A Physics-informed Deep Learning Approach for
Minimum Effort Stochastic Control of Colloidal Self-Assembly

Iman Nodozi, Jared O'Leary, Ali Mesbah, Abhishek Halder

Abstract— We propose formulating the finite-horizon stochastic optimal control problem for colloidal self-assembly in the space of probability density functions (PDFs) of the underlying state variables (namely, order parameters). The control objective is formulated in terms of steering the state PDFs from a prescribed initial probability measure towards a prescribed terminal probability measure with minimum control effort. For specificity, we use a univariate stochastic state model from the literature. Both the analysis and the computational steps for control synthesis as developed in this paper generalize for multivariate stochastic state dynamics given by generic nonlinear in state and non-affine in control models. We derive the conditions of optimality for the associated optimal control problem. This derivation yields a system of three coupled partial differential equations together with the boundary conditions at the initial and terminal times. The resulting system is a generalized instance of the so-called Schrödinger bridge problem. We then determine the optimal control policy by training a physics-informed deep neural network, where the "physics" are the derived conditions of optimality. The performance of the proposed solution is demonstrated via numerical simulations on a benchmark colloidal self-assembly problem.

I. INTRODUCTION

Colloidal self-assembly (SA) is the process by which discrete components (e.g., micro-/nano-particles in solution) spontaneously organize into an ordered state [1]. The spontaneous self-organization central to colloidal SA enables "bottom-up" materials synthesis, which can allow for manufacturing advanced, highly-ordered crystalline structures in an inherently parallelizable and cost-effective manner [2], [3]. The fact that colloidal SA can begin with micro- and/or nano-scale building blocks of varying complexity indicates that this bottom-up engineering approach can be used to synthesize novel metamaterials with unique optical, electrical, or mechanical properties [2]–[4].

Colloidal SA is an inherently stochastic (i.e., random) process prone to kinetic arrest due to particle Brownian motion [2], [3], [5], [6]. This leads to variability in materials manufacturing and possibly high defect rates, which can severely compromise the viability of using colloidal SA to reproducibly manufacture advanced materials. This lack of reproducibility in turn prevents colloidal SA from achieving

Iman Nodozi is with the Department of Electrical and Computer Engineering, University of California, Santa Cruz, CA 95064, USA, `inodozi@ucsc.edu`.
Jared O'Leary and Ali Mesbah are with the Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA, {`jared.oleary,mesbah`}`@berkeley.edu`.
Abhishek Halder is with the Department of Applied Mathematics, University of California, Santa Cruz, CA 95064, USA, `ahalder@ucsc.edu`.

cost-effective and scalable manufacturing of such materials [2], [3], [7], [8]. Thus, the thermodynamic and kinetic driving forces that govern colloidal SA, will need to be precisely and systematically modulated to consistently and efficiently direct colloidal SA systems towards high-value mass-producible structures and materials.

To more reproducibly drive collodial SA systems towards desired structures, it has been proposed to design a model-based feedback control policy wherein a global actuator (e.g., electric field voltage) is manipulated based on currently available information on the system state and a dynamical system model [9]–[11]. Work in [9] presents a model predictive control (MPC) method for controlling colloidal SA. These authors consider the dynamical model for the stochastic colloidal assembly process based on actuator-parameterized Langevin equations. In [10], [11], the system is guided towards the desired highly-ordered structure based on a Markov decision process optimal control policy.

In comparison, the perspective and approach taken in this paper towards optimal control of a colloidal SA process are significantly different, as we seek to control the time evolution of the joint probability distribution supported over the states of colloidal SA. Our technical approach and contributions are as follows.

(1) We show that the problem of controlling colloidal SA over a finite-time horizon can be naturally formulated in terms of steering of the probability distributions of the underlying stochastic system states, namely the order parameters. This leads to a two-point boundary value problem over the manifold of state probability measures, which from a control viewpoint is a non-traditional stochastic optimal control problem known as the Schrödinger bridge problem (SBP), see e.g., [12]. The notion of lifting the stochastic control of colloidal SA directly onto the space of state probability measure as an SBP, is novel.

(2) While there exists a growing literature on numerically solving the SBPs with nonlinear prior dynamics [13]–[17], these works leverage specific structures of the underlying drift and diffusion terms in a stochastic differential equation (SDE). In contrast, the typical colloidal SA application, as considered here, requires more generic considerations since the drift and diffusion coefficients can be nonlinear w.r.t to the states as well as non-affine in control. We show that, unlike the existing SBP conditions of optimality, our setting leads to a system of three coupled partial differential equations (PDEs) with endpoint PDF constraints. These three

PDEs are the controlled Fokker-Planck-Kolmogorov (FPK) PDE, the Hamilton-Jacobi-Bellman (HJB) PDE, and a policy PDE.

(3) The resulting system of three coupled PDEs is not amenable to existing computational approaches for the SBPs, such as the contractive fixed point recursions on the so-called Schrödinger factors. To address this challenge, we employ the notion of physics-informed neural networks (PINN) (e.g., [18], [19]) to train a deep neural network approximating the solution of our coupled PDE system and the boundary conditions.

The distinct feature of the proposed control methodology is that it derives an optimal control policy that steers a given probability distribution of the order parameters of a colloidal SA system to a desired one over a finite-time horizon. The computation does not require making parametric approximation of the statistics such as the Gaussian mixture or exponential family. It also avoids approximating the nonlinearities in the drift and diffusion *a priori*, e.g., via Taylor series.

Our technical contribution is a new control methodology. To make the exposition concrete, we use a specific univariate SDE model from the literature [9], [11] to demonstrate the proof-of-concept for our proposed method. The specificity of the model is used from the outset as a didactic writing style, even though our proposed method is general, i.e., *not* contingent on the nonlinearity structure or dimensionality of the SDE model.

Organization: This paper is structured as follows. Sec. II details our proposed stochastic optimal control problem formulation. In Sec. III, we derive the first order conditions for optimality in the form of a system of coupled PDEs. We then learn the solutions for this system of equations by training a PINN, as detailed in Sec. IV. The numerical simulation results are reported in Sec. V, followed by concluding remarks in Sec. VI.

Notations: We use $\mathbb{E}_{\mu^\pi}[\cdot]$ to denote the expectation w.r.t. the controlled state probability measure or distribution μ^π, that is, $\mathbb{E}_{\mu^\pi}[\cdot] := \int (\cdot) \, \mathrm{d}\mu^\pi$. The superscript π in μ^π indicates that the probability measure depends on the choice of the control policy π. When the probability measure μ^π is absolutely continuous, it admits a PDF ρ^π. The symbol $\sim$ is used as a shorthand for "follows the probability distribution".

II. PROBLEM FORMULATION

A. Colloidal Self-Assembly Sample Path Model

The specific model we consider for the sample path dynamics for SA of colloidal particles is given by [9], [11] the Itô SDE

$$\mathrm{d}\langle C_6\rangle = D_1(\langle C_6\rangle, \pi) \, \mathrm{d}t + \sqrt{2D_2(\langle C_6\rangle, \pi)} \, \mathrm{d}w \quad (1)$$

where t denotes time, and the state variable $\langle C_6\rangle \in [0,6]$ is an order parameter denoting the average number of hexagonally close packed particles around each particle. We consider the dynamics (1) over a fixed time horizon $[0,T]$. The control input $u := \pi(\langle C_6\rangle, t) \in \mathbb{R}$ denotes electric field

voltage [20] that results from a Markovian control policy $\pi : [0,6] \times [0,T] \mapsto \mathbb{R}$, and w denotes the standard Wiener process in $\mathbb{R}$. In (1), the functionals $D_1(\cdot, \cdot)$ and $D_2(\cdot, \cdot)$ are referred to as the *drift* and the *diffusion landscapes*, respectively. In the colloidal SA context, both D_1, D_2 are nonlinear in state and non-affine in control.

Typically, the drift landscape D_1 is expressed in terms of the so-called *free energy landscape* F and the *diffusion landscape* D_2, as

$$\begin{aligned} D_1(\langle C_6\rangle, \pi) = & \frac{\partial}{\partial\langle C_6\rangle} D_2(\langle C_6\rangle, \pi) \\ & - \frac{D_2(\langle C_6\rangle, \pi)}{k_{\mathrm{B}}\,\theta} \frac{\partial}{\partial\langle C_6\rangle} F(\langle C_6\rangle, \pi), \quad (2) \end{aligned}$$

where the Boltzmann constant $k_{\mathrm{B}} = 1.38066 \times 10^{-23}$ Joules per Kelvin, and θ denotes a suitable temperature in Kelvin. In Sec. V, we will give illustrative numerical results for specific choices of the diffusion and the free energy landscapes.

For an admissible Markovian policy $\pi(\cdot, t)$, we assume that the landscapes D_1, D_2 satisfy

(A1) non-explosion and Lipschitz conditions: there exist constants c_1, c_2 such that

$$\|D_1(\cdot, \pi(\cdot, t))\|_2 + \|\sqrt{D_2(\cdot, \pi(\cdot, t))}\|_2 \le c_1(1 + \|\cdot\|_2),$$

and that

$$\|D_1(x, \pi(x, t)) - D_1(\widetilde{x}, \pi(\widetilde{x}, t))\|_2 \le c_2\|x - \widetilde{x}\|_2$$

for all $x, \widetilde{x} \in [0,6]$, $t \in [0,T]$,

(A2) uniformly lower bounded diffusion: there exists constant c_3 such that $(\cdot)^\top D_2(\cdot, \pi(\cdot, t))(\cdot) \ge c_3\|\cdot\|_2^2$ for all $t \in [0,T]$.

The assumption **A1** guarantees [21, p. 66] existence-uniqueness for the sample path of the SDE (1). The assumptions **A1, A2** together guarantee [22, Ch. 1] that the generator associated with (1) yields absolutely continuous probability measures $\mu^\pi(\langle C_6\rangle, t)$ for all $t > 0$ provided the initial measure $\mu_0 := \mu^\pi(\cdot, t = 0)$ is absolutely continuous. In other words, the PDFs $\rho^\pi(\langle C_6\rangle, t)$ exist such that $\mathrm{d}\mu^\pi(\langle C_6\rangle, t) = \rho^\pi(\langle C_6\rangle, t)\mathrm{d}\langle C_6\rangle$.

B. Controlled Self-Assembly as Distribution Steering

We propose reformulating the problem of designing a control policy $\pi(\langle C_6\rangle, t)$ for the controlled self-assembly subject to (1), to that of steering the *statistics* of the stochastic state $\langle C_6\rangle$ from a prescribed initial probability measure μ_0 at $t = 0$ to a prescribed terminal probability measure μ_T at $t = T$. This is motivated by the fact that $\langle C_6\rangle \approx 0$ implies disordered crystalline structure while $\langle C_6\rangle \approx 5.1$ implies an ordered structure. So steering the stochastic state from disordered to ordered naturally transcribes to steering a high concentration of probability mass around 0 to the same around 5.1. Note the target value of $\langle C_6\rangle$ is 5.1 and not 6 due to edge effects in the lattice structure.

We emphasize here that we use the term "statistics" in nonparametric sense, i.e., we allow arbitrary probability measures μ_0, μ_T supported on the compact set $[0,6]$, and
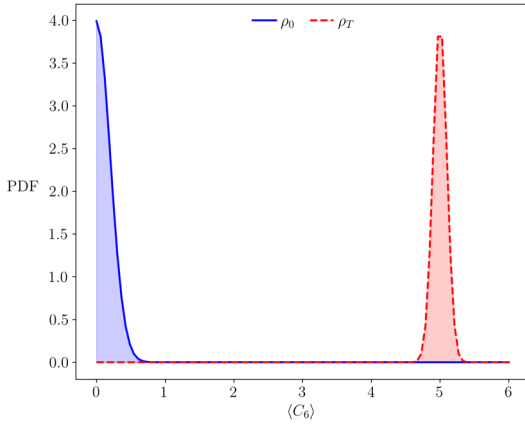
Fig. 1: The prescribed initial PDF ρ_0 (solid line) at the initial time $t = 0$, and the prescribed terminal PDF ρ_T (dashed line) at the final time $t = T$. Both PDFs are supported over $[0,6]$, which is the range of values for the state variable $\langle C_6 \rangle$ denoting a crystallinity order parameter. In particular, $\langle C_6 \rangle \approx 0$ implies a disordered state and $\langle C_6 \rangle \approx$ 5-6 implies a highly ordered state.

ask for provable steering of μ_0 to μ_T via control, not just steering of first few statistical moments of μ_0 to μ_T such as mean and variance. Even if μ_0, μ_T have finite dimensional sufficient statistics, the transient $\langle C_6 \rangle$ probability measures induced by the nonlinear SDE (1) for a given control policy π, may not have so. This motivates formulating the control synthesis as a two point boundary value problem over the (infinite dimensional) manifold of state probability measures.

Specifically, we consider the minimum control effort steering of μ_0 to μ_T, i.e., solving

$$\inf_{\pi \in \mathcal{U}} \int_0^T \mathbb{E}_{\mu^\pi} \left[\frac{1}{2} \pi^2 \right] \, \mathrm{d}t$$

$$\text{subject to} \quad \mathrm{d}\langle C_6 \rangle = D_1(\langle C_6 \rangle, \pi)\mathrm{d}t + \sqrt{2D_2(\langle C_6 \rangle, \pi)} \, \mathrm{d}w,$$

$$\langle C_6 \rangle (t = 0) \sim \mu_0 \text{ (given)},$$

$$\langle C_6 \rangle (t = T) \sim \mu_T \text{ (given)},$$

(3)

where $\mu^\pi \equiv \mu^\pi(\langle C_6 \rangle, t)$ denotes the controlled state probability measure at time t. In other words, we design state feedback for dynamically reshaping uncertainties subject to the dynamical constraint (1), endpoint statistical constraints, and the deadline constraint.

In (3), the set of feasible controls $\mathcal{U}$ comprises of the finite energy Markovian inputs, i.e.,

$$\mathcal{U} := \left\{ \pi : [0,6] \times [0,T] \mapsto \mathbb{R} \mid \int_0^T \mathbb{E}_{\mu^\pi}[\pi^2] \, \mathrm{d}t < \infty \right\}. \quad (4)$$

Recall that the input is the electrical field voltage, and the minimum effort objective is a natural candidate to promote control parsimony.

We suppose that the endpoint probability measures μ_0, μ_T are absolutely continuous with respective PDFs ρ_0, ρ_T; see Fig. 1. Then, we can rewrite (3) as the variational problem:

$$\inf_{(\rho^\pi, \pi)} \int_0^T \int_{\mathbb{R}} \frac{1}{2} \pi^2(\langle C_6 \rangle, t) \rho^\pi(\langle C_6 \rangle, t) \, \mathrm{d}\langle C_6 \rangle \, \mathrm{d}t \quad (5a)$$

$$\text{subject to} \quad \frac{\partial \rho^\pi}{\partial t} = -\frac{\partial}{\partial \langle C_6 \rangle}(D_1 \rho^\pi) + \frac{\partial^2}{\partial \langle C_6 \rangle^2}(D_2 \rho^\pi),$$

(5b)

$$\rho^\pi(\langle C_6 \rangle, 0) = \rho_0, \quad \rho^\pi(\langle C_6 \rangle, T) = \rho_T, \quad (5c)$$

where (5b) is the *controlled* Fokker-Planck-Kolmogorov's forward (FPK) PDE. The feasible pair $(\rho^\pi, \pi) \in \mathcal{P}_{0T} \times \mathcal{U}$ where $\mathcal{P}_{0T}$ denotes PDF-valued trajectories connecting ρ_0, ρ_T, i.e.,

$$\mathcal{P}_{0T} := \left\{ \rho(\cdot, t) \geq 0 \mid \int \rho(\cdot, t)\mathrm{d}(\cdot) = 1 \text{ for all } t \in [0, T], \right.$$

$$\left. \rho(\cdot, t = 0) = \rho_0, \quad \rho(\cdot, t = T) = \rho_T \right\}, \quad (6)$$

and $\mathcal{U}$ is given by (4).

The variational problem (5) is an instance of the SBP that concerns with most likely stochastic evolution that transports ρ_0 to ρ_T over $[0, T]$. The topic originated in the works of Erwin Schrödinger [23], [24], and recognizing its connection with stochastic control [12], [25]–[28] has unfolded a rapid development in the control literature including when the drift is nonlinear in state [13]–[15], [17], [29]. The computational approach in these works rely on a contractive fixed point recursion [30] over the so-called Schrödinger factors (see e.g., [13, Sec. II]) resulting from certain change of variables related to the Fleming's logarithmic transform [31], [32] or the Hopf-Cole transform [33], [34].

With respect to the existing literature on SBP with nonlinear state-dependent drift, the additional difficulty in (5) is that D_1, D_2 for colloidal SA are *non-affine* in the control policy. To the best of our knowledge, an SBP with this level of generality has not been investigated before. Beyond analytical difficulties, the computation also becomes challenging since non-affine control precludes the aforesaid fixed point recursion approach via the Schrödinger factors.

III. OPTIMALITY

In this work, we will not pursue the existence-uniqueness proofs for the solution of (5). Instead, we next formally derive the first order conditions for optimality for problem (5) in the form of three coupled PDEs with endpoint boundary conditions while tacitly assuming the existence-uniqueness. In Sec. IV, we will numerically solve this system via PINN.

Theorem 1. *(First order conditions for optimality) The pair* $(\rho^\pi_{\mathrm{opt}}(\langle C_6 \rangle, t), \pi_{\mathrm{opt}}(\langle C_6 \rangle, t))$ *that solves* (5), *must satisfy the system of coupled PDEs*

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}(\pi_{\mathrm{opt}})^2 - D_1 \frac{\partial \psi}{\partial \langle C_6 \rangle} - D_2 \frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2}, \quad (7a)$$

$$\frac{\partial \rho^\pi_{\mathrm{opt}}}{\partial t} = -\frac{\partial}{\partial \langle C_6 \rangle}(D_1 \rho^\pi_{\mathrm{opt}}) + \frac{\partial^2}{\partial \langle C_6 \rangle^2}(D_2 \rho^\pi_{\mathrm{opt}}), \quad (7b)$$

$$\pi_{\mathrm{opt}}(\langle C_6 \rangle, t) = \frac{\partial \psi}{\partial \langle C_6 \rangle} \frac{\partial D_1}{\partial \pi_{\mathrm{opt}}} + \frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2} \frac{\partial D_2}{\partial \pi_{\mathrm{opt}}}, \quad (7c)$$

in unknowns $(\psi(\langle C_6 \rangle, t), \rho^\pi_{\mathrm{opt}}(\langle C_6 \rangle, t), \pi_{\mathrm{opt}}(\langle C_6 \rangle, t))$ *with boundary conditions*

$$\rho^\pi_{\mathrm{opt}}(\langle C_6 \rangle, 0) = \rho_0, \quad \rho^\pi_{\mathrm{opt}}(\langle C_6 \rangle, T) = \rho_T. \quad (8)$$
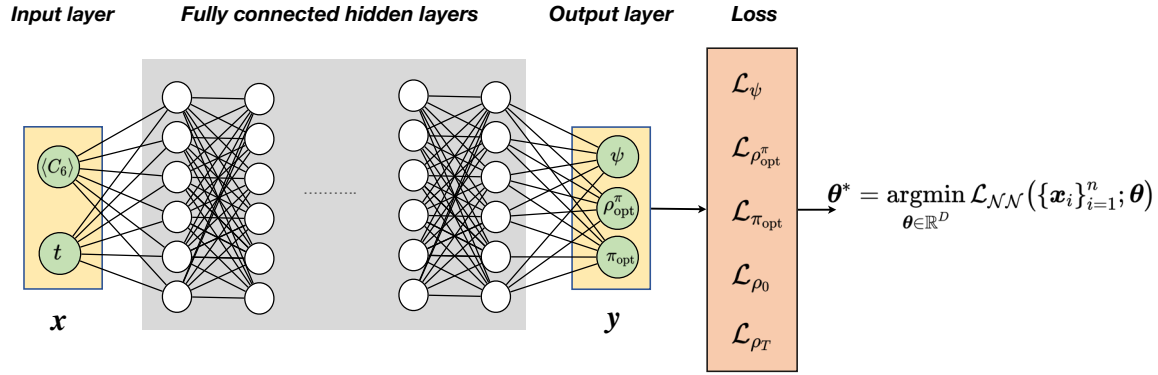
Fig. 2: The architecture of the physics-informed neural network with the system order parameter and time as the input features $\boldsymbol{x} := (\langle C_6 \rangle, t)$. The output $\boldsymbol{y}$ comprises of the value function, optimally controlled PDF, and optimal control policy, i.e., $\boldsymbol{y} := (\psi, \rho_{\text{opt}}^{\pi}, \pi_{\text{opt}})$.

Proof. Consider the Lagrangian associated with (5):

$$\mathcal{L}(\rho^{\pi}, \pi, \psi) := \int_0^T \int_{\mathbb{R}} \left\{ \frac{1}{2}\pi^2\rho^{\pi} + \psi \times \left(\frac{\partial \rho^{\pi}}{\partial t} \right. \right.$$
$$\left. \left. + \frac{\partial}{\partial \langle C_6 \rangle}(D_1\rho^{\pi}) - \frac{\partial^2}{\partial \langle C_6 \rangle^2}(D_2\rho^{\pi}) \right) \right\} \mathrm{d}x\, \mathrm{d}t$$
(9)

where $\psi(\langle C_6 \rangle, t)$ is a $C^1(\mathbb{R}; \mathbb{R}_{\geq 0})$ Lagrange multiplier.

Performing integration by parts, the Lagrangian (9) can be written as

$$\mathcal{L}(\rho^{\pi}, \pi, \psi) = \int_0^T \int_{\mathbb{R}} \left(\frac{1}{2}\pi^2 - \frac{\partial \psi}{\partial t} - D_1\frac{\partial \psi}{\partial \langle C_6 \rangle} \right.$$
$$\left. - D_2\frac{\partial^2}{\partial \langle C_6 \rangle^2}\psi \right) \rho^{\pi}\, \mathrm{d}x\, \mathrm{d}t.$$
(10)

For ρ^{π} fixed, pointwise minimization of (10) with respect to π yields (7c).

By substituting (7c) in (10) and equating the resulting expression to zero, we get the dynamic programming equation

$$\int_0^T \int_{\mathbb{R}} \left(\frac{1}{2}\left(\frac{\partial \psi}{\partial \langle C_6 \rangle}\frac{\partial D_1}{\partial \pi} + \frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2}\frac{\partial D_2}{\partial \pi} \right)^2 - \frac{\partial \psi}{\partial t} \right.$$
$$\left. - D_1\frac{\partial \psi}{\partial \langle C_6 \rangle} - D_2\frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2} \right) \rho^{\pi}(\langle C_6 \rangle, t)\mathrm{d}\langle C_6 \rangle\, \mathrm{d}t = 0.$$
(11)

For (11) to be satisfied for arbitrary ρ^{π}, we must have

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}\left(\frac{\partial \psi}{\partial \langle C_6 \rangle}\frac{\partial D_1}{\partial \pi} + \frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2}\frac{\partial D_2}{\partial \pi} \right)^2$$
$$- D_1\frac{\partial \psi}{\partial \langle C_6 \rangle} - D_2\frac{\partial^2 \psi}{\partial \langle C_6 \rangle^2},$$
(12)

which upon using (7c), gives the HJB PDE (7a). The associated FPK PDE (7b) and the boundary conditions (8) follow from (5b) and (5c), respectively. ■

Remark 1. *The equation* (7a) *in Theorem 1 is the HJB PDE, and the* $\psi(\langle C_6 \rangle, t)$ *is referred to as the value function.*

Remark 2. *Unlike the conditions for optimality for control-affine SBP [12, eq. (5.7)-(5.8)], [13, eq. (20)-(21)], [14, eq. (4)] where we get two coupled PDEs, one being HJB and another being FPK, the system of equations* (7) *has three coupled PDEs because the policy equation* (7c) *is implicit*

in $\boldsymbol{u}_{\text{opt}}$. *Due to the non-affine control, we can no longer express* $\boldsymbol{u}_{\text{opt}}$ *as the scaled gradient of the value function* ψ. *Instead, we now need to solve the coupled system* (7)-(8).

IV. SOLVING THE CONDITIONS FOR OPTIMALITY USING PINN

We propose leveraging recent advances in neural network based computational frameworks to jointly learn the solutions of (7)-(8). In the following, we discuss training of a PINN [18], [19] to numerically solve (7)-(8), which is a system of three coupled PDEs together with the endpoint PDF boundary conditions.

The proposed architecture of the PINN is shown in Fig. 2. In our problem, $\boldsymbol{x} := (\langle C_6 \rangle, t)$ comprises of the features given to the DNN, and the DNN output $\boldsymbol{y} := (\psi, \rho_{\text{opt}}^{\pi}, \pi_{\text{opt}})$ comprises of the value function, optimally controlled PDF, and optimal policy. We parameterize the output of the fully connected feed-forward DNN via $\boldsymbol{\theta} \in \mathbb{R}^D$, i.e.,

$$\boldsymbol{y}(\boldsymbol{x}) \approx \mathcal{N}\mathcal{N}(\boldsymbol{x}; \boldsymbol{\theta}), \tag{13}$$

where $\mathcal{N}\mathcal{N}(\cdot; \boldsymbol{\theta})$ denotes the neural network approximant parameterized by $\boldsymbol{\theta}$, and D is the dimension of the parameter space (i.e., the total number of to-be-trained weight, bias and scaling parameters for the DNN).

The overall loss function for the network, denoted as $\mathcal{L}_{\mathcal{N}\mathcal{N}}$, consists of the sum of the losses associated with the three equations in (7) and the losses associated with the boundary conditions (8). Specifically, let $\mathcal{L}_{\psi}$ be the loss term for the HJB PDE (7a), let $\mathcal{L}_{\rho_{\text{opt}}^{\pi}}$ be the loss term for the FPK PDE (7b), and let $\mathcal{L}_{\pi_{\text{opt}}}$ be the loss term for the control policy equation (7c). Likewise, let $\mathcal{L}_{\rho_0}$ and $\mathcal{L}_{\rho_T}$ denote the loss terms for the corresponding endpoint constraints (8). Then

$$\mathcal{L}_{\mathcal{N}\mathcal{N}} := \mathcal{L}_{\psi} + \mathcal{L}_{\rho_{\text{opt}}^{\pi}} + \mathcal{L}_{\pi_{\text{opt}}} + \mathcal{L}_{\rho_0} + \mathcal{L}_{\rho_T}, \tag{14}$$

where each summand loss term in (14) is evaluated on a set of n collocation points $\{\boldsymbol{x}_i\}_{i=1}^n$ in the domain of the feature space $\Omega := [0, 6] \times [0, T]$, i.e., $\{\boldsymbol{x}_i\}_{i=1}^n \subset \Omega$, and

$$\mathcal{L}_{\psi} := \frac{1}{n}\sum_{i=1}^n \left(\frac{\partial \psi}{\partial t}\bigg|_{\boldsymbol{x}_i} - \frac{1}{2}(\pi_{\text{opt}})^2\bigg|_{\boldsymbol{x}_i} + D_1\frac{\partial}{\partial \langle C_6 \rangle}\psi\bigg|_{\boldsymbol{x}_i} \right.$$
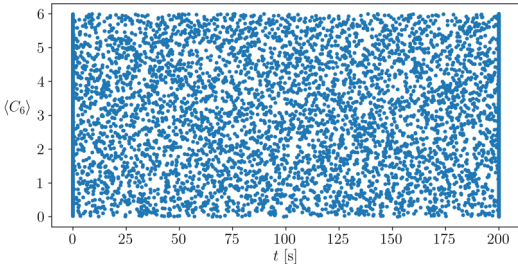
Fig. 3: Training data in the domain $\Omega = [0,6] \times [0,200]$.

$$\left. +D_2 \frac{\partial^2}{\partial \langle C_6 \rangle^2} \psi \right|_{\boldsymbol{x}_i} \right)^2,$$

$$\mathcal{L}_{\rho_{\mathrm{opt}}^\pi} := \frac{1}{n} \sum_{i=1}^n \left(\left. \frac{\partial \rho_{\mathrm{opt}}^\pi}{\partial t} \right|_{\boldsymbol{x}_i} + \left. \frac{\partial}{\partial \langle C_6 \rangle}(D_1 \rho_{\mathrm{opt}}^\pi) \right|_{\boldsymbol{x}_i} \right.$$
$$\left. - \left. \frac{\partial^2}{\partial \langle C_6 \rangle^2}(D_2 \rho_{\mathrm{opt}}^\pi) \right|_{\boldsymbol{x}_i} \right)^2,$$

$$\mathcal{L}_{\pi_{\mathrm{opt}}} := \frac{1}{n} \sum_{i=1}^n \left(\left. \pi_{\mathrm{opt}} \right|_{\boldsymbol{x}_i} - \left. \frac{\partial}{\partial \langle C_6 \rangle} \psi \frac{\partial}{\partial \pi_{\mathrm{opt}}} D_1 \right|_{\boldsymbol{x}_i} \right.$$
$$\left. - \left. \frac{\partial^2}{\partial \langle C_6 \rangle^2} \psi \frac{\partial}{\partial \pi_{\mathrm{opt}}} D_2 \right|_{\boldsymbol{x}_i} \right)^2,$$

$$\mathcal{L}_{\rho_0} := \frac{1}{n} \sum_{i=1}^n \left(\left. \rho_{\mathrm{opt}}^\pi(\cdot, t=0) \right|_{\boldsymbol{x}_i} - \left. \rho_0(\cdot) \right|_{\boldsymbol{x}_i} \right)^2,$$

$$\mathcal{L}_{\rho_T} := \frac{1}{n} \sum_{i=1}^n \left(\left. \rho_{\mathrm{opt}}^\pi(\cdot, t=T) \right|_{\boldsymbol{x}_i} - \left. \rho_T(\cdot) \right|_{\boldsymbol{x}_i} \right)^2,$$

for each collocation point $\boldsymbol{x}_i$, $i = 1, \ldots, n$. For training the PINN, we minimize the overall loss (14) by solving

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \mathbb{R}^D}{\mathrm{argmin}} \, \mathcal{L}_{\mathcal{NN}}(\{\boldsymbol{x}_i\}_{i=1}^n; \boldsymbol{\theta}). \tag{15}$$

In the next Section, we detail the simulation set up and report the numerical results.

V. SIMULATION RESULTS

We consider the colloidal SA from [9], [11], where the free energy and the diffusion landscapes are

$$F(\langle C_6 \rangle, \pi) = a \, k_{\mathrm{B}} \, \theta \left(\langle C_6 \rangle - b - c\pi \right)^2, \tag{16}$$

$$D_2(\langle C_6 \rangle, \pi) = d \, \exp\left(-(\langle C_6 \rangle - b - c\pi)^2 \right) + f, \tag{17}$$

with known parameters $a = 10, b = 2.1, c = 0.75, d = 4.5 \times 10^{-3}, f = 0.5 \times 10^{-3}, k_{\mathrm{B}} = 1.38066 \times 10^{-23}$ Joules per Kelvin, and $\theta = 293$ Kelvin.

We use the DeepXDE library [19] for training the PINN. In particular, we choose a neural network with 3 hidden layers with 70 neurons in each layer. The activation functions are chosen to be $\tanh(\cdot)$. The input-output structure of the network are as explained in Sec. IV. For solving (15), we use the Adam optimizer [35].

We fix the final time $T = 200$ s, and consider the endpoint PDFs ρ_0, ρ_T shown in Fig. 1, represented as truncated normal PDFs (see e.g., [36, Sec. 2.2]):

$$\rho_i(x) := \begin{cases} \dfrac{1}{\sigma_i} \dfrac{\phi\left(\frac{x-\mu_i}{\sigma_i}\right)}{\Phi\left(\frac{b-\mu_i}{\sigma_i}\right) - \Phi\left(\frac{a-\mu_i}{\sigma_i}\right)} & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise,} \end{cases}$$

where $i \in \{0, T\}$, $\mu_0 = 0, \mu_T = 5, \sigma_0 = 0.2, \sigma_T = 0.1, a = 0, b = 6$. The functions $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF, and the standard normal cumulative distribution function, respectively. Recall from Sec. II-B that our proposed method is applicable for arbitrary compactly supported endpoint PDFs.

As shown in Fig. 3, we choose 1000 training points at each of the initial ($t = 0$ s) and terminal ($t = 200$ s) times, and another 5000 state-time points inside the domain $\Omega := [0, 6] \times [0, 200]$. We used the residual-based adaptive refinement method in [19] for generating these training points. As shown in Fig. 4, after 15,000 training epochs, the residuals for all loss functions go below 10^{-3}. The training was performed in Python 3 on a MacBook Air with 1.1 GHz Quad-Core Intel Core i5 processor and 8 GB memory. The computational time for training was 2097.40 seconds.

Fig. 5 shows the optimal policy $\pi_{\mathrm{opt}}(\langle C_6 \rangle, t)$ obtained as the output of the trained PINN. The value function $\psi(\langle C_6 \rangle, t)$ obtained as another output of the trained PINN is shown in Fig. 6. In Fig. 7, the snapshots of the optimally controlled PDFs ρ_{opt}^π obtained from the trained PINN are shown as solid curves with grey filled areas. To verify these results, we sampled 1000 initial states from the given ρ_0 using the Metropolis-Hastings [37] Markov Chain Monte Carlo algorithm, and then performed closed-loop simulations using the optimal policy $\pi_{\mathrm{opt}}(\langle C_6 \rangle, t)$ (shown in Fig. 5) obtained from the PINN. The stem plots shown in Fig. 7 depict the kernel density estimates (KDE) of these closed-loop trajectories. The KDE stems provide empirical approximations for the closed-loop optimally controlled PDFs, which match very well with the learnt solutions from PINN.

Fig. 8 shows the 1000 random sample paths for the closed loop simulation using the learnt optimal policy $\pi_{\mathrm{opt}}(\langle C_6 \rangle, t)$ that provably steers the given ρ_0 from $t = 0$ to the given ρ_T at $t = T$ over $[0, T] \equiv [0, 200]$ s.

VI. CONCLUSIONS

We propose formulating the finite horizon stochastic control problem for colloidal SA as a two point boundary value problem in the space of PDFs supported over the underlying state space. We develop the idea in detail for a univariate state model from the literature. We show that the resulting problem leads to a "nonlinear in state and non-affine in control" variant of the Schrödinger bridge problem (SBP), and derive the conditions for optimality for the same. We point out how the resulting system of equations differ from the existing SBP literature, and the difficulty in applying the standard computational approach of solving the SBP
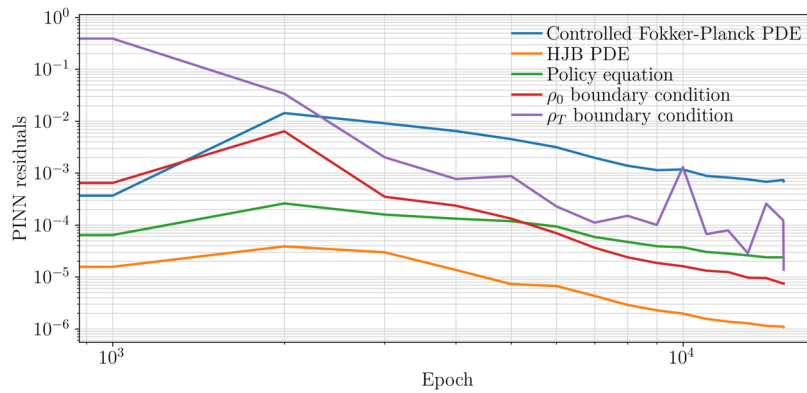
Fig. 4: The PINN residuals for solving the conditions for optimality (7)-(8).
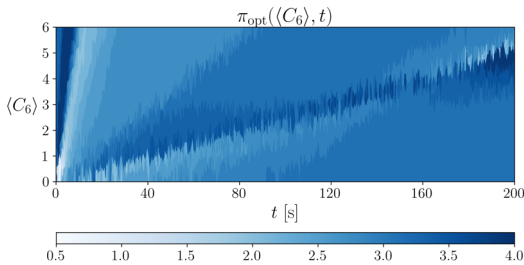

Fig. 5: The optimal policy $\pi_{\mathrm{opt}}(\langle C_6 \rangle, t)$ obtained as an output of the trained PINN solving the conditions for optimality (7)-(8).
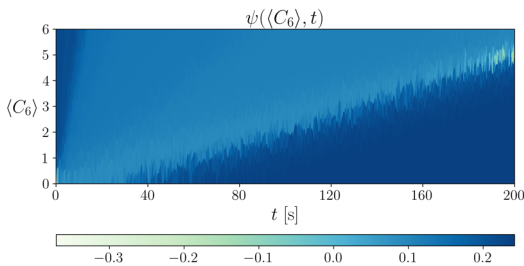

Fig. 6: The value function $\psi(\langle C_6 \rangle, t)$ obtained as an output of the trained PINN solving the conditions for optimality (7)-(8).

via fixed point recursions in our setting. These difficulties are fundamentally caused by both the drift and diffusion landscapes being control non-affine–a situation typical for the colloidal SA application. To circumvent these challenges, we adopt a learning based approach, and train a PINN to jointly learn the optimal control policy, the value function and the optimally controlled state PDFs. Our numerical experiments on a simple benchmark from the literature show that the proposed method performs very well.

REFERENCES

[1] G. M. Whitesides and B. Grzybowski, "Self-assembly at all scales," *Science*, vol. 295, no. 5564, pp. 2418–2421, 2002.
[2] J. A. Paulson, A. Mesbah, X. Zhu, M. C. Molaro, and R. D. Braatz, "Control of self-assembly in micro-and nano-scale systems," *Journal of Process Control*, vol. 27, pp. 38–49, 2015.
[3] X. Tang and M. A. Grover, "Control of microparticle assembly," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 491–514, 2022.
[4] J. J. Juárez and M. A. Bevan, "Feedback controlled colloidal self-assembly," *Advanced Functional Materials*, vol. 22, no. 18, pp. 3833–3839, 2012.
[5] D. T. Gillespie *et al.*, "Stochastic simulation of chemical kinetics," *Annual review of physical chemistry*, vol. 58, no. 1, pp. 35–55, 2007.
[6] D. T. Gillespie, A. Hellander, and L. R. Petzold, "Perspective: Stochastic algorithms for chemical kinetics," *The Journal of chemical physics*, vol. 138, no. 17, p. 05B201_1, 2013.
[7] E. M. Furst, "Directed self-assembly," *Soft Matter*, vol. 9, no. 38, pp. 9039–9045, 2013.
[8] J. A. Liddle and G. M. Gallatin, "Nanomanufacturing: a perspective," *ACS nano*, vol. 10, no. 3, pp. 2995–3014, 2016.
[9] X. Tang, Y. Xue, and M. A. Grover, "Colloidal self-assembly with model predictive control," in *2013 American Control Conference*. IEEE, 2013, pp. 4228–4233.
[10] X. Tang, B. Rupp, Y. Yang, T. D. Edwards, M. A. Grover, and M. A. Bevan, "Optimal feedback controlled assembly of perfect crystals," *ACS nano*, vol. 10, no. 7, pp. 6791–6798, 2016.
[11] Y. Xue, D. J. Beltran-Villegas, X. Tang, M. A. Bevan, and M. A. Grover, "Optimal design of a colloidal self-assembly process," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1956–1963, 2014.
[12] Y. Chen, T. T. Georgiou, and M. Pavon, "Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrodinger bridge," *SIAM Review*, vol. 63, no. 2, pp. 249–313, 2021.
[13] K. F. Caluya and A. Halder, "Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift," *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1163–1178, 2021.
[14] K. F. Caluya and A. Halder, "Reflected Schrödinger bridge: Density control with path constraints," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 1137–1142.
[15] I. Nodozi and A. Halder, "Schrödinger meets Kuramoto via Feynman-Kac: Minimum effort distribution steering for noisy nonuniform Kuramoto oscillators," *2022 IEEE Conference on Decision and Control (CDC), arXiv:2202.09734*, 2022.
[16] K. F. Caluya and A. Halder, "Finite horizon density control for static state feedback linearizable systems," *arXiv preprint arXiv:1904.02272*, 2019.
[17] K. F. Caluya and A. Halder, "Finite horizon density steering for multi-input state feedback linearizable systems," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3577–3582.
[18] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational physics*, vol. 378, pp. 686–707, 2019.
[19] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, "Deepxde: A deep learning library for solving differential equations," *SIAM Review*, vol. 63, no. 1, pp. 208–228, 2021.
[20] S.-R. Yeh, M. Seul, and B. I. Shraiman, "Assembly of ordered colloidal aggregates by electric-field-induced fluid flow," *Nature*, vol. 386, no. 6620, pp. 57–59, 1997.
[21] B. Oksendal, *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
[22] A. Friedman, *Partial differential equations of parabolic type*. Courier Dover Publications, 2008.
[23] E. Schrödinger, "Über die umkehrung der naturgesetze," *Sitzungsberichte der Preuss. Phys. Math. Klasse*, vol. 10, pp. 144–153, 1931.
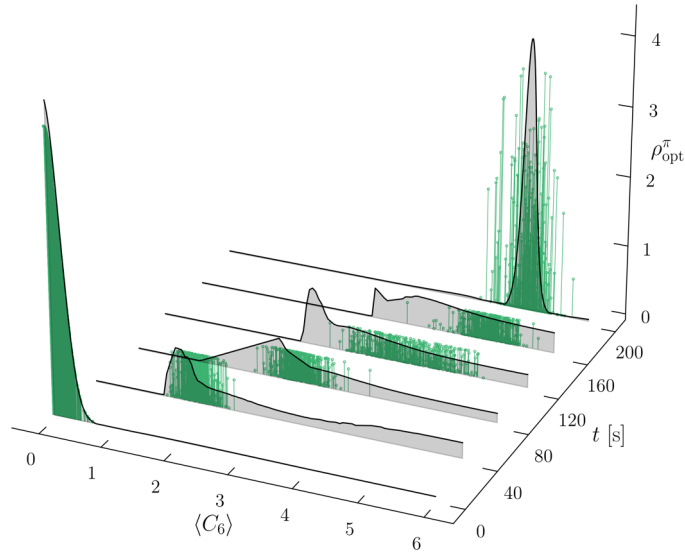
Fig. 7: Snapshots of the optimally controlled joint PDFs ρ_{opt}^{π} steering the state $\langle C_6 \rangle$ distribution from the given ρ_0 to ρ_T, as in Fig. 1, over the given time horizon $[0, T] \equiv [0, 200]$ s subject to the controlled noisy nonlinear sample path dynamics (1). The solid black curves with grey filled areas are obtained from the PINN. The stem plots are the KDE approximants of the optimally controlled PDF snapshots obtained from the closed-loop sample paths, as explained in Sec. V.
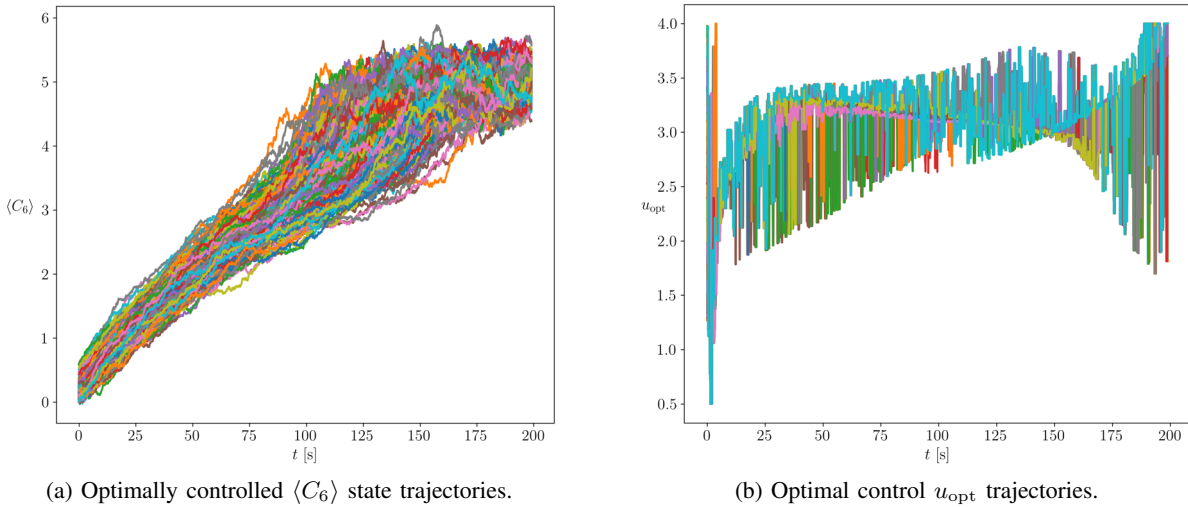

(a) Optimally controlled $\langle C_6 \rangle$ state trajectories.

(b) Optimal control u_{opt} trajectories.

Fig. 8: The 1000 random sample paths resulting from the closed loop simulation using the learnt optimal policy $\pi_{\text{opt}}\left(\langle C_6 \rangle, t\right)$.

[24] E. Schrödinger, "Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique," in *Annales de l'institut Henri Poincaré*, vol. 2, no. 4, 1932, pp. 269–310.

[25] B. Jamison, "The Markov processes of Schrödinger," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 32, no. 4, pp. 323–331, 1975.

[26] P. D. Pra and M. Pavon, "On the Markov processes of Schrödinger, the Feynman-Kac formula and stochastic control," in *Realization and Modelling in System Theory*. Springer, 1990, pp. 497–504.

[27] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.

[28] Y. Chen, T. T. Georgiou, and M. Pavon, "Controlling uncertainty," *IEEE Control Systems Magazine*, vol. 41, no. 4, pp. 82–94, 2021.

[29] S. Haddad, K. F. Caluya, A. Halder, and B. Singh, "Prediction and optimal feedback steering of probability density functions for safe automated driving," *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 2168–2173, 2020.

[30] Y. Chen, T. Georgiou, and M. Pavon, "Entropic and displacement interpolation: a computational approach using the Hilbert metric,"

SIAM Journal on Applied Mathematics, vol. 76, no. 6, pp. 2375–2396, 2016.

[31] W. H. Fleming, "Logarithmic transformations and stochastic control," in *Advances in Filtering and Optimal Stochastic Control*. Springer, 1982, pp. 131–141.

[32] H. J. Kappen, "Path integrals and symmetry breaking for optimal control theory," *Journal of statistical mechanics: theory and experiment*, vol. 2005, no. 11, p. P11011, 2005.

[33] E. Hopf, "The partial differential equation $u_t + uu_x = \mu_{xx}$," *Communications on Pure and Applied mathematics*, vol. 3, no. 3, pp. 201–230, 1950.

[34] J. D. Cole, "On a quasi-linear parabolic equation occurring in aerodynamics," *Quarterly of Applied Mathematics*, vol. 9, no. 3, pp. 225–236, 1951.

[35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[36] C. P. Robert, "Simulation of truncated normal variables," *Statistics and computing*, vol. 5, no. 2, pp. 121–125, 1995.

[37] W. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, pp. 97–109, 1970.