^2}{2} \nabla \ln p_t(x) \right) p_t(x) \right]. \tag{19}$$

For a known density $p_t(x)$, this equation describes the evolution of an ensemble of independent systems with each ensemble member following the deterministic dynamics:

$$\frac{dX(t)}{dt} = f(X(t),t) - \frac{\sigma^2}{2} \nabla \ln p_t(X(t)). \tag{20}$$

Note that here individual trajectories $\{X(t)\}_{t=0}^T$ are distinct from solutions of the underlying SDE, since each particle follows pure deterministic dynamics.

To obtain a solution of the Fokker-Planck equation, we approximate the density $p_t(x)$ by an empirical distribution of N ensemble members (*"particles"*) $\{X^{(i)}(t)\}_{i=1}^N$ via

$$\hat{p}_t(x) \approx \frac{1}{N} \sum_{i=1}^{N} \delta(x - X^{(i)}(t)). \tag{21}$$

Based on this empirical representation of the density $p_t(x)$, we approximate its logarithmic gradient with a statistical *estimator* $\mathcal{S}(x, \hat{p}_t) \approx \nabla \ln p_t(x)$, obtained from the solution of a variational formulation of the score function (see Ref. [60], Sec. VI). Thus we express the resulting dynamics of individual particles in terms of a system of ordinary differential equations (ODEs)

$$\frac{dX^{(i)}(t)}{dt} = f(X^{(i)}(t), t) - \frac{\sigma^2}{2}\mathcal{S}(X^{(i)}(t), \hat{p}_t). \tag{22}$$

While this approach is sufficient to solve control problems without path costs ($U(x,t) \equiv 0$), the extra sink term $-U(x,t)\varrho_t(x)$ in the forward filtering PDE (Eq. (11)) in the presence of path constraints requires an additional numerical technique. Thus, to incorporate path costs, we employ the formulation of the two stage process given by Eq. (12), and combine our Fokker-Planck deterministic particle solver with a deterministic particle filter method, the *ensemble transform particle filter* [72]. To simulate such a two stage process for each small time interval δt, we first propagate the particles following the dynamics of Eq. (22) to auxiliary positions $Y_t^{(i)}$ and assign to each particle i a *weight* Ω_i:

$$\Omega_i(t) \propto e^{-\delta t U(Y_t^{(i)}, t)}. \tag{23}$$

To transform the weighted particles to unweighted ones, we employ the *ensemble transform particle filter* [72]. This method solves an optimal transport [76] problem to provide the minimal necessary *deterministic* shifts required to transform an ensemble of weighted particles into an ensemble of uniformly weighted ones, minimizing the expected distance between the two ensembles (see Ref. [60], Sec. V).

D. Guiding probability flows to extreme terminal states

In settings where the terminal target state lies outside of the typical values of the uncontrolled system, the sampled forward flow $\varrho_t(x)$ fails to provide sufficient evidence in the vicinity of the terminal point. Thereby the ensuing logarithmic gradient estimation $\nabla \ln \varrho_t(x)$ is inaccurate.

For conservative systems, i.e., when the drift is a gradient of a potential $f(x) = -\nabla V(x)$, and for terminal constraints defined by a delta function, i.e., $\chi(x) = \delta(x - x^*)$, we address this issue by proposing an additional modified forward sampling that incorporates the extreme terminal constraint in the forward dynamics. To that end, we employ a *d-dimensional Brownian Bridge (B)* dynamics. Brownian bridges are Brownian motions, i.e., diffusions with vanishing drift $f(x) \equiv 0$, conditioned on the terminal state x^* (see Ref. [60], Sec. III Eq. S.39).

To maintain the correct path statistics, we employ the Girsanov's change of measure formula to reweight the biased forward paths. More precisely, we obtain the correct path probability measure of the controlled process $\mathbb{P}_f^B$ by reweighting the Brownian bridge path measure $\mathbb{P}_0^B$ with the likelihood (Ref. [60], Sec. III)

$$\frac{d\mathbb{P}_f^B}{d\mathbb{P}_0^B}(X_{0:T}) \propto \exp\left[-\int_0^T U_B(X_t)dt\right], \tag{24}$$

with

$$U_B(x) = \frac{1}{2\sigma^2}(f^2(x) + \sigma^2 \nabla \cdot f(x)). \tag{25}$$

Hence, to simulate constrained paths of an SDE with drift $f(x) = -\nabla V(x)$ and extreme terminal constraints, we transform the extreme terminal constraint to a path constraint $U_B(x)$ for an appropriate Brownian bridge process that already incorporates the terminal state x^*. In particular, we employ the modified forward equation

$$\frac{\partial \varrho_t(x)}{\partial t} = \mathcal{L}_{f_0}\varrho_t(x) - U_B(x)\varrho_t(x), \tag{26}$$

that generates paths with correct statistics that reach the terminal target by imposing the path constraint $U_B(x)$ to the Brownian bridge forward dynamics with drift f_0 [Eq. (S.39)]. We term this variant of our framework *guided deterministic particle flow* (gDPF) control.

III. EVALUATION OF OPTIMAL DYNAMICAL INTERVENTIONS

To illustrate our formalism in action, we computed optimal intervention protocols for biologically inspired systems by employing the proposed deterministic particle framework, and compared the obtained controls to those computed with the *path integral cross entropy method* (PICE) (see Refs. [60], Sec. VII, and [46]). We tested our method on systems of increasing complexity and dimensionality, with conservative and nonconservative forces, as well as in settings with terminal, path, or collective state constraints.

To design the optimal interventions $u(x,t)$ we employed the presented method in two alternative variants: (i) the *deterministic particle flow control* (DPF), where the forward density follows the dynamics of Eq. (11) and (ii) the *guided deterministic flow control* (gDPF), in which the forward density evolves according to an appropriately reweighted Brownian bridge dynamics, as described in Sec. II D.

To evaluate the quality of the obtained controls, we considered the design of optimal interventions for inducing state transitions to multistable conservative and nonconservative systems (Sec. III B), for artificially manipulating molecular phenotypes on adaptive landscapes (Sec. III C), and for synchronizing finite-size networks of Kuramoto phase oscillators (Sec. III D). We quantified the quality of the identified interventions in terms of employed control energy $[\|u(x,t)\|_2^2]$, reflecting the optimality of the computed control, as well as in terms of deviations from terminal $[(x^* - X_T)^2]$ and path constraints $[U(X_t,t)]$, characterizing thereby the effectiveness of the devised interventions to enforce the intended constraints. Unless explicitly mentioned otherwise, all metrics were evaluated by considering 1000 independent stochastic trajectories controlled with each method.

A. Incorporating terminal and path constraints on conservative systems.

We considered a two dimensional nonlinear system evolving in the presence of deterministic forces derived from a potential gradient $f(\mathbf{x}) = -\nabla V(\mathbf{x})$, with associated potential $V(x,y) = ((1-x)^2 + (y-x^2)^2)$ (Fig. 2), while white noise

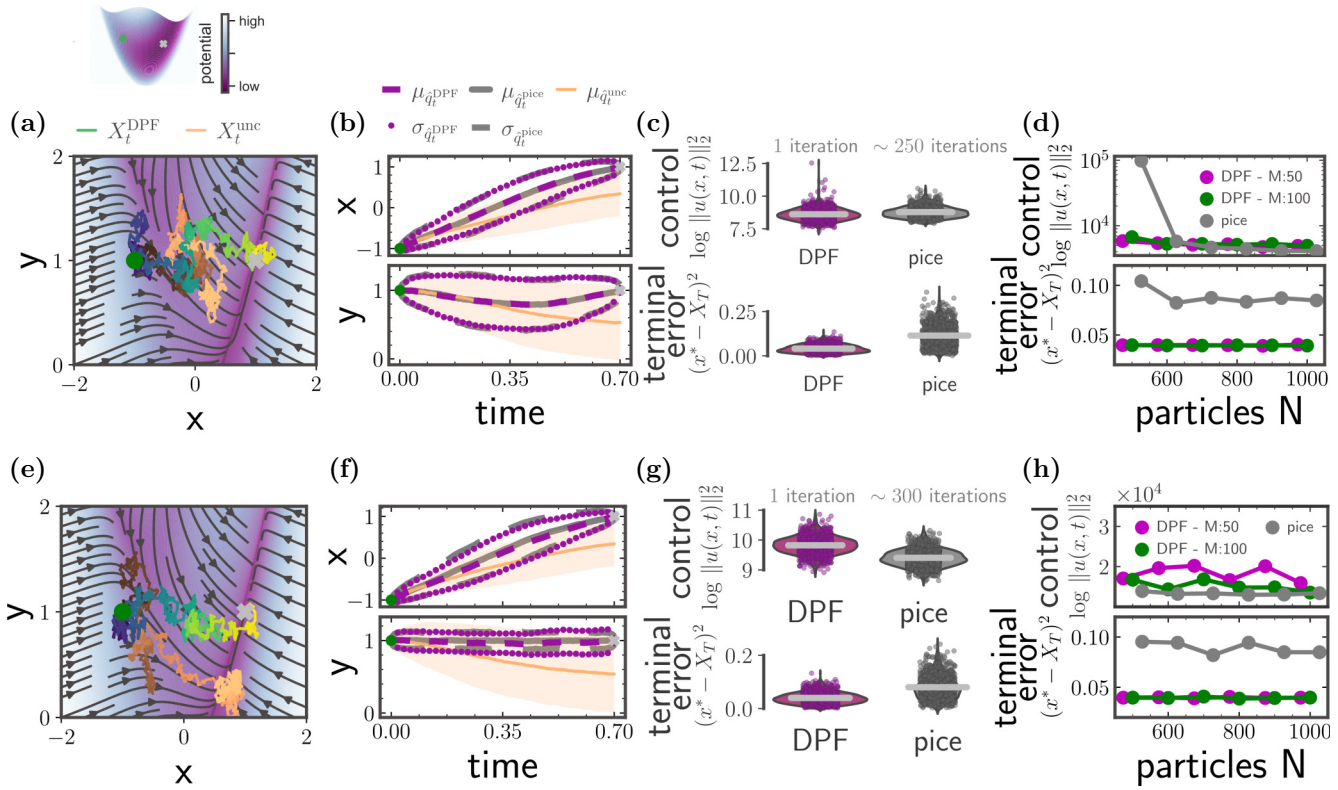

FIG. 2. Deterministic particle flow (DPF) control provides optimal interventions to drive the controlled system to target state (*grey cross*). (a) A controlled trajectory starting from state $\mathbf{x}_0 = (-1, 1)$ reaches the target $\mathbf{x}^* = (1, 1)$ at time $T = 0.7$ (*blue-yellow line*), while an uncontrolled trajectory remains in the vicinity of initial condition $\mathbf{x}_0$ during the same time interval (*orange line*). (b) Mean and standard deviation of the marginal densities of 1000 controlled trajectories employing interventions computed with our framework DPF (purple) and with PICE (grey). Orange line indicates mean of 1000 uncontrolled trajectories, while shaded area captures the associated standard deviation. For estimating the controls, we employed $N = 400$ particles for DPF, and $N_{\text{pice}} = 500$ for PICE. (c.) Comparison of (upper) (logarithmic) control energy $\|u(\mathbf{x}, t)\|_2^2$, and (lower) deviation from the terminal point $\|\mathbf{X}_T - \mathbf{x}^*\|^2$ for each controlled trajectory with interventions computed according to DPF (magenta) and PICE (grey). (d) (Logarithmic) control energy (upper) improves moderately, and terminal error (lower) remains constant for increasing particle number N. The number of inducing points in the logarithmic gradient estimation conveys negligible difference in control energy and terminal error (inducing point number magenta: $M = 50$, green: $M = 100$). Grey line indicates the performance of PICE in the same setting. [(e)–(h)] Same as (a)–(d) with additional path constraint $U((x, y), t) = \beta(y - 1)^2$, with $\beta = 10^3$. Average iteration number required for PICE was 250 in the non path-constrained setting, and 300 for the path constrained one.

fluctuations perturb its deterministic evolution. The governing system equation was

$$d\mathbf{x} = -\nabla V(\mathbf{x})dt + \sigma d\mathbf{W}. \tag{27}$$

Changes along one dimension of Eq. (27) are often accompanied by changes along covarying dimensions as dictated by the landscape gradient. Thus attempts to modify the system state only along a single axis, lead to undesired variations along the covarying dimensions. To demonstrate this, we controlled stochastic trajectories initiated at state $\mathbf{x}_0 = (-1, 1)$ with target state $\mathbf{x}^* = (1, 1)$. The designed interventions by DPF with only terminal constraints $\chi(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}^*)$ successfully drove the system to the intended target [Figs. 5(a) and 5(b)]. Nevertheless, although the initial and terminal state along the second dimension remained the same, without introducing additional path constraints, the system underwent considerable transient fluctuations along the second dimension (y) as indicated both by the nonconstant mean of the simulated paths ($\mu_{\hat{q}_t}$), as well as by the increasing dispersion of paths from the mean, captured by the standard deviation $\sigma_{\hat{q}_t}$ [Fig. 5(b)].

To limit the fluctuations along the covarying second state β, we introduced path constraints

$$U((x, y), t) = \beta\,(y - 1)^2$$

with $\beta = 10^3$, that penalized transient deviations from the intended value of y [Fig. 5(e)]. The necessary interventions, identified by DPF with path constraints, successfully steered the system towards the predefined target, reducing thereby considerably the fluctuations along the second axis (Figs. 5(e) and 5(f) and supplementary Fig. S-5 in Ref. [60]). Compared to the path integral control framework (PICE), for both scenarios, our method delivered results with comparable dissipated control energy, and was considerably more precise in terms of terminal errors [Figs. 5(e) and 5(f)].

Evaluating the performance of both methods for increasing particle number N employed for the estimation of the required interventions $u(x, t)$, revealed that DPF provided controls on par with PICE already for $N = 500$ particles and,

by design, only with a *single* iteration. In contrast, PICE required on average 250 and 300 iterations for the path unconstrained and the path constrained settings, respectively. The proposed method presented a relatively stable performance for increasing number of particles, with small improvement in terms of the exerted control energy, whereas the path integral cross entropy method improved substantially when more particles where employed in the computations, and consequently matched the performance our approach. Considering the deviation from the terminal state, DPF was consistently more precise and accurate in reaching the target as indicated both by considerably smaller terminal errors [Figs. 5(d) and 5(h)], and by smaller deviations of individual terminal states around the average (see also supplementary Fig. S8 in Ref. [60]). Both the exerted control energy and terminal error do not show considerable improvement for increasing inducing point number employed in the logarithmic gradient estimator (magenta for $M = 50$; green for $M = 100$).

B. Controlling state transitions of conservative and nonconservative systems

We further employed the proposed framework to devise interventions that reliably induced switching between equilibrium states in a time constrained scheme for an one-dimensional conservative system, and for a two-dimensional nonconservative one. For both settings, the unconstrained system either performs the transition between the two equilibrium states at not prescribed instances, or completely fails to reach the target, when the transition paths to that state strongly deviate from typical system trajectories (Fig. 3 and supplementary Figs. S-3 and S-4 in Ref. [60]).

For the one-dimensional bistable system ($f(x) = 4x - x^3$) starting from the stable state at $x_0 = -1$, we provided optimal interventions under the objective of driving the system towards a predefined target x^* at time $T = 1$. We applied both variants of the presented method [DPF (Sec. II C) and gDPF (Sec. II D)], and explored two complementary scenarios: one with typical, $x^* = 0$ [Figs. 3(a)–3(d)], and one with extreme, $x^* = 1$ [Figs. 3(e)–3(h)], target states for the uncontrolled system at time T. Notice that the state $x^* = 0$ is an unstable equilibrium.

For the typical target state $x^* = 0$, all three employed methods (DPF: *magenta*, gDPF: *yellow*, PICE: *grey*) successfully biased the controlled system towards the target x^* [Fig. 3(a)]. In fact, the distributions of simulated independent controlled stochastic trajectories delivered by each framework strongly agreed throughout the entire time interval $[0,\ T]$ [Fig. 3(b)].

Considering the control energy dissipated by each method, DPF, on average, provided slightly larger interventions [Fig. 3(c)]. Yet, both variants of our approach (DPF and gDPF) induced control trajectories that were consistently more *precise* in reaching the terminal state [Fig. 3(c)].

Comparing the distributions of terminal errors, DPF was both more accurate in reaching the target, as mediated by a smaller average value over the 1000 realizations (grey bar), but also more precise, as indicated by the smaller dispersion of terminal errors around the average. This demonstrates that although DPF slightly overestimated the required controls, it provided sufficient force to lead the trajectories faithfully onto the target. In contrast, PICE relatively underestimated the necessary interventions, resulting in more energy efficient controls that nevertheless moderately deviated from the target.

Importantly, these results suggest that a *single* iteration of either variants of the proposed method (DPF or gDPF) provide comparable controls to the *iterative* PICE framework (grey).

For the extreme (not typical) terminal state $x^* = 1$, the guided probability flow deterministic control (gDPF) and the path integral cross entropy method (PICE) successfully pushed the system to the target [Fig. 3(e)]. The distributions of controlled trajectories [Fig. 3(f)] from the two frameworks showed strong agreement, while the control costs and terminal state precision were qualitatively similar to those obtained for the typical target state.

Notice that the deterministic particle flow control (DPF), the simple variant of our method, is inappropriate for this setting for a reasonable number of employed particles representing the forward flow. Since the target x^* is an extreme system state for time T, it is highly unlikely that particles representing the forward flow $\varrho_t(x)$ will reach it. Thus the estimation of logarithmic gradients of the forward flow $\varrho_t(x)$ in the vicinity of x^* will be inaccurate, since the particles will not provide sufficient evidence for gradient estimation in that region.

Departing from gradient systems, onwards we consider triggering transitions between stable states in a time-reliable way for a nonconservative system. To that end, we employed DPF for controlling a two dimensional phenomenological model of the cell fate division module of a gene regulatory network [77] with self-excitation and cross inhibition (see Appendix A). We applied our method to induce transitions to the system between its coexisting stable states.

Similar to the conservative setting examined previously, the DPF successfully mediated the necessary interventions for the transition between the stable states [Fig. 4(a)]. We considered three noise conditions with $\sigma = \{1.0, 1.25, 1.5\}$ and compared again the deterministic particle flow control (DPF) with the path integral cross entropy method (PICE). The transient statistics computed over 1000 controlled trajectories for each framework agreed for all noise conditions (Fig. 4(b) and supplementary Fig. S-9 in Ref. [60]).

Control costs and terminal error precision were comparable for the two approaches, with DPF performing slightly better in terms of dissipated control energy for increasing noise strength [Fig. 4(c)]. Both methods had comparable control accuracy in reaching the target that deteriorated moderately for increasing noise amplitude. However, while DPF was considerable more accurate and precise for the low noise conditions, for larger noise amplitudes the accuracy of both methods in reaching the terminal point x^* became comparable [Fig. 4(d)].

Taken together, DPF (and gDPF where necessary for conservative systems) successfully provided the necessary controls for reaching the targets in both conservative and nonconservative systems and under various noise conditions. The provided interventions were on par with the established iterative PICE framework in terms of dissipated control energy, and moderately more precise in reaching the target.

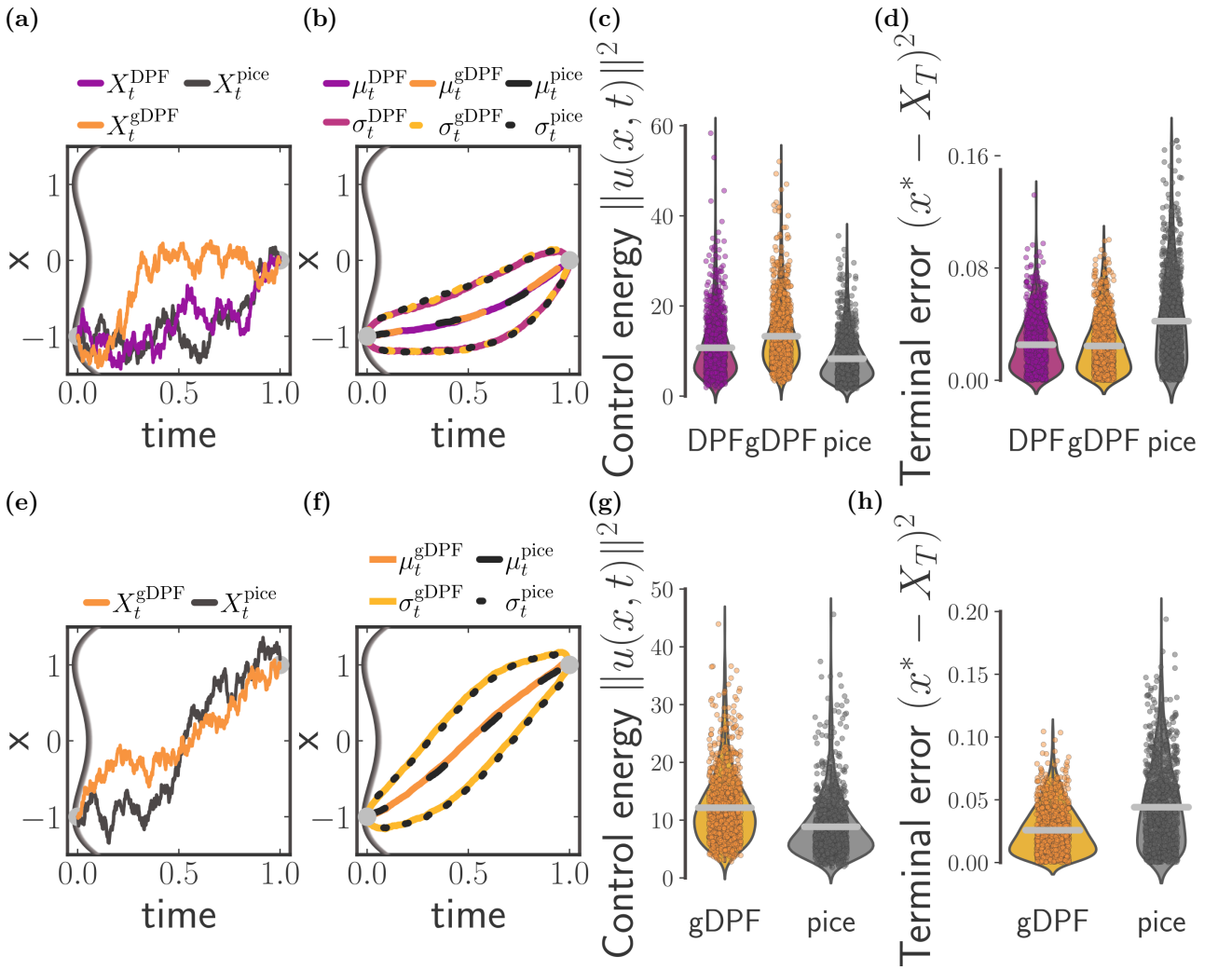

FIG. 3. Equivalence of dynamical interventions delivered by deterministic particle flow control (*DPF*), guided deterministic particle flow control (gDPF), and PICE, for typical (top) and atypical (bottom) terminal conditions. (Top) (a) Controlled trajectories simulated by employing dynamical interventions delivered by deterministic particle flow control (*magenta*), guided deterministic particle flow control (*orange*), and path integral cross entropy method (*grey*). All trajectories started from initial point $x_0 = -1$ (left silver circle) and reached the target $x^* = 0$ at time $T = 1$. (b) Transient mean μ_t and standard deviation σ_t over 1000 independent trajectories controlled with each framework. (c) Total control energy $\|u(x,t)\|^2$, and (d) deviation from terminal point for 1000 independent controlled trajectories with interventions computed according to each framework. Grey horizontal lines in (c) and (d) denote mean values over all realizations. The proposed methods result in slightly more expensive control costs, but are more consistent in precisely reaching the terminal state. (Bottom) Same as upper row for target $x^* = 1$ only for gDPF and PICE. Here DPF is not applicable since the forward probability flow does not reach the atypical target point $x^* = 1$.

C. Evolutionary control through artificial selection

Building upon the evolutionary stochastic control formalism recently introduced by Nourmohammad *et al.* [29], we employed the deterministic particle flow control to devise artificial selection protocols for molecular phenotypes. The obtained interventions optimally drove the simulated evolutionary process to desired phenotypic states. In this setting, experimentally imposed path constraints become relevant for preventing undesirable outcomes on covarying phenotypes.

For an evolving population, the main evolutionary drivers comprise fitness and mutation forces that continuously adjust the composition of phenotypes in the population, while genetic drift perturbs the whole process stochastically. We described the evolution of the *mean phenotypes* $d\mathbf{x}$ of the population by the overdamped Langevin equation [78]

$$d\mathbf{x} = C \cdot \nabla F(\mathbf{x})dt + \Sigma d\mathbf{W}, \tag{28}$$

with $F(\mathbf{x})$ denoting the adaptive fitness landscape in the presence of natural selection [78]. With Σ we denote the noise amplitude arising from genetic drift, that equals the trait covariance matrix C rescaled by the inverse effective population size n, $\Sigma = C^{1/2}n^{-1}$. The gradient of the landscape $f(\mathbf{x}) = C \cdot \nabla F(\mathbf{x})$ defines the adaptive pressure under natural selection. For quadratic phenotypic landscapes [58] the resulting selection forces have the form $f(\mathbf{x}) = -C \cdot L\mathbf{x}$, where $L \in \mathcal{R}^{n\times n}$ is the matrix of the selection coefficients.

Equation (28) describes the evolutionary dynamics of populations in the presence of natural selection towards an

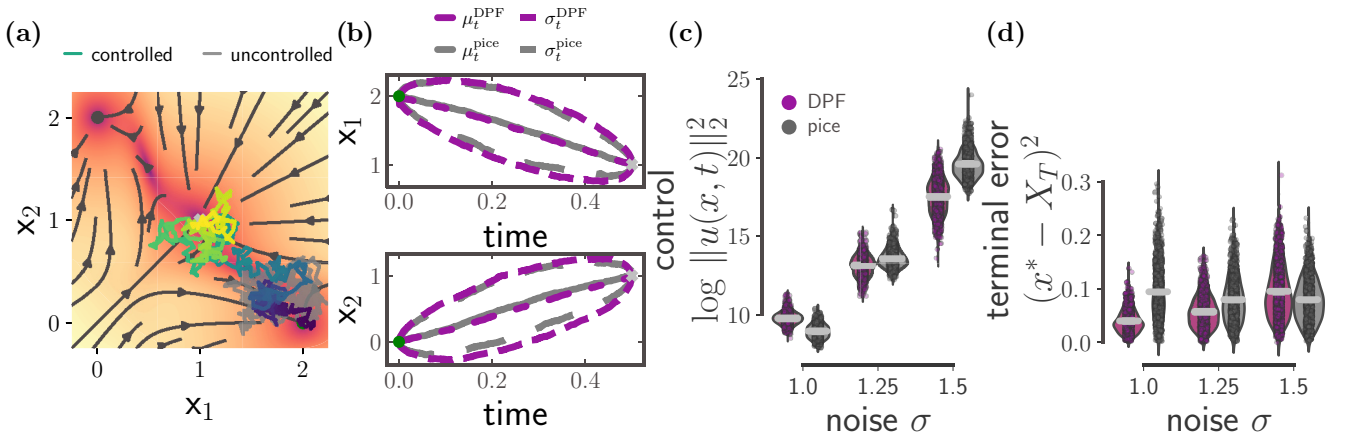

FIG. 4. Optimal interventions effectively drive the nonconservative system to the target state for different noise amplitudes. (a) Individual trajectory controlled by DPF (green-yellow) starting from stable state $x_0 = (1.996, 0)$ successfully reaches the target $x^* = (1, 1)$ at $T = 0.5$, while the uncontrolled trajectory (grey) fails to leave the basin of attraction of $x_0 = (1.996, 0)$. (b) Agreement between controlled densities of 1000 independent controlled trajectories driven by DPF (magenta) and PICE (grey) for noise amplitude $\sigma = 1.5$. (c) Dissipated control energy and (d) terminal error for increasing noise strength σ for the two control frameworks. For increasing noise DPF delivers more efficient, but moderately less precise controls than PICE. Further parameters: particle number: $N = 600$ inducing point number: $M = 20$, discretization time step $dt = 10^{-3}$.

evolutionary optimum identified by the landscape peak, adhering thereby to physiological and environmental constraints. Yet, to study the outcomes and dynamics of adaptive evolution, intervention protocols are required that drive phenotypes towards evolutionary nonoptimum states $\mathbf{x}^*$, or through evolutionary trajectories that deviate from landscape gradients. These interventions are implemented through artificial selection, which we formulate here as a time- and state-dependent perturbation $\mathbf{u}(x, t)$ on the dynamics under natural selection, and apply the proposed method to obtain the necessary controls.

For most evolutionary processes, changes along one phenotypic axis are often accompanied with changes along covarying phenotypic traits. Thus attempts to bias and enhance selective forces towards a specific direction, may lead to possibly undesired variations along the covarying traits. However, cross-phenotype correlations are often underestimated in experimental setups, where individual traits are assumed independent. Here, we apply the proposed method on an evolutionary process of two covarying phenotypes with the target to modify only the prevalence of a single trait (x) in the population, while keeping the prevalence of the second trait (y) unchanged. To that end, we employed controllers that either neglect or consider the actual cross-phenotype correlations of the uncontrolled process.

We steered phenotypic trajectories initiated from the evolutionary optimum at $\mathbf{x}_0 = (0, 0)$ towards a target at $\mathbf{x}^* = (0.5, 0)$. The designed interventions by DPF assuming both the correct and the misspecified model with only terminal constraints $\chi(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}^*)$ successfully drove the system to the intended target (Figs. 5(a) and 5(b) and supplementary Fig. S-1 in Ref. [60]). We found that both the average transient deviations from the optimal y value [Fig. 5(d)], and from the target (Fig S-1) were comparable for the processes steered by a correct and a misspecified controller independent of the cross-phenotype correlations. However as indicated by the marginal mean of the controlled trajectories [Figs. 5(a) and 5(b), purple lines] the transient densities of the two controlled processes did not necessarily overlap, especially for large correlation values. Yet, the exerted control efforts of both controller variants were on the same level [Fig. 5(c)].

Similarly to the numerical experiments of Sec. III A, although the initial and terminal state are equal along the second dimension, as expected, without introducing additional path constraints, the phenotypic trait along the second dimension (y) considerably covaried with the first trait [Fig. 5(b)]. To limit the fluctuations along the covarying second phenotypic trait, we introduced path constraints

$$U((x, y), t) = \beta\, y^2$$

with $\beta = 10^3$, that penalized transient transient deviations from the second trait's target value [Fig. 5(e)]. The necessary interventions, identified by DPF with path constraints, reduced considerably the fluctuations along the second axis for processes controlled by either the misspecified or the correct controller [Figs. 5(c) and 5(d)]. The terminal error, the deviations from the optimal y value, and the dissipated control energy were on the same level for both controller variants.

D. Controlling collective states: synchronization control of stochastic Kuramoto oscillators

To further demonstrate the generalizability of the proposed framework, we considered a system where the constraints do not explicitly penalize regions of the state space, but rather pertain the *collective state* of a system of interacting stochastic units. Specifically, we applied our method for synchronizing finite size networks of stochastic Kuramoto phase oscillators (see Appendix B for the evolution equations and Ref. [60], Sec. IX for further details). We performed systematic studies on a prototypical network of two interacting oscillators (Figs. 6–8), and a network of $K = 6$ (Figs. 9 and 10) heterogeneous oscillators with all-to-all uniform coupling (Fig. S-2 and Ref. [60], Sec. IX).

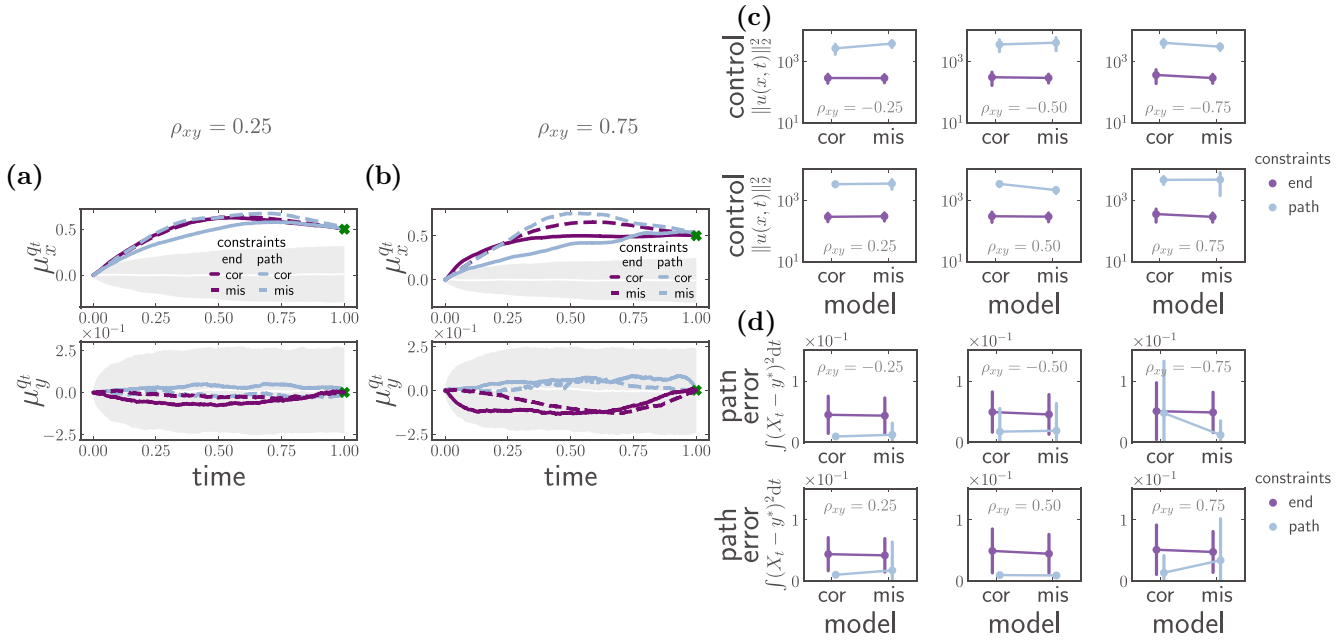

FIG. 5. Deterministic particle flow (DPF) control steers an evolutionary process of molecular phenotypes to a non (evolutionary) optimum state. [(a) and (b)] Transient mean of marginal densities of 1000 controlled trajectories, $\mu_x^{q_t}$ and $\mu_y^{q_t}$, of controlled trajectories of an evolutionary process with covarying phenotypes with correlation (a) $\varrho_{xy} = 0.25$, and (b) $\varrho_{xy} = 0.75$ starting from phenotypic state $\mathbf{x}_0 = (0, 0)$ and reaching the target $\mathbf{x}^* = (0.5, 0)$ (*green cross*) at time $T = 1.0$. The employed interventions were computed with the correct model (*solid lines*) or with a misspecified model that assumed zero cross-phenotype correlations $\varrho_{xy} = 0$. (*dashed line*) imposing: (i) only terminal constraints (*purple lines*) and (ii) both terminal and path constraints (*light blue lines lines*). The path constraint $U((x, y), t) = \beta y^2$, with $\beta = 10^3$ was set to limit the fluctuations around the y axis. The white line and grey shaded region indicate the mean and standard deviation of 1000 uncontrolled trajectories evolving under natural selection from the same initial state $\mathbf{x}_0$. [(c) and (d)] Comparison of (c) dissipated control energy $\|u(x, t)\|_2^2$, and (d) path deviations from target along the second (y) phenotype axis $\int ((X_t)_y - y^*)dt$ when steering the evolutionary processes with (top) anticorrelated and (bottom) correlated phenotypes with increasing (in absolute value) correlation coefficient ϱ_{xy} from left to right, with $\varrho_{xy} = \pm\{0.25, 0.50, 0.75\}$. The employed interventions followed the same model and constraint assumptions as in (a) and (b) with "cor" and "mis" denoting results from computed controls employing the correct and the misspecified dynamical model, respectively. Further parameters: particle number: $N = 1000$, inducing point number: $M = 50$, discretization time step: $dt = 10^{-3}$.

To accommodate synchronization control with our framework, we considered a time constrained setting where we applied DPF for an interval $[0,\ T]$ (time units). An alternative "online" approach described in Ref. [60], Sec. IX alternates computation of controls within small time intervals with simulation of controlled trajectories over those intervals.

We implemented the synchronization constraint as a path constraint that promotes synchrony without any further requirement for the terminal state, and characterized the level of synchronization in terms of the Kuramoto order parameter for phase coherence $R(\boldsymbol{\theta}_t, t)$ ([60], Sec. IX):

$$R(\boldsymbol{\theta}_t, t) = \frac{1}{K}\left|\sum_{j=1}^{K} e^{i\theta_t^{(j)}}\right|. \tag{29}$$

To that end, we employed the following path constraint that promotes order parameter values closer to 1, i.e., closer to a synchronous state

$$U(\boldsymbol{\theta}_t) = \beta(1 - R(\boldsymbol{\theta}_t, t)),$$

with $\boldsymbol{\theta}_t \in \mathbb{R}^K$ denoting the vector of oscillator phases, and $\beta \in \mathbb{R}$ a scaling constant.

For all considered networks of interacting Kuramoto oscillators, we applied interventions $u_i(\boldsymbol{\theta}_t, t)$ on the phases $\boldsymbol{\theta} = \{\theta^{(i)}\}_{i=1}^K$ of all network nodes. Further considerations of optimally selecting a subset of nodes to control were out of the scope of the current article.

For a prototypical network of $K = 2$ interacting oscillators with weak coupling ($J = 1.2$), DPF induced rapid synchrony, whereas the uncontrolled oscillators became progressively incoherent as indicated both by observing their phases [Fig. 6(a)] and the transient values of their phase-coherent order parameter R [Fig. 6(b)]. DPF provided fairly strong control inputs at the beginning of the simulation to fully align the phases of the oscillators [Fig. 6(c)] and subsequently delivered only moderate controls to maintain synchrony counteracting the effect of noise.

DPF successfully induced synchrony for different levels of coupling and under two different noise conditions $\sigma = \{0.5, 1.0\}$ [Fig. 7(a)]. For strongly coupled oscillators where the uncontrolled network progressively synchronizes solely due to the interactions, DPF synchronized the oscillators faster—indicated by earlier reaching the order parameter $R = 1$ value compared to their uncontrolled counterparts—and prevented spontaneous desynchronsation events that

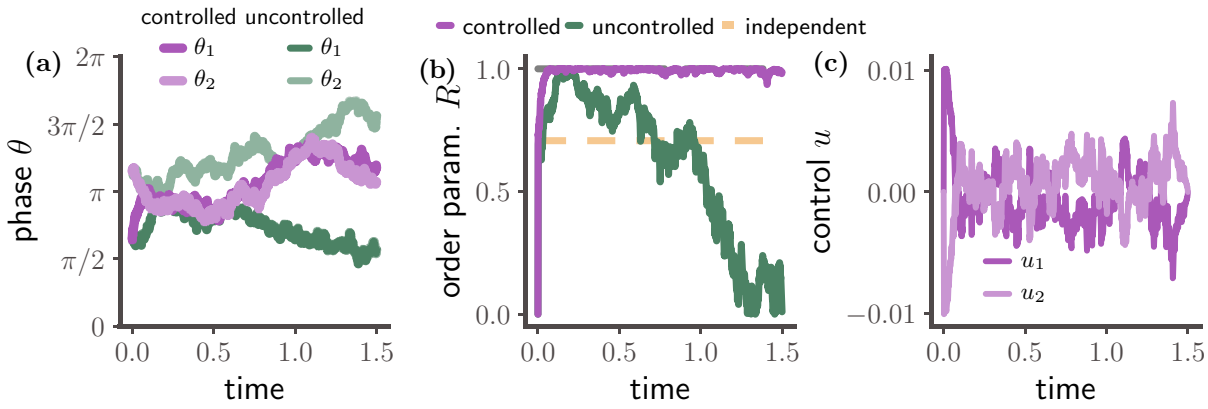

FIG. 6. Synchronization control of two coupled Kuramoto phase oscillators. (a) Evolution of phases (θ_i) of two controlled (*purple*) and two uncontrolled (*green*) Kuramoto oscillators mutually coupled with coupling $J = 1.2$ and noise amplitude $\sigma = 1.0$. (b) Evolution of Kuramoto phase-coherence order parameter R for the controlled (*purple*) oscillators indicates a fast transition to synchrony ($R = 1$), while the identical uncontrolled oscillators become progressively incoherent, indicated by a strongly fluctuating order parameter (*green*). The orange line denotes the expected long time average value of the order parameter for noninteracting oscillators considering finite size scaling effects. The grey line marks the level of $R = 1$ indicating a completely synchronous state. (c) Control input provided to each oscillator. Further parameters: particle number: $N = 2000$, inducing point number: $M = 80$, natural frequencies: $\omega_i = 0$, initial condition: $\theta^{(i)} \sim \mathcal{N}(3, 0.5^2)$, and $T = 1.5$ (time units).

occurred in the uncontrolled networks especially in the presence of strong noise [Fig. 7(b)].

To quantify these effects further, we analysed the onset of synchronization t^{syn} and the percentage of time spent in the synchronized state of the examined controlled and uncontrolled networks for increasing coupling J.

For each network realization, we defined the onset of synchronization t^{syn} as the first time point when the phase-coherence order parameter exceeded the value $R \geqslant 0.99$ and remained above that value for a minimal duration of $\tau^s = 20 \times dt = 0.02$ time units. For uncontrolled networks with weak couplings, we considered only the subset of realizations that reached the synchronous state [indicated as a ratio by the grey annotations in Fig. 8(a)]. To quantify the robustness of synchronization in each network, we estimated the percentage of time the network remained synchronized after

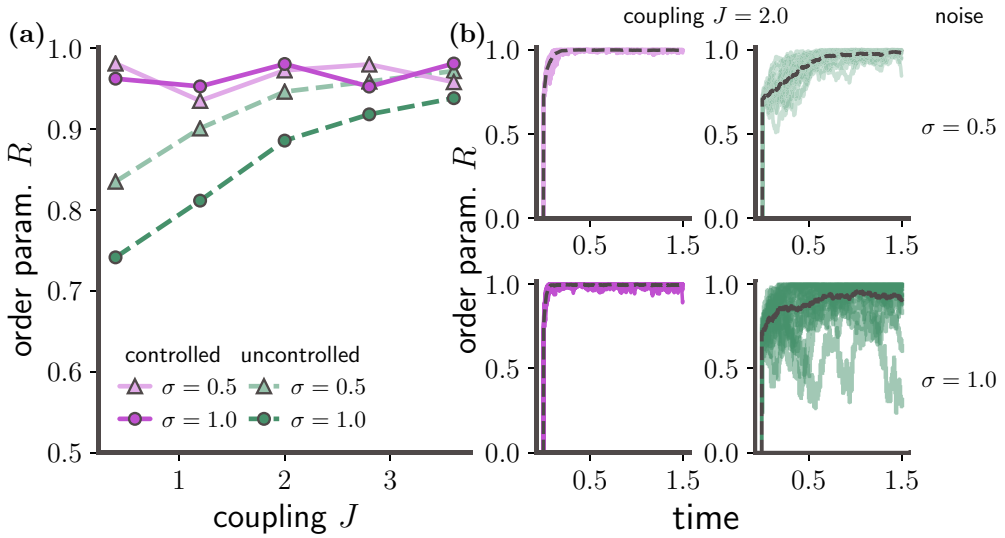

FIG. 7. Synchronization control of a network of two coupled Kuramoto phase oscillators for different coupling strengths. (a) Time averaged phase-coherence order parameter (R) of controlled (*purple*) and uncontrolled (*green*) networks under different noise conditions [$\sigma = 0.5$ (triangles) and 1.0 (circles)]. The proposed method (DPF) effectively synchronizes the controlled oscillators already for vanishing coupling strength between them. (b) Evolution of Kuramoto order parameter R for networks with coupling $J = 2.0$ and two noise conditions [$\sigma = 0.5$ (top) and 1.0 (bottom)] for controlled (*purple*) and uncontrolled (*green*) oscillators. The control induces fast transition to synchrony ($R = 1$), while the identical uncontrolled oscillators either synchronize slower (for low noise), or become only partially synchronized (for strong noise). Individual lines indicate evolution of the order parameter in 20 realizations of the network starting from same initial conditions and from a single computation of the required controls for each setting (where relevant). Dotted black lines denote the mean over the 20 realizations. Further parameters: particle number $N = 2000$, inducing point number: $M = 80$, natural frequencies: $\omega_i = 0$, initial condition: $\theta^{(i)} \sim \mathcal{N}(3, 0.5^2)$ and $T = 1.5$ (time units).

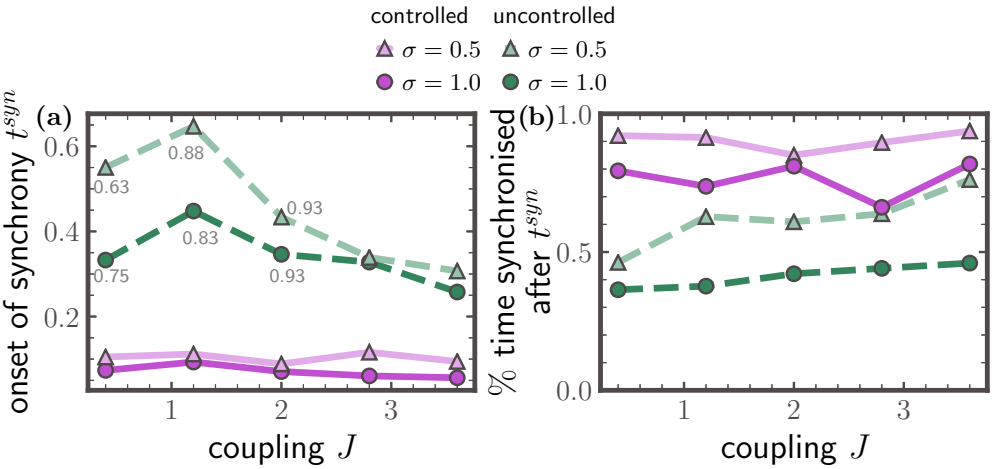

FIG. 8. Onset of synchronization and percentage of time in synchronized state after synchrony onset reveal the effectiveness of deterministic particle flow control to induce robust synchronization on a network of $K = 2$ oscillators. (a) Onset of synchronization for controlled (*purple*) and uncontrolled (*green*) networks quantified as the first time t^{syn} the phase-coherence order parameter exceeds $R \geqslant 0.99$ and remains over that value for duration $\tau^s = 20 \times dt = 0.02$ time units. Grey annotations denote the percentage of the examined networks that reached the synchronous state for duration τ^s. Absence of annotation indicates that all examined networks reached synchrony. (b) Percentage of simulation time the networks spontaneously spent in synchronized state ($R \geqslant 0.99$) after synchrony onset t^{syn}. Both figures consider two different noise conditions [$\sigma = 0.5$ (triangles) and 1.0 (circles)]. Further parameters: particle number $N = 2000$, inducing point number: $M = 80$ and $T = 1.5$ (time units). For each noise condition dots denote average over 3 control computations with different initial conditions with 20 controlled trajectories for each (60 total controlled trajectories for each point).

the synchronization onset by counting the time points when the order parameter *spontaneously* exceeded the $R = 0.99$ threshold after t^{syn}.

For both noise conditions and independent of coupling strength J, networks controlled by DPF reached the synchronous state considerably faster than their uncontrolled counterparts [Fig. 8(a)], and consistently remained synchronized for the entire simulation, as mediated by the percentage of time spent in the synchronized state after the synchronization onset t^{syn}. More precisely, while a subset of uncontrolled weakly coupled networks synchronized for at least $\tau^s = 20$ time units, as expected due to the presence of noise they failed to remain in that synchronized state as indicated in Fig. 8(b). For stronger couplings desynchronization was less pronounced, yet still more frequent than in their controlled counterparts.

For $K = 6$ interacting heterogeneous oscillators and for network characteristics (coupling strength $J = 1.0$ and noise amplitude $\sigma = 1.0$) that render the uncontrolled system only partially synchronizable [Fig. 9(b), bottom], state feedback control delivered by DPF successfully drove the oscillators to a fully synchronized state [Fig. 9(b), top]. As indicated by the evolution of the Kuramoto order parameter for each of the 20 realizations shown in Fig. 9(b), our framework not only delivered sufficient controls to rapidly synchronize the oscillators, but also provided the necessary interventions to maintain the phase synchronization [Fig. 9(c)]. In fact, as indicated by the nonfluctuating order parameter R for most realizations, only 2 network instances underwent spontaneous noise-induced desynchronizations late in the simulations, which nevertheless were partially recovered.

Similar to the smaller network, the six oscillator network was successfully synchronized for a range of coupling strengths [Fig. 10(a)] and for both noise conditions. Compared

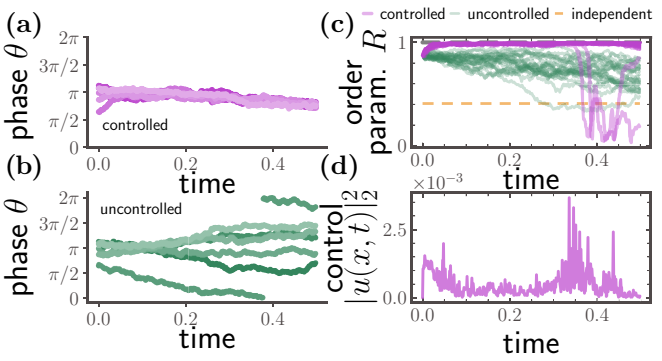

FIG. 9. Synchronization control of a finite-size network of $K = 6$ interacting Kuramoto phase oscillators. (a) Evolution of phases $\theta^{(i)}$ of a controlled network and (b) an identical uncontrolled network. The phases of the oscillators quickly synchronize when controlled by DPF and remain synchronized throughout the entire simulation interval $[0,\ T = 0.5]$. In the absence of control the phases of the oscillators become increasingly incoherent. (c) Evolution of Kuramoto order parameter capturing phase coherence R for the controlled (*purple*) oscillator network indicates a rapid transition to complete synchrony ($R = 1$) for all 20 realizations (individual purple lines). The order parameter of identical uncontrolled networks fluctuates strongly indicating partial incoherence (*green*). The orange line denotes the expected long time average value of the order parameter if the oscillators were noninteracting considering finite size scaling effects. For visual clarity, the grey line marks the level of $R = 1$ indicating a completely synchronous state. (d) Control energy spent on all $K = 6$ oscillators for a single control realization. Further parameters: coupling strength $J = 1.0$, noise amplitude $\sigma = 1$, particle number $N = 3000$, inducing point number $M = 300$, initial condition: $\theta^{(i)} \sim \mathcal{N}(3, 0.5^2)$, natural frequencies: $\omega_i \sim \mathcal{N}(0, 1)$ discretization time step $dt = 10^{-3}$ and $T = 0.5$ (time units).

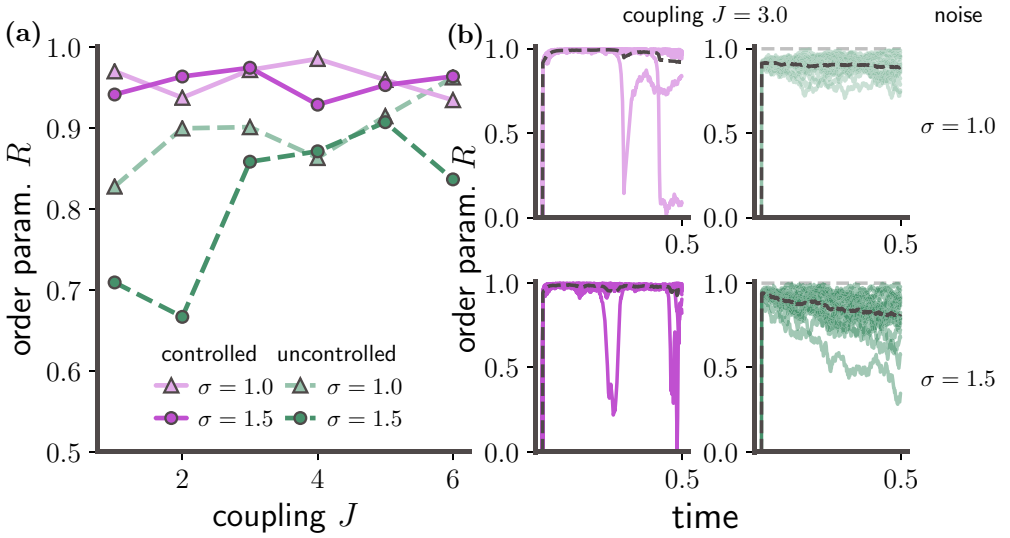

FIG. 10. Synchronization control of a network of six heterogeneous Kuramoto phase oscillators for different coupling strengths. (a.) Time averaged phase-coherence order parameter (R) of controlled (*purple*) and uncontrolled (*green*) networks under different noise conditions [$\sigma = 1.0$ (triangles) and 1.5 (circles)]. The proposed method (DPF) effectively synchronizes the controlled oscillators already for weak coupling. (b.) Evolution of Kuramoto order parameter R for networks with coupling $J = 3.0$ and two noise conditions [$\sigma = 1.0$ (top) and 1.5 (bottom)] for controlled (*purple*) and uncontrolled (*green*) oscillators. The control induces fast transition to synchrony ($R = 1$), while the identical uncontrolled oscillators either synchronize slower (for low noise), or become only partially synchronized (for strong noise). Individual lines indicate evolution of the order parameter in 20 realizations of the network starting from same initial conditions and from a single computation of the required controls for each setting (where relevant). Dashed black lines denote the mean over the 20 realizations. Further parameters: particle number $N = 3000$, inducing point number: $M = 300$, initial condition: $\theta^{(i)} \sim \mathcal{N}(3, 0.5^2)$, natural frequencies: $\omega_i \sim \mathcal{N}(0, 1)$ and $T = 0.5$ (time units). Each point in (a) denotes the mean over 20 realizations of a single control computation.

to uncontrolled networks with identical characteristics, the networks controlled by DPF exhibited through all examined settings larger phase coherence order parameter values. Although in some controlled network realizations the oscillators desynchronized spontaneously towards the end of the simulation interval, the phase coherence was quickly recovered in most cases [Fig. 10(b)]. These desynchronization events occurred when the state of the system reached the boundaries of the state space volume covered by the particles. They can be prevented by either increasing the number N of employed particles, or by resorting to an online control strategy with interleaved control computations and state advancement as described in Ref. [60], Sec. IX.

IV. DISCUSSION

In this paper, we introduced a novel methodological framework for estimating optimal dynamical interventions for constraining stochastic nonlinear systems. Distinctively from previous work [27,28,46,79] that devises optimal control protocols by employing iterative optimization procedures, here, we obtained the required interventions in a *deterministic* and *noniterative* way. We showed that splitting the time-resolved constraining information into retrospective and prospective parts, allows for a representation of the optimal controls in terms of the difference of logarithmic gradients (*scores*) of two forward probability flows. By introducing statistical estimators for the logarithmic gradients of the empirical probability densities, and by employing novel advances for deterministic evolution of sample based probability flows [54,72], we proposed an efficient, nonparametric approximation of the optimal controls.

We demonstrated the feasibility and potential of our framework on a battery of diverse, biologically inspired systems, and challenging settings of increasing complexity and dimensionality. More precisely, we employed the proposed method to induce switches between equilibrium states on multistable systems (Sec. III B), to devise artificial selection protocols on phenotypic landscapes by implementing constraints for covarying phenotypes (Sec. III C), and to induce synchronization on networks of stochastic phase oscillators (Sec. III D).

We compared our approach against the recently proposed Path integral cross entropy method [46], which approximates time dependent controls by iterative optimization based on stochastic path sampling. Our results suggest that our *one-shot, deterministic* framework is on par with the iterative path integral cross entropy method in terms of control efficiency, and more precise and accurate in terms of deviation from terminal target states.

We propagated probability densities by employing recent advances for solving Fokker-Planck equations in terms of deterministic particle dynamics. In principle, any particle filtering algorithm employing stochastic particle dynamics [80] may be combined with the logarithmic gradient density estimator of Eq. S.56 to obtain a numerical approximation of the time-reversed drift of Eq. (18) at each time step. Yet, numerical experiments of such a method (not shown here) showed, that the stochastic fluctuations of the particles lead to fairly noisy control estimates over time. Hence, here, taking advantage of the fact that the *deterministic sampling* framework of [54] already employs the logarithmic gradient

estimator of Eq. S.56 as a building block, we integrated this method into the computation of the optimal interventions.

Although the representation of constrained densities with particles is computationally more efficient compared to solutions of discretized PDEs, control representations for regions of the state space where the particles do not provide sufficient evidence for the underlying density, will be inaccurate. While the probability density functions for the regions of the state space unoccupied by particles are expected to have small values, with insufficient particle number, the estimated logarithmic gradient of the related densities might be inaccurate in low probability regions. However, due to the deterministic nature of our approach, our framework provides better representation of the underlying densities compared to stochastic path sampling.

While the proposed method assumes access to and controllability of all dynamical variables, such a scenario might be unrealistic for some practical problems. Interventions to biological systems may have limited access to the system's state variables, or only system parameters might be accessible for control. In these settings, a proper modification of the methodology that additionally considers these constraints is necessary.

We successfully employed the proposed method for synchronization control of networks of coupled stochastic Kuramoto oscillators. However, when path constraints are required, the deterministic particle flow control framework solves an optimal transport problem at every time step to implement the path constraints as a deterministic particle reweighting. This computation employs the Earth Mover's distance for an ensemble of N particles, to solve the linear program required for the optimal transport reweighting. This computation scales rather unfavorably for increasing particle number as $\mathcal{O}(N^3 \ln N)$. Here, for the networked systems we employed an alternative solution for computing the Earth Mover's distance, the network simplex solver [81,82] (see also Ref. [60], Sec. V), which has computational complexity that scales as $\mathcal{O}(N^2)$. Future developments of both theoretical and computational interest will focus on a dynamical reformulation of the reweighting problem.

Considering further the topic of network synchronization, we regarded out of the scope of the present paper to explore the possibility of controlling only a subset of network nodes or by considering mean-field couplings. Previous attempts to solve the same problem in a stochastic setting have considered only the synchronization of two coupled identical Kuramoto oscillators [83], while, here, we considered networks up to six oscillators with differing natural frequencies. Insights from network control theory for nonlinear systems and mean-field control coupled with the proposed framework may provide a more energy-efficient approach for network synchronization. Additionally, further systematic studies that will explore various network topologies, coupling schemes, and intrinsic parameter heterogeneities, will provide additional insight on the properties of our method to induce robust synchronization to networks of interacting stochastic (phase) oscillators.

Given the broad applicability of the proposed method, future numerical developments may further improve the computational complexity of the method, allowing thereby efficient computation of interventions for high dimensional systems. Candidate directions for improvement involve applying sequential solutions for the optimization problems, employing alternative functional representations of the logarithmic gradients, and generalizing the guided intervention variant of our method (gDPF) to nonconservative systems. Major computational steps of the algorithm result from optimization problems: the variational formulation of the logarithmic gradient, and the optimal transport treatment of the particle reweighting. In both cases, iterative solvers starting from the optimization result of the previous time step may provide more computationally efficient solutions. Moreover, neural network function representations for estimating the logarithmic gradients may result in more economic computations in settings, where the intrinsic dimensionality of invariant density of the stochastic dynamics is considerably lower than the dimensionality of the state space.

The proposed framework is moreover relevant for various computational or applied settings, where marginal densities of constrained diffusive systems are required, i.e., for parameter or state estimation of such systems, or for computing the transition probabilities in extreme event calculations, and more broadly for simulation based inference [84,85]. In those settings, only the constrained path distribution $q_t(x)$ is required instead of the precise dynamical interventions. Although not explicitly demonstrated here, the proposed method is also applicable for computing averages over constrained densities or functionals over constrained paths. Since the reverse time sampled flow $\tilde{q}_t(x)$ already provides a good representation of the constrained density, averages over functions evaluated on the paths of $q_t(x)$ already provide accurate estimates of the computed quantities for equilibrium processes.

We employed the proposed framework on a prototypical scenario of devising artificial selection protocols for molecular phenotypes inspired by [29]. The nascent field of continuous culturing [86,87] for studying adaptive evolution has created a growing demand for devising efficient and precise stochastic control frameworks that may be integrated in advanced open-source platforms like EVOLVER [30]. These platforms enable real time monitoring of cell cultures and administer *exact* custom perturbations in the form of selection pressures. By providing accurate interventions that implement arbitrary state constraints the proposed method is well suited for such a platform.

Taken together, the clearest advantage of our framework is its noniterative, and deterministic nature, providing thereby computational advantages compared to existing control methods [46,88]. Moreover, the proposed formulation for the optimal interventions generalizes beyond particle systems and may be easily implemented by neural networks.

A repository with a PYTHON implementation of the presented framework can be found in Ref. [100].

ACKNOWLEDGMENTS

We thank Sebastian Reich for insightful discussions during the early development of this work. This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG)-SFB1294/1-318763901. Preliminary account of this

work appeared in the *Machine Learning, and the Physical Sciences* Workshop at the 35th Conference on Neural Information Processing Systems (NeurIPS).

APPENDIX A: PHENOMENOLOGICAL MODEL OF CELL FATE DIFFERENTIATION

For the controlled numerical experiment with the nonconservative system we employed the model

$$dX_t = \left(\frac{{X_t}^4}{p_1^4 + {X_t}^4} + \frac{p_2^4}{p_2^4 + {Y_t}^4} - X_t\right) dt + \sigma dW_t^{(1)} \quad \text{(A1)}$$

$$dY_t = \left(\frac{{Y_t}^4}{p_3^4 + {Y_t}^4} + \frac{p_4^4}{p_4^4 + {X_t}^4} - Y_t\right) dt + \sigma dW_t^{(2)}, \quad \text{(A2)}$$

with parameters $p_1 = p_2 = p_3 = p_4 = 0.5$, and $W^{(1)}, W^{(2)}$ independent Wiener processes.

APPENDIX B: KURAMOTO MODEL

The Kuramoto model [101] employed in the numerical experiments comprises a population of $K \geqslant 1$ phase oscillators, with dynamics evolving according to the following system of nonlinear coupled stochastic equations

$$d\theta_t^{(i)} = \left(\omega_i + \frac{J}{K}\sum_{j=1}^{K} \sin\left(\theta_t^{(j)} - \theta_t^{(i)}\right)\right) dt + \sigma dW_t^{(i)}, \quad \text{(B1)}$$

with initial conditions $\theta_0^{(i)}$, and $\{W^{(i)}\}_{i=1}^K$ K independent Wiener processes. In Eq. (B1), $\theta_t^{(i)}$ denotes the phase of the i-th oscillator, ω_i its natural frequency. We consider the natural frequencies ω_i to arise from a Gaussian distribution $g(\omega)$ with mean zero, and variance w^2, $g(w) = \mathcal{N}(0, w^2)$.

[1] J. M. Raser and E. K. O'shea, Noise in gene expression: Origins, consequences, and control, Science **309**, 2010 (2005).

[2] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, Proc. Natl. Acad. Sci. USA **99**, 12795 (2002).

[3] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins, Noise-based switches and amplifiers for gene expression, Proc. Natl. Acad. Sci. USA **97**, 2075 (2000).

[4] T. B. Kepler and T. C. Elston, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations, Biophys. J. **81**, 3116 (2001).

[5] J. García-Ojalvo and J. Sancho, *Noise in Spatially Extended Systems* (Springer, New York, 2012).

[6] A. Eldar and M. B. Elowitz, Functional roles for noise in genetic circuits, Nature (London) **467**, 167 (2010).

[7] M. L. Simpson, C. D. Cox, M. S. Allen, J. M. McCollum, R. D. Dar, D. K. Karig, and J. F. Cooke, Noise in biological circuits, WIREs Nanomed. Nanobiotechnol **1**, 214 (2009).

[8] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, Noise in eukaryotic gene expression, Nature (London) **422**, 633 (2003).

[9] W. Horsthemke, Noise induced transitions, in *Non-Equilibrium Dynamics in Chemical Systems* (Springer, Berlin, Heidelberg, 1984), pp. 150–160.

[10] T. Zhou, L. Chen, and K. Aihara, Molecular Communication through Stochastic Synchronization Induced by Extracellular Fluctuations, Phys. Rev. Lett. **95**, 178103 (2005).

[11] M. Rullan, D. Benzinger, G. W. Schmidt, A. Milias-Argeitis, and M. Khammash, An optogenetic platform for real-time, single-cell interrogation of stochastic transcriptional regulation, Mol. Cell **70**, 745 (2018).

[12] D. Del Vecchio, A. J. Dy, and Y. Qian, Control theory meets synthetic biology, J. R. Soc. Interface **13**, 20160380 (2016).

[13] P. Nguyen, N. A. Pease, and H. Y. Kueh, Scalable control of developmental timetables by epigenetic switching networks, J. R. Soc. Interface **18**, 20210109 (2021).

[14] D. A. Sivak and G. E. Crooks, Thermodynamic Metrics and Optimal Paths, Phys. Rev. Lett. **108**, 190602 (2012).

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

[16] A. Gomez-Marin, T. Schmiedl, and U. Seifert, Optimal protocols for minimal work processes in underdamped stochastic thermodynamics, J. Chem. Phys. **129**, 024114 (2008).

[17] T. Schmiedl and U. Seifert, Optimal Finite-Time Processes In Stochastic Thermodynamics, Phys. Rev. Lett. **98**, 108301 (2007).

[18] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, Whither the future of controlling quantum phenomena? Science **288**, 824 (2000).

[19] A. N. Pechen and D. J. Tannor, Are there Traps in Quantum Control Landscapes? Phys. Rev. Lett. **106**, 120402 (2011).

[20] D. A. Hendrix and C. Jarzynski, A "fast growth" method of computing free energy differences, J. Chem. Phys. **114**, 5974 (2001).

[21] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, Equilibrium Free Energies from Nonequilibrium Measurements Using Maximum-Likelihood Methods, Phys. Rev. Lett. **91**, 140601 (2003).

[22] C. Hartmann and C. Schütte, Efficient rare event simulation by optimal nonequilibrium forcing, J. Stat. Mech.: Theory Exp. (2012) P11004.

[23] R. Chetrite and H. Touchette, Variational and optimal control representations of conditioned and driven processes, J. Stat. Mech.: Theory Exp. (2015) P12001.

[24] J. W. Kim and P. G. Mehta, An optimal control derivation of nonlinear smoothing equations, in *Advances in Dynamics, Optimization and Computation. SON 2020. Studies in Systems, Decision and Control*, edited by O. Junge, O. Schültze, G. Froyland, S. Ober-Blöbaum, and K. Padberg-Gehle (Springer, Cham, 2020), Vol. 304.

[25] J. Casadiego, D. Maoutsa, and M. Timme, Inferring Network Connectivity from Event Timing Patterns, Phys. Rev. Lett. **121**, 054101 (2018).

[26] E. Todorov, Efficient computation of optimal actions, Proc. Natl. Acad. Sci. USA **106**, 11478 (2009).

[27] H. J. Kappen, Linear Theory for Control of Nonlinear Stochastic Systems, Phys. Rev. Lett. **95**, 200201 (2005).

[28] D. K. Wells, W. L. Kath, and A. E. Motter, Control of Stochastic and Induced Switching in Biophysical Networks, Phys. Rev. X **5**, 031036 (2015).

[29] A. Nourmohammad and C. Eksin, Optimal Evolutionary Control for Artificial Selection on Molecular Phenotypes, Phys. Rev. X **11**, 011044 (2021).

[30] Z. Zhong, B. G. Wong, A. Ravikumar, G. A. Arzumanyan, A. S. Khalil, and C. C. Liu, Automated continuous evolution of proteins in-vivo, ACS Synthetic Biology **9**, 1270 (2020).

[31] T.-C. Kao, M. S. Sadabadi, and G. Hennequin, Optimal anticipatory control as a theory of motor preparation: A thalamo-cortical circuit model, Neuron **109**, 1567 (2021).

[32] M. Lässig and V. Mustonen, Eco-evolutionary control of pathogens, Proc. Natl. Acad. Sci. USA **117**, 19694 (2020).

[33] A. Iolov, S. Ditlevsen, and, A. Longtin, Stochastic optimal control of single neuron spike trains, J. Neural Eng. **11**, 046004 (2014).

[34] S. H. Scott, Optimal feedback control and the neural basis of volitional motor control, Nat. Rev. Neurosci. **5**, 532 (2004).

[35] E. Bernton, J. Heng, A. Doucet, and P. E. Jacob, Schrödinger bridge samplers, arXiv:1912.13170.

[36] F. Vargas, P. Thodoroff, A. Lamacraft, and N. Lawrence, Solving Schrödinger bridges via maximum likelihood, Entropy **23**, 1134 (2021).

[37] I. Exarchos and E. A. Theodorou, Stochastic optimal control via forward and backward stochastic differential equations and importance sampling, Automatica **87**, 159 (2018).

[38] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, in *Proceedings of the 9th International Conference on Learning Representations* (ICLR, Vienna, Austria, 2021).

[39] R. Bellman, Dynamic programming and Lagrange multipliers, Proc. Natl. Acad. Sci. USA **42**, 767 (1956).

[40] J. Garcke and A. Kröner, Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids, J. Sci. Comput. **70**, 1 (2017).

[41] M. Annunziato and A. Borzì, A Fokker–Planck control framework for multidimensional stochastic processes, J. Comput. Appl. Math. **237**, 487 (2013).

[42] M. B. Horowitz, A. Damle, and J. W. Burdick, Linear Hamilton Jacobi Bellman equations in high dimensions, in *53rd IEEE Conference on Decision and Control, Los Angeles, CA, USA* (IEEE, Piscataway, NJ, 2014), pp. 5880–5887.

[43] H. J. Kappen, Path integrals and symmetry breaking for optimal control theory, J. Stat. Mech.: Theory Exp. (2005) P11011.

[44] B. Van den Broek, W. Wiegerinck, and B. Kappen, Graphical model inference in optimal control of stochastic multi-agent systems, J. Artif. Intell. Res. **32**, 95 (2008).

[45] K. Rawlik, M. Toussaint, and S. Vijayakumar, Path integral control by reproducing kernel Hilbert space embedding, in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, edited by F. Rossi (IJCAI/AAAI, Beijing, China, 2013).

[46] H. J. Kappen and H. C. Ruiz, Adaptive importance sampling for control and inference, J. Stat. Phys. **162**, 1244 (2016).

[47] W. Zhang, H. Wang, C. Hartmann, M. Weber, and C. Schuette, Applications of the cross-entropy method to importance sampling and optimal control of diffusions, SIAM J. Sci. Comput. **36**, A2654 (2014).

[48] E. Theodorou, F. Stulp, J. Buchli, and Stefan Schaal, An iterative path integral stochastic optimal control approach for learning robotic tasks, IFAC Proc. **44**, 11594 (2011).

[49] S. Thijssen and H. J. Kappen, Path integral control and state-dependent feedback, Phys. Rev. E **91**, 032104 (2015).

[50] E. Todorov, General duality between optimal control and estimation, in *2008 47th IEEE Conference on Decision and Control* (IEEE, Piscataway, NJ, 2008), pp. 4286–4292.

[51] H. J. Kappen, V. Gómez, and M. Opper, Optimal control as a graphical model inference problem, Mach. Learn. **87**, 159 (2012).

[52] S. Levine, Reinforcement learning and control as probabilistic inference: Tutorial and review, arXiv:1805.00909.

[53] H. Attias, Planning by probabilistic inference, in *International Workshop on Artificial Intelligence and Statistics* (PMLR, 2003), pp. 9–16.

[54] D. Maoutsa, S. Reich, and M. Opper, Interacting particle solutions of Fokker–Planck equations through gradient–log–density estimation, Entropy **22**, 802 (2020).

[55] B. D. O. Anderson, Reverse-time diffusion equation models, Stochastic Processes Applications **12**, 313 (1982).

[56] D. T. Gillespie and L. R. Petzold, Improved leap-size selection for accelerated stochastic simulation, J. Chem. Phys. **119**, 8229 (2003).

[57] A. Saarinen, M.-L. Linne, and O. Yli-Harja, Stochastic differential equation model for cerebellar granule cell excitability, PLoS Comput. Biol. **4**, e1000004 (2008).

[58] R. Lande, S. Engen, and B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, Oxford, England, UK, 2003).

[59] N. Takahata, K. Ishii, and H. Matsuda, Effect of temporal fluctuation of selection coefficient on gene frequency in a population, Proc. Natl. Acad. Sci. USA **72**, 4541 (1975).

[60] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.043035 for detailed derivations, additional information on the numerical experiments, and supplementary figures, which includes Refs. [89–99].

[61] E. Todorov, Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system, Neural Comput. **17**, 1084 (2005).

[62] E. Todorov, Optimality principles in sensorimotor control, Nat. Neurosci. **7**, 907 (2004).

[63] E. Todorov, Linearly-solvable Markov decision problems, in *Proceedings of the Twentieth Annual Conference on Advances in Neural Information Processing Systems*, edited by B. Schölkopf, J. C. Platt, and T. Hoffman (MIT Press, 2007), pp. 1369–1376.

[64] W. H. Fleming, Exit probabilities and optimal stochastic control, Appl. Math. Optim. **4**, 329 (1977).

[65] H. Orland, Generating transition paths by Langevin bridges, J. Chem. Phys. **134**, 174114 (2011).

[66] A. Mazzolo, Constrained Brownian processes and constrained Brownian bridges, J. Stat. Mech.: Theory Exp. (2017) 023203.

[67] J. Szavits-Nossan and M. R. Evans, Inequivalence of nonequilibrium path ensembles: the example of stochastic bridges, J. Stat. Mech.: Theory Exp. (2015) P12008.

[68] S. N. Majumdar and H. Orland, Effective Langevin equations for constrained stochastic processes, J. Stat. Mech.: Theory Exp. (2015) P06039.

[69] N. Macris and R. Marino, Solving non-linear Kolmogorov equations in large dimensions by using deep learning: a numerical comparison of discretization schemes, arXiv:2012.07747.

[70] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, Fourier neural operator for parametric partial differential equations, *International Conference on Learning Representations* (ICLR, Vienna, Austria, 2020).

[71] In the second equality, we considered that for small δt the commutator of the two operators $\mathcal{L}_f$ and $U(x,t)$ is negligible.

[72] S. Reich, A nonparametric ensemble transform method for Bayesian inference, SIAM J. Sci. Comput. **35**, A2013 (2013).

[73] U. von Toussaint, Bayesian inference in physics, Rev. Mod. Phys. **83**, 943 (2011).

[74] V. De Bortoli, J. Thornton, J. Heng, and A. Doucet, Diffusion Schrödinger bridge with applications to score-based generative modeling, Advances in Neural Information Processing Systems **34**, 17695 (2021).

[75] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, Heidelberg, 2009).

[76] C. Villani, *Optimal Transport: Old and New* (Springer, Berlin, Heidelberg, 2009), Vol. 338.

[77] S. Huang, Y.-P. Guo, G. May, and T. Enver, Bifurcation dynamics in lineage-commitment in bipotent progenitor cells, Dev. Biol. **305**, 695 (2007).

[78] R. Lande, Natural selection and random genetic drift in phenotypic evolution, Evolution, **30** 314 (1976).

[79] N. Nüsken and L. Richter, Solving high-dimensional Hamilton–Jacobi–Bellman PDEs using neural networks: Perspectives from the theory of controlled diffusions and measures on path space, Partial Differ. Equ. Appl. **2**, 48 (2021).

[80] H. Moradkhani, G. S. Nearing, P. Abbaszadeh, and S. Pathiraja, Fundamentals of data assimilation and theoretical advances, *Handbook of Hydrometeorological Ensemble Forecasting* (Springer, Berlin, Heidelberg, 2019), pp. 675–699.

[81] N. Bonneel, M. V. De Panne, S. Paris, and W. Heidrich, Displacement interpolation using Lagrangian mass transport, in *Proceedings of the 2011 SIGGRAPH Asia Conference* (ACM, New York, 2011), pp. 1–12.

[82] R. Flamary, N. Courty, A. Gramfort, M. Z. Alaya, A. Boisbunon, S. Chambon, L. Chapel, A. Corenflos, K. Fatras, N. Fournier *et al.*, POT: Python optimal transport, J. Mach. Learn. Res. **22**, 1 (2021).

[83] J.-S. Li, W. Zhang, and S. Wang, Optimal control and stochastic synchronization of phase oscillators, IFAC-PapersOnLine **48**, 83 (2015).

[84] J. Brehmer, Simulation-based inference in particle physics, Nat. Rev. Phys. **3**, 305 (2021).

[85] K. Cranmer, J. Brehmer, and G. Louppe, The frontier of simulation-based inference, Proc. Natl. Acad. Sci. USA **117**, 30055 (2020).

[86] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, 2019).

[87] D. Gresham and M. J. Dunham, The enduring utility of continuous culturing in experimental evolution, Genomics **104**, 399 (2014).

[88] M. Hairer, A. M. Stuart, and J. Voss, Sampling conditioned diffusions, Trends in Stochastic Analysis **353**, 159 (2009).

[89] I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, Theory Probab. Appl. **5**, 285 (1960).

[90] K. D. Elworthy, *Stochastic Differential Equations on Manifolds* (Springer, New York, 1982), Vol. 70.

[91] E. Bayraktar and M. Sirbu, Stochastic Perron's Method for Hamilton–Jacobi–Bellman Equations, SIAM J. Control Optim. **51**, 4274 (2013).

[92] R. Marques and G. Storvik, Reweighting schemes based on particle methods, in *The Contribution of Young Researchers to Bayesian Statistics* (Springer, Cham, 2014), pp. 73–76.

[93] O. Pele and M. Werman, Fast and robust earth mover's distances, in *2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan* (IEEE, Piscataway, NJ, 2009), pp. 460–467.

[94] A. Hyvärinen, Estimation of non-normalized statistical models by score matching, J. Mach. Learn. Research **6**, 695 (2005).

[95] M. L. Waskom, Seaborn: Statistical data visualization, J. Open Source Software **6**, 3021 (2021).

[96] A. Nourmohammad, S. Schiffels, and M. Lässig, Evolution of molecular phenotypes under stabilizing selection, J. Stat. Mech.: Theory Exp. (2013) P01012.

[97] T. Held, A. Nourmohammad, and M. Lässig, Adaptive evolution of molecular phenotypes, J. Stat. Mech.: Theory Exp. (2014) P09029.

[98] G. A. Gottwald, Finite-size effects in a stochastic Kuramoto model, Chaos **27**, 101103 (2017).

[99] A. T. Winfree, *The Geometry of Biological Time* (Springer, Berlin, Heidelberg, 2001), Vol. 12.

[100] A repository with a python implementation of the presented framework can be found in https://github.com/dimitramaoutsa/DeterministicParticleFlowControl.

[101] Y. Kuramoto, Self-entrainment of a population of coupled non-linear oscillators, in *International Symposium on Mathematical Problems in Theoretical Physics* (Springer, Berlin, Heidelberg, 1975), pp. 420–422.