Supplementary Information for: 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, United Kingdom*

CONTENTS

I. EQUIVALENCE BETWEEN CONTROLLED DIFFUSIONS AND PATH REWEIGHTING

We consider a stochastic system whose dynamics is governed by the SDE

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

with initial condition $X_0 = x_0$. In Eq. (S.1) the drift $f(\cdot,\cdot) : \mathcal{R}^d \times \mathcal{R}^+ \to \mathcal{R}^d$ is a smooth nonlinear function representing the deterministic part of the driving forces, while W stands for a d–dimensional vector of independent Wiener processes acting as white noise sources, modelling the contributions from unaccounted degrees of freedom, thermal fluctuations, or external perturbations. We denote the noise strength by $\sigma \in \mathcal{R}$.

We are interested in introducing path and terminal constraints into the system, which we denote by $U(x,t)$ and $\chi(X_T)$. Therefore we define the constrained process by a reweighting of the paths induced by the unconstrained process of Eq. (S.1).The Radon-Nykodym derivative of the constrained path measure $\mathbb{Q}^*$ with respect to the unconstrained one $\mathbb{P}_f$ is

$$\mathscr{L} = \frac{d\mathbb{Q}^*}{d\mathbb{P}_f}(X_{0:T}) = \frac{\chi(X_T)}{Z} \exp\left[-\int_0^T U(X_t, t)dt\right]. \tag{S.2}$$

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

In Eq. (S.2) Z denotes the partition function that renders $\mathbb{Q}^*$ properly normalised i.e. $\int d\mathbb{Q}^* = 1$, and is given by

$$Z = \left\langle \chi(X_T) \exp\left(- \int_0^T U(X_t, t) dt \right) \right\rangle_{\mathbb{P}_f} . \tag{S.3}$$

Eq. (S.2) implies that path probabilities with respect to $\mathbb{Q}^*$ are obtained by those of $\mathbb{P}_f$ by a reweighting with the factor on the right and side of Eq. (S.2).

We show that we can eliminate the constraints by introducing a new process with drift

$$g(x,t) = f(x,t) + \sigma^2 \nabla_x \mathcal{J}(x,t), \tag{S.4}$$

and same diffusion, i.e.

$$dX_t = g(X_t, t)dt + \sigma dW_t, \tag{S.5}$$

that induces the path measure $\mathbb{P}_g$.

In Eq. (S.4), $\mathcal{J}(\cdot)$ denotes the solution to the *Hamilton–Jacobi–Bellman* partial differential equation

$$\frac{1}{2}\|\nabla_x \mathcal{J}(x,t)\|^2_{\sigma^2} + \mathcal{L}_f^\dagger \mathcal{J}(x,t) + \frac{\partial \mathcal{J}(x,t)}{\partial t} = U(x,t)\mathcal{J}(x,t) \qquad \text{with} \quad \mathcal{J}(x,T) = \ln \chi(x). \tag{S.6}$$

Here the square norm is defined with respect to the diffusion as $\|A\|^2_{\sigma^2} \doteq A^\top \cdot \sigma^2 A$, while we have also introduced the adjoint Fokker–Planck operator

$$\mathcal{L}_f^\dagger \mathcal{J}(x,t) = f(x,t) \cdot \nabla_x \mathcal{J}(x,t) + \frac{1}{2}\sigma^2 \nabla^2_{xx} \mathcal{J}(x,t). \tag{S.7}$$

We will show that the path measures $\mathbb{Q}^*$ and $\mathbb{P}_g$ induced by the constrained process and by the process of Eq. (S.5) respectively are the same. The proof follows from Girsanov's change of measure theorem [89] for the two path measures. This yields

$$\begin{aligned} \frac{d\mathbb{Q}^*}{d\mathbb{P}_g}(X_{0:T}) = \frac{d\mathbb{Q}^*}{d\mathbb{P}_f}\frac{d\mathbb{P}_f}{d\mathbb{P}_g}(X_{0:T}) = \frac{\chi(X_T)}{Z} \exp\Big(- \mathcal{J}(X_T, T) + \mathcal{J}(X_0, 0)\Big) \\ \times \exp\left(\int_0^T \left(-U(X_t,t) + \frac{1}{2}\|\nabla_x \mathcal{J}(X_t,t)\|^2_{\sigma^2} + \mathcal{L}_f^\dagger \mathcal{J}(X_t,t) + \frac{\partial \mathcal{J}(X_t,t)}{\partial t} \right) dt \right). \end{aligned} \tag{S.8}$$

In Eq. (S.8) we have used that for the likelihood ratio of $\mathbb{P}_g$ with respect to $\mathbb{P}_f$ we have

$$\begin{aligned} -\ln \frac{d\mathbb{P}_g}{d\mathbb{P}_f}(X_{0:T}) &= \frac{1}{2}\int_0^T \Big(\|g(X_t,t)\|^2_{\sigma^{-2}} - \|f(X_t,t)\|^2_{\sigma^{-2}} \Big) dt - \int_0^T \Big(g(X_t,t) - f(X_t,t) \Big)^\top \sigma^{-2} dX_t & \text{(S.9)} \\ &= \frac{1}{2}\int_0^T \Big(\|\nabla_x \mathcal{J}(X_t,t)\|^2_{\sigma^2} + 2 f(X_t,t) \cdot \nabla_x \mathcal{J}(X_t,t) \Big) dt - \int_0^T \nabla_x \mathcal{J}(X_t,t) \cdot dX_t. & \text{(S.10)} \end{aligned}$$

To eliminate the stochastic integral $\nabla_x \mathcal{J}(X_t,t) \cdot dX_t$ in Eq.(S.10) we use Ito's lemma

$$d\mathcal{J}(X_t,t) = \frac{\partial \mathcal{J}(X_t,t)}{\partial t} dt + \nabla_x \mathcal{J}(X_t,t) \cdot dX_t + \frac{1}{2}\sigma^2 \nabla^2_{xx} \mathcal{J}(X_t,t) dt \tag{S.11}$$

to obtain

$$\begin{aligned} -\ln \frac{d\mathbb{P}_g}{d\mathbb{P}_f}(X_{0:T}) &= \int_0^T \left\{ \frac{1}{2}\|\nabla_x \mathcal{J}(X_t,t)\|^2_{\sigma^2} + \mathcal{L}_f^\dagger \mathcal{J}(X_t,t) + \frac{\partial \mathcal{J}(X_t,t)}{\partial t} \right\} dt - \mathcal{J}(X_T,T) + \mathcal{J}(X_0,0) & \text{(S.12)} \\ &= \int_0^T U(X_t,t) dt - \ln \chi(X_T) + \mathcal{J}(X_0,0). & \text{(S.13)} \end{aligned}$$

In the last equality we have inserted the Hamilton-Jacobi-Bellman equation (Eq.(S.6)) and its terminal condition at time T to arrive at the term $\int_0^T U(X_t,t)dt - \ln \chi(X_T)$. Comparing Eq. (S.13) with Eq. (S.8), to show the two path measures are the same, i.e.

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}_g}(X_{0:T}) = 1,$$

it remains to show that

$$\mathcal{J}(x_0, 0) = \ln Z. \tag{S.14}$$

Indeed by employing the Hopf-Cole logarithmic transformation [64] we have

$$\mathcal{J}(x, t) = -\ln \varphi_t(x), \tag{S.15}$$

where φ fulfils the backward PDE

$$\frac{\partial \varphi_t(x)}{\partial t} + \mathcal{L}_f^\dagger \varphi_t(x) - U(x,t)\varphi_t(x) = 0\ , \tag{S.16}$$

where the operator $\mathcal{L}_f^\dagger$ is defined by

$$\mathcal{L}_f^\dagger \varphi_t(x) = f(x)\cdot\nabla_x\varphi_t(x) + \frac{1}{2}\sigma^2\nabla^2_{xx}\varphi_t(x) \qquad \text{with} \quad \varphi_T(x) = \chi(x). \tag{S.17}$$

Since Eq. (S.16) is a parabolic partial differential equation, by employing the *Feynman–Kac theorem* (see [90]) its solutions $\varphi_t(x)$ can be written as the following path integral/conditional expectation over the path measure $\mathbb{P}_f$

$$\varphi_t(x) = \left\langle \chi(X_T)\exp\left[-\int_t^T U(X_s, s)ds\right] | X_t = x \right\rangle_{\mathbb{P}_f}. \tag{S.18}$$

By setting now $t = 0$ in Eq.(S.18), we obtain $\varphi_0(X_0) = Z$ and thus conclude that

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}_g}(X_{0:T}) = 1,$$

and therefore the two path measures are the same.

II. FROM CONSTRAINTS TO THE HAMILTON-JACOBI-BELLMAN EQUATION

We consider here stochastic systems in a more general setting compared to the one employed in the main text, where the system dynamics is captured by the following stochastic differential equation

$$dX_t = f(X_t, t)dt + \sigma(X_t, t)dW_t, \tag{S.19}$$

with state variables $X_t \in \mathcal{R}^d$, drift $f(\cdot,\cdot) : \mathcal{R}^d \times \mathcal{R}^+ \to \mathcal{R}^d$, state and time dependent noise matrix $\sigma(x,t) \in \mathcal{R}^d \times \mathcal{R}^+ \to \mathcal{R}^{d\times d}$, and W_t denoting a d–dimensional vector of independent Wiener processes acting as white noise sources.

We intend to impose dynamical constraints $\mathcal{C}$ onto the system that may pertain its terminal state X_T as dictated by a constraining function $\chi(x) \in \mathcal{R}^d \to \mathcal{R}$, or its transient behaviour within a time interval $[0,\ T]$ through a path-penalising function $U(x,t) \in \mathcal{R}^d \times \mathcal{R}^+ \to \mathcal{R}$

$$\mathcal{C} := \begin{cases} \chi(X_T), & \text{terminal constraint} \\ U(X_t, t), \text{for } t \le T, & \text{path constraint.} \end{cases} \tag{S.20}$$

To that end, we introduce the constraints into the evolution equation of the system Eq.(S.19) as a time- and space-dependent perturbation of the uncontrolled dynamics

$$dX_t = \Big(f(X_t) + u(X_t, t)\Big)dt + \sigma(X_t, t)dW_t. \tag{S.21}$$

To identify the necessary dynamical adjustments $u(x,t)$, we consider the problem of Eq.(7) in a more general form by introducing the cost functional $J(X_t, t)$ as the time integrated optimal expected cost from time t to the end of the interval T

$$\mathcal{J}(X_t, t) = \min_u \left\langle \int_t^T \frac{1}{2}u(X_{t'}, t')^\top H(X_{t'}, t')u(X_{t'}, t') + U(X_{t'}, t')\,dt' - \ln\chi(X_T)\right\rangle_{\mathbb{Q}}. \tag{S.22}$$

The term $u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t')$ accounts for the minimisation of the dissipated energy through the controls, with H determining the control cost along each state dimension, while the function $U(X_{t'},t')$ captures the cost from deviating from a predefined path (or from more general collective path costs in multi-unit systems). The angular brackets denote the expectation over the path probability measure $\mathbb{Q}$ for paths of the controlled system (Eq.(S.21)) within the interval $[t,\ T]$, i.e.

$$\Big\langle C(X_{t:T})\Big\rangle_{\mathbb{Q}} = \int C(X_{t:T})d\mathbb{Q}. \tag{S.23}$$

In Eq. (S.23) $C(\cdot)$ denotes a function of continuous trajectories (paths) starting from the point x_t, $X_{t:T} := (X_s)_{s\in(t,T]}|x_t$, and $\mathbb{Q}$ the path probability measure induced by the controlled dynamics.

To arrive to the Hamilton-Jacobi-Bellman equation that determines the necessary and sufficient optimality conditions for the control problem, we rewrite the optimal expected cost $\mathcal{J}(x,t)$ recursively

$$\begin{aligned}
\mathcal{J}(X_t,t) &= \min_u \left\langle \int_t^T \left(\frac{1}{2}u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t') + U(X_{t'},t')\right)dt' - \ln\chi(X_T)\right\rangle_{\mathbb{Q}} \\
&= \lim_{\delta t\to 0}\min_u \left\langle \int_t^{t+\delta t} \left(\frac{1}{2}u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t') + U(X_{t'},t')\right)dt' + \right. \\
&\qquad \left. \int_{t+\delta t}^T \left(\frac{1}{2}u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t') + U(X_{t'},t')\right)dt' - \ln\chi(X_T)\right\rangle_{\mathbb{Q}} \\
&= \lim_{\delta t\to 0}\min_u \left\langle \int_t^{t+\delta t} \left(\frac{1}{2}u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t') + U(X_{t'},t')\right)dt' + \mathcal{J}(X_{t+\delta t},t+\delta t)\right\rangle_{\mathbb{Q}} \\
&= \lim_{\delta t\to 0}\min_u \left\langle \int_t^{t+\delta t} \left(\frac{1}{2}u(X_{t'},t')^\top H(X_{t'},t')u(X_{t'},t') + U(X_{t'},t')\right)dt' + \mathcal{J}(X_t,t) + \delta t\frac{\partial\mathcal{J}(X_t,t)}{\partial t} + \right. \\
&\qquad \left. + g(X_t,t)^\top\nabla_x\mathcal{J}(X_t,t)\delta t + \frac{\delta t}{2}\mathrm{Tr}\left[D(X_t,t)\nabla^2_{xx}\mathcal{J}(X_t,t)\right]\right\rangle_{\mathbb{Q}},
\end{aligned} \tag{S.24}$$

with $D(x,t) = \sigma(x,t)^\top\sigma(x,t)$ denoting the noise covariance, and $\nabla^2_{xx}\mathcal{J}(X_t,t)$ indicating the matrix of second derivatives of $\mathcal{J}(X_t,t)$ with respect to the spatial variable.

In the third line we have isolated the expected cost-to-go for the time interval $[t,t+\delta t]$ from the optimal cost-to-go of subsequent times $\mathcal{J}(X_{t+\delta t},t+\delta t)$. This expression encompasses the Bellman's principle of optimality, according to which an optimal solution over an interval $[0,\ T]$ consists of optimal sub-solutions over the respective sub-intervals $[t',\ T]$ with later starting times t', and appropriate initial conditions [39].

In fourth line in Eq.(S.24) we have employed a Taylor approximation for $\mathcal{J}(X_{t+\delta t},t+\delta t)$ around t up to $\mathcal{O}(dt)$ and $\mathcal{O}(dX_t^2)$ that yields

$$\mathcal{J}(X_{t+\delta t},t+\delta t) \approx \mathcal{J}(X_t,t) + \delta t\frac{\partial\mathcal{J}(X_t,t)}{\partial t} + \mathcal{O}(\delta t^2) + dX_t\nabla_x\mathcal{J}(X_t,t) + \frac{1}{2}dX_t^\top\nabla^2_{xx}\mathcal{J}(X_t,t)dX_t. \tag{S.25}$$

We proceed by taking into account that X_t arises from an SDE of the form

$$dX_t = g(X_t,t)dt + \sigma(X_t,t)dW_t = (f(X_t,t)+u(X_t,t))dt + \sigma(X_t,t)dW_t,$$

and by considering that the expectation over $\mathbb{Q}$ of the last term of Eq. (S.25) equals

$$\Big\langle dX_t^\top\nabla^2_{xx}\mathcal{J}(X_t,t)dX_t\Big\rangle_{\mathbb{Q}} = \mathrm{Tr}\Big[D(X_t,t)\nabla^2_{xx}\mathcal{J}(X_t,t)\Big]\delta t. \tag{S.26}$$

In the fourth term of Eq. (S.25) the expectation of the stochastic part vanishes $\Big\langle\sigma(X_t,t)dW_t\nabla_x\mathcal{J}(X_t,t)\Big\rangle_{\mathbb{Q}} = 0$, where $\mathrm{Tr}[\cdot]$ stands for the trace of the matrix.

By dividing by δt and taking the limit in Eq. (S.24) we arrive at the Hamilton–Jacobi–Bellman equation

$$-\frac{\partial}{\partial t}\mathcal{J}(X_t,t) = \min_u\left[\frac{1}{2}u(X_t)^\top H(X_t,t)u(X_t) + U(X_t,t) + g(X_t,t)^\top\nabla_x\mathcal{J}(X_t,t) + \frac{1}{2}\mathrm{Tr}[D(X_t,t)\nabla^2_{xx}\mathcal{J}(X_t,t)]\right], \tag{S.27}$$

which admits at most one unique (viscosity) solution [91]. If such a solution exists, then the expected cost $\mathcal{J}(X_t,t)$ is the optimal one. Minimising Eq.(S.27) yields the optimal state- and time-dependent interventions

$$u^*(x,t) = -H^{-1}(x,t)\nabla_x\mathcal{J}(x,t). \tag{S.28}$$

Substituting the optimal controls $u^*(x,t)$ into the Hamilton-Jacobi-Bellman equation (Eq.(S.27)), we arrive at

$$-\frac{\partial}{\partial t}\mathcal{J}(X_t,t) = -\frac{1}{2}(\nabla_x \mathcal{J}(X_t,t))^\top H^{-1}(x,t)\nabla_x \mathcal{J}(X_t,t) + U(X_t,t) + f(X_t,t)^\top \nabla_x \mathcal{J}(X_t,t) + \frac{1}{2}\mathrm{Tr}[D(X_t,t)\nabla^2_{xx}\mathcal{J}(X_t,t)], \tag{S.29}$$

which is a nonlinear PDE in $\mathcal{J}(X_t,t)$.

To simplify the last equation, we introduce the variable transformation $\mathcal{J}(x,t) = -\ln(\varphi_t(x))$, and thus $\varphi_t(x) = -\exp(\mathcal{J}(x,t))$, thereby obtaining

$$\frac{1}{\varphi_t(X_t)}\frac{\partial}{\partial t}\varphi_t(X_t) = -\frac{1}{2\varphi_t^2(X_t)}(\nabla_x \varphi_t(X_t))^\top H^{-1}(x,t)\nabla_x \varphi_t(X_t) + U(X_t,t) \tag{S.30}$$

$$-\frac{1}{\varphi_t(X_t)} f(X_t,t)^\top \cdot \nabla_x \varphi_t(X_t) + \frac{1}{2\varphi_t^2(X_t)}(\nabla_x \varphi_t(X_t))^\top D(X_t,t)(\nabla_x \varphi_t(X_t)) \tag{S.31}$$

$$-\frac{1}{2\varphi_t(X_t)}\mathrm{Tr}[D(X_t,t)\nabla^2_{xx}\varphi_t(X_t)].$$

Eq. (S.30) becomes linear in $\varphi_t(x)$ if we assume that noise covariance $D(x,t)$ is inversely proportional to the intervention cost H, i.e. by assuming that $H^{-1}(x,t) = D(x,t)$. This relationship characterises *Path Integral control problems* [27]. Thereby the transformed Hamilton–Jacobi–Bellman equation simplifies into

$$\frac{\partial}{\partial t}\varphi_t(X_t) = -f(X_t,t)^\top \cdot \nabla_x \varphi_t(X_t) - \frac{1}{2}\mathrm{Tr}[D(X_t,t)\nabla^2_{xx}\varphi_t(X_t)] + U(X_t,t)\varphi_t(X_t) \tag{S.32}$$

$$\doteq \mathcal{L}_f^\dagger \varphi_t(X_t) + U(X_t,t)\varphi_t(X_t),$$

where $\mathcal{L}_f^\dagger$ denotes the *adjoint Fokker–Planck operator* associated with the uncontrolled system.

III. TACKLING PROBLEMS WITH EXTREME TERMINAL CONSTRAINTS

With a small modification the presented framework is also able to tackle problems with extreme (not typical) terminal constraints, i.e. with terminal constraints that are not reached by the typical paths of the unconstrained system. In settings where the terminal point lies outside the typical values of the uncontrolled system, the forward flow $\varrho_t(x)$ fails to cover the state space in the vicinity of this extreme terminal point. Thereby the particles do not provide sufficient evidence for the ensuing score estimation which results in inaccurate estimates.

To tackle this issue, we modify the forward sampling to already incorporate the extreme terminal constraint. To that end, we employ a *d-dimensional Brownian Bridge (B)* for forward sampling. Brownian bridges are derived from Brownian motion with vanishing drift $f(x) \equiv 0$ and terminal constraint $\chi(x) = \delta(x-y)$. Therefore they are sufficiently simple to admit an analytic closed form solution, facilitating thereby succeeding calculations for correcting the path statistics.

To maintain the correct path statistics, we employ the Girsanov's change of measure formula to reweight the biased forward paths. In particular, 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

$$\frac{d\mathbb{P}_f^B}{d\mathbb{P}_0^B}(X_{(0:T)}) \propto \exp\left[-\int_0^T \frac{1}{2\sigma^2}\left(f^2(X_t)dt - 2f(X_t)\cdot dX_t\right)\right]. \tag{S.33}$$

Since the path measure $\mathbb{P}_0^B$ contains the correct terminal constraints of the targeted path measure $\mathbb{P}_f^B$, the common terminal-condition $\chi(X_T)$ of both bridges has dropped out of the likelihood ratio, and we are left with the likelihood ratio between the unconstrained process with drift f and drift 0.

To transform the expression for the weight Eq.(S.33) into the form of Eq.(3) for which our filtering approach applies, we employ Ito's formula to eliminate the stochastic integral in Eq.(S.33) assuming thereby that the drift comes as a potential gradient $f(x) = -\nabla V(x)$, and therefore the underlying system is conservative. In particular, to eliminate the stochastic integral, we apply Ito's lemma in the form

$$dV(X_t) = \nabla V(X_t)\cdot dX_t + \frac{\sigma^2}{2}\nabla^2 V(X_t)dt. \tag{S.34}$$

Replacing $-\nabla V(x)$ with $f(x)$ yields

$$-2\int_0^T f(X_t)\cdot dX_t = \sigma^2\int_0^T \nabla\cdot f(X_t)dt + 2(V(y)-V(x_0)), \tag{S.35}$$

where we have used that $X_0 = x_0$ and $X_T = y$.

Hence, the path weight reduces to

$$\frac{d\mathbb{P}_f^B}{d\mathbb{P}_0^B}(X_{0:T}) \propto \exp\left[-\int_0^T U_B(X_t)dt\right], \tag{S.36}$$

with

$$U_B(x) = \frac{1}{2\sigma^2}\left(f^2(x) + \sigma^2\nabla\cdot f(x)\right), \tag{S.37}$$

where we have omitted the normalisation constant.

Hence, we can simulate constrained paths of an SDE with drift $f(x)$ and extreme terminal constraints following the forward filtering and backward smoothing approach described in the main article for settings with path constraints. The filtering equation is given by

$$\frac{\partial \varrho_t(x)}{\partial t} = \mathcal{L}_{f_0}\varrho_t(x) - U_B(x)\varrho_t(x), \tag{S.38}$$

where $U_B(x)$ is defined in Eq.(S.37). The drift f_0 is that of a d-dimensional Brownian bridge that is designed to reach the target state $x^* \in \mathcal{R}^d$ exactly after T time units and defined as

$$f_0(x,t) = \frac{x^* - x}{T-t}, \tag{S.39}$$

which corresponds to d-independent one–dimensional Brownian bridges, one for each dimensional component of the vector x.

IV. CONTROL FORMALISM FOR STATE DEPENDENT DIFFUSIONS

In the main article, for simplicity, we considered a *state independent* diffusion, and presented our formalism in terms of this simplified assumption. Here, we briefly outline the central results for settings where the noise term is *state dependent*, and thus has multiplicative noise effects, i.e. when the SDE admits the formulation (in the Ito sense)

$$dX_t = f(X_t,t)dt + \sigma(X_t,t)dW_t, \tag{S.40}$$

with $\sigma(x,t) \in \mathcal{R}^{n\times n} \times \mathcal{R}^+$.

We outline here the equations that require adaptation to consider the state dependent effects of the diffusion. The deterministic Fokker–Planck particle flow with a state dependent diffusion admits an additional $-\frac{1}{2}\nabla\cdot D(x,t)$ term

$$\frac{\partial \varrho_t(x)}{\partial t} = -\nabla\cdot\left[\left(f(x,t) - \frac{D(x,t)}{2}\nabla\ln\varrho_t(x) - \frac{1}{2}\nabla\cdot D(x,t)\right)\varrho_t(x)\right], \tag{S.41}$$

with $D(x,t) = \sigma(x,t)^\top\cdot\sigma(x,t)$ denoting the noise covariance matrix. Here we used the following notation for the gradient of the noise covariance

$$\left(\nabla\cdot D(x,t)\right)_i = \sum_j \frac{\partial D_{ij}(x,t)}{\partial x_j}. \tag{S.42}$$

The marginal constrained density $q_t(x)$ fulfills the Fokker–Planck equation

$$\frac{\partial q_t(x)}{\partial t} = \mathcal{L}_f q_t(x) - \nabla\cdot\left(D(x,t)\ln\varphi_t(x)q_t(x)\right), \tag{S.43}$$

with the Fokker–Planck operator (Eq.(2))

$$\mathcal{L}_f q_t(x) = \nabla\cdot\Big[-f(x,t)q_t(x) + \frac{1}{2}\nabla\cdot\left(D(x,t)q_t(x)\right)\Big]. \tag{S.44}$$

The time-reversed Fokker–Planck reads

$$\frac{\partial \tilde{q}_\tau(x)}{\partial \tau} = -\nabla \cdot \left[\left(D(x, T-\tau)\nabla \ln \varrho_{T-\tau}(x) + \nabla D(x, T-\tau) - f(x, T-\tau) \right) \tilde{q}_\tau(x) \right] + \frac{1}{2}\nabla^2 \cdot \Big(D(x, T-\tau)\tilde{q}_\tau(x) \Big). \quad \text{(S.45)}$$

This represents a diffusion process with drift

$$g(x,\tau) = D(x, T-\tau)\nabla \ln \varrho_{T-\tau}(x) + \nabla D(x, T-\tau) - f(x, T-\tau), \quad \text{(S.46)}$$

which is the drift of the reverse time process.

The resulting optimal controls are obtained from

$$u^*(x,t) = D(x,t) \cdot \left(\nabla \ln \tilde{q}_{T-t}(x) - \nabla \ln \varrho_t(x) \right). \quad \text{(S.47)}$$

V. DETERMINISTIC PARTICLE SOLUTION TO THE FILTERING EQUATION

For settings with path constraints $U(x,t) \neq 0$, the forward filtering flow $\varrho_t(x)$ (Eq.(11)) is distinct from a Fokker–Planck flow. In order to account for the extra sink term $-U(x,t)\varrho_t(x)$ in the forward PDE (Eq.(11)), we incorporate the (deterministic) *ensemble transform particle filter* [72] into the Fokker–Planck probability flow.

By formally integrating the filtering equation Eq.(11) at time t over a small time interval δt we obtain

$$\varrho_{t+\delta t}(x) = \exp\Big(\delta t(\mathcal{L}_f - U)\Big)\varrho_t(x) = \exp\Big(-U(x,t)\delta t\Big) \exp\left(\delta t \mathcal{L}_f\right)\varrho_t(x) + O((\delta t)^2) \quad \text{(S.48)}$$

in terms of operator exponentials. In the second equality, the non–commutativity of the two operators $\mathcal{L}_f$ and $U(x,t)$ has been neglected for small δt. This result can be understood as the concatenation of two processes: the propagation of a density described by the uncontrolled Fokker–Planck equation (Eq.(2)) followed by a multiplication of the resulting density by a factor $e^{-\delta t U(x,t)}$.

To simulate such a two stage process during the time interval δt, we first propagate the particles using the dynamics Eq.(22) at intermediate positions $Y^{(i)}$. For the second process, we assign *importance weights* $\Omega^{(i)}(t)$ to each particle i according to

$$\Omega^{(i)}(t) \propto e^{-\delta t U(Y^{(i)},t)}. \quad \text{(S.49)}$$

To obtain uniformly weighted particles required for the next integration step, we need to compute new particle positions that will represent the weighted particle distribution with a uniform weighted one, i.e with particle weights $\tilde{\Omega}^{(i)}(t) = \frac{1}{N}$.

Particle reweighting is commonly resolved with random resampling according to the particle weights to create a new *unweighted* particle ensemble [92]. Thus resampling is inherently stochastic. Here, instead, to maintain the deterministic nature of our framework, we employ the *deterministic ensemble transform particle filter* [72], that identifies the minimum *deterministic shift* of particle positions to transform the ensemble of weighted particles to an ensemble of uniformly weighted ones by maximising the covariance between the two particle ensembles. In particular, the algorithm is based on optimal transport problem with transport costs informed by the particles' spatial distribution.

The particle shifts are defined by a linear transformation mediated through the transform matrix $T^* \in \mathcal{R}^{N\times N}$

$$X^{(j)}_{t+\delta t} = N \sum_{i=1}^{N} T^*_{ij} Y^{(i)}, \quad \text{(S.50)}$$

where the optimal transport matrix T^* results from the solution of a linear program

$$T^* = \arg\min_T \sum_{i,j=1}^{N} T_{ij} \|Y^{(i)} - Y^{(j)}\|^2 \quad \text{(S.51)}$$

under the conditions

$$\sum_{i=1}^{N} T^*_{ij} = \frac{1}{N} \qquad \sum_{j=1}^{N} T^*_{ij} = \Omega^{(i)}. \quad \text{(S.52)}$$

In the limit of $N \to \infty$ expectations computed over the weighted and the shifted unweighted ensembles become equal. More details on this framework can be found in [72]. The cost to be minimised in Eq. (S.51) is the Wasserstein distance for discrete

distributions, that is alternatively known as the Earth Mover's Distance (EMD). For the numerical experiments performed in this paper, for computing the transport matrix we have employed the implementation of FastEMD [93] for the numerical experiment on molecular phenotypes and the implementation of Earth Mover's Distance provided by the python toolbox POT:Python Optimal Transport [82] for synchronising the Kuramoto networks. The latter implementation provides considerable speedup in the EMD computation compared to FastEMD, reducing the computational complexity from $\mathcal{O}(N^3 \log(N))$ to $\mathcal{O}(N^2)$. However it is slightly less accurate. Interestingly, experiments with entropy regularised solutions for the optimal transport problem introduced considerable contractions in the particle distributions resulting in unreliable representations of the underlying densities (even when the optimal transport problem was conditioned to provide a sparse solution).

VI. SCORE ESTIMATION FROM EMPIRICAL DENSITIES

Both the dynamics of the controlled system (Eq.(15-17)) and the propagation of particles according to Eq.(22) require knowledge of the logarithmic gradient of a density , $\nabla_x \ln \varrho_t(x)$, known as the *score function* of $\varrho_t(x)$ in Machine Learning [94]. Here, we resort to a kernel based non-parametric estimation of $\nabla_x \ln \varrho_t(x)$ inspired by the score matching objective introduced by Hyvärinen in [94]. The framework is described in detail by Maoutsa et. al in [54].

To obtain the logarithmic gradient of an unknown density $\varrho(x)$, we resort to a component-wise *variational* formulation of the problem identifying the unknown $\nabla_x \ln \varrho(x)$ by the function that minimizes functional $\mathcal{I}_\alpha[h, \varrho]$ for each dimensional component $\alpha \in [1, \dots, d]$

$$\partial_\alpha \ln \varrho(x) = \arg\min_h \ \mathcal{I}_\alpha[h, \varrho](x), \tag{S.53}$$

where the functional $\mathcal{I}_\alpha[h, \varrho]$ is

$$\mathcal{I}_\alpha[h, \varrho] = \int \varrho(x) \left(2\partial_\alpha h(x) + h^2(x)\right) dx.$$

This result may be verified using integration by parts, assuming that the density $\varrho(x)$ vanishes sufficiently fast at the boundaries.

Consider an empirical representation of the density $\varrho(x)$ in terms of samples $\{X^{(i)}\}_{i=1}^N$, we approximate the functional $\mathcal{I}$ by

$$\mathcal{I}_\alpha[h, \hat{\varrho}] = \frac{1}{N} \sum_{i=1}^{N} \left(h^2(X^{(i)}) + 2\partial_\alpha h(X^{(i)})\right).$$

For practical computations with a finite number of particles, to regularise the optimisation, we restrict the complexity of the function class h over which we minimise, by employing a *kernel approximation* (for more details, see [54]). More precisely, we assume that h belongs to a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{F}$ associated with a Gaussian (radial base function) kernel $K(\cdot, \cdot)$

$$K(x, x') = \exp\left[-\frac{1}{2l^2}\|x - x'\|^2\right] \tag{S.54}$$

with a length scale l. To regularise the empirical functional, we add a penalty term proportional to the RKHS norm of $\|h\|_{RKHS}$ to Eq.(S.54)

$$\partial_\alpha \ln \varrho(x) \approx \arg\min_{h \in \mathcal{F}} \left\{\mathcal{I}_\alpha[h, \hat{\varrho}] + \frac{\lambda}{N}\|h\|^2_{\text{RKHS}}\right\}(x). \tag{S.55}$$

To keep the computational complexity tractable for large particle number N, we employ a sparse kernel approximation assuming that h may be expressed in terms of a linear combination of the kernel evaluated at $M << N$ *inducing points* Z_i $h(x) = \sum_{i=1}^M a_i K(Z_i, x)$. These inducing points are selected randomly from a uniform distribution spanning the state space volume covered by the samples of $\varrho(x)$. Thus we obtain an empirical formulation of the score estimator

$$\partial_\alpha \ln \varrho(x) \approx \sum_{i=1}^{M} \left(\sum_{k=1}^{M} A_{ik}(x) \sum_{l=1}^{N} \left\{\nabla_{(X_l)_\alpha} K(X_l, Z_k)\right\}\right), \tag{S.56}$$

with $\nabla_{(x)_\alpha} K(x, Z_k) \doteq \frac{\partial}{\partial x_\alpha} K(x, Z_k)$ denoting the partial derivative of the first kernel argument across the α-th dimensional component, and with A_{ik} denoting the i-th row, and k-th column of the matrix valued function $A(x)$ defined as

$$A(x) \doteq K(x, \mathcal{Z}) \left[\lambda I + K_{zz}^{-1} K_{xz}^\top K_{xz}\right]^{-1} K_{zz}^{-1}, \tag{S.57}$$

where $\mathcal{X} = \{X_i\}_{i=1}^N$ and $\mathcal{Z} = \{Z_i\}_{i=1}^M$ denote the sets of samples and inducing points respectively, while $K_{zz} := K(\mathcal{Z}, \mathcal{Z})$ and $K_{xz} := K(\mathcal{X}, \mathcal{Z})$ denote the $M \times M$ and $N \times M$ matrices resulting from evaluation of the kernel at the inducing points, and at the samples and inducing points respectively. With I we denote the $M \times M$ identity matrix. Throughout our numerical experiments we set the regularising constant as $\lambda = 10^{-3}$. To ensure numerical stability we add a diagonal component of the order of 10^{-5} before the matrix inversions.

In the main text we denote the score estimator for the density $\hat{\varrho}(x)$ evaluated at point x across all dimensions by

$$\mathcal{S}(x, \hat{\varrho}) = \nabla \ln \hat{\varrho}(x) = [\partial_1 \ln \varrho(x), \partial_2 \ln \varrho(x), \ldots, \partial_d \ln \varrho(x)]^\top. \tag{S.58}$$

We remark here, that the score estimator of Eq. (S.56)- Eq. (S.57) was employed in [54] solely for estimation of the logarithmic gradient of a density *at* the samples $\mathcal{X}$. In the present work, for propagating the time-reversed flow $\tilde{q}_t(x)$ (Eq.17), we additionally require estimation of the logarithmic gradient of the forward flow $\nabla \ln \varrho_t(x)$ in the vicinity of the samples of $\varrho_t(x)$, but not *exactly* on them.

VII. PATH INTEGRAL CROSS ENTROPY CONTROL

Path integral control problems are a class of stochastic control problems for which the Hamilton-Jacobi-Bellman equation admits a linearisation through the introduction of the Hopf-Cole logarithmic transformation already presented in the main text, i.e. when the expected cost is transformed as $\mathcal{J}(x,t) = -\log(\varphi_t(x))$. As stated in the main text, to linearise the nonlinear Hamilton-Jacobi-Bellman equation through this transformation, the external interventions should be additive in the dynamics, appear in a quadratic form in the cost function ($\|u(x,t)\|^2$), while the control costs should be inversely proportional to the diffusion matrix (in the cases where a scalar diffusion does not suffice to capture the problem at hand). No further restrictions apply, and thus the system dynamics and control costs are allowed to have arbitrary nonlinear relationship to the state of the system.

An information theoretic approach to compute the optimal interventions for this subclass of problems is the Path Integral Cross Entropy (PICE) method, recently introduced in [46]. PICE is an iterative method, that approximates the optimal interventions $u^*(x,t)$ within a parametric family of functions by minimising the relative entropy between the path measures $\mathbb{Q}^*$ and $\mathbb{P}_{\hat{u}}$ induced by optimally controlled process with control $u^*(x,t)$ and a process with provisional control $\hat{u}(x,t)$ which we want to optimise within the selected parametric family of functions. The minimisation follows an iterative procedure, where simulated paths of the system with some provisional control $u(x,t)$ are employed to compute the path integral required for minimising the relative entropy between the two path measures. Since the optimal path measure $\mathbb{Q}^*$ is for nontrivial settings unknown, the method employs importance sampling to compute expectations over the optimal path measure $\mathbb{Q}^*$.

Disclaimer: For uncluttered notation, in contrast to the rest of the article, in this section the subscripts of path measures refer to the control term in the drift and not to the entire drift function, i.e. $\mathbb{P}_u$ refers to the path measure induced by a process with drift $f(x) + \sigma u(x,t)$.

More precisely, we denote by $\mathbb{P}_{\hat{u}}$ the path measure induced by employing the provisional control $\hat{u}(x,t)$ into the system dynamics, and $\mathbb{Q}^*$ the path measure induced by optimally controlled system. The optimal control is obtained as the minimiser of the relative entropy (KL divergence) between the optimally and nearly optimally controlled path measures , and may be expressed as

$$\begin{aligned}
u^*(x,t) = \arg\min_{\hat{u}_{t:T}} KL(\mathbb{Q}^* || \mathbb{P}_{\hat{u}}) &= \arg\min_{\hat{u}_{t:T}} \int \log \frac{d\mathbb{Q}^*}{d\mathbb{P}_{\hat{u}}} d\mathbb{Q}^* = \arg\min_{\hat{u}_{t:T}} \int \log \left(\frac{d\mathbb{Q}^*}{d\mathbb{P}_0} \frac{d\mathbb{P}_0}{d\mathbb{P}_{\hat{u}}} \right) d\mathbb{Q}^* \\
&= \arg\min_{\hat{u}_{t:T}} \Big\langle -\log \left(\frac{d\mathbb{P}_{\hat{u}}}{d\mathbb{P}_0} \right) \Big\rangle_{\mathbb{Q}^*} \\
&= \arg\min_{\hat{u}_{t:T}} \Big\langle -\log \left(\frac{d\mathbb{P}_{\hat{u}}}{d\mathbb{P}_0} \right) \frac{d\mathbb{Q}^*}{d\mathbb{P}_u} \Big\rangle_{\mathbb{P}_u} \\
&= \arg\min_{\hat{u}_{t:T}} \Big\langle \int_t^T ds \left(\frac{1}{2} \|\hat{u}(x,s)\|^2 - \frac{\hat{u}(x,s)}{\sigma^2} \left(\frac{dX_s}{ds} - f(X_s, s) \right) \right) \frac{d\mathbb{Q}^*}{d\mathbb{P}_u} \Big\rangle_{\mathbb{P}_u} \\
&= \arg\min_{\hat{u}_{t:T}} \frac{1}{\psi_t(x)} \Big\langle e^{-\mathcal{S}(X_t,t,u)} \int_t^T ds \left(\frac{1}{2} \|\hat{u}(x,s)\|^2 - \hat{u}(x,s) \left(u(x,s) + \frac{dW_s}{ds} \right) \right) \Big\rangle_{\mathbb{P}_u}.
\end{aligned} \tag{S.59}$$

In Eq. (S.59) in the first line, we wrote the relative entropy between the path measures $\mathbb{Q}^*$ and $\mathbb{P}_{\hat{u}}$ in terms of the Radon-Nikodym derivative of $\mathbb{Q}^*$ with respect to $\mathbb{P}_{\hat{u}}$, and then, in the second equality, rewrote $\frac{d\mathbb{Q}^*}{d\mathbb{P}_{\hat{u}}}$ in terms of Radon-Nikodym derivatives of $\mathbb{Q}^*$ and $\mathbb{P}_{\hat{u}}$ with respect to the path measure $\mathbb{P}_0$ induced by the uncontrolled process. In the second line we dropped the term $\int \log \frac{d\mathbb{Q}^*}{d\mathbb{P}_0} d\mathbb{Q}^*$, since this is unrelated to the optimisation with respect to $\hat{u}_{t:T}$. For the third line, we changed

the expectation over $\mathbb{Q}^*$ to an expectation over the path measure $\mathbb{P}_u$ induced by a diffusion process with drift $f(x,t)+\sigma u(x,t)$, by multiplying with the relevant Radon-Nikodym derivative $\frac{d\mathbb{Q}^*}{d\mathbb{P}_u}$. For the fourth line, we substituted the Radon-Nikodym derivative

$$\frac{d\mathbb{P}_{\hat{u}}}{d\mathbb{P}_0} = \exp\left\{-\frac{1}{2}\int_t^T \|\hat{u}(x,s)\|^2 ds + \int_t^T \hat{u}(x,t)\left(dX_s - f(X_s,s)\right) ds\right\}, \tag{S.60}$$

while for the final line we expressed the Radon-Nikodym derivative $\frac{d\mathbb{Q}^*}{d\mathbb{P}_u}$ with (see also [46] for details)

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}_u} = \frac{\exp\left(-\mathcal{S}(t,x,u)\right)}{\langle e^{-\mathcal{S}(t,x,u)}\rangle_{\mathbb{P}_u}}, \tag{S.61}$$

where the $\mathcal{S}$ denoting the cost of introducing the control $u(x,t)$ into the free dynamics is given by

$$\mathcal{S}(t,x,u) = -\log\chi(X_T) + \int_t^T ds\left(U(X_s,s) + \frac{1}{2\sigma^2}u(X_s,s)^2\right) + \int_t^T u(X_s,s)dW_s. \tag{S.62}$$

Moreover $\psi_t(x)$ in the last line of Eq. (S.59) can be identified as $\psi_t(x) = \langle e^{-\mathcal{S}(t,x,u)}\rangle_{\mathbb{P}_u}$.

Expanding both $\hat{u}(x,t)$ and $u(x,t)$ on a set of K basis functions $\{h_k(x,t)\}_{k=1}^K$, with $k = 1,\ldots,K$ we obtain

$$u(x,t) = \sum_{k=1}^K a_k(t)h_k(x,t) = A^t \cdot h(x,t),$$

and

$$\hat{u}(x,t) = \sum_{k=1}^K \hat{a}_k(t)h_k(x,t) = \hat{A}^t \cdot h(x,t),$$

with $A, \hat{A} \in \mathbb{R}^{K\times(T/dt)}$. Here, we introduced the notation A^t to denote the set of basis coefficients corresponding to time t, i.e. $\hat{A}^t \doteq [\hat{a}_1(t), \hat{a}_2(t), \ldots, \hat{a}_k(t)]$.

Introducing now the explicit dependency of the controls on the parameters $\hat{A}$, $\hat{u}(x,t|\hat{A}) = \sum_{k=1}^K a_k(t)h_k(x,t)$, we compute the gradient of the relative entropy with respect to the parameters $\hat{A}$ as

$$\frac{\partial KL(\mathbb{Q}^*||\mathbb{P}_{\hat{u}})}{\partial \hat{A}} = \left\langle\left\langle \int_t^T \left(\hat{u}(X_s,s)ds - u(X_s,s)ds - dW_s\right)\frac{\partial \hat{u}(X_s,s)}{\partial \hat{A}}\right\rangle\right\rangle_{\mathbb{P}_u}, \tag{S.63}$$

where the double angular brackets denote a weighted expectation $\langle\langle\mathcal{F}\rangle\rangle_{\mathbb{P}_u} = \langle\frac{e^{-\mathcal{S}(t,x,u)}}{\psi_t(x)}\mathcal{F}(X_{t:T})\rangle_{\mathbb{P}_u}$. Eq.(S.63) computes the gradient of the $KL(\mathbb{Q}^*||\mathbb{P}_{\hat{u}})$ with the help of the importance sampling density P_u. In an iterative optimisation scheme for identifying the parameters $\hat{A}$ for the optimal control $\hat{u}(x,t)$, it is sensible to use as importance sampling density the path density $\mathbb{P}_{\hat{u}}$ induced by the best estimate of the control we have in our disposal. The best estimate in our disposal is the estimate of $\hat{u}(x,t)$ of the current iteration. Thus in Eq. (S.63) we replace the arbitrary control $u(x,t)$ with $\hat{u}(x,t)$ and obtain

$$\frac{\partial KL(\mathbb{Q}^*||\mathbb{P}_{\hat{u}})}{\partial \hat{A}} = -\langle\langle \int_t^T dW_s \frac{\partial \hat{u}(X_s,s)}{\partial \hat{A}}\rangle\rangle_{\mathbb{P}_{\hat{u}}}. \tag{S.64}$$

Thus, for a learning rate $\eta > 0$, the gradient descent parameter update for step $n \in [1,2,\ldots,r]$ reads

$$\hat{A}^t_{n+1} = \hat{A}^t_n + \eta\left\langle\left\langle \frac{dW_s}{ds}\frac{\partial \hat{u}(X_s,s)}{\partial \hat{A}^t}\right\rangle\right\rangle_{\mathbb{P}_{\hat{u}}}, \tag{S.65}$$

where we have dropped the time integral in order to optimise for each time step independently.

Thus starting from a set of initial coefficients $\hat{A}_0 \in \mathbb{R}^{K\times(T/dt)}$, denoting by n the index of the current iteration, we iterate over the following computations:

- Draw the noise realisations for N trajectories and T/dt timesteps: $\zeta = \mathcal{N}(0, \sqrt{(dt)})$.

- Simulate and store N stochastic trajectories following the controlled dynamics with the current estimation of the control for each time step, i.e. the time discretised version of

$$dX_t = \big(f(X_t, t) + A_n^t \cdot h(X_t, t)\big)\, dt + \sigma dW_t.$$

- Compute the action $\mathcal{S}$ according to Eq. (S.62), and the weight $w_n = e^{-\mathcal{S}}$ for each trajectory individually, and normalise the weights $w_n^{(i)} := \frac{w_n^{(i)}}{\sum_{i=1}^N w_n^{(i)}}$.

- The update of the basis function coefficients for each time step is given by

$$A_{n+1}^t = A_n^t + \eta \frac{dQ(h_t)}{dt} B^{-1}(t), \tag{S.66}$$

where

$$dQ(h_t) = \sum_{i=1}^{N} w_n^{(i)}\, \zeta^{(i)}(t) \otimes h_t(X_t^{(i)}, t),$$

and

$$B(t) = \sum_{i=1}^{N} w_n^{(i)}\, h_t(X_t^{(i)}, t) \otimes h_t(X_t^{(i)}, t).$$

Here, $\otimes$ denotes the outer product. In the numerical experiments we employ a learning rate $\eta = 0.05$ for the presenting settings with only terminal constraints, and rate $\eta = 0.025$ in the settings where both terminal and path constraints are considered. In the latter case, employing a larger learning rate, results in numerical instabilities especially for small number N of simulated trajectories.

In our comparisons we employ a slightly modified version of PICE to render it comparable to our framework. More precisely, we found that when the typical trajectories of the uncontrolled dynamics differ considerably from the trajectories required to meet the terminal and path constraints, the PICE method converges to some local minimum and fails to fulfill the required terminal constrains sufficiently. On the other hand, our method is considerably more precise in reaching the terminal targets, since the time reversed probability flow is initiated from that point. Thus, to obtain controlled trajectories with PICE comparable to our approach, where terminal constraints are met precisely, we introduced weighting coefficients in front of the constraints in the action of the path integral. In particular, we use the modified action

$$\mathcal{S} = \beta_0 \cdot \log \chi(X_T) + \int_t^T ds \left(\beta_1 U(X_s, s) + \frac{1}{2\sigma^2} u(X_s, s)^2 \right) + \int_t^T u(X_s, s) dW_s, \tag{S.67}$$

where $\beta_0 = diag(\beta_0^{(1)}, \ldots, \beta_0^{(d)})$, with $\beta_0^{(i)} \in \mathbb{R}$ for $i = 1, \ldots, d$, and $\beta_1 \in [0,\, 1]$.

We found that by progressively increasing the values of these coefficients when the optimisation gets stale, while the constraints are not yet sufficiently fulfilled, drives the trajectory ensemble towards the intended regions of the state space that satisfy the constraints. We note here, that by initiating the optimisation with already large values for the coefficients β_0 and β_1 leads very fast to numerical instabilities. Therefore we proceeded to progressively increase their values, when otherwise the optimisation does not result in improvements in terms of fulfilling the imposed constraints.

We denote here, that in our framework, although we can set the coefficient for the path constraint β_1, the coefficients β_0 are not explicitly defined, since for propagating the forward and time reversed flows, we initiate the simulations from delta functions centered at the initial and terminal points. Since our approach results in particularly precise reaching of terminal target states, we can consider this equivalent to having large β_0 coefficients. Small β_0 coefficients could be implemented, if necessary for the problem at hand, by initiating the simulation of the time reversed flow from a small Gaussian distribution centered around the terminal state.

For the computations without path constraints, we initiate the simulations with $\beta_0 = \mathcal{I}_d$. At each time step we track the mean terminal state of the simulated trajectories. We terminate the optimisation when the average distance (over all trajectories) to the target x^* across each dimension is smaller than $5 \cdot 10^{-3}$, i.e. $||\mathcal{D}(X_T)||_\infty = ||\langle x^* - X_T \rangle_N||_\infty < 5 \cdot 10^{-3}$, where $||\cdot||_\infty$ denotes the maximum norm.

We consider the optimisation frozen across dimension d_i, if within 50 consecutive iterations the magnitude of first derivative of average distance to the target across dimension d_i, $\nabla_n^{(d_i)} \mathcal{D}(X_T)$ falls below a threshold $\theta_0 = 10^{-3}$, where n indicates that the derivatives are computed across iterations. In that case we double the current value of the relevant weighting coefficient, $\beta_0^{(d)} := 2 \cdot \beta_0^{(d)}$.

VIII. EVOLUTIONARY FEEDBACK CONTROL

We demonstrated the performance of our framework in the context of evolutionary feedback control of molecular phenotypes. Here we provide detailed information and a brief introduction on the formalism.

The prevalence of phenotypic traits in a population of individuals can be construed as a diffusion process evolving on a topographic (or 'adaptive') landscape that characterises the fitness of a population, i.e their average reproductive success. Following the formalism introduced by Lande [78], we consider a population consisting of n individuals, where each individual i is characterised by a vector of d phenotypic traits $\boldsymbol{p}^{(i)}(t) = \left(p_1^{(i)}(t), p_2^{(i)}(t), \ldots, p_d^{(i)}(t)\right)$. These traits are quantitative characteristics describing each member of the population, and have time dependent values with time scales in the range of geneological time scales. Each time step of the evolutionary process represents one step forward in the geneology. The *mean phenotype* of the population is a vector with the average of phenotypic traits among the individuals comprising the population $\mathbf{x}(t) = (\bar{p}_1(t), \bar{p}_2(t), \ldots, \bar{p}_d(t))$. We denote the *fitness* of a population with mean phenotype $\bar{\mathbf{x}}(t)$ by a smooth real valued function $F : \mathcal{R}^d \to \mathcal{R}$ that captures the average reproductive success in the population. This function can be considered as a potential function. The peak of the potential landscape identifies the evolutionary optimum of the population for a certain environment in the presence of natural selection.

Considering that natural selection drives the population towards phenotypic traits with higher fitness, the trajectory of the mean phenotype thus follows the gradient of the landscape *uphill* towards the nearest local maximum. In that sense, the deterministic part of the stochastic process that describes the evolution of phenotypic traits is given by landscape gradient $\nabla F(x)$ that denotes the selection pressure from the environment. The finite size of the population causes random genetic drift that is modeled as white Gaussian noise with amplitude that scales inversely proportional to the effective population size.

Here, we are interested in devising artificial selection protocols for molecular phenotypes that optimally drive evolution to desired phenotypic states adhering to the adaptive pressures as dictated by the landscape gradient. We build on the formalism recently introduced by Nourmohammad et al.[29].

For Gaussian distributed phenotypes, fitness and mutation forces drive the evolution of mean phenotypes induced by the fitness potential, while genetic drift perturbs the process stochastically. We describe the evolution of *mean phenotypes* $\mathbf{x}$ by the overdamped Langevin equation

$$d\mathbf{x} = C \cdot \nabla F(\mathbf{x})dt + \Sigma d\mathbf{W}, \tag{S.68}$$

with $F(\mathbf{x})$ denoting the adaptive landscape capturing the effects of natural selection and mutation forces, while Σ stands for the noise amplitude. The gradient of the landscape $f(\mathbf{x}) = C \cdot \nabla F(\mathbf{x})$ captures the adaptive pressures during evolution. The covariance matrix C captures the covariance among different traits, and influences both the deterministic and stochastic part of the forces acting on the system. The genetic drift, i.e. the noise amplitude of the process is determined by the square root of the covariance matrix rescaled by the inverse of the effective population size, i.e. $\Sigma = C^{1/2}n^{-1}$.

In general, the covariance matrix C follows a similar Langevin equation [29] that can be nevertheless be considered as remaining constant over time, given the much smaller timescales its fluctuations unfold [96], and the weak dependency on the evolutionary selection strength [97].

The dynamics of Eq.(S.68) describe the evolution of populations in the presence of natural selection towards an evolutionary optimum, captured by the minimum of the adaptive landscape. Yet, in order to study and understand the outcomes and dynamics of adaptive evolution, we desire to devise intervention protocols that will drive the evolutionary process towards an non-evolutionary optimum state $\mathbf{x}^*$ of the adaptive landscape. This intervention is implemented through artificial selection, which we formulate here as a time- and state- dependent perturbation $\mathbf{u}(\mathbf{x}, t)$ to the natural selection dynamics

$$d\mathbf{x} = \Big(f(\mathbf{x}) + \mathbf{u}(\mathbf{x}, t)\Big)dt + \Sigma d\mathbf{W}. \tag{S.69}$$

This defines an optimal control problem, where we seek to identify the optimal intervention protocol $\mathbf{u}(\mathbf{x}, t)$ in order to attain a target state $\mathbf{x}^*$ at a predefined deterministic time T, minimising thereby the cost function of Eq. (**??**) over the entire trajectory. To design an optimal artificial selection protocol as defined in Eq.(**??**), we employ the presented sampling scheme in terms of simulated probability flows.

For the numerical experiments we consider a two-dimensional evolutionary process with covarying phenotypes, and apply controllers that either assume the correct cross-phenotype correlations (**'correct'** controller), or consider the two phenotypes uncorrelated (**'misspecified'** controller). More precisely, we consider an evolutionary process with dynamics

$$d\mathbf{x} = C \cdot \nabla F(\mathbf{x})dt + \Sigma d\mathbf{W}, \tag{S.70}$$

with covariance matrix $C = \begin{bmatrix} c_{1,1} & \varrho_{xy}\sigma_1\sigma_2 \\ \varrho_{xy}\sigma_1\sigma_2 & c_{2,2}, \end{bmatrix}$ with $\sigma_s = \sqrt{c_{s,s}}$ for $s \in \{1, 2\}$, and $c_{1,1} = 0.2$, $c_{2,2} = 0.5$. The landscape gradient we employed was given by $\nabla F(\mathbf{x}) = -L \cdot \mathbf{x}$, with the matrix L indicating the selection coefficient across each

dimension being set to $L = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ while the diffusion matrix was set to the square root of the covariance $\Sigma = C^{1/2}$. For our experiments we employed 6 different correlation levels, namely $\varrho_{xy} = \pm\{0.25, 0.50, 0.75\}$. While this particular setting is analytically solvable, the computation of controls with our method generalises also for less trivial settings.

To design the controllers for each cross-phenotype correlation level ϱ_{xy}, we sampled the forward and time-reversed flows with the correctly correlated dynamics ('correct' controller), and with uncorrelated dynamics, i.e. with $\hat{\varrho}_{xy} = 0$ set to zero ('misspecified' controller). We applied both controllers to steer a process with the predefined correlation value ϱ_{xy}. Notice, that by modifying the misspecification of the correlation, we alter both the drift and the diffusion term of the process assumed by the controller, since both terms depend on the covariance matrix C.

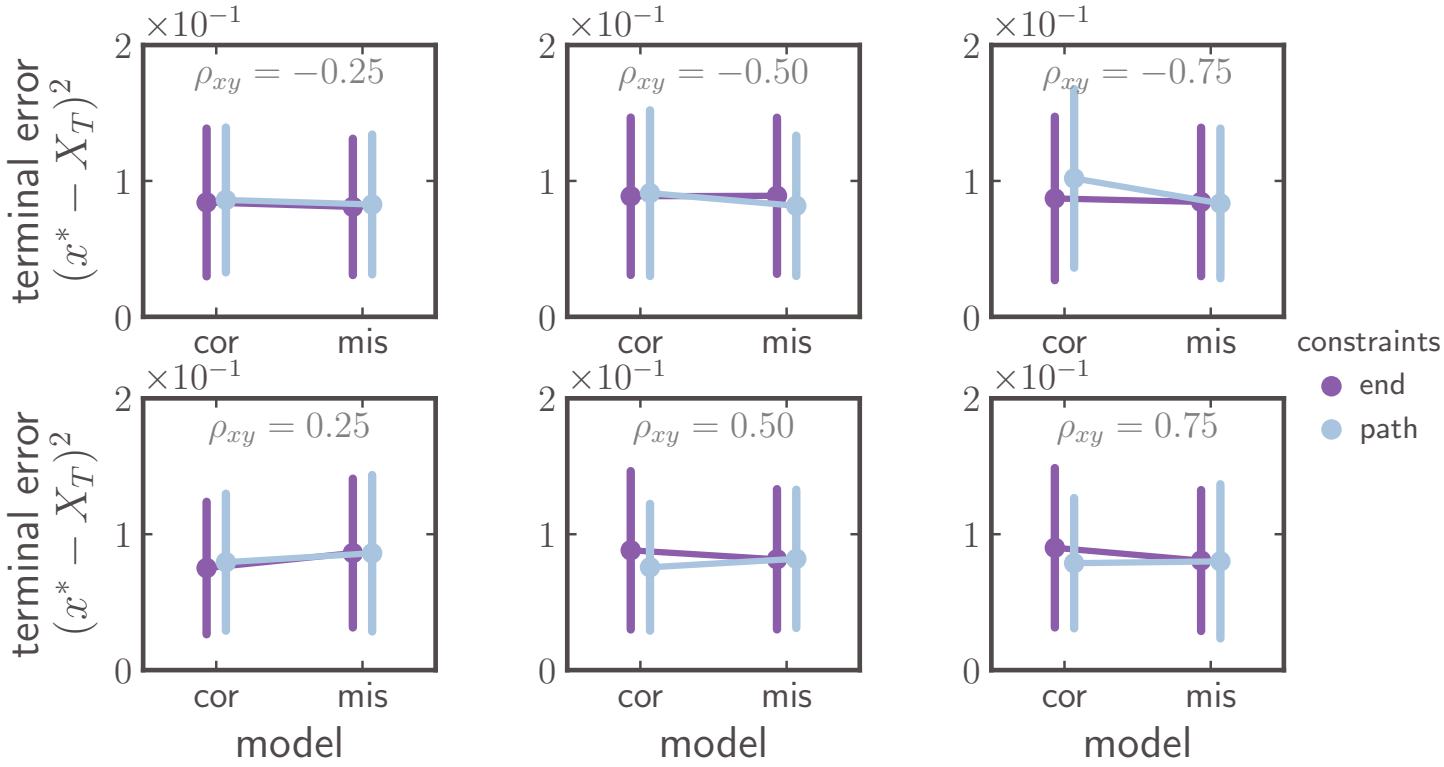

FIG. S-1. **Precision of the controller for misspecified evolutionary models.** Terminal error of interventions for covarying phenotypes with cross phenotype correlation $\varrho_{xy} = \pm\{0.25, 0.50, 0.75\}$, when computed controls assumed the correct ('*cor*'), or a misspecified ('*mis*') dynamical model that neglects the correlations, i.e. $\varrho_{xy} = 0$. Purple indicates only terminally constrained processes, while light blue stands for systems with both terminal and path constraints.

IX. SYNCHRONISATION OF KURAMOTO OSCILLATORS

The Kuramoto model [100] has been extensively used in the literature to study synchronisation phenomena in many disciplines, including biology, physics, and engineering. In its original formulation, it comprises a population of $K \geq 1$ phase oscillators, with dynamics evolving according to the following system of nonlinear coupled ordinary differential equations

$$\dot{\theta}_i(t) = \omega_i + \frac{J}{K} \sum_{j=1}^{K} \sin\left(\theta_j(t) - \theta_i(t)\right), \quad i = 1, \ldots, K \tag{S.71}$$

with initial conditions $\theta_i(0) = \theta_i^0$. In Eq.(S.71) θ_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, $g(w) = \mathcal{N}(0, w^2)$. Here, we adopt an all-to-all coupling topology where each oscillator interacts with all the other oscillators with uniform coupling strength J.

In the numerical experiments we consider noisy phase oscillators, thus we extend the dynamics of Eq.(S.71) by perturbing each oscillator with additive white Gaussian noise independent for each oscillator [98]. The noise amplitude σ is set to be uniform among all oscillators, resulting in

$$d\theta_t^{(i)} = \left(\omega_i + \frac{J}{K} \sum_{j=1}^{K} \sin\left(\theta_t^{(j)} - \theta_t^{(i)}\right) \right) dt + \sigma dW_t^{(i)}, \tag{S.72}$$

with initial conditions $\theta_0^{(i)}$, and $\{W^{(i)}\}_{i=1}^{K}$ K independent Wiener processes.

We denote here that an alternative formulation of the Kuramoto model perturbed by noise entails multiplicative noise involving the oscillator phase response curve $Z(\theta) = \sin(\theta)$ [99] as a prefactor of the noise term. This formulation does not pose an issue for our framework, since, as mentioned in Section IV, our method is applicable also for systems with multiplicative noise.

For systems of Kuramoto oscillators with heterogeneous dynamics, e.g. with differing natural frequencies akin to the one we employed here, we employ the notion of synchronisation that requires phase alignment between oscillation pairs (instead of complete phase agreement)

$$\lim_{t \to +\infty} |d\theta_t^{(i)} - d\theta_t^{(j)}| = 0 \tag{S.73}$$

We characterise the level of synchronisation among the oscillators with the Kuramoto order parameter for phase coherence

$$R(\boldsymbol{\theta}_t, t) = \frac{1}{K} \left| \sum_{j=1}^{K} e^{i\theta_t^{(j)}} \right|, \tag{S.74}$$

where i denotes the imaginary unit. In the thermodynamic limit $K \to \infty$, the order parameter converges to values between 0 - in settings where the oscillators' phases are entirely incoherent - and 1 when all the oscillators synchronise and rotate in unison. In finite size networks, incoherent dynamics are indicated by a time averaged order parameter of the order of the inverse square root of network size, i.e. $\bar{R} = \langle R(\boldsymbol{\theta}_t, t) \rangle_t = \mathcal{O}(1/\sqrt{K})$.

In the absence of control and when the coupling between the oscillators exceeds a critical strength J_c, the phases of the oscillators synchronise for $t \to \infty$. Below that critical coupling strength, the oscillator ensemble is in incoherent or partially synchronised state. Here, we provide continuous time state control to induce synchronisation to oscillator networks.

In a similar vein with the systems treated previously, we introduce an external intervening force for each oscillator

$$d\theta_t^{(i)} = \left(\omega_i + \frac{J}{K} \sum_{j=1}^{K} \sin\left(\theta_t^{(j)} - \theta_t^{(i)}\right) + u_i(\boldsymbol{\theta}_t, t) \right) dt + \sigma dW_t^{(t)}. \tag{S.75}$$

We formulate the synchronisation constraint as a path constraint. More precisely, we consider as path constraint the following term that considers the collective state of the network by rewarding paths that result in order parameter R closer to 1 (complete synchrony)

$$U(\boldsymbol{\theta}_t) = \beta \left(1 - R(\boldsymbol{\theta}_t, t)\right) dt, \tag{S.76}$$

with $\beta = 10^3$.

We considered the problem as a finite horizon optimisation problem of time T without terminal constraint, but with path constraint $U(x, t)$. More precisely, we initially computed the forward flow $\varrho_t(x)$ within the interval $[0, T]$ by employing at every time step the necessary reweighting that accounts for the path constraint. Since we did not define a terminal target for time T, we set for the probability flow $\tilde{q}_0(x) = \varrho_T(x)$ i.e. we match the terminal state of $\varrho_t(x)$ with the initial condition of the time reversed probability flow $\tilde{q}_t(x)$, and perform the time reversed sampling. As in the previous settings we applied the resulting interventions obtained from the difference of the logarithmic gradients of the two simulated flows for controlling 20 different realisations of the same network.

In the main article we considered synchronisation of networks where the natural frequency distribution $g(\omega)$ is a zero-centered Gaussian. To demonstrate that our approach is not limited to Gaussian natural frequency distributions we additionally considered networks with frequency distribution comprising two delta functions centered at $-\omega^0$ and ω^0 for $\omega^0 \in \{0.25, 0.5, 1\}$, i.e., $g(\omega) = \frac{1}{2}\left(\delta(\omega - \omega^0) + \delta(\omega + \omega^0)\right)$ (Fig. S-2) with initial phases drawn from Gaussian distributions of two different variances ($\theta_0 \sim \{\mathcal{N}(\pi, 0.5^2), \mathcal{N}(\pi, 1)\}$). The remaining network and simulation characteristics were similar to numerical experiments of Fig. 9 in the main article with noise amplitude $\sigma = 1$. For all examined frequency distributions the controlled networks exhibited larger phase coherence order parameter values than their uncontrolled counterparts. The magnitude of the natural

frequency ω^0 employed in the network did not considerably influence the resulting phase coherence of the oscillators. However, the initial dispersion of the oscillator phases $\boldsymbol{\theta}_0$ had a stronger influence in the amount of synchronisation achieved within the predefined time interval T. Networks initiated from more dispersed initial conditions (lower row in Fig.. S-2) achieved only partial synchronisation within the time interval T both for the uncontrolled and controlled networks and for all frequency and coupling conditions. Nevertheless, the networks controlled with **DPF** showed consistently larger order parameter values than their uncontrolled counterparts. A more detailed study of the performance of our control framework to synchronise (larger) networks with connections to mean-field solutions will be discussed in a future publication.

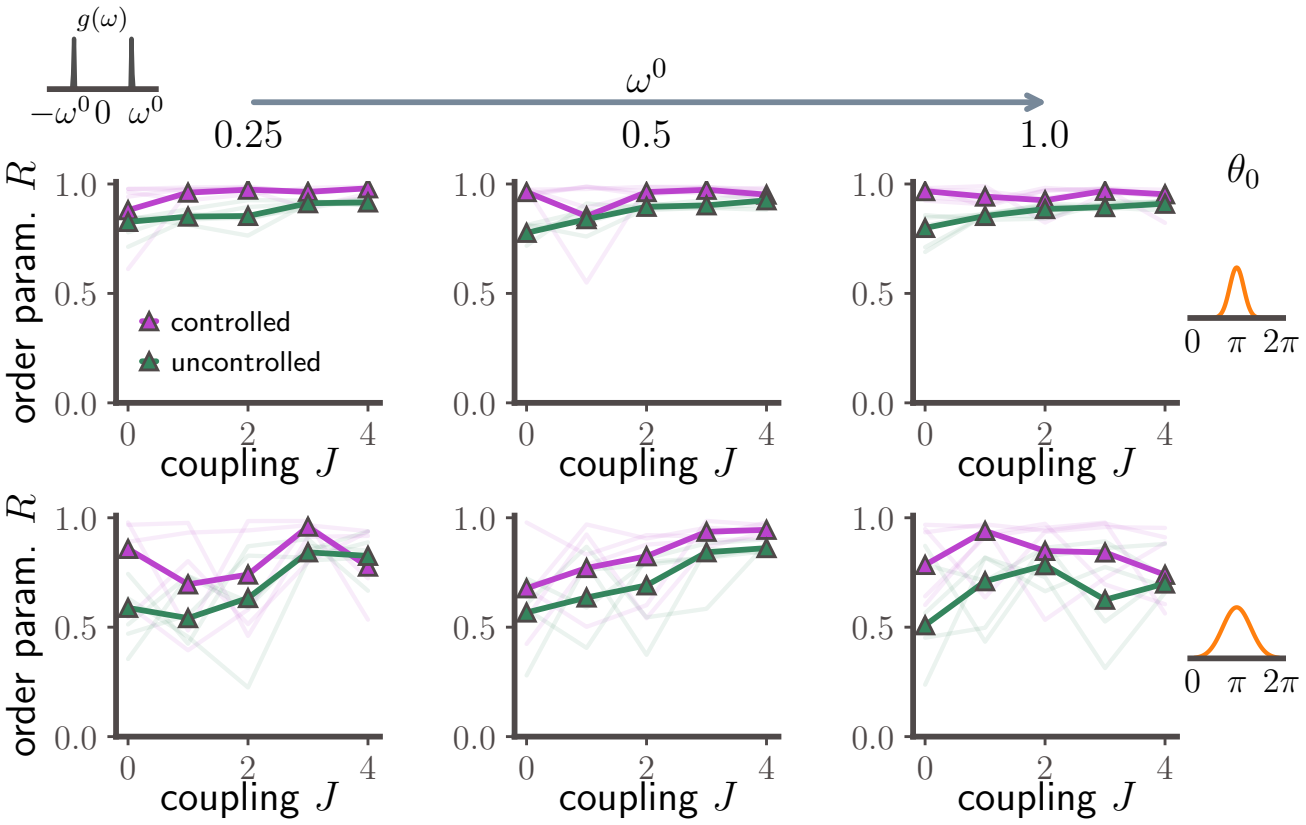

FIG. S-2. **Synchronisation control of finite-size networks of $K = 6$ oscillators with bimodal natural frequency distribution.** Time averaged phase coherence order parameter (R) of controlled (*purple*) and uncontrolled (*green*) networks with natural frequency distributions drawn from a bimodal distribution $g(\omega)$ comprising two delta functions centered at $+\omega^0$ and $-\omega^0$ with $\omega^0 \in \{0.25, 0.5, 1\}$, and initial phases θ_0 drawn from Gaussian distributions of different variance: (*upper*) $\mathcal{N}(\pi, 0.5^2)$, (*lower*) $\mathcal{N}(\pi, 1)$. (marker denote average over 5 different control computations with 20 different controlled trajectories within each realisation, and thin lines averages over 20 realisations of each control computation.) [a]

[a] Further parameters: noise amplitude $\sigma = 1$, particle number $N = 3000$, inducing point number $M = 300$, discretisation time step $dt = 10^{-3}$, temporary finite horizon $T = 0.5$.

An alternative approach for applying **DPF** for synchronisation control, is to treat the problem as an infinite horizon optimisation problem, and thus optimise the controls 'online', i.e. within finite small sub-intervals. This approach can be performed with substantially smaller number of particles compared to the scheme presented in the main text. Nevertheless, to remain realistic it requires rapid computations of controls, since the control calculations should happen in parallel with the simulation.

More precisely, the 'online' scheme alternates estimation of the constrained densities in sub-intervals $T_{sub} < T$ starting from the current state of the system, and evolution of the controlled stochastic trajectory within this same time interval T_{sub}. Thus, for this alternative computation of controls, we computed the controlled density for the time interval $[0,\ T_{sub}]$, and subsequently propagated the controlled stochastic system within that interval until the state $x_{T_{sub}}$. Subsequently, the estimation of the controlled density followed from the initial condition $x_{T_{sub}}$ and run over the interval $[T_{sub},\ 2\,T_{sub} - 1]$.

To accommodate this control setting with our framework, we optimised the state interventions $u(\boldsymbol{\theta}_t, t)$ within sub-intervals of length $T_{sub} = 50\,dt$, where $dt = 10^{-3}$ denotes the discretisation time-step, and employed a path constraint involving the order parameter $U(\boldsymbol{\theta}_t) = \beta(1 - R(\boldsymbol{\theta}_t, t))$.

For numerical experiments with the six oscillator network in Figure S-3 we employed natural frequencies from a Gaussian $\mathcal{N}(0, 1.)$, where we rejected the samplings with unbalanced natural frequency signs, i.e. we accepted only natural frequency sampling where 3 oscillators had positive, and 3 oscillators had negative frequency. The T_{sub} was set at 0.05.

According to Figure S-3, the online version of our method has a paradoxical impact on the effect of noise on the critical coupling for synchronisation (Figure S-3d.). Whereas in the uncontrolled setting, the critical coupling increases for increasing noise amplitude, employing our control framework seems to render the networks with stronger noise more easily synchronisable through stronger feedback control. However further research in necessary for pinpointing the exact relationship between the critical coupling and noise strength in the presence of control.

To identify the effect of noise and coupling strength on the synchronisability of the network when controlled by our method, we simulated different instances of the same network with different coupling strengths under two noise conditions $\sigma = \{0.1, 0.5\}$, and observed both the controlled and the uncontrolled systems. As expected [98], the uncontrolled system synchronises in the absence of external interventions for strong coupling, whereas the noise strength has a hindering effect on synchronisation, pushing the critical for synchronisation coupling strength to slightly larger values. In the presence of external interventions mediated through our framework, the network synchronised already for weak couplings for both noise conditions.

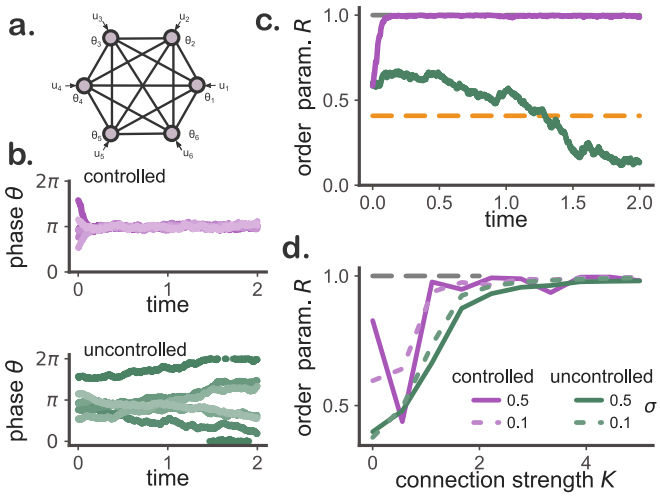

FIG. S-3. **Synchronisation control of a finite-size network of $N = 6$ interacting Kuramoto phase oscillators controlled over sub-intervals. a.)** Schematic of oscillator network depicting state variables (phases) θ_i at each network node, controls u_i, and network interactions (edges). **b.) (upper)** The phases of the oscillators quickly synchronise when controlled by the deterministic particle flow control method and remain synchronised throughout the entire simulation interval $[0,\ T = 2]$. **b.) (lower)** The same network with same connection strength ($K = 1.6$) fails to synchronise in the absence of control. **c.)** Evolution of Kuramoto order parameter R for the controlled (*purple*) oscillator network indicates a rapid transition to synchrony (R=1), while for the identical uncontrolled network the order parameter fluctuates strongly, indicating an phase incoherence (*green*). The orange line denotes the expected long time average value of the order parameter if the oscillators were non-interacting considering finite size scaling effects. For visual clarity, the grey line marks the level of $R = 1$ indicating a completely synchronous state. **d.)** Time averaged order parameter against coupling strength K for different noise amplitudes $\sigma = 0.1$ (*dashed*) and $\sigma = 0.5$ (*solid*) for a controlled (*purple*) and an uncontrolled (*purple*) network. Our control framework effectively shifts the critical coupling strength for synchronisation to smaller values. (Lines denote average over 3 different network realisations.) [a]

[a] Further parameters: noise amplitude $\sigma = \{0.25, 0.5\}$, particle number $N = 500$, inducing point number $M = 40$, discretisation time step $dt = 10^{-3}$, temporary finite horizon $T_{sub} = 0.05$.

X. IMPLEMENTATION DETAILS

Here we provide the outline algorithm for computing optimal interventions. For a concrete implementation we defer the reader to the official repository located at https://github.com/dimitra-maoutsa/DeterministicParticleFlowControl. In the comments, we denote the computational complexity of each operation in the gradient–log–density estimation in terms of big-$\mathcal{O}$ notation. Since the inducing point number M employed in the gradient–log–density estimation is considerably smaller than sample number N, i.e., $M \ll N$, the overall computational complexity of a *single* gradient-log-density evaluation amounts to $\mathcal{O}\left(N\,M^2\right)$. As it becomes obvious from the simulation at the first step of each forward simulation we perform a stochastic Euler-Maruyama step, in order to overcome the singularity of having to compute the gradient log density of a delta function that represents the initial condition.

Algorithm A1: Deterministic Particle Flow (**DPF**) control algorithm

Input: N, M: scalars, number of particles and number of inducing points
t_0, t_1, dt: scalars, initial and final timepoints, and discretisation step
x_0, x_1: $1 \times d$, $1 \times d$ initial and target state
f: drift function
σ: noise amplitude
$U(x,t)$: function, path constraint (optional)

Output: Z, B: $d \times N \times (t_1 - t_0)/dt$, samples from forward flows $\varrho_t(x)$ and $q_t(x)$
$u(x,t)$: functions from $\mathbb{R}^d \to \mathbb{R}^d$ for each $(t_1 - t0)/dt$ time step, time- and state-dependent controls

$k = (t_1 - t_0)/dt$ // number of timesteps

// Forward propagation of flow $\varrho_t(x)$

$Z_{ti=0} = x_0$ // initialise particles' positions $\mathcal{O}()$

$Z_{ti=1} = Z_0 + dt f(Z_0, t_0) + \sigma\mathcal{N}(0, \sqrt{dt})$ // 1st step is stochastic $\mathcal{O}()$

For $ti = 2 : k$ // deterministic propagation

  $Z_{ti+1} = Z_{ti} + dt\left(f(Z_{ti}, t) - \frac{1}{2}\sigma^2\nabla \log \varrho(Z_{ti})\right)$

  If $\exists$ path cost:

    $W = \exp\left(-U(Z_{ti+1}, t)\, dt\right)$

    $T^* = \text{EnsembleTransformParticleFilter}(Z_{ti+1}, W)$

    $Z_{ti+1} = Z_{ti+1} \cdot T^*$

// Time-reversed propagation of flow $q_t(x)$

$B_{ti=k} = x_1$ // initialise particles' positions

// 1st step is stochastic

$B_{ti=k-1} = B_k - dt\left(f(B_k, t_1) + \frac{1}{2}\sigma^2\nabla \log \varrho(Z_k)\right) + \sigma\mathcal{N}(0, \sqrt{dt})$

For $ti = k - 2 : 0$ // deterministic propagation

  $B_{ti-1} = B_{ti} - dt\left(f(B_{ti}, t) - \frac{1}{2}\sigma^2\nabla \log \varrho(Z_{ti}) + \frac{1}{2}\sigma^2\nabla \log q(B_{ti})\right)$

// Compute $u(x\,t)$

For $ti = 2 : k$

  $u(x, ti) = \sigma^2\nabla \log q(B_{ti}) - \sigma^2\nabla \log \varrho(Z_{ti})$

For generating the figures in the main text and supplement we used the Seaborn [95] python package.

XI. SUPPLEMENTARY FIGURES

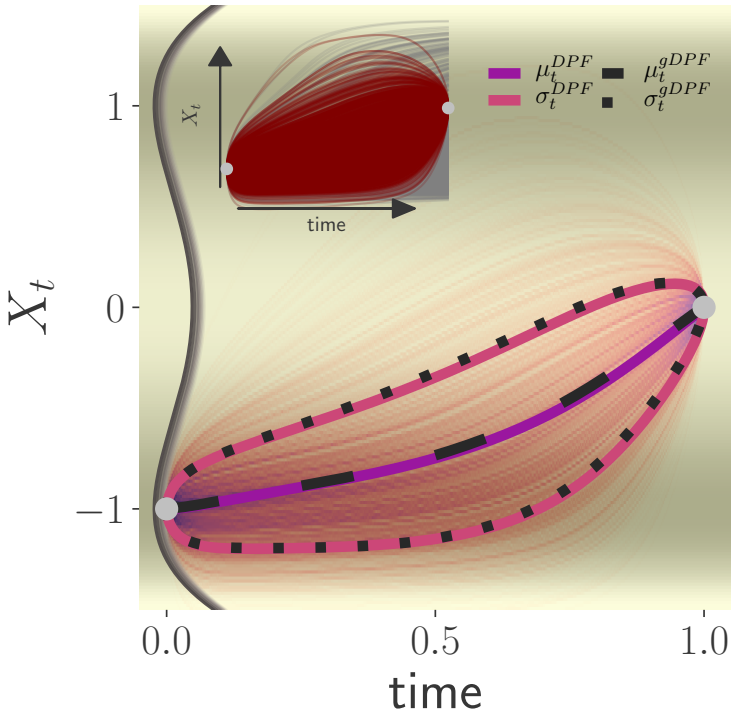

FIG. S-4. **Agreement of constrained probability flows resulting from DPF and gDPF for the typical terminal condition in a double well potential.** Mean μ_t and standard deviation σ_t of the constrained time-reversed probability flows ($\tilde{q}_{DPF}$ and $\tilde{q}_{gDPF}$) computed by employing **DPF** (*solid pink/purple*) and **gDPF** (*dashed black*). The background coloring indicates the density of the particle trajectories of the time-reversed constrained flow, while the two *silver* circles denote the initial and terminal points. **(inset.)** Deterministic particle trajectories capturing the forward unconstrained flow $\varrho_t(x)$ (*grey*) and resulting time-reversed constrained flow $\tilde{q}_t(x)$ (*magenta*) as computed by **DPF**.

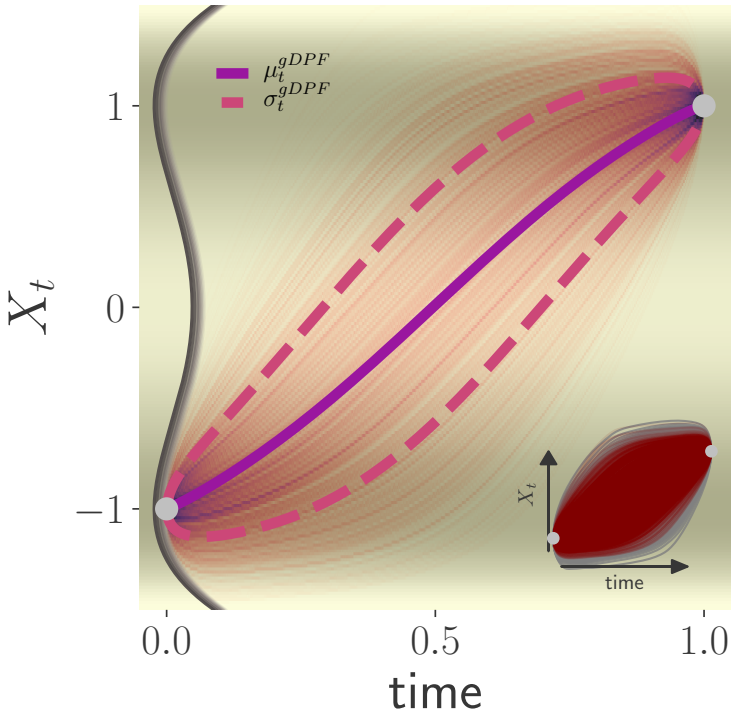

FIG. S-5. **Constrained probability flow resulting from guided particle flow (gDPF) control for extreme terminal condition in a double well potential.** Mean μ_t and standard deviation σ_t of the constrained flows ($\tilde{q}_t$). The background coloring indicates the density of the particle trajectories of the time-reversed constrained flow, while the two *silver* circles denote the initial and terminal points. **(inset.)** Deterministic particle trajectories capturing the forward $\varrho_t^{BB}(x)$ (*grey*) and the time-reversed constrained $\tilde{q}_t(x)$ (*magenta*) probability flows as computed during **gDPF**control.

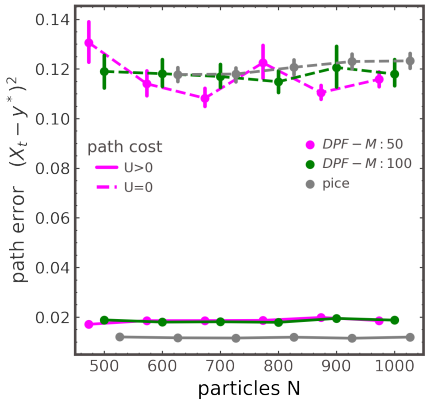

FIG. S-6. **Total path cost for path constrained and path constrained unconstrained systems.** Adding the path constraint to the condition considerably reduces the deviation from the $y^* = 1$ axis for both frameworks.

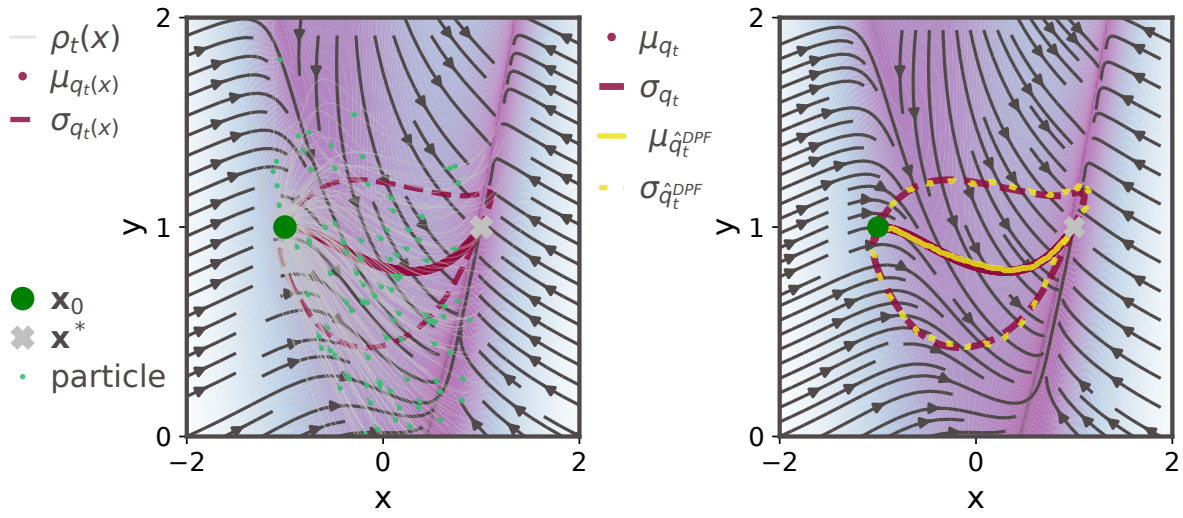

FIG. S-7. **Deterministic trajectories of forward probability flow and statistics of reverse time sampled constrained flow. (upper) (a.) (b.)** Complete agreement of mean μ (solid lines) and standard deviation σ (dashed lines) between the time-reversed controlled density $q_t(x)$ (*maroon*) and the density generated from 1000 independent controlled forward trajectories employing the estimated control dynamics $u(x,t)$, $\hat{q}_t(x)$ (*yellow*).

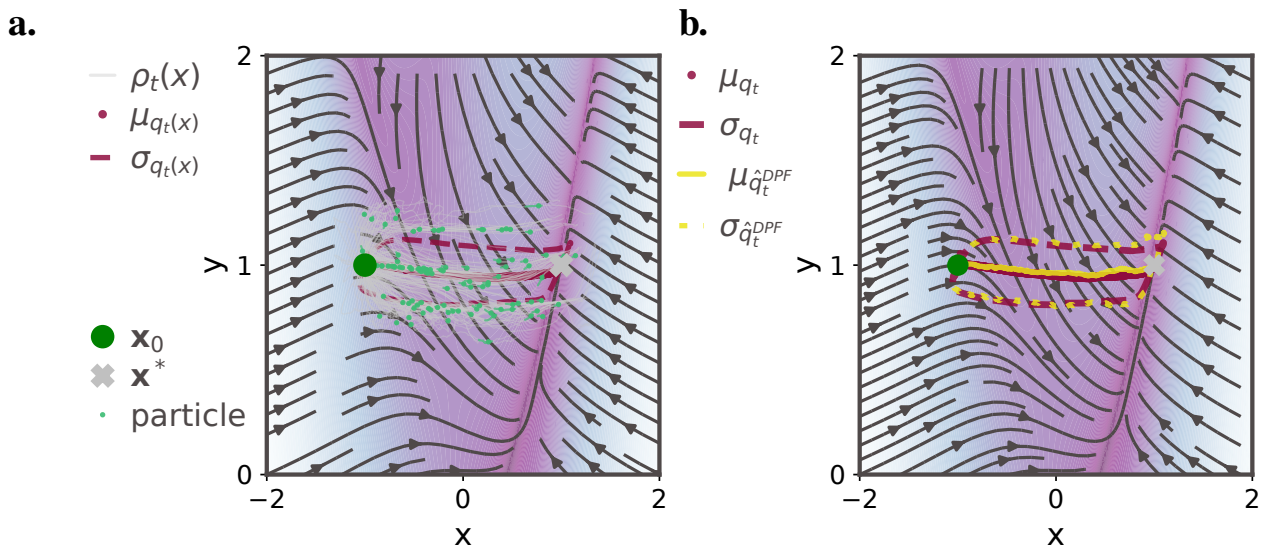

FIG. S-8. **Deterministic trajectories of forward probability flow $\varrho_t(x)$ and statistics of reverse time sampled constrained flow $q_t(x)$ for a path constrained system. (a.)** Transient mean (*solid maroon line*) and standard deviation (*dashed maroon lines*) of the time reversed probability flow $q_t(x)$, and deterministic trajectories of the forward flow (*grey lines*). **(b.)** Complete agreement of mean μ (solid lines) and standard deviation σ (dashed lines) between the time-reversed controlled density $q_t(x)$ (*maroon*) and the density generated from 1000 independent controlled forward trajectories employing the estimated control dynamics $u(x,t)$, $\hat{q}_t(x)$ (*yellow*).

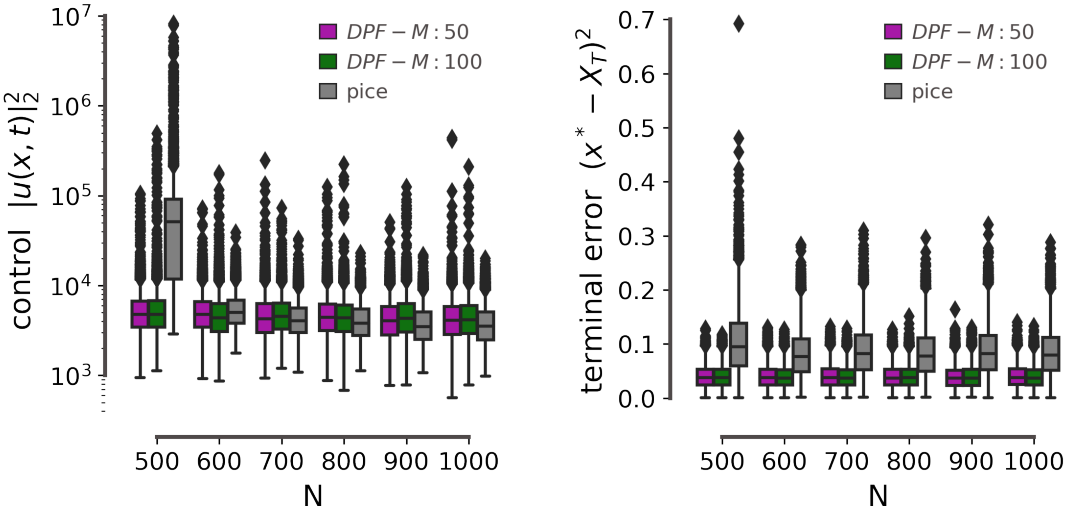

FIG. S-9. **Scaling of employed control energy and terminal error with particle number for deterministic particle control (DPF) and path integral cross entropy method (PICE) for the terminally constrained system.** Inducing point number $M = 50$ (magenta) and $M = 100$ (green).

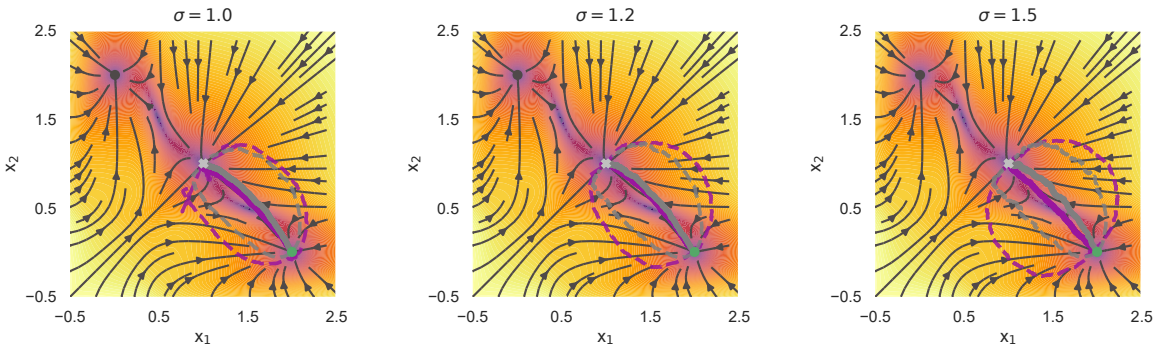

FIG. S-10. **Transient statistics of controlled densities over 1000 trajectories controlled with our method (magenta) and with the path integral cross entropy framework (grey) for increasing noise strength.**

[1] Jonathan M Raser and Erin K O'shea. Noise in gene expression: origins, consequences, and control. *Science*, 309(5743):2010–2013, 2005.

[2] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002.

[3] Jeff Hasty, Joel Pradines, Milos Dolnik, and James J Collins. Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences*, 97(5):2075–2080, 2000.

[4] Thomas B Kepler and Timothy C Elston. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophysical journal*, 81(6):3116–3136, 2001.

[5] Jordi García-Ojalvo and José Sancho. *Noise in spatially extended systems*. Springer Science & Business Media, 2012.

[6] Avigdor Eldar and Michael B Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, 2010.

[7] Michael L Simpson, Chris D Cox, Michael S Allen, James M McCollum, Roy D Dar, David K Karig, and John F Cooke. Noise in biological circuits. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 1(2):214–225, 2009.

[8] William J Blake, Mads Kærn, Charles R Cantor, and James J Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.

[9] Werner Horsthemke. Noise induced transitions. In *Non-equilibrium dynamics in chemical systems*, pages 150–160. Springer, 1984.

[10] Tianshou Zhou, Luonan Chen, and Kazuyuki Aihara. Molecular communication through stochastic synchronization induced by extracellular fluctuations. *Physical Review Letters*, 95(17):178103, 2005.

[11] Marc Rullan, Dirk Benzinger, Gregor W Schmidt, Andreas Milias-Argeitis, and Mustafa Khammash. An optogenetic platform for real-time, single-cell interrogation of stochastic transcriptional regulation. *Molecular cell*, 70(4):745–756, 2018.

[12] Domitilla Del Vecchio, Aaron J Dy, and Yili Qian. Control theory meets synthetic biology. *Journal of The Royal Society Interface*, 13 (120):20160380, 2016.

[13] Phuc Nguyen, Nicholas A Pease, and Hao Yuan Kueh. Scalable control of developmental timetables by epigenetic switching networks. *Journal of the Royal Society Interface*, 18(180):20210109, 2021.

[14] David A Sivak and Gavin E Crooks. Thermodynamic metrics and optimal paths. *Physical Review Letters*, 108(19):190602, 2012.

[15] Patrick R Zulkowski, David A Sivak, Gavin E Crooks, and Michael R DeWeese. Geometry of thermodynamic control. *Physical Review E*, 86(4):041148, 2012.

[16] Alex Gomez-Marin, Tim Schmiedl, and Udo Seifert. Optimal protocols for minimal work processes in underdamped stochastic thermodynamics. *The Journal of chemical physics*, 129(2):024114, 2008.

[17] Tim Schmiedl and Udo Seifert. Optimal finite-time processes in stochastic thermodynamics. *Physical Review Letters*, 98(10):108301, 2007.

[18] Herschel Rabitz, Regina de Vivie-Riedle, Marcus Motzkus, and Karl Kompa. Whither the future of controlling quantum phenomena? *Science*, 288(5467):824–828, 2000.

[19] Alexander N Pechen and David J Tannor. Are there traps in quantum control landscapes? *Physical Review Letters*, 106(12):120402, 2011.

[20] DA Hendrix and C Jarzynski. A "fast growth" method of computing free energy differences. *The Journal of Chemical Physics*, 114 (14):5974–5981, 2001.

[21] Michael R Shirts, Eric Bair, Giles Hooker, and Vijay S Pande. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Physical Review Letters*, 91(14):140601, 2003.

[22] Carsten Hartmann and Christof Schütte. Efficient rare event simulation by optimal nonequilibrium forcing. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(11):P11004, 2012.

[23] Raphaël Chetrite and Hugo Touchette. Variational and optimal control representations of conditioned and driven processes. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(12):P12001, 2015.

[24] Jin Won Kim and Prashant G Mehta. An optimal control derivation of nonlinear smoothing equations. In *Proceedings of the Workshop on Dynamics, Optimization and Computation held in honor of the 60th birthday of Michael Dellnitz*, pages 295–311. Springer, 2020.

[25] Jose Casadiego, Dimitra Maoutsa, and Marc Timme. Inferring network connectivity from event timing patterns. *Physical review letters*, 121(5):054101, 2018.

[26] Emanuel Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28):11478–11483, 2009.

[27] Hilbert J Kappen. Linear theory for control of nonlinear stochastic systems. *Physical Review Letters*, 95(20):200201, 2005.

[28] Daniel K Wells, William L Kath, and Adilson E Motter. Control of stochastic and induced switching in biophysical networks. *Physical Review X*, 5(3):031036, 2015.

[29] Armita Nourmohammad and Ceyhun Eksin. Optimal evolutionary control for artificial selection on molecular phenotypes. *Physical Review X*, 11(1):011044, 2021.

[30] Ziwei Zhong, Brandon G Wong, Arjun Ravikumar, Garri A Arzumanyan, Ahmad S Khalil, and Chang C Liu. Automated continuous evolution of proteins in-vivo. *ACS synthetic biology*, 9(6):1270–1276, 2020.

[31] Ta-Chu Kao, Mahdieh S Sadabadi, and Guillaume Hennequin. Optimal anticipatory control as a theory of motor preparation: a thalamo-cortical circuit model. *Neuron*, 109(9):1567–1581, 2021.

[32] Michael Lässig and Ville Mustonen. Eco-evolutionary control of pathogens. *Proceedings of the National Academy of Sciences*, 117 (33):19694–19704, 2020.

[33] Alexandre Iolov, Susanne Ditlevsen, and André Longtin. Stochastic optimal control of single neuron spike trains. *Journal of Neural Engineering*, 11(4):046004, 2014.

[34] Stephen H Scott. Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7): 532–545, 2004.

[35] Espen Bernton, Jeremy Heng, Arnaud Doucet, and Pierre E Jacob. Schrödinger Bridge Samplers. *arXiv preprint arXiv:1912.13170*, 2019.

[36] Francisco Vargas, Pierre Thodoroff, Neil D Lawrence, and Austen Lamacraft. Solving Schrödinger Bridges via Maximum Likelihood. *arXiv preprint arXiv:2106.02081*, 2021.

[37] Ioannis Exarchos and Evangelos A Theodorou. Stochastic optimal control via forward and backward stochastic differential equations and importance sampling. *Automatica*, 87:159–165, 2018.

[38] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

[39] Richard Bellman. Dynamic programming and Lagrange multipliers. *Proceedings of the National Academy of Sciences of the United States of America*, 42(10):767, 1956.

[40] Jochen Garcke and Axel Kröner. Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids. *Journal of Scientific Computing*, 70(1):1–28, 2017.

[41] Mario Annunziato and Alfio Borzì. A Fokker–Planck control framework for multidimensional stochastic processes. *Journal of Computational and Applied Mathematics*, 237(1):487–507, 2013.

[42] Matanya B Horowitz, Anil Damle, and Joel W Burdick. Linear Hamilton Jacobi Bellman equations in high dimensions. In *53rd IEEE Conference on Decision and Control*, pages 5880–5887. IEEE, 2014.

[43] Hilbert J Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, 2005(11):P11011, 2005.

[44] Bart Van Den Broek, Wim Wiegerinck, and Bert Kappen. Graphical model inference in optimal control of stochastic multi-agent systems. *Journal of Artificial Intelligence Research*, 32:95–122, 2008.

[45] Konrad Rawlik, Marc Toussaint, and Sethu Vijayakumar. Path integral control by reproducing kernel Hilbert space embedding. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

[46] Hilbert Johan Kappen and Hans Christian Ruiz. Adaptive importance sampling for control and inference. *Journal of Statistical Physics*, 162(5):1244–1266, 2016.

[47] Wei Zhang, Han Wang, Carsten Hartmann, Marcus Weber, and Christof Schuette. Applications of the cross-entropy method to importance sampling and optimal control of diffusions. *SIAM Journal on Scientific Computing*, 36(6):A2654–A2672, 2014.

[48] Evangelos Theodorou, Freek Stulp, Jonas Buchli, and Stefan Schaal. An iterative path integral stochastic optimal control approach for learning robotic tasks. *IFAC Proceedings Volumes*, 44(1):11594–11601, 2011.

[49] Sep Thijssen and HJ Kappen. Path integral control and state-dependent feedback. *Physical Review E*, 91(3):032104, 2015.

[50] Emanuel Todorov. General duality between optimal control and estimation. In *2008 47th IEEE Conference on Decision and Control*, pages 4286–4292. IEEE, 2008.

[51] Hilbert J Kappen, Vicenç Gómez, and Manfred Opper. Optimal control as a graphical model inference problem. *Machine Learning*, 87 (2):159–182, 2012.

[52] Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*, 2018.

[53] Hagai Attias. Planning by probabilistic inference. In *International Workshop on Artificial Intelligence and Statistics*, pages 9–16. PMLR, 2003.

[54] Dimitra Maoutsa, Sebastian Reich, and Manfred Opper. Interacting particle solutions of Fokker–Planck equations through gradient–log–density estimation. *Entropy*, 22(8):802, 2020.

[55] Brian DO Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

[56] Daniel T Gillespie and Linda R Petzold. Improved leap-size selection for accelerated stochastic simulation. *The Journal of Chemical Physics*, 119(16):8229–8234, 2003.

[57] Antti Saarinen, Marja-Leena Linne, and Olli Yli-Harja. Stochastic differential equation model for cerebellar granule cell excitability. *PLoS Computational Biology*, 4(2):e1000004, 2008.

[58] Russell Lande, Steinar Engen, and Bernt-Erik Saether. *Stochastic population dynamics in ecology and conservation*. Oxford University Press, 2003.

[59] Naoyuki Takahata, Kazushige Ishii, and Hirotsugu Matsuda. Effect of temporal fluctuation of selection coefficient on gene frequency in a population. *Proceedings of the National Academy of Sciences*, 72(11):4541–4545, 1975.

[60] . See Supplemental Material at Supplementary Information for detailed derivations, additional information on the numerical experiments, and Supplementary Figures.

[61] Emanuel Todorov. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural computation*, 17(5):1084–1108, 2005.

[62] Emanuel Todorov. Optimality principles in sensorimotor control. *Nature neuroscience*, 7(9):907–915, 2004.

[63] Emanuel Todorov. Linearly-solvable Markov decision problems. In *Advances in neural information processing systems*, pages 1369–1376, 2007.

[64] Wendell H Fleming. Exit probabilities and optimal stochastic control. *Applied Mathematics and Optimization*, 4(1):329–346, 1977.

[65] Henri Orland. Generating transition paths by Langevin bridges. *The Journal of chemical physics*, 134(17):174114, 2011.

[66] Alain Mazzolo. Constrained Brownian processes and constrained Brownian bridges. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(2):023203, 2017.

[67] Juraj Szavits-Nossan and Martin R Evans. Inequivalence of nonequilibrium path ensembles: the example of stochastic bridges. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(12):P12008, 2015.

[68] Satya N Majumdar and Henri Orland. Effective Langevin equations for constrained stochastic processes. *Journal of Statistical Mechan-*

ics: Theory and Experiment, 2015(6):P06039, 2015.
[69] Nicolas Macris and Raffaele Marino. Solving non-linear Kolmogorov equations in large dimensions by using deep learning: a numerical comparison of discretization schemes. *arXiv preprint arXiv:2012.07747*, 2020.
[70] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *International Conference on Learning Representations*, 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] Sebastian Reich. A nonparametric ensemble transform method for Bayesian inference. *SIAM Journal on Scientific Computing*, 35(4): A2013–A2024, 2013.
[73] Udo Von Toussaint. Bayesian inference in physics. *Reviews of Modern Physics*, 83(3):943, 2011.
[74] Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger Bridge with Applications to Score-Based Generative Modeling. *arXiv preprint arXiv:2106.01357*, 2021.
[75] Crispin W.. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, 2009.
[76] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.
[77] Sui Huang, Yan-Ping Guo, Gillian May, and Tariq Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental Biology*, 305(2):695–713, 2007.
[78] Russell Lande. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, pages 314–334, 1976.
[79] Nikolas Nüsken and Lorenz Richter. Solving high-dimensional Hamilton–Jacobi–Bellman PDEs using neural networks: perspectives from the theory of controlled diffusions and measures on path space. *Partial Differential Equations and Applications*, 2(4):1–48, 2021.
[80] Hamid Moradkhani, Grey S Nearing, Peyman Abbaszadeh, and Sahani Pathiraja. Fundamentals of data assimilation and theoretical advances. *Handbook of Hydrometeorological Ensemble Forecasting. Springer Berlin Heidelberg, Berlin, Heidelberg*, pages 675–699, 2019.
[81] Nicolas Bonneel, Michiel Van De Panne, Sylvain Paris, and Wolfgang Heidrich. Displacement interpolation using Lagrangian mass transport. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, pages 1–12, 2011.
[82] Rémi Flamary, Nicolas Courty, Alexandre Gramfort, Mokhtar Zahdi Alaya, Aurélie Boisbunon, Stanislas Chambon, Laetitia Chapel, Adrien Corenflos, Kilian Fatras, Nemo Fournier, et al. POT: Python optimal transport. *Journal of Machine Learning Research*, 22(78): 1–8, 2021.
[83] Jr-Shin Li, Wei Zhang, and Shuo Wang. Optimal Control and Stochastic Synchronization of Phase Oscillators. *IFAC-PapersOnLine*, 48 (18):83–88, 2015.
[84] Johann Brehmer. Simulation-based inference in particle physics. *Nature Reviews Physics*, 3(5):305–305, 2021.
[85] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences of the United States of America*, 117(48):30055–30062, 2020.
[86] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2019.
[87] David Gresham and Maitreya J Dunham. The enduring utility of continuous culturing in experimental evolution. *Genomics*, 104(6): 399–405, 2014.
[88] Martin Hairer, Andrew M Stuart, and Jochen Voss. Sampling conditioned diffusions. *Trends in stochastic analysis*, 353:159–186, 2009.
[89] Igor Vladimirovich Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*, 5(3):285–301, 1960.
[90] Kenneth David Elworthy. *Stochastic differential equations on manifolds*, volume 70. Cambridge University Press, 1982.
[91] Erhan Bayraktar and Mihai Sirbu. Stochastic Perron's Method for Hamilton–Jacobi–Bellman Equations. *SIAM Journal on Control and Optimization*, 51(6):4274–4294, 2013.
[92] Reinaldo Marques and Geir Storvik. Reweighting schemes based on particle methods. In *The Contribution of Young Researchers to Bayesian Statistics*, pages 73–76. Springer, 2014.
[93] Ofir Pele and Michael Werman. Fast and robust earth mover's distances. In *2009 IEEE 12th International Conference on Computer Vision*, pages 460–467. IEEE, September 2009.
[94] Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(Apr): 695–709, 2005.
[95] Michael L Waskom. Seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021.
[96] Armita Nourmohammad, Stephan Schiffels, and Michael Lässig. Evolution of molecular phenotypes under stabilizing selection. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01012, 2013.
[97] Torsten Held, Armita Nourmohammad, and Michael Lässig. Adaptive evolution of molecular phenotypes. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(9):P09029, 2014.
[98] Georg A Gottwald. Finite-size effects in a stochastic Kuramoto model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27 (10):101103, 2017.
[99] Arthur T Winfree. *The geometry of biological time*, volume 12. Springer Science & Business Media, 2001.
[100] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975.