2021 American Control Conference (ACC)
New Orleans, USA, May 25-28, 2021

Reflected Schrödinger Bridge: Density Control with Path Constraints

Kenneth F. Caluya, and Abhishek Halder

***Abstract*— How to steer a given joint state probability density function to another over finite horizon subject to a controlled stochastic dynamics with hard state (sample path) constraints? In applications, state constraints may encode safety requirements such as obstacle avoidance. In this paper, we perform the feedback synthesis for minimum control effort density steering (a.k.a. Schrödinger bridge) problem subject to state constraints. We extend the theory of Schrödinger bridges to account the reflecting boundary conditions for the sample paths, and provide a computational framework building on our previous work on proximal recursions, to solve the same.**

I. Introduction

We consider finite horizon feedback steering of an ensemble of trajectories subject to a controlled stochastic differential equation (SDE) with endpoint joint state probability density function (PDF) constraints – a topic of growing interest in the systems-control literature. Motivating applications include belief space motion planning for vehicular autonomy, and the steering of robotic or biological swarms via decentralized feedback. While early contributions focused on the covariance control [1]–[3], more recent papers [4]–[6] addressed the optimal feedback synthesis for steering an arbitrary prescribed initial joint state PDF to another prescribed terminal joint state PDF subject to controlled linear dynamics, and revealed the connections between the associated stochastic optimal control problem, the theory of optimal mass transport [7], and the Schrödinger bridge [8], [9]. Follow up works have accounted terminal cost [10], input constraints [11], [12], output feedback [13], and some nonlinear dynamics [14]–[16]. The research front is fast moving and the mentioned references are only a representative sampler, far from comprehensive. As for the state or path constraints, prior work [17] incorporated the same in soft probabilistic sense. The contribution of the present paper is to account *hard deterministic path constraints* in the problem of minimum effort finite horizon PDF steering via feedback synthesis. This can be intuitively phrased as the "hard safety with soft endpoint" problem.

The proposed idea underlying the ensuing development is to modify the unconstrained Itô SDEs to the "reflected Itô SDEs" [18]–[21], i.e., the controlled sample paths in the state space (in addition to the control-affine deterministic drift) are driven by two stochastic processes: a Wiener process, and a local time stochastic process. The latter enforces the sample paths in the state space to satisfy the deterministic non-strict[1] path containment constraints at all times. These considerations engender a Schrödinger bridge-like formulation–referred hereafter as the *Reflected Schrödinger Bridge Problem (RSBP)*–which unlike its classical counterpart, has extra boundary conditions involving the gradients of the so-called Schrödinger factors. We show how recent developments in contraction mapping w.r.t. the Hilbert metric, and the proximal recursion over the Schrödinger factors can be harnessed to solve the RSBP.

II. Reflected Schrödinger Bridge Problem

A. Formulation

Consider a connected, smooth[2] and bounded domain $\mathcal{X} \subset \mathbb{R}^n$. Let $\overline{\mathcal{X}} := \mathcal{X} \cup \partial\mathcal{X}$ denote the closure of $\mathcal{X}$. For time $t \in [0,1]$, consider the stochastic control problem

$$\inf_{\boldsymbol{u}\in\mathcal{U}} \quad \mathbb{E}\left\{\int_0^1 \frac{1}{2}\|\boldsymbol{u}(t,\boldsymbol{x}_t^{\boldsymbol{u}})\|_2^2 \,\mathrm{d}t\right\} \tag{1a}$$

$$\text{subject to} \quad \mathrm{d}\boldsymbol{x}_t^{\boldsymbol{u}} = \boldsymbol{f}(t,\boldsymbol{x}_t^{\boldsymbol{u}})\,\mathrm{d}t + \boldsymbol{u}(t,\boldsymbol{x}_t^{\boldsymbol{u}})\,\mathrm{d}t + \sqrt{2\theta}\,\mathrm{d}\boldsymbol{w}_t + \boldsymbol{n}(\boldsymbol{x}_t^{\boldsymbol{u}})\mathrm{d}\gamma_t, \tag{1b}$$

$$\boldsymbol{x}_0^{\boldsymbol{u}} := \boldsymbol{x}_t^{u}\,(t=0) \sim \rho_0, \quad \boldsymbol{x}_1^{\boldsymbol{u}} := \boldsymbol{x}_t^{u}\,(t=1) \sim \rho_1, \tag{1c}$$

where $\boldsymbol{w}_t$ is the standard Wiener process in $\mathbb{R}^n$, the controlled state $\boldsymbol{x}_t^{\boldsymbol{u}} \in \overline{\mathcal{X}}$, and the endpoint joint state PDFs ρ_0, ρ_1 are prescribed[3] such that their supports are in $\overline{\mathcal{X}}$, both are everywhere nonnegative, have finite second moments, and $\int \rho_0 = \int \rho_1 = 1$. The parameter $\theta > 0$ is referred to as the thermodynamic temperature, and the expectation operator $\mathbb{E}\{\cdot\}$ in (1a) is w.r.t. the law of the controlled state $\boldsymbol{x}_t^{\boldsymbol{u}}$. The set $\mathcal{U}$ consists of all admissible feedback policies $\boldsymbol{u}(t,\boldsymbol{x}_t^u)$, given by $\mathcal{U} := \{\boldsymbol{u} : [0,1]\times\overline{\mathcal{X}} \mapsto \mathbb{R}^n \mid \|\boldsymbol{u}\|_2^2 < \infty, \boldsymbol{u}(t,\cdot) \in \text{Lipschitz}\left(\overline{\mathcal{X}}\right) \text{ for all } t \in [0,1]\}$. We assume that the prior drift vector field $\boldsymbol{f}$ is bounded Borel measurable in $(t,\boldsymbol{x}_t^{\boldsymbol{u}}) \in [0,1]\times\overline{\mathcal{X}}$, and Lipschitz continuous w.r.t. $\boldsymbol{x}_t^{\boldsymbol{u}} \in \overline{\mathcal{X}}$. The vector field $\boldsymbol{n}$ is set to be the inward unit normal to the boundary $\partial\mathcal{X}$, and gives the direction of reflection. Furthermore, for $t \in [0,1]$, γ_t is *minimal local time*: a continuous, nonnegative and nondecreasing stochastic process [22]–[24] that restricts $\boldsymbol{x}_t^u$ to the domain $\mathcal{X}$, with $\gamma_0 \equiv 0$. Specifically, letting $\mathbb{1}_{\{\}}$ denote the indicator function of the subscripted set, we have

$$\gamma_t = \int_0^t \mathbb{1}_{\{\boldsymbol{x}_s^{\boldsymbol{u}}\in\partial\mathcal{X}\}}\,\mathrm{d}\gamma_s, \quad \int_0^1 \mathbb{1}_{\{\boldsymbol{x}_t^{\boldsymbol{u}}\notin\partial\mathcal{X}\}}\,\mathrm{d}\gamma_t = 0, \tag{2}$$

[1]There is no loss of generality in allowing the sample paths to satisfy *non-strict* path containment in given $\mathcal{X} \subset \mathbb{R}^n$ since *strict* containment can be enforced by reflecting them from ϵ-inner boundary layer of $\partial\mathcal{X}$ for ϵ small enough.

[2]More precisely, there exists $\xi \in C_b^2(\mathbb{R}^n)$ such that $\mathcal{X} \equiv \{\boldsymbol{x} \in \mathbb{R}^n \mid \xi(\boldsymbol{x}) > 0\}$ with boundary $\partial\mathcal{X} \equiv \{\boldsymbol{x} \in \mathbb{R}^n \mid \xi(\boldsymbol{x}) = 0\}$.

[3]The notation $\boldsymbol{x} \sim \rho$ means that the random vector $\boldsymbol{x}$ has joint PDF ρ.

Kenneth F. Caluya, and Abhishek Halder are with the Department of Applied Mathematics, University of California, Santa Cruz, CA 95064, USA, {kcaluya,ahalder}@ucsc.edu. This research was partially supported by NSF award 1923278.

which is to say that the process γ_t only increases at times $t \in [0,1]$ when $\boldsymbol{x}_t^{\boldsymbol{u}}$ hits the boundary, i.e., when $\boldsymbol{x}_t^{\boldsymbol{u}} \in \partial\mathcal{X}$. Thus, (1b) is a controlled reflected SDE, and the tuple $(\boldsymbol{x}_t^{\boldsymbol{u}}, \gamma_t)$ solves the *Skorokhod problem* [25]–[27]. We point the readers to [20, Sec. 3] for proofs of existence-uniqueness of solutions to (1b) under the stated regularity assumptions.

To formalize the probabilistic setting of the problem at hand, let Ω be the space of continuous functions $\omega : [0,1] \mapsto \overline{\mathcal{X}}$. We view Ω as a complete separable metric space endowed with the topology of uniform convergence on compact time intervals. With Ω, we associate the σ-algebra $\mathscr{F} = \sigma\{\omega(s) \mid 0 \leq s \leq 1\}$. Consider the complete filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with filtration $\mathscr{F}_t = \sigma\{\omega(s) \mid 0 \leq s \leq t \leq 1\}$ wherein "complete" means that $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets, and $\mathscr{F}_t$ is right continuous. The processes $\boldsymbol{w}_t$, $\boldsymbol{x}_t^{\boldsymbol{u}}$ (for a given feedback policy $\boldsymbol{u}$) and γ_t are $\mathscr{F}_t$-adapted (i.e., non-anticipating) for $t \in [0,1]$. In (1c), the random vectors $\boldsymbol{x}_0^{\boldsymbol{u}}$ and $\boldsymbol{x}_1^{\boldsymbol{u}}$ are respectively $\mathscr{F}_0$-measurable and $\mathscr{F}_1$-measurable.

Denote the Euclidean gradient operator as ∇, the inner product as $\langle \cdot, \cdot \rangle$, and the Laplacian as Δ. Letting

$$\mathcal{L} := \theta\Delta + \langle \boldsymbol{f} + \boldsymbol{u}, \nabla \rangle,$$

the law of the sample path of (1b) can be characterized [28] as follows: for each $\boldsymbol{x} \in \overline{\mathcal{X}}$, there is a *unique* probability measure $\mathbb{P}_{\boldsymbol{x}}^{\mu}$ on Ω such that (i) $\mathbb{P}_{\boldsymbol{x}}^{\mu}\left(\boldsymbol{x}_t^{\boldsymbol{u}}(t=0) = \boldsymbol{x}\right) = 1$, (ii) for any $\phi \in C_c^{1,2}\left([0,1]; \overline{\mathcal{X}}\right)$ whose inner normal derivative on $\partial\mathcal{X}$ is nonnegative,

$$\phi(t, \boldsymbol{x}_t^{\boldsymbol{u}}) - \int_0^t \left(\frac{\partial \phi}{\partial s} + \mathcal{L}\phi\right)(s, \boldsymbol{x}_s^{\boldsymbol{u}})\, \mathrm{d}s$$

is $\mathbb{P}_{\boldsymbol{x}}^{\mu}$-submartingale, and (iii) there is a continuous, nonnegative, nondecreasing stochastic process γ_t satisfying (2). As a consequence [28, p. 196] of this characterization it follows that the process $\boldsymbol{x}_t^{\boldsymbol{u}}$ is Feller continuous and strongly Markov. In particular, the measure-valued trajectory $\mathbb{P}_{\boldsymbol{x}_t^{\boldsymbol{u}}}^{\mu(t)}$ comprises of absolutely continuous measures w.r.t. Lebesgue measure.

The objective in problem (1) is to perform the minimum control effort steering of the given initial state PDF ρ_0 at $t = 0$ to the given terminal state PDF ρ_1 at $t = 1$ subject to the controlled sample path dynamics (1b). In other words, the data of the problem consists of the domain $\overline{\mathcal{X}}$, the prior dynamics data $\boldsymbol{f}, \theta$, and the two endpoint PDFs ρ_0, ρ_1.

Formally, we can transcribe (1) into the following variational problem [29]:

$$\inf_{(\rho,\boldsymbol{u}) \in \mathcal{P}_2(\overline{\mathcal{X}}) \times \mathcal{U}} \int_0^1 \int_{\overline{\mathcal{X}}} \frac{1}{2} \|\boldsymbol{u}(t, \boldsymbol{x}_t^u)\|_2^2\, \rho(t, \boldsymbol{x}_t^u)\, \mathrm{d}\boldsymbol{x}_t^u \mathrm{d}t \tag{3a}$$

$$\text{subject to} \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho(\boldsymbol{u} + \boldsymbol{f})) = \theta\Delta\rho, \tag{3b}$$

$$\langle -(\boldsymbol{u} + \boldsymbol{f})\rho + \theta\nabla\rho, \boldsymbol{n} \rangle\big|_{\partial\mathcal{X}} = 0, \tag{3c}$$

$$\rho(0, \boldsymbol{x}_t^u) = \rho_0, \quad \rho(1, \boldsymbol{x}_t^u) = \rho_1, \tag{3d}$$

where a PDF-valued curve $\rho(t, \cdot) \in \mathcal{P}_2(\overline{\mathcal{X}})$ if for each $t \in [0,1]$, the PDF ρ is supported on $\overline{\mathcal{X}}$, and has finite second moment. In this paper, we will not focus on the rather technical direction of establishing the existence of minimizer for (3), which can be pursued along the lines of [7, p. 243–245]. Instead, we will formally derive the conditions of optimality, convert them to the so-called Schrödinger system, and argue the existence-uniqueness of solutions for the same.

B. Necessary Conditions of Optimality

The following result (see [30, Appendix A] for proof) summarizes how the optimal pair $(\rho^{\text{opt}}, \boldsymbol{u}^{\text{opt}})$ for problem (3) can be obtained.

*Theorem 1 (**Optimal control and optimal state PDF**):* A pair $(\rho^{\text{opt}}, \boldsymbol{u}^{\text{opt}})$ solving the variational problem (3) must satisfy the system of coupled nonlinear PDEs:

$$\frac{\partial \rho^{\text{opt}}}{\partial t} + \nabla \cdot \left(\rho^{\text{opt}}(\nabla\psi + \boldsymbol{f})\right) = \theta\Delta\rho^{\text{opt}}, \tag{4a}$$

$$\frac{\partial \psi}{\partial t} + \frac{1}{2}\|\nabla\psi\|_2^2 + \langle \nabla\psi, \boldsymbol{f} \rangle = -\theta\Delta\psi, \tag{4b}$$

where

$$\boldsymbol{u}^{\text{opt}}(t, \cdot) = \nabla\psi(t, \cdot), \tag{5}$$

subject to the boundary conditions

$$\langle \nabla\psi, \boldsymbol{n} \rangle\big|_{\partial\mathcal{X}} = 0, \quad \text{for all} \quad t \in [0,1], \tag{6a}$$

$$\rho^{\text{opt}}(0, \cdot) = \rho_0, \quad \rho^{\text{opt}}(1, \cdot) = \rho_1, \tag{6b}$$

$$\langle \rho^{\text{opt}}(\nabla\psi + \boldsymbol{f}) - \theta\nabla\rho^{\text{opt}}, \boldsymbol{n} \rangle\big|_{\partial\mathcal{X}} = 0, \quad \text{for all} \quad t \in [0,1]. \tag{6c}$$

The PDE (4a) is a controlled Fokker-Planck-Kolmogorov (FPK) equation, and (4b) is a Hamilton-Jacobi-Bellman (HJB) equation. Because the equations (4a)-(4b) have one way coupling, and the boundary conditions (6a)-(6c) are atypical, solving (4) is a challenging task in general. In the following, we show that it is possible to transform the *coupled nonlinear* system (4) into a boundary coupled *linear* system of PDEs which we refer to as the *Schrödinger system*. We will see that the resulting system paves way to a computational pipeline for solving the density steering problem with path constraints.

Remark 1: Obviously, it is possible to derive (4) by performing a change of variable $\widetilde{\boldsymbol{u}} := \boldsymbol{f} + \boldsymbol{u}$ for the drift in (1b), and modifying (1a) accordingly. This does not trivialize the development because it will turn out that the optimal feedback computation will require us to propagate the uncontrolled version of the nonlinear SDE (1b) with $\boldsymbol{f}$ as well as the minimal local time term unaltered. Since our proposed *gridless* computation in Sec. IV-B will be based on performing proximal recursion on weighted point clouds, it is nontrivial how to numerically enforce the reflecting boundary conditions for these sample paths. See Remark 4.

C. Schrödinger System

We now apply the Hopf-Cole transform [31], [32] to the system of nonlinear PDEs (4).

*Theorem 2 (**Schrödinger system**):* (see [30, Appendix B] for proof) Given $\overline{\mathcal{X}}, \boldsymbol{f}, \theta, \rho_0, \rho_1$ for problem (3), consider the Hopf-Cole transform $(\rho^{\text{opt}}, \psi) \mapsto (\varphi, \hat{\varphi})$ given by

$$\varphi(t, \cdot) := \exp\left(\psi(t, \cdot)/2\theta\right), \tag{7a}$$

$$\hat{\varphi}(t, \cdot) := \rho^{\text{opt}}(t, \cdot)\exp\left(-\psi(t, \cdot)/2\theta\right), \tag{7b}$$

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
\hat{\varphi}_0(\boldsymbol{x}) \arrow[r, "\int \text{with b.c. } \langle \boldsymbol{f}\hat{\varphi} - \theta\nabla\hat{\varphi}, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = 0"] & \hat{\varphi}_1(\boldsymbol{x}) \arrow[d, "\rho_1(\boldsymbol{x}) \oslash \hat{\varphi}_1(\boldsymbol{x})"] \\
\varphi_0(\boldsymbol{x}) \arrow[u, "\rho_0(\boldsymbol{x}) \oslash \varphi_0(\boldsymbol{x})"] & \varphi_1(\boldsymbol{x}) \arrow[l, "\int \text{with b.c. } \langle \nabla\varphi, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = 0"]
\end{tikzcd}
$$

Fig. 1: Schematic of the fixed point recursion for the Schrödinger system (8)-(9). The abbreviation "b.c." stands for boundary condition, the symbol $\oslash$ denotes the Hadamard division.

applied to (4) where $t \in [0,1]$. For $k \in \{0,1\}$, introduce the notation $\varphi_k := \varphi(k,\cdot)$, $\hat{\varphi}_k := \hat{\varphi}(k,\cdot)$. Then the pair $(\varphi,\hat{\varphi})$ satisfies the system of linear PDEs

$$\frac{\partial \varphi}{\partial t} = -\langle \nabla\varphi, \boldsymbol{f}\rangle - \theta\Delta\varphi, \tag{8a}$$

$$\frac{\partial \hat{\varphi}}{\partial t} = -\nabla\cdot(\boldsymbol{f}\hat{\varphi}) + \theta\Delta\hat{\varphi}, \tag{8b}$$

subject to the boundary conditions

$$\varphi_0\hat{\varphi}_0 = \rho_0, \quad \varphi_1\hat{\varphi}_1 = \rho_1, \tag{9a}$$

$$\langle \nabla\varphi, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = \langle \boldsymbol{f}\hat{\varphi} - \theta\nabla\hat{\varphi}, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = 0. \tag{9b}$$

For all $t \in [0,1]$, the pair $(\rho^{\text{opt}}, \boldsymbol{u}^{\text{opt}})$ can be recovered as

$$\rho^{\text{opt}}(t,\cdot) = \varphi(t,\cdot)\hat{\varphi}(t,\cdot), \;\; \boldsymbol{u}^{\text{opt}}(t,\cdot) = 2\theta\nabla\log\varphi(t,\cdot). \tag{10}$$

Remark 2: From (7), both $\varphi,\hat{\varphi}$ are nonnegative by definition, and strictly positive if ψ is bounded and ρ^{opt} is positive.

Remark 3: Under the regularity assumptions on $\boldsymbol{f}$ and $\overline{\mathcal{X}}$ stated in Section II-A, the process $\boldsymbol{x}_t$ satisfying the *uncontrolled* reflected Itô SDE

$$\mathrm{d}\boldsymbol{x}_t = \boldsymbol{f}(t,\boldsymbol{x}_t)\,\mathrm{d}t + \sqrt{2\theta}\,\mathrm{d}\boldsymbol{w}_t + \boldsymbol{n}(\boldsymbol{x}_t)\,\mathrm{d}\gamma_t, \; t \in [0,1], \tag{11}$$

is a Feller continuous strongly Markov process. Therefore, the theory of semigroups applies and the transition density of (11) satisfies Kolmogorov's equations. Notice that the transition density or Green's function will depend on the domain $\overline{\mathcal{X}}$. In particular, we point out that (8a) is the *backward Kolmogorov equation* in unkonwn φ with the corresponding *Neumann boundary condition* $\langle \nabla\varphi, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = 0$ in (9b). On the other hand, (8b) is the *forward Kolmogorov equation* in unkonwn $\hat{\varphi}$ with the corresponding *Robin boundary condition* $\langle \boldsymbol{f}\hat{\varphi} - \theta\nabla\hat{\varphi}, \boldsymbol{n}\rangle\big|_{\partial\mathcal{X}} = 0$ in (9b). These "backward Kolmogorov with Neumann" and "forward Kolmogorov with Robin" system of PDE boundary value problems are coupled via the atypical boundary conditions (9a).

Theorem 2 reduces finding the optimal pair $(\rho^{\text{opt}}, \boldsymbol{u}^{\text{opt}})$ for the RSBP to that of finding the pair[4] $(\varphi(t,\boldsymbol{x}_t), \hat{\varphi}(t,\boldsymbol{x}_t))$ associated with the uncontrolled SDE (11). To do so, we need to compute the terminal-initial condition pair $(\varphi_1,\hat{\varphi}_0)$, which can be obtained by first making an initial guess for $(\varphi_1,\hat{\varphi}_0)$ and then performing time update by integrating the system (8)-(9b). Using (9a), this then sets up a fixed point recursion over the pair $(\varphi_1,\hat{\varphi}_0)$ (see Fig. 1). If this recursion converges to a unique pair, then the converged pair $(\varphi_1,\hat{\varphi}_0)$ can be used to compute the transient factors $(\varphi(t,\boldsymbol{x}_t), \hat{\varphi}(t,\boldsymbol{x}_t))$, and we can recover $(\rho^{\text{opt}}, \boldsymbol{u}^{\text{opt}})$ via (10). This computational pipeline will be pursued in this paper.

Since the PDEs in (8) are linear, and the boundary couplings in (9a) are in product form, the nonnegative function pair $(\varphi_1,\hat{\varphi}_0)$ can only be unique in the projective sense, i.e., if $(\varphi_1,\hat{\varphi}_0)$ is a solution then so is $(\alpha\varphi_1, \hat{\varphi}_0/\alpha)$ for any $\alpha > 0$. In [33], it was shown that the aforesaid fixed point recursion is in fact *contractive* on a suitable cone in Hilbert's projective metric, and hence guaranteed to converge to a unique pair $(\varphi_1,\hat{\varphi}_0)$, provided that the transition density for (11) is positive and continuous[5] on $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$ for all $t \in [0,1]$, and ρ_0, ρ_1 are supported on compact subsets of $\overline{\mathcal{X}}$.

III. Case Study: RSBP in 1D without Prior Drift

To illustrate the ideas presented thus far, we now consider a simple instance of problem (3) over the state space $\overline{\mathcal{X}} = [a,b] \subset \mathbb{R}$, and with the prior drift $f \equiv 0$. That is to say, we consider the finite horizon density steering subject to the controlled two-sided reflected Brownian motion. Using some properties of the associated Markov kernel, we will show that the Schrödinger system (8)-(9) corresponding to this particular RSBP has a unique solution which can be obtained by the kind of fixed point recursion mentioned toward the end of Section II-C.

In this case, the Schrödinger system (8)-(9) reduces to

$$\frac{\partial \varphi}{\partial t} = -\theta\frac{\partial^2\varphi}{\partial x^2}, \tag{12a}$$

$$\frac{\partial \hat{\varphi}}{\partial t} = \theta\frac{\partial^2\hat{\varphi}}{\partial x^2}, \tag{12b}$$

$$\varphi_0\hat{\varphi}_0 = \rho_0, \quad \varphi_1\hat{\varphi}_1 = \rho_1, \tag{12c}$$

$$\frac{\partial \varphi}{\partial x}\bigg|_{x=a,b} = \frac{\partial \hat{\varphi}}{\partial x}\bigg|_{x=a,b} = 0. \tag{12d}$$

Notice that (12a)-(12b) are the backward and forward heat PDEs, respectively, which subject to (12d), have solutions

$$\varphi(x,t) = \int_{[a,b]} K_\theta(x,y,1-t)\varphi_1(y)\,\mathrm{d}y, \quad t \leq 1, \tag{13a}$$

$$\hat{\varphi}(x,t) = \int_{[a,b]} K_\theta(y,x,t)\hat{\varphi}_0(y)\,\mathrm{d}y, \qquad t \geq 0, \tag{13b}$$

where

$$K_\theta(x,y,t) := \frac{1}{b-a} + \frac{2}{b-a}\sum_{m=1}^{\infty}\exp\left(-\frac{\theta\pi^2m^2}{(b-a)^2}t\right) \times \cos\left(\frac{m\pi(x-a)}{b-a}\right)\cos\left(\frac{m\pi(y-a)}{b-a}\right) \tag{14}$$

[4]We refer to $\varphi(t,\boldsymbol{x}_t), \hat{\varphi}(t,\boldsymbol{x}_t)$ as the *Schrödinger factors*.

[5]Under the regularity assumptions on $\boldsymbol{f}$ and $\overline{\mathcal{X}}$ stated in Section II-A, the transition density for (11) indeed satisfies these conditions.

is the Markov kernel or transition density [34, Sec. 4.1], [35, p. 410-411] associated with the uncontrolled reflected SDE

$$\mathrm{d}x_t = \sqrt{2\theta}\,\mathrm{d}w_t + \mathrm{d}L_t - \mathrm{d}U_t, \quad t \in [0,1]. \tag{15}$$

In (15), L_t, U_t are the two local time stochastic processes [22], [23] at the lower and upper boundaries respectively, which restrict x_t to the interval $[a,b]$; see [30, Fig. 2].

Combining (13) and (12c), we get a system of coupled nonlinear integral equations in unknowns $(\varphi_1, \hat{\varphi}_0)$, given by

$$\rho_0(x) = \hat{\varphi}_0(x) \int_{[a,b]} K_\theta(x,y,1)\varphi_1(y)\,\mathrm{d}y, \tag{16a}$$

$$\rho_1(x) = \varphi_1(x) \int_{[a,b]} K_\theta(y,x,1)\hat{\varphi}_0(y)\,\mathrm{d}y. \tag{16b}$$

Clearly, solving (16) is equivalent to solving (12). The pair $(\varphi_1, \hat{\varphi}_0)$ can be solved from (16) iteratively as a fixed point recursion with guaranteed convergence established through contraction mapping in Hilbert's projective metric; see [33]. The Lemma 1 stated next will be used in the Proposition 1 that follows, showing the existence and uniqueness of the pair $(\varphi_1, \hat{\varphi}_0)$ in (16) as well as the fact that the aforesaid fixed point recursion is guaranteed to converge to that pair.

Lemma 1: (see [30, Appendix C] for proof) For $0 < \theta$, $a < b$, consider the transition probability density $K_\theta(x,y,t)$ in (14). Then,

(i) $K_\theta(x,y,t=1)$ is continuous on the set $[a,b]\times[a,b]$.
(ii) $K_\theta(x,y,t=1) > 0$ for all $(x,y) \in [a,b]\times[a,b]$.

Proposition 1: (see [30, Appendix D] for proof) Given $0 < \theta$, $a < b$, and the endpoint PDFs ρ_0, ρ_1 having compact supports $\subseteq [a,b]$. There exists a unique pair $(\varphi_1, \hat{\varphi}_0)$ that solves (16), and equivalently (12). This unique pair can be computed by the fixed point recursion shown in Fig. 1.

To illustrate how the above results can be used for practical computation, consider solving the RSBP (1) with $f \equiv 0$, $\theta = 0.5$, $\overline{\mathcal{X}} = [a,b] \equiv [-4,4]$, and ρ_0, ρ_1 as (see Fig. 2)

$$\rho_0(x) \propto 1 + (x^2-16)^2 \exp(-x/2), \tag{17a}$$

$$\rho_1(x) \propto 1.2 - \cos(\pi(x+4)/2), \tag{17b}$$

where the supports of (17) are restricted to $[-4,4]$, and the proportionality constants are determined accordingly. For state feedback synthesis enabling this unimodal to bimodal steering over $t \in [0,1]$, we performed the fixed point recursion over the pair $(\varphi_1, \hat{\varphi}_0)$ using (16) with ρ_0, ρ_1 as in (17), and K_θ given by (14). For numerical implementation, we truncated the infinite sum in (14) after the first 100 terms. The converged pair $(\varphi_1, \hat{\varphi}_0)$ is used to compute the transient Schrödinger factors $(\varphi(t,\boldsymbol{x}_t), \hat{\varphi}(t,\boldsymbol{x}_t))$ via (13), and then the pair $(\rho^{\mathrm{opt}}(t,\boldsymbol{x}_t^u), \boldsymbol{u}^{\mathrm{opt}}(t,\boldsymbol{x}_t^u))$ via (10). Fig. 3 depicts the evolution of the optimal controlled transient joint state PDFs $\rho^{\mathrm{opt}}(t,x_t^u)$ as well as 100 sample paths x_t^u of the optimal closed-loop reflected SDE. These sample paths were computed by applying the Euler-Maruyama scheme with time-step size 10^{-3}. Notice from Fig. 3 that (i) the closed-loop sample paths satisfy $-4 \leq x_t^u \leq 4$ for all $t \in [0,1]$, and (ii) in the absence of feedback, the terminal constraint $\rho(1,x_1^u) = \rho_1$ (given by (17b)) cannot be satisfied.

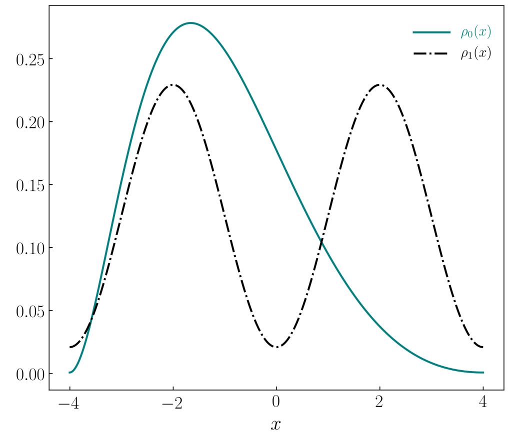

Fig. 2: The endpoint PDFs ρ_0, ρ_1 shown above are supported on $[-4,4]$, and are given by (17).

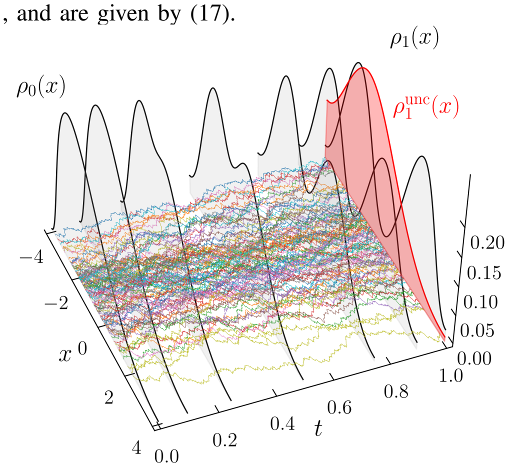

Fig. 3: Shown as the black curves are the optimal controlled transient joint state PDFs $\rho^{\mathrm{opt}}(t,x_t^u)$ for steering the two-sided reflecting Brownian motion with endpoint PDFs ρ_0, ρ_1 as in Fig. 2. The red curve ρ_1^{unc} is the uncontrolled state PDF at $t=1$, i.e., obtained by setting $u \equiv 0$. Also depicted are the 100 sample paths of the optimally controlled (i.e., closed-loop) reflected SDE. This simulation corresponds to the RSBP (1) with problem data $f \equiv 0, [a,b] = [-4,4], \theta = 0.5$, and ρ_0, ρ_1 given by (17).

IV. RSBP with Prior Drift

For generic $\boldsymbol{f}$, $\overline{\mathcal{X}}$, there is no closed-form expression of the Markov Kernel associated with (8)-(9b). Hence, unlike the situation in Section III, we cannot explicitly set up coupled integral equations of the form (16), thus preventing the numerical implementation of the fixed point recursion (Fig. 1) via direct matrix-vector recursion. In this Section, we will show that if $\boldsymbol{f}$ is gradient of a potential, then we can reformulate (8)-(9) in a way that leads to a variational recursion which in turn enables us to implement the fixed point recursion (Fig. 1) in an implicit manner.

A. Reformulation of the Schrödinger System

Let $\boldsymbol{f}$ be a gradient vector field, i.e., $\boldsymbol{f} = -\nabla V$ for some potential $V \in C^2(\overline{\mathcal{X}})$. The associated Schrödinger system (8)-(9) becomes

$$\frac{\partial \varphi}{\partial t} = \langle \nabla\varphi, \nabla V \rangle - \theta\Delta\varphi, \tag{18a}$$

$$\frac{\partial \hat{\varphi}}{\partial t} = \nabla \cdot (\nabla V \hat{\varphi}) + \theta \Delta \hat{\varphi}, \tag{18b}$$

$$\varphi_0 \hat{\varphi}_0 = \rho_0, \quad \varphi_1 \hat{\varphi}_1 = \rho_1, \tag{18c}$$

$$\langle \nabla \varphi, \boldsymbol{n} \rangle\big|_{\partial \mathcal{X}} = \langle \nabla V \hat{\varphi} + \theta \nabla \hat{\varphi}, \boldsymbol{n} \rangle\big|_{\partial \mathcal{X}} = 0. \tag{18d}$$

The idea now is to exploit the structural nonlinearities in (18) to design an algorithm that allows computing the Schrödinger factors $(\varphi, \hat{\varphi})$. To that end, the following result (proof in [30, Appendix E]) is a crucial step.

Theorem 3: Given $V \in C^2(\overline{\mathcal{X}}), \theta > 0$, and $t \in [0,1]$, consider $\varphi(t, \boldsymbol{x}_t)$ in (18). Let $s := 1 - t$, and define the mappings $\varphi \mapsto q \mapsto p$ given by $q(s, \boldsymbol{x}_s) := \varphi(t, \boldsymbol{x}_t) = \varphi(1-s, \boldsymbol{x}_{1-s})$, $p(s, \boldsymbol{x}_s) := q(s, \boldsymbol{x}_s) \exp(-V(\boldsymbol{x}_s)/\theta)$. Then $p(s, \boldsymbol{x}_s)$ solves the PDE initial boundary value problem:

$$\frac{\partial p}{\partial s} = \nabla \cdot (p \nabla V) + \theta \Delta p, \tag{19a}$$

$$p(0, \boldsymbol{x}) = \varphi_1(\boldsymbol{x}) \exp(-V(\boldsymbol{x})/\theta), \tag{19b}$$

$$\langle \nabla V p + \theta \nabla p, \boldsymbol{n} \rangle\big|_{\partial \mathcal{X}} = 0. \tag{19c}$$

Thanks to Theorem 3, solving (18) is equivalent to solving

$$\frac{\partial p}{\partial s} = \nabla \cdot (p \nabla V) + \theta \Delta p, \tag{20a}$$

$$\frac{\partial \hat{\varphi}}{\partial t} = \nabla \cdot (\nabla V \hat{\varphi}) + \theta \Delta \hat{\varphi}, \tag{20b}$$

$$p(s=1, \boldsymbol{x}) \exp(V(\boldsymbol{x})/\theta) \hat{\varphi}_0(\boldsymbol{x}) = \rho_0,$$

$$p(s=0, \boldsymbol{x}) \exp(V(\boldsymbol{x})/\theta) \hat{\varphi}_1(\boldsymbol{x}) = \rho_1, \tag{20c}$$

$$\langle \nabla V p + \theta \nabla p, \boldsymbol{n} \rangle\big|_{\partial \mathcal{X}} = \langle \nabla V \hat{\varphi} + \theta \nabla \hat{\varphi}, \boldsymbol{n} \rangle\big|_{\partial \mathcal{X}} = 0. \tag{20d}$$

From (20a)-(20b), φ and p satisfy the exact same FPK PDE with different initial conditions and integrated in different time coordinates t and s. From (20d), φ and p satisfy the same Robin boundary condition. Therefore, a single FPK initial boundary value problem solver can be used to set up the fixed point recursion to solve for $(p_1, \hat{\varphi}_0)$, and hence $(p(s, \boldsymbol{x}_s), \hat{\varphi}(t, \boldsymbol{x}_t))$. From p, we can recover φ as

$$\varphi(t, \boldsymbol{x}_t) = \varphi(1-s, \boldsymbol{x}_{1-s}) = p(s, \boldsymbol{x}_s) \exp(-V(\boldsymbol{x}_s)/\theta).$$

B. Computation via Wasserstein Proximal Recursion

Building on our previous works [16], [36], [37], we propose proximal recursions to numerically time march the solutions of the PDE initial boundary value problems (20) by exploiting certain infinite dimensional gradient descent structure. This enables us to perform the computation associated with the horizontal arrows in Fig. 1, and hence the fixed point recursions to solve for $(p, \hat{\varphi})$, and consequently for $(\varphi, \hat{\varphi})$. We give here a brief outline of the ideas behind these proximal recursions.

It is well-known [38], [39] that the flows generated by (20a),(20b),(20d) can be viewed as the gradient descent of the Lyapunov functional

$$F(\varrho) := \int_{\overline{\mathcal{X}}} V(\boldsymbol{x}) \varrho(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} + \theta \int_{\overline{\mathcal{X}}} \varrho(\boldsymbol{x}) \log \varrho(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \tag{21}$$

w.r.t. the distance metric W referred to as the (quadratic) Wassertein metric [7] on $\mathcal{P}_2(\overline{\mathcal{X}})$. For chosen time-steps τ, σ, this allows us to set up a variational recursion over the discrete time pair $(t_{k-1}, s_{k-1}) := ((k-1)\tau, (k-1)\sigma)$ as

$$\begin{pmatrix} \hat{\phi}_{t_k} \\ \varpi_{s_k} \end{pmatrix} = \begin{pmatrix} \operatorname{prox}_{\tau F}^{W^2}(\hat{\phi}_{t_{k-1}}) \\ \operatorname{prox}_{\sigma F}^{W^2}(\varpi_{s_{k-1}}) \end{pmatrix}, \quad k \in \mathbb{N}, \tag{22}$$

wherein the Wasserstein proximal operator

$$\operatorname{prox}_{hF}^{W^2}(\cdot) := \underset{\varrho \in \mathcal{P}_2(\overline{\mathcal{X}})}{\arg\inf} \frac{1}{2} W^2(\cdot, \varrho) + hF(\varrho), \quad h > 0. \tag{23}$$

The sequence of functions generated by the proximal recursions (22) approximate the flows $(p(s, \boldsymbol{x}_s), \hat{\varphi}(t, \boldsymbol{x}_t))$ for (20a),(20b),(20d) in the small time step limit, i.e.,

$$\hat{\phi}_{t_{k-1}} \to \hat{\varphi}(t = (k-1)\tau, \boldsymbol{x}_t) \quad \text{in } L^1(\overline{\mathcal{X}}) \text{ as } \tau \downarrow 0,$$

$$\varpi_{s_{k-1}} \to p(s = (k-1)\sigma, \boldsymbol{x}_s) \quad \text{in } L^1(\overline{\mathcal{X}}) \text{ as } \sigma \downarrow 0.$$

In the numerical example provided next, we solved (22) using the algorithm developed in [37].

Remark 4: An additional novelty in implementing (23) compared to our prior works [16], [37] is that a Skorokhod map such as [27] is needed to update the scattered samples in the state space. Specifically, to time march the supports of the functional updates (22), we composed the Euler-Maruyama updates in [37, Fig. 2] with the Skorokhod map corresponding to the reflection constraints.

C. Numerical Example

We consider an instance of the RSBP with $\overline{\mathcal{X}} = [-4, 4]^2$, $\boldsymbol{f} = -\nabla V$, $V(x_1, x_2) := (x_1^2 + x_2^3)/5$. For

$$\rho_0(x_1, x_2) \propto \prod_{i=1,2} \left(1 + (x_i^2 - 16)^2 \exp(-x_i/2)\right), \tag{24a}$$

$$\rho_1(x_1, x_2) \propto \prod_{i=1,2} \left(1.2 - \cos(\pi(x_i + 4)/2)\right), \tag{24b}$$

the optimal controlled joint state PDFs $\rho^{\text{opt}}(t, \boldsymbol{x}_t^{\boldsymbol{u}})$ are shown in Fig. 4. The corresponding uncontrolled joint state PDFs $\rho^{\text{unc}}(t, \boldsymbol{x}_t)$ are shown in Fig. 5. These results were obtained by solving (22) via [37, Sec. III.B] with $\tau = \sigma = 10^{-3}$ to perform the fixed point recursion (Fig. 1) applied to (20).

References

[1] A. Hotz and R. E. Skelton, "Covariance control theory," *International Journal of Control*, vol. 46, no. 1, pp. 13–32, 1987.

[2] R. Skelton and M. Ikeda, "Covariance controllers for linear continuous-time systems," *International Journal of Control*, vol. 49, no. 5, pp. 1773–1785, 1989.

[3] K. M. Grigoriadis and R. E. Skelton, "Minimum-energy covariance controllers," *Automatica*, vol. 33, no. 4, pp. 569–578, 1997.

[4] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part I," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1158–1169, 2015.

[5] ——, "Optimal steering of a linear stochastic system to a final probability distribution, Part II," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1170–1180, 2015.

[6] ——, "Optimal transport over a linear dynamical system," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2137–2152, 2016.

[7] C. Villani, *Topics in optimal transportation*. American Mathematical Soc., 2003, no. 58.

[8] E. Schrödinger, "Über die umkehrung der naturgesetze," *Sitzungsberichte der Preuss. Phys. Math. Klasse*, vol. 10, pp. 144–153, 1931.

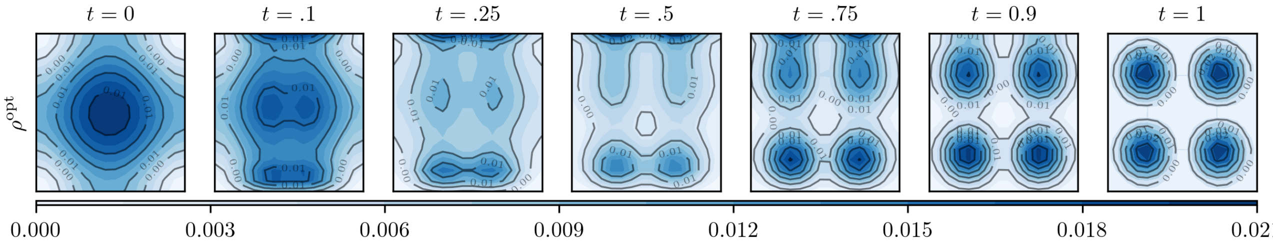

Fig. 4: For the RSBP in Section IV-B, shown here are the contour plots of the optimal controlled joint state PDFs $\rho^{\mathrm{opt}}(t, \boldsymbol{x}_t^{\boldsymbol{u}})$ over $\overline{\mathcal{X}} = [-4, 4]^2$. Each subplot corresponds to a different snapshot of ρ^{opt} in time. The color denotes the joint PDF value; see colorbar (dark hue = high, light hue = low).

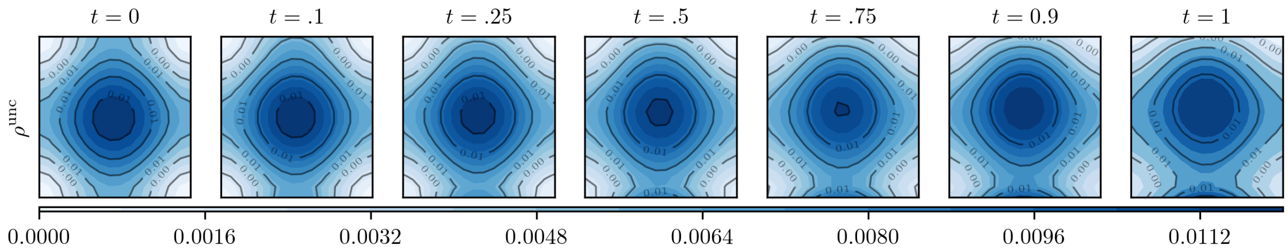

Fig. 5: For the RSBP in Section IV-B, shown here are the contour plots of the uncontrolled joint state PDFs $\rho^{\mathrm{unc}}(t, \boldsymbol{x}_t)$ over $\overline{\mathcal{X}} = [-4, 4]^2$ starting from (24a). Each subplot corresponds to a different snapshot of ρ^{unc} in time. The color denotes the joint PDF value; see colorbar (dark hue = high, light hue = low).

[9] ——, "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.

[10] A. Halder and E. D. Wendel, "Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 7249–7254.

[11] E. Bakolas, "Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints," *Automatica*, vol. 91, pp. 61–68, 2018.

[12] K. Okamoto and P. Tsiotras, "Input hard constrained optimal covariance steering," *arXiv preprint arXiv:1903.10964, IEEE Conference on Decision and Control, Nice, France*, 2019.

[13] E. Bakolas, "Covariance control for discrete-time stochastic linear systems with incomplete state information," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 432–437.

[14] K. F. Caluya and A. Halder, "Finite horizon density control for static state feedback linearizable systems," *arXiv preprint arXiv:1904.02272*, 2019.

[15] ——, "Finite horizon density steering for multi-input state feedback linearizable systems," *arXiv preprint arXiv:1909.12511, to appear in 2020 American Control Conference*, 2019.

[16] ——, "Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift," *arXiv preprint arXiv:1912.01244*, 2019.

[17] K. Okamoto, M. Goldshtein, and P. Tsiotras, "Optimal covariance control for stochastic systems under chance constraints," *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 266–271, 2018.

[18] N. Ikeda, "On the construction of two-dimensional diffusion processes satisfying Wentzells boundary conditions and its application to boundary value problems," *Memoirs of the College of Science, University of Kyoto. Series A: Mathematics*, vol. 33, no. 3, pp. 367–427, 1961.

[19] S. Watanabe, "On stochastic differential equations for multi-dimensional diffusion processes with boundary conditions," *Journal of Mathematics of Kyoto University*, vol. 11, no. 1, pp. 169–180, 1971.

[20] P.-L. Lions and A.-S. Sznitman, "Stochastic differential equations with reflecting boundary conditions," *Communications on Pure and Applied Mathematics*, vol. 37, no. 4, pp. 511–537, 1984.

[21] J. M. Harrison and R. J. Williams, "Multidimensional reflected Brownian motions having exponential stationary distributions," *The Annals of Probability*, pp. 115–137, 1987.

[22] P. W. Glynn and R. J. Wang, "On the rate of convergence to equilibrium for reflected Brownian motion," *Queueing Systems*, vol. 89, no. 1-2, pp. 165–197, 2018.

[23] J. M. Harrison, *Brownian models of performance and control*. Cambridge University Press, 2013.

[24] A. Pilipenko, *An introduction to stochastic differential equations with reflection*. Universitätsverlag Potsdam, 2014, vol. 1.

[25] A. V. Skorokhod, "Stochastic equations for diffusion processes in a bounded region," *Theory of Probability & Its Applications*, vol. 6, no. 3, pp. 264–274, 1961.

[26] ——, "Stochastic equations for diffusion processes in a bounded region. II," *Theory of Probability & Its Applications*, vol. 7, no. 1, pp. 3–23, 1962.

[27] L. Kruk, J. Lehoczky, K. Ramanan, S. Shreve *et al.*, "An explicit formula for the Skorokhod map on [0, a]," *The Annals of Probability*, vol. 35, no. 5, pp. 1740–1768, 2007.

[28] D. W. Stroock and S. S. Varadhan, "Diffusion processes with boundary conditions," *Communications on Pure and Applied Mathematics*, vol. 24, no. 2, pp. 147–225, 1971.

[29] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem," *Numerische Mathematik*, vol. 84, no. 3, pp. 375–393, 2000.

[30] K. F. Caluya and A. Halder, "Reflected schrödinger bridge: Density control with path constraints," 2020.

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

[32] 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.

[33] 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.

[34] V. Linetsky, "On the transition densities for reflected diffusions," *Advances in Applied Probability*, vol. 37, no. 2, pp. 435–460, 2005.

[35] R. N. Bhattacharya and E. C. Waymire, *Stochastic processes with applications*. SIAM, 2009, vol. 61.

[36] K. F. Caluya and A. Halder, "Proximal recursion for solving the Fokker-Planck equation," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 4098–4103.

[37] K. Caluya and A. Halder, "Gradient flow algorithms for density propagation in stochastic systems," *IEEE Transactions on Automatic Control*, 2019.

[38] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker-Planck equation," *SIAM journal on mathematical analysis*, vol. 29, no. 1, pp. 1–17, 1998.

[39] F. Santambrogio, "{Euclidean, metric, and Wasserstein} gradient flows: an overview," *Bulletin of Mathematical Sciences*, vol. 7, no. 1, pp. 87–154, 2017.