Schrödinger Approach to Optimal Control of Large-Size Populations
Kaivalya Bakshi, David D. Fan and Evangelos A. Theodorou

Abstract—Large-size populations consisting of a continuum of identical and non-cooperative agents with stochastic dynamics are useful in modeling various biological and engineered systems. This paper addresses the problem of designing mean-field type optimal state-feedback controllers which guarantee closed-loop stability of the stationary density of such agents with nonlinear Langevin dynamics, under the action of their individual steady state controls.

We represent the corresponding optimality system which consists of coupled forward-backward PDEs as decoupled Schrödinger equations, by introducing a novel variable transform. Spectral properties of the Schrödinger operator which underlie the stability analysis are used to obtain explicit control design constraints, in the case that the agents do not interact explicitly via density dependent dynamics or cost function. Our interpretation of the Schrödinger potential as the cost function of a closely related optimal control problem motivates a quadrature based algorithm to compute the finite time optimal control. We show the deep connection between the nonlinear Schrödinger equation and Mean Field Games for agents with nonlinear Langevin dynamics and explicit interactions.

I. INTRODUCTION

Dynamics and control of multi-agent populations consisting of a large number of identical and non-cooperative agents are of interest in various applications including robotic swarms, macro-economics, traffic and neuroscience. Prior works on optimal open-loop or closed-loop ensemble (broadcast) control consider several copies of a particular deterministic [1] or stochastic ( [2], [3], [4]) systems and have applications in quantum control [5] and neuroscience [6]. A standard idea in engineering, economics and biology is regulation using local feedback information, which is used to model decision making in large-size populations of rational agents with limited information. Optimal feedback control applications to large-size populations of small robots with individual state-feedback controllers have been proposed for inspection of industrial machinery [7], centralized control of hybrid automata [8] and decentralized control of robotic bee swarms for pollinating crops [9].

The mean-field approach provides a tractable framework for describing collective behavior of a continuum of agents, by approximating their individual actions [10] as the oblivious control [11] of a single representative agent. Mean field games (MFGs) ( [12], [13]) utilize PDE representation to model such continuum systems and provide a game-theoretic optimal control interpretation of emergent behavior in self-organized systems. Most works on MFGs consider explicit interactions between agents through the dependence of their dynamics or cost function on the population density ( [14], [15], [16], [17], [18]). The corresponding optimality system consists of a backward-in-time semilinear Hamilton-Jacobi-Bellman (HJB) equation governing the value function and a forward-in-time Fokker-Planck (FP) equation governing the density, wherein the HJB equation depends on the density and the FP equation depends on the value function. However, even if the individual dynamics or cost functions are independent of the density, the agents implicitly interact with each other since their controls optimize a utility which depends on the population density. In this case, the HJB equation is independent of the density but the FP equation depends on the value function. MFGs lacking explicit agent interactions have been applied in macroeconomics [19] and robotics ( [8], [9]). In physical systems such as robot swarms, if the dimensions of individual agents are small compared to their region of operation, then it can be assumed that the agents do not locally interact with each other.

In this paper we consider the finite and infinite time optimal control problem (OCP) of a density of identical and non-cooperative agents which have individual state-feedback controllers with no explicit dependence of the agent dynamics or cost functions on the population density. An important question is whether the steady state controls can be used to stabilize an initial (perturbed) density of agents to the corresponding steady state density. We address this question for large-size populations wherein agents obey nonlinear dynamics and provide explicit control design constraints required for stability. For the finite time case, we present a computationally efficient quadrature based control algorithm and demonstrate it for a population of agents with nonlinear dynamics.

Stability of fixed points of MFG models, which involves analysis of the forward-backward HJB and FP equations has been analyzed previously ( [19], [20], [21]). A common limitation of most prior works on this topic is that individual agent dynamics are assumed to be simple integrator systems. On the other hand, mean-field representations of large-scale systems with nonlinear agent mobilities are used to model crowds [22], flocks [23] and robotic systems [8], stochastic gradient descent optimization of neural network parameters [24] and Brownian particles in non-equilibrium thermodynamics [25]. Specifically, nonlinear agent dynamics of overdamped

Kaivalya Bakshi is a PhD Candidate in the Department of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA (email: k bakshi@gatech.edu)
David D. Fan is a PhD candidate of Robotics in the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA
Evangelos A. Theodorou is in the Department of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA
1Code publicly available at https://github.com/ddfan/pi_quadrature
Langevin type appear in several of these works.

In the recent works (17, 18) by some of the authors, local (linear) stability results were presented for certain MFGs wherein agents obey nonlinear Langevin dynamics. The approach in these works was based on exploiting spectral properties of the closed-loop generator of the agent dynamics, which governed the linear perturbation PDEs to analyze forward-backward stability. In this work we take a different approach. Since we assume that the cost function has no explicit density dependence, the stability analysis corresponds to the forward-time FP equation which depends on the steady state controls. However, we present more general nonlinear stability results which do not rely on linearization of the HJB-FP equations.

The history of variable transforms in PDEs, particularly for obtaining linear representations of parabolic equations, dates back to the independent papers by Cole (26) and Hopf (27). The role of this transform in the connection between optimal control and quantum mechanics was known by Schrödinger (28). More recently, this transform has been successfully used to obtain sampling control algorithms (29). In section III-A we introduce a novel Cole-Hopf type transform in order to obtain a decoupled representation of the coupled HJB-FP optimality system of the considered large-scale OCP. The motivation for this transform is three-fold.

First, the decoupled system consists of linear imaginary-time Schrödinger equations so that spectral properties of the corresponding Schrödinger operator can be used to analyze the stability properties of the stationary (steady state) density. In section IV, we provide explicit stability constraints on the control design which guarantee closed-loop stability of the steady state density. Second, in section III-C the Schrödinger potential of this operator is interpreted as the cost function of a closely related optimal control problem subject to simple integrator dynamics. We use this fact to obtain a sampling representation of a nonlinear OCP using trajectories sampled from linear dynamics. In the path integral approach, sampling algorithms rely on simulating noisy trajectories of nonlinear dynamics (30), (31) to numerically approximate trajectories of nonlinear dynamics, a step which introduces high computational complexity and inaccuracy. Since we use samples from linear dynamics, this step can be replaced by analytic knowledge of the required distribution, motivating a quadrature method to compute the control. We explain this algorithm in section V. Finally, we observe that given an (uncontrolled) Langevin system there exists a corresponding equivalent control problem with simple integrator dynamics, such that the optimal control of this simplified system recovers the given uncontrolled dynamics of the original system. Third, it allows us to show an interesting connection between the Schrödinger equation and MFGs.

The connection between the imaginary-time Schrödinger equation and optimal control has been explored previously in the context of OMT (32), Schrödinger bridges (33) and in (34) which showed an interesting connection between a specific class of MFG models and the nonlinear Schrödinger (NLS) equation. However, this connection was shown only for MFGs wherein agents have very simple integrator dynamics in (34). In section III-B we show that this connection is true for the broader class of MFGs in which agents obey nonlinear Langevin dynamics. Our conclusions and directions for future research are presented in section VI.

II. CONTROL OF LARGE-SIZE POPULATIONS

We first introduce some notation and then describe the large scale stochastic control problem considered in this work. We denote vector inner products by $a \cdot b$, the induced Euclidean norm by $|a|$ and its square by $a^2 = |a|^2$. $\partial_t$ denotes partial derivative with respect to $t$ while $\nabla$, $\nabla \cdot$ and $\Delta$ denote the gradient, divergence and Laplacian operations respectively. $L^2(\mathbb{R}^d)$ denotes the class of square integrable functions of $\mathbb{R}^d$. The norm of a function $f$ and inner product of functions $f_1$, $f_2$ in this class is denoted by $\|f\|_{L^2(\mathbb{R}^d)}$ and $\langle f_1, f_2 \rangle_{L^2(\mathbb{R}^d)}$ respectively.

Consider a set of $1 \leq N$ agents indexed $1 \leq i \leq N$ with model for the $i^{th}$ agent:
\[
\text{dx}_i^s = -\nabla \nu(x_i^s) ds + u_i^s(s) ds + \sigma dw_i^s
\]
where $x_i^s$, $u_i^s(s) \in \mathbb{R}^d$ are the state and control inputs and $w_i^s$ is a standard $\mathbb{R}^d$ Brownian motion. Suppose that the $i^{th}$ agent minimizes its individual performance objective given by
\[
J_i^s(u) := \lim_{T \to +\infty} \frac{1}{T} E \left[ \int_0^T q(x_i^s) ds + \frac{R}{2} (u_i^s)^2 ds \right],
\]
then under certain standard conditions, the equivalent stationary HJB PDE problem is
\[
0 = q - \frac{(\nabla v_i^s)^2}{2R} - \nabla v_i^s \cdot \nabla \nu + \frac{\sigma^2}{2} \Delta v_i^s
\]
with the optimal control $u_i^{s, \infty}(x) = -\nabla v_i^s(x)/R$. Since the noise driving each agent is mutually independent and Brownian, the states of each agent $x_i^s$ are i.i.d. random variables independent of $w_i^s$. The set of the states $\{x_i^s\}_{1 \leq i \leq N}$ represents the population of agents. Next, we assume that the number of agents is infinite, $N \to +\infty$. We use the mean-field approach to represent the problem as a standard OCP of a representative agent with state $x_s \sim p(s, \cdot)$ obeying dynamics (4) and the distribution of the continuum of agents’ states being modeled by the density $p(s, \cdot)$. Assuming that there exists a constant $k$ such that $\sup_{1 \leq h \leq N} E[|x_h^0|^2] < k + \infty$, the initial distribution is approximated by the empirical density $p_N(0, x) = \sum_{i=1}^N \delta(x - E[x_i^0])$ where $\delta$ is the Dirac delta function. We assume that $p_N(0, x)$ converges weakly to $p(0, x) \in C^{1,2}(0 \times \mathbb{R}^d)$, that is $\lim_{N \to +\infty} \int \gamma(x)p_N(0, x) = p(0, x)$ for any bounded continuous function $\gamma(x)$ on $\mathbb{R}^d$.

A. Control Problem

Let $x_s, u(s) \in \mathbb{R}^d$ denote the state and control inputs of a representative agent which obeys the controlled first order dynamics:
\[
dx_s = -\nabla \nu(x_s) ds + u(s) ds + \sigma dw_s
\]
for every $s \geq 0$, driven by standard $\mathbb{R}^d$ Brownian motion, with noise intensity $0 < \sigma$ on the filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}\}$. These dynamics are the controlled version of a Langevin system in the overdamped case. The smooth
function $\nu : \mathbb{R}^d \to \mathbb{R}$ is called the Langevin potential and the control $u \in \mathcal{U} := \mathcal{U}[t,T]$ where $\mathcal{U}$ is the class of admissible controls [36] containing functions $u : [t,T] \times \mathbb{R}^d \to \mathbb{R}^d$. Consider the following optimal control problem (OCP)

$$\min_{u \in \mathcal{U}} J(u) := \lim_{T \to +\infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T q(x_s) ds + \frac{R}{2} u^2 ds \right]$$

subject to [4], where we denote the probability density of $x_s$ by $p(s, x)$ for every $s \geq 0$ which represents the density of all agents, with initial density being $x_0 \sim p(0, x)$, $q : \mathbb{R}^d \to \mathbb{R}$ is a known deterministic function which has at most quadratic growth and $R > 0$ is the control cost. We assume that $\nabla \nu(\cdot), q(\cdot)$ and functions in the class $\mathcal{U}$ are measurable. We refer to the OCP [5] subject to dynamics [4] as problem (P1). Since the cost function does not explicitly depend on the density, the above OCP does not model systems in which agents interact explicitly. We therefore do not refer to the above large-scale OCP as a MFG. We will briefly return to the case of MFGs in section III-B.

B. PDE Optimality System

Standard application of dynamic programming [28] as in [35, 12], implies that under certain regularity conditions [20], problem (P1) is equivalent to the following HJB-FP PDE optimality system governing the value and density functions respectively:

$$q - c - \frac{(v^\infty)^2}{2R} - \nabla v^\infty \cdot \nabla \nu + \frac{\sigma^2}{2} \Delta v^\infty = 0$$  

$$\nabla((\nabla \nu + \frac{v^\infty}{R})p) + \frac{\sigma^2}{2} \Delta p = 0$$

with the constraint $\int p^\infty dx = 1$, where $c$ is the optimal cost. The optimal control is given by $u^\infty(x) = -\nabla v^\infty/R$. Under certain regularity conditions [20] which we assume to be true, the time-varying relative value [37] function and density corresponding to problem (P1) are governed by the optimality system:

$$-\partial_t v = q - c - \frac{(v^2)}{2R} - \nabla v \cdot \nabla \nu + \frac{\sigma^2}{2} \Delta v$$  

$$\partial_t p = \nabla \cdot ((\nabla \nu + \frac{v}{R})p) + \frac{\sigma^2}{2} \Delta p$$

with the constraint $\int p(t, x)dx = 1$ for all $t \geq 0$. In this work, we assume to be true, the additional conditions [36] which are required to show that the HJB PDEs (6) and (8) have unique solutions. Note that steady state and time varying HJB PDEs are both semilinear.

Remark 1. The finite time OCP analogous to the infinite time OCP (P1) given by:

$$\min_{u \in \mathcal{U}} J(u) := \mathbb{E} \left[ \int_0^T q(x_s) ds + \frac{R}{2} u^2 ds \right].$$

subject to the dynamics [4] has the optimality system given by equations [9, 9] with $c = 0$, initial density given by $p(0, x)$ and constraint $\int p(t, x)dx = 1$.

C. Stationary Solution

The FP equation governing the density of an overdamped Langevin system is called the Smoluchowski PDE. The SPDE (7), can be interpreted as the Smoluchowski PDE for such a Langevin system with the restoring potential $\nu + v^\infty/R$. The analytical solution to the SPDE can be obtained as a Gibbs distribution using this interpretation, under certain conditions given below, on the fixed point pair $(v^\infty, p^\infty)$ of the optimality system (6) and the Langevin potential $\nu$. We denote $w(x) := \nu(x) + \frac{v^\infty}{R}$.

(A0) There exist $(v^\infty(x), p^\infty(x)) \in (C^2(\mathbb{R}^d))^2$ satisfying (6) such that $\lim_{|x| \to \infty} w(x) = +\infty$ and $\exp \left( -\frac{\sigma^2}{2} w(x) \right) \in L^1(\mathbb{R}^d)$.

Lemma II.1. Let (A0) be true. If $\nu(x)$ is a smooth functions satisfying (A0), then the unique stationary solution to the density given by the Fokker Planck equation (7) is

$$p^\infty(x) := \frac{1}{Z} \exp \left( -\frac{\sigma^2}{2} \left( w(x) \right) \right),$$

where $Z = \int \exp \left( -\frac{\sigma^2}{2} w(x) \right) dx$.

Proof. We observe that the (7) is the Smoluchowski equation for an overdamped Langevin system given by

$$dx_s = -\nabla (\nu + v^\infty/R)(x_s) \, ds + \sigma dw_s.$$  

Under the assumptions above, the proof then follows directly from proposition 4.2, pp 110 in [38].

III. SCHRODINGER APPROACH

The HJB PDEs above have a linear representation in the time-varying and steady state case. In the time varying case this representation is obtained using a Cole-Hopf (26, 27) transform

$$\phi := \exp(-v/\sigma^2 R)$$

which was applied in stochastic control theory by Kappen [29]:

$$-\partial_t \phi = -\frac{q\phi}{\sigma^2 R} - \nabla \phi \cdot \nabla \nu + \frac{\sigma^2}{2} \Delta \phi.$$  

The advection-diffusion equation above has a path integral solution [39] which is useful in computing the control [29], [31, 40]. In what follows we will introduce two transforms providing a diffusion PDE representation of the semilinear HJB and linear FP equations. This transform facilitates a stability analysis of the fixed point of the optimality system [8, 9] based on the spectral properties of a Schrödinger operator in section IV. Further, in this section we interpret the corresponding Schrödinger potential as the cost function of a fictitious but intimately related OCP with integrator dynamics. This motivates a quadrature based algorithm to solve the transformed HJB equation and thus compute the control in the section [IV].
A. Cole-Hopf Type transform

We introduce a Cole-Hopf type transform:

\[ f(t, x) := \exp(-(v(t, x) + R\nu(x))/\sigma^2 R), \]  

which leads to the following representation of equation (9):

\[ -\partial_t f = \frac{cf}{\sigma^2 R} - \frac{V f}{\sigma^2 R} + \sigma^2 \Delta_f = \frac{cf}{\sigma^2 R} - H f, \]  

where we denote the modified cost function \( V := q + (R/2)(\nabla v)^2 - (\sigma^2 R/2)\Delta v \) and the operator \( H := \frac{\sigma^2 R}{\sigma^2 R} \). The transformed PDE can be verified by using the calculations \( \partial_t v = -\sigma^2 R \frac{c f}{f} \), \( \Delta_f = -\sigma^2 \frac{\nabla f}{f^2} \), \( \nabla(v + R\nu) = -\frac{1}{\sigma^2 R} \nabla(v + R\nu) \), \( \Delta_f = -\frac{\nabla f}{f^2} \).

The introduced Cole-Hopf transform combined with hermitization of the density corresponds to a diagonalization of the coupled optimality system (8), (9) which is not coupled with the HJB equation (8).

Analogously, it can be shown that the stationary value and density functions satisfying the stationary nonlinear optimality system (6, 7) can be represented by the transformation variables \( f^\infty := \exp(-(v^\infty + R\nu)/\sigma^2 R) \) and \( g^\infty := p^\infty/f^\infty \), which both satisfy the following eigenvalue problem

\[ He(x) = \frac{c}{\sigma^2 R} e(x) \]  

subject to the normalizing constraint \( \int f^\infty(x)g^\infty(x)dx = 1 \).

Theorem III.1. \((f^\infty(x), g^\infty(x))\) are both solutions to the eigenvalue problem (22) such that \( \int f^\infty(x)g^\infty(x)dx = 1 \) if and only if

\begin{align*}
  v^\infty(x) & = -\sigma^2 R \ln(f^\infty)(x) - R\nu(x) \\
p^\infty(x) & = f^\infty(x)g^\infty(x)
\end{align*}

is a solution to the nonlinear optimality system (6, 7). Further, the optimal control is given by \( u^\infty = -\nabla v^\infty/R = \sigma^2 \nabla f^\infty/f^\infty \).

Given a solution pair \((v^\infty, p^\infty)\) to the optimality system (6, 7) it is possible to obtain explicit solutions to functions \((f^\infty, g^\infty)\) satisfying equation (22) such that \( \int f^\infty g^\infty dx = 1 \). The result in theorem III.1 and the introduced Cole-Hopf transform can be used to verify the following corollary to theorem III.2.

Corollary III.2.1. Let \( p^\infty := \frac{1}{\sigma^2} \exp\left(-\frac{2}{\sigma^2}\left(w(x)\right)\right) \) (x) with \( w(x) := (v(x) + e^\infty(x))/\sigma \) and \( Z \) the normalizing constant where \((v^\infty, p^\infty)\) is a pair satisfying (41). Then \( f^\infty := \sqrt{Z} p^\infty \) and \( g^\infty := f^\infty/Z \) both satisfy equation (22) such that \( \int f^\infty g^\infty dx = 1 \).

B. Nonlinear Schrödinger Equation and Mean Field Games

MFGs model large-scale stochastic systems which permit interaction among agents. In the continuum case, the simplest version of such a MFG for agents with nonlinear Langevin dynamics can be expressed as the OCP (P1) with a density dependent cost function \( q := q[p] := q(x, p(x, s)), \) the mean-field time-varying optimality system (8, 9) for this MFG is given by equations (8, 9) and \( q = \bar{q} \).

In (41) by Ullmo et al., it was shown that there is a deep connection between the imaginary time nonlinear Schrödinger (NLS) equation a specific class of MFGs. A major limitation of the work (35) which relied on the previously known variable transform (14) is that this connection was shown only for MFG models in which agent dynamics are restricted to be simple integrator systems. We apply the results presented in this section to extend the class of MFGs exhibiting the connection with the NLS equation.

From the preceding discussion, it can be easily verified that using the transforms (15, 16), the corresponding MFG model constituted by the time-varying optimality system (8, 9) and \( q := \bar{q} \) has the following NLS representation:

\[ -\partial_t f = \frac{c f}{\sigma^2 R} - \frac{V f}{\sigma^2 R} + \frac{\sigma^2}{2} \Delta_f \]
\[ -\partial_t g = -\frac{c g}{\sigma^2 R} + \frac{V g}{\sigma^2 R} - \frac{\sigma^2}{2} \Delta g \]

where \( \bar{V}[f]g = \bar{q}[fg] + (R/2)(\nabla v)^2 - (\sigma^2 R/2)\Delta v \). Thus, we have generalized the connection between MFGs and the
imaginary time NLS equation introduced in [11], to MFG models in which agent dynamics lie in the general class of nonlinear Langevin dynamics.

C. Interpretation

The Schrödinger potential \( \frac{V(x)}{\sigma^2} \) defined earlier can be interpreted in terms of the cost function of the following fictitious OCP with simple integrator dynamics which has an intimate connection with the original OCP in section II-A:

\[
\min_{u \in \mathcal{U}} J(u) := \lim_{T \to +\infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T V(\hat{x}_s) ds + \frac{R}{2} \hat{u}^2 \right] \tag{27}
\]

subject to the simple integrator dynamics

\[
d\hat{x}_s = \hat{u}(s) ds + \sigma dw_s. \tag{28}
\]

We refer to the OCP (27) subject to (28) as problem (P2). The time-varying optimality system associated with problem (P2) is given by:

\[
-\partial_t \hat{v} = V - \hat{c} - \frac{(\nabla \hat{v})^2}{2R} + \frac{\sigma^2}{2} \Delta \hat{v} \tag{29}
\]

\[
\partial_t \hat{p} = \nabla \cdot \left( \frac{\nabla \hat{v}}{R} \right) p + \frac{\sigma^2}{2} \Delta \hat{p} \tag{30}
\]

where \( \hat{c} \) is the optimal cost.

It is easily observed that if \( v \) is the solution to the HJB equation (8), then \( \hat{v} = v + R \nu \) is a solution to the HJB equation (29). Therefore, the time-varying optimal controls: \( u^* \) of the OCP (P1) and \( \hat{u}^* \) of the OCP (P2), are related as \( \hat{u}^* = u^* - \nabla \nu \). Similarly, by substituting \( \nabla \hat{v} = \nabla v + R \nabla \nu \) into equation (30), we can see that the PDEs (9), (30) satisfied by the densities \( p(s, x), \hat{p}(s, x) \) respectively, are identical. Therefore, given identical initial conditions \( \hat{p}(0, x) = p(0, x) \), lemma [11] implies that \( \hat{p}(s, x) = p(s, x) \) for all \( s \geq 0 \) where \( p(s, x) \) is the density of optimally controlled agents associated with the OCP (P1). To summarize, solving the optimality system (30), (31) corresponding to the OCP (P1) (subject to nonlinear passive dynamics) is equivalent to solving the optimality system (29), (30) corresponding to the OCP (P2) (subject to simple integrator dynamics). This fact will be used in section V to synthesize a solver to compute the finite time optimal control.

The steady state optimality system corresponding to problem (P2) given by:

\[
V - \frac{(\nabla \nu^\infty)^2}{2R} + \frac{\sigma^2}{2} \Delta \nu^\infty = 0 \tag{31}
\]

\[
\nabla \cdot \left( \frac{\nabla \nu^\infty}{R} \right) \nu^\infty + \frac{\sigma^2}{2} \Delta \nu^\infty = 0, \tag{32}
\]

can be similarly shown to be connected to the solutions of the optimality system (9), (7) by \( \nabla \hat{v}^\infty = \nabla v^\infty + R \nabla \nu \) and \( \hat{p}^\infty(s, x) = p^\infty(s, x) \) for all \( s \geq 0 \), given that the initial densities are equal \( \hat{p}(0, x) = p(0, x) \). The steady state control \( u^\infty \) of OCP (4), (4) and \( \hat{u}^\infty \) of OCP (27), (28) are related as \( \hat{u}^\infty = u^\infty - \nabla \nu^\infty \).

Further, setting \( q(x) = 0 \) in the cost function \( V(x) \) of the OCP (P2) results in an optimal control \( \hat{u}^\infty(s) \) which recovers the passive Langevin dynamics (4) with \( u(s) = 0 \).

It can be proved that if \( q(x) = 0 \), then \( \hat{u}^\infty = -\nabla \nu^\infty/R = -\nabla \nu \) by verifying that \( R \nu(x) \) is a solution to the stationary HJB equation (31). This can also be proved by observing that if \( q(x) = 0 \) in the OCP (P1), then the steady state optimal control is \( u^\infty = 0 \), so that from the relationship in the previous paragraph \( \hat{u}^\infty = u^\infty - \nabla \nu = -\nabla \nu \). In conclusion, given certain uncontrolled Langevin dynamics (4) with smooth Langevin potential \( v(x) \), the steady state optimal control corresponding to the OCP (P2) with cost function \( V := (R/2)(\nabla \nu)^2 - (\sigma^2 R/2) \Delta \nu, \) recovers the uncontrolled dynamics as \( \hat{u}^\infty(x) = -\nabla \nu(x) \).

IV. CONTROL DESIGN

The decay of an initial density of particles under open loop (or uncontrolled) overdamped Langevin dynamics to a stationary density is a classic topic in statistical physics [42]. In this section we analyze the decay of a perturbed density of agents under the action of the steady state controller to the corresponding steady state density. Since the HJB-FP (8, 9) optimality system is coupled one-way, the perturbation analysis corresponds to that of the FP equation. Evolution of a perturbed density governed by the FP PDE (9) is analyzed through evolution of the hermitized density (17) governed by equation (18). Diagonalization of the coupled PDEs constituting the optimality system as in equation (21) facilitates stability analysis based on the spectral properties of the Schrödinger operator. Based on the analysis we obtain explicit analytical design constraints on the cost function \( q(x) \) and control parameter \( R \) which guarantee stability of the fixed point density.

A. Perturbation System

Consider a controlled large-size non-interacting population expressed by problem (P1), which is controlled by the optimal steady state control \( u^\infty = -\nabla \nu^\infty/R \) corresponding to the optimality system (6, 7) with a unique fixed point \((v^\infty, \hat{p}^\infty)\) satisfying assumption (A0). Theorem III.2 implies that in this case the steady state value and density functions can be written as \( f^\infty, g^\infty \) in terms of a pair of functions \((f^\infty, g^\infty)\) both satisfying equation (22) and \( \int f^\infty g^\infty dx = 1 \). Corollary III.2.1 gives formulae for the function pair \((f^\infty(x), g^\infty(x))\) in terms of the steady state solution \((v^\infty, \hat{p}^\infty)\). Time varying value and density functions can be written as \( f(t, x), g(t, x) \) in terms of the corresponding transformation variables \( (f(t,x), g(t,x)) \).

Time varying densities, perturbed from the steady state density of agents can therefore be written using the hermitization transform (17) as \( p(t,x) = p^\infty(x) + \hat{p}(t,x) = f^\infty(x)g^\infty(t,x) + f^\infty(x)\hat{g}(t,x) \). Since we are studying stability of the steady state controller, there are no perturbations in the value function \( v^\infty \) nor consequently, in the transformation variable \( f^\infty \). Here, the function \( \hat{g}(t,x) \) corresponds to a perturbation in the hermitized density given as \( g(t,x) = g^\infty(x) + \hat{g}(t,x) \), which obeys the time-varying PDE (18). In this section we study the decay of a perturbed density \( p^\infty + \hat{p} \) to its steady state density \( p^\infty \). We state the following corollary.
Corollary IV.01. If \( g^\infty(x) \) is a solution to the stationary PDE \((22)\) and \( \tilde{g}(t,x) \) is a solution to the PDE \((18)\) where \( \tilde{g}(t,x) \in C^{1,2}(0,+,\mathbb{R}^d) \), then \( \tilde{g}(t,x) \) is governed by the linear PDE

\[
\partial_t \tilde{g} = -H \tilde{g}.
\]

Theorem IV.3. Let \((A0, A1, A2)\) be true. Let \((v^\infty(x), p^\infty(x))\) be the unique stationary solution to the optimality system \((6, 7)\) and denote by \((f^\infty, g^\infty)\) the two solutions to problem \((22)\) given in corollary III.2.1. If \( \tilde{g}(0,x) \in \mathcal{S}_0 \) and \( \{g_n\}_{0\leq n \leq +\infty} \) are determined by

\[
g_n(t) = -\lambda_n t.
\]

then \( \tilde{g}(t,x) = \sum_{n=1}^{+\infty} g_n(t) e_n(t) \) is the unique \( \mathcal{H} \) solution to the perturbation equation \((33)\). \( p^\infty(x) \) is asymptotically stable with respect to \( \mathcal{S}(\epsilon) \).

Proof. Since \( \tilde{g}(0,x) \in \mathcal{H} \) we have the unique representation \( \tilde{g}(t,x) = \sum_{n=0}^{+\infty} g_n(t) e_n(x) \) where \( g_n(0) = (\tilde{g}(0,x), e_n(x))_\mathcal{H} < +\infty \) for all \( n \). Since \( \{e_n\}_{0\leq n < +\infty} \) is a complete basis on \( \mathcal{H} \), any solution in \( \mathcal{H} \) to the PDE \((33)\) must have the form \( \sum_{n=0}^{+\infty} g_n(t) e_n(x) \) where \( \{g_n\}_{0\leq n \leq +\infty} \) are finite for all \( t \in [0, +\infty) \). Substituting the selected form of the solution in the perturbation equation \((33)\) and using the eigenproperty \( H e_n = \lambda_n e_n \), we obtain the ODEs \((34)\). Due to assumption \((A1)\) the eigenproperties of the Shrödinger operator given in lemmas \((V.1)\) \((IV.2)\) hold. Using the eigenproperty yields the ODEs \((34)\) with the unique solutions \( g_n(t) = g_n(0) e^{-\lambda_n t} \). Therefore \( \tilde{g}(t,x) = \sum_{n=0}^{+\infty} g_n(t) e_n(x) \) where \( g_n(t) = g_n(0) e^{-\lambda_n t} \) is the unique \( \mathcal{H} \) solution to the perturbation equation \((33)\). From the Krein-Rutman theorem \((44)\) under the assumption that \( V(x) \geq 0 \) given by \((A2)\), the first eigenvalue is \( \lambda_0 = \lambda_0 \) and the first eigenfunction is \( v^\infty(x) = e_0(x) \) corresponding to the eigenvalue problem \((22)\). Further, \( \tilde{g}(0,x) \in \mathcal{S}_0 \) implies that \( g_0(0) = (\tilde{g}(0,x), e_0(x))_\mathcal{H} = (\tilde{g}(0,x), f^\infty(x))_\mathcal{H} = 0 \) implying \( g_0(t) = 0 \) for all \( t \geq 0 \). This completes the first part of the proof.

Using integration by parts we have that \( (H e_0, e_0)_{L^2(\mathbb{R})} \geq \lambda_0 = (\int_{-R}^R \frac{\partial}{\partial\nu} e_0, e_0)_{L^2(\mathbb{R})} + \frac{\lambda_0}{2} \|\nabla e_0\|_{L^2(\mathbb{R})}^2 \). Since \( V(x) \geq 0 \) from assumption \((A2)\) and \( \lambda_0 < \lambda_1 < \cdots \) due to assumption \((A1)\), we conclude that \( \lambda_n > 0 \) for all \( n > 1 \). Using Parseval’s identity \( \|g(t,x)\|_{L^2(\mathbb{R})}^2 = \sum_{n=0}^{+\infty} g_n(t)^2 \), noting that \( g_0(t) = 0, g_n(t)^2 = g_n(0)^2 e^{-2\lambda_n t} \) where \( \lambda_n > 0 \) for all \( n > 1 \) and using the Lebesgue dominated convergence theorem for the limit \( t \to +\infty \), we have that \( p^\infty(x) \) is nonlinearly asymptotically stable with respect to \( \mathcal{S}(\epsilon) \).

From the theorem above, we note that assumptions \((A1,A2)\) provide the explicit design constraints on the cost function \( q(x) \) and control parameter \( R \), which guarantee stability of an initially perturbed density of agents to the corresponding steady state density, under the action of the steady state controller. In figure 1 we show stabilization of an initially (perturbed) uniform density of agents to the stationary density corresponding to the steady state controls. The agent dynamics are unstable with the Langevin potential \( \nu(x) = -x^3/3 \) and the system is stabilized using a cost function \( q(x) = (5/2) \cdot x^2 \) such that conditions \((A1,A2)\) are satisfied. Equation \((22)\) is solved using a spectral solver \((45)\) for the parameters \( \sigma = R = 1/2 \) and the steady state density is obtained using equation \((12)\). Initial states of agents are sampled from a uniform density over the interval \([-2,2] \). Trajectories for \( N = 500 \) agents are simulated with 100 stochastic realizations.
with \( x_n \) governed by the discrete dynamical system:
\[
x_{n+1} = x_n + \sigma \sqrt{\delta t} \epsilon, \quad x_0 = \bar{x}, \quad \epsilon \sim \mathcal{N}(0, I)
\]
with the associated transition probability \( p(x_{n+1} | x_n) \sim \mathcal{N}(x_n, \sigma^2 \delta t \mathbf{I}) \). Letting
\[
w_n(x_n) := \exp \left( -\frac{V}{\sigma^2 R}(x_n) \delta t \right) \quad n = 0, \ldots, N - 1
\]
\[
w_N(x_N) := f(T, x_N)
\]
\[
w := \prod_{n=0}^{N} w_n(x_n)
\]
from equation (37) we have
\[
f(t, \bar{x}) = \mathbb{E}_\tau \left[ \prod_{n=0}^{N} w_n(x_n) \right]
\]
\[
= \cdots \int w_N(x_N) \left[ \prod_{n=2}^{N-1} w_n(x_n)p(x_{n+1} | x_n) \right] \times \int w_0(\bar{x})p(x_1 | x_0 = \bar{x})w_1(x_1)p(x_2 | x_1)dx_1 \cdots dx_N.
\]

The second integral above is approximated by Gaussian quadrature with \( M \) grid points \( \{\xi^i\}_{i=1}^{M} \) and weights \( \alpha^i \) as
\[
\int w_0(\bar{x})p(x_1 | x_0 = \bar{x})w_1(x_1)p(x_2 | x_1)dx_1 \approx \sum_{i=1}^{M} \left( p(x_2 | x_1 = \xi^i) \alpha^i w_1(x_1 = \xi^i) w_0(\bar{x})p(x_1 = \xi^i | x_0 = \bar{x}) \right).
\]

Defining the \( M \) dimensional vectors \( \Phi_1(x_2), \gamma_1 \), and \( \Phi_0(\bar{x}) \) to have elements \( \phi^i_1(x_2), \gamma^i_1, \phi^0_0(\bar{x}) \), respectively and define \( \Gamma_1 = \text{diag}(\gamma_1) \), can be written as a set of vector products:
\[
\int w_0(\bar{x})p(x_1 | x_0 = \bar{x})w_1(x_1)p(x_2 | x_1)dx_1 = \Phi_1(x_2)^\top \Gamma_1 \Phi_0(\bar{x}).
\]

Recall that \( p(x_1 = \xi^i | x_0 = \bar{x}) \) is a Gaussian PDF, so each element of \( \Phi_0(\bar{x}) \) is Gaussian weighted by \( w_0(\bar{x}) \). Plugging this back into (42) yields:
\[
= \cdots \int w_N(x_N) \left[ \prod_{n=3}^{N-1} w_n(x_n)p(x_{n+1} | x_n) \right] \times \int w_2(x_2)p(x_3 | x_2)\Phi_1(x_2)^\top \Gamma_1 \Phi_0(\bar{x})dx_2 \cdots dx_N.
\]
Take the integral within the brackets and perform another quadrature, this time at points \(\{\xi_2^i\}_{i=1}^M\) and weights \(\alpha_2^i\). We have:

\[
\int w_2(x_2)p(x_3|x_2)\Phi_1(x_2)^T\Gamma_1\Phi_0(\bar{x})dx_2
\approx \sum_{i=1}^{M} p(x_3|x_2 = \xi_2^i) \frac{\alpha_2^i w_2(x_2 = \xi_2^i) \Phi_1(x_2 = \xi_2^i)^T \Gamma_1 \Phi_0(\bar{x})}{\gamma_2^i}
\]

(46)

Let \(\bar{\Phi}_n\) be an \(M \times M\) transition matrix with elements \(\bar{\Phi}_{ij} = p(x_{n+1} = \xi_{n+1} | x_n = \xi_n^i)\). Then we can write (46) as:

\[
\Phi_2(x_3)^T \Gamma_2 \bar{\Phi} \Gamma_1 \Phi_0(\bar{x})
\]

(47)

Plugging this back into (45), we can perform the nested integrals recursively. At each timestep \(x_n\), we use a different quadrature grid, with points \(\{\xi_n^i\}_{i=1}^M\) and weights \(\alpha_n^i\). The entire integral will therefore be:

\[
f(t, \bar{x}) \approx \gamma_n^T \left[ \prod_{n=1}^{N-1} (\Phi_n \Gamma_n) \right] \Phi_0(\bar{x})
\]

(48)

where we have used the definitions:

\[
\gamma_n = \left\{ \alpha_n w_n(\xi_n^i) \right\}_{i=1}^M
\]

(49)

\[
\Gamma_n = \text{diag}(\gamma_n)
\]

(50)

\[
(\Phi_n)_{ij} = p(x_{n+1} = \xi_{n+1}^i | x_n = \xi_n^j)
\]

(51)

\[
\phi_n^i(\bar{x}) = w_0(\bar{x}) p(x_1 = \xi_1^i | x_0 = \bar{x})
\]

(52)

\[
\Phi_0(\bar{x}) = \left[ \phi_0(\bar{x}) \right]_{i=1}^M
\]

(53)

Since \(V(x)\) is time invariant and one chooses the same quadrature grid points at each timestep, \(\gamma_n\) and \(\Phi_n\) are the same for all \(n = 1, \cdots, N - 1\). So (48) can be simplified to:

\[
f(t, \bar{x}) \approx \gamma_N^T (\Phi \Gamma)^{N-1} \Phi_0(\bar{x})
\]

(54)

We consider a 2-dimensional problem with the following Langevin potential:

\[
\nu = 1/2 \cos(x_1 x_2)^2 - 1/24(x_1^4 + x_2^4)
\]

(55)

This results in dynamics:

\[
\begin{align*}
dx_1 &= (\cos(x_1 x_2) \sin(x_1 x_2) x_2 - 1/6 x_3 + u_1(s)) ds + \sigma dw_1 \\
dx_2 &= (\cos(x_1 x_2) \sin(x_1 x_2) x_1 - 1/6 x_3 + u_2(s)) ds + \sigma dw_2.
\end{align*}
\]

(56)

(57)

In Figure 2 we plot the potential \(\nu\) along with several uncontrolled trajectories of agents initialized at random locations. The agents collect into 4 stable and attracting equilibria. We design a cost function \(q(x) = 1/2 \left[ (x_1 - 1)^2 + (x_2 - 1)^2 \right] + (x_1 + 1)^2 + (x_2 + 1)^2\) to encourage the agents to move into two locations at \((-1, -1)\) and \((1, 1)\). We let \(R = 1, Q = 0.1, \sigma = 0.2\) and \(T = 4.0s\), with a time discretization step size of \(dt = 0.1\). We solve for \(f(t, x)\) at each timestep using our quadrature method with a fixed 2-d Gauss-Hermite grid spanning \([-2, 2]\) in both \(x_1\) and \(x_2\). We found 20 grid points in each dimension to yield good results (for a total of 400 grid points). We then plot the modified value \(\hat{v}(x, t) = -\sigma^2/2 \log(f(x, t))\). With this method we are able to find an optimal feedback control law for the entire domain of integration. We simulate 40 agents under this feedback control which have been initialized randomly (see Figure 3). Note that we are also able to solve the problem by calculating controls for each agent locally and independently using our quadrature method, modified to use a smaller grid (with width \(4\sigma(T - t)/\sqrt{dt}\) in each dimension), centered at the agent’s current position. Unlike with PDE solver-based solutions, we are able to find optimal controls for each agent locally. This is advantageous when the size of the state space is large and the number of agents is small. (We observed no difference between the optimal controls calculated with the global fixed grid quadrature method and those calculated locally.) The results of the simulation show that early on (\(t = 1.0s\)), the agents are pushed towards the center of the space. As time progresses, the agents are controlled towards the goal position at \((1, 1)\) and \((-1, -1)\) for \((t = 2.0s, 3.0s)\). At the final time, the agents are mainly concentrated around the goal regions \((t = 4.0s)\). The modified value \(\hat{v}\) is smallest at the goal state but also has valleys around the four stable equilibria.

We make our code publicly available at [GitHub repository](https://github.com/ddfan/pi_quadrature) on an Intel(R) Core(TM) i7-4980HQ CPU @ 2.80GHz machine, calculating the value function and simulating the agents took 21.7 seconds, using Python and Numpy’s linear algebra library. The code was written without any significant optimization or parallelization, with which the solver could be made far more efficient.

VI. CONCLUSIONS

In this paper we study the design of optimal controllers for large scale systems in which agents obey multidimensional nonlinear Langevin dynamics and provide a framework for closed-loop stability analysis of the fixed point density.
utilize an imaginary time Schrödinger PDE representation of the original optimality system, obtained by introducing a novel variable transform (section III-A), to facilitate the stability analysis. It is observed that spectral properties of the Schrödinger operator underlie the stability of fixed point density of the optimality system. In section IV we provide explicit control design constraints which guarantee closed-loop stability of the steady state density using these spectral properties.

The potential corresponding to the Schrödinger PDE is interpreted as the cost function of a related OCP with simple integrator dynamics (section III-C). This motivates a quadrature based algorithm, explained in section V, to compute the finite time optimal control and is demonstrated on a two dimensional large scale control example. It is observed that given an (uncontrolled) Langevin system there exists a corresponding control problem with simple integrator dynamics, such that the optimal control recovers the given passive dynamics.

The soliton theory used in [33] to study MFGs was based on a connection between NLS and MFGs for agents with simple integrator dynamics. In section III-B this connection was generalized to include MFG models in which agent dynamics lie in the general class of nonlinear Langevin dynamics. A topic of future work is therefore to extend and apply the theory of solitons to create a reduced order computational tool for this broader class of MFGs. These tools can then be used to design phase transitions (operating regimes) in multi-agent networked systems such as agile swarms [46] and electrical micro-grids [47]. Generalization of the presented approach to the case of second order Langevin systems is a natural extension which we intend to work on in the future. Finally, we will introduce sparse grids [48] in the proposed quadrature based finite time optimal control solver in a forthcoming publication, with the goals of speeding up computation and scaling to high dimensional systems.

References

[1] R. Brockett. Notes on the Control of the Liouville Equation. Springer-Verlag Berlin Heidelberg, 2012.
[2] R. Brockett and N. Khaneja. On the Stochastic Control of Quantum Ensembles. In: System Theory by Dijferis T.E., Schick I.C. (eds), volume 518 of The Springer International Series in Engineering and Computer Science, vol 518. Springer, Boston, MA, 2000.
[3] A. Palmer and D. Milutinovic. A hamiltonian approach using partial differential equations for open-loop stochastic optimal control. In American Control Conference (ACC), 2011, pages 2056–2061, June 2011.
[4] K. Bakshi and E. Theodorou. Infinite dimensional control of doubly stochastic jump diffusions. In Proceedings of 2016 IEEE 55th Conference on Decision and Control (CDC), pages 1145–1152, Las Vegas, USA, Dec 2016.
[5] Je-Shin Li, Justin Ruths, Tsyr-Yan Yu, Haribabu Arthanari, and Gerhard Wagner. Optimal pulse design in quantum control: A unified computational method. Proceedings of the National Academy of Sciences, 108(5):1879–1884, 2011.
[6] J. Li, I. Dasanayake, and J. Ruths. Control and synchronization of neuron ensembles. IEEE Transactions on Automatic Control, 58(8):1919–1930, Aug 2013.
[7] Nikolaos Correll and Alcherio Martinoli. Towards multi-robot inspection of industrial machinery - from distributed coverage algorithms to experiments with miniature robotic swarms. IEEE Robotics and Automation Magazine, 16(1):103–112, 2009.
[8] D. Milutinović. Utilizing Stochastic Processes for Computing Distributions of Large Size Robot Population Optimal Centralized Control, volume 83 of Springer Tracts in Advanced Robotics. Springer, Berlin, Heidelberg, 2013.
[9] Hanjun Li, Chunhan Feng, Henry Ehrhard, Yijun Shen, Bernardo Cobos, Fangbo Zhang, Karthik Elamvazhuthi, Spring Berman, Matt Haberland, and Andrea L. Bertozzi. Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment. In IROS 2017 - IEEE/RSJ International Conference on Intelligent Robots and Systems, volume 2017-September, pages 4344–4347, United States, 12 2017. Institute of Electrical and Electronics Engineers Inc.
[10] Gabriel Y. Weintraub, C. Lanier Benkard, and Benjamin Van Roy. Oblivious equilibrium: A mean field approximation for large-scale dynamic games. In Advances in Neural Information Processing Systems, 2005.
[11] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. IEEE Transactions on Automatic Control, 57(4):920–935, April 2012.
[12] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. Japanese Journal of Mathematics, 2(1):229–260, Mar 2007.
[13] M. Huang, P. E. Caines, and R. P. Malhame. Large-population cost-coupled lag problems with nonuniform agents: Individual-mass behavior and decentralized 949-nash equilibria. IEEE Transactions on Automatic Control, 52(9):1560–1571, Sept 2007.
[14] Pushkin Kachroo, Shaurya Agarwal, and Shankar Sastry. Inverse problem for non-viscous mean field control: Example from traffic. IEEE Transactions on Automatic Control, 61(11):3412–3421, 2016.
[15] Rene Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. Mean field games and systemic risk. arXiv:1306.2172, 2013.
[16] Romain Couillet, Samir M Perlaza, Hamidou Tembine, and Mérouane Debbah. Electrical vehicles in the smart grid: A mean field game analysis. IEEE Journal on Selected Areas in Communications, 30(6):1086–1096, 2012.
[17] Kaivalya Bakshi, Piyush Grover, and Evangelos A Theodorou. On mean field games with langevin dynamics. Transactions on Control of Networked Systems (to appear), page 061103, 2018.
[18] Piyush Grover, Kaivalya Bakshi, and Evangelos A Theodorou. A mean-field game model for homogeneous flocking. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28(6):061103, 2018.
[19] O. Guéant. A reference case for mean field games models. Journal de Mathématiques Pures et Appliquées, 92(3):276–294, 2009.
[20] Huibing Yin, Prashant G. Mehta, Sean P. Meyn, and Uday V. Shanbhag. Synchronization of coupled oscillators is a game. In American Control Conference, 2010. ACC ’10., June 2010. Submitted for publication, IEEE Trans. Auto. Control.

[21] M. Nourian, P. E. Caines, and R. P. Malhame. A mean field game synthesis of initial mean consensus problems: A continuum approach for non-gaussian behavior. IEEE Transactions on Automatic Control, 59(2):449–455, Feb 2014.

[22] Martin Burger, Marco Di Francesco, Peter A. Markowich, and Marie-Therese Wolfram. Mean field games with nonlinear mobilities in pedestrian dynamics. Discrete & Continuous Dynamical Systems - B, 19(1531-3492-2014-5-1311):1311, 2014.

[23] Alethea BT Barbaro, José A Cañizo, José A Carrillo, and Pierre Degond. Phase transitions in a kinetic flocking model of Cucker–Smale type. Multiscale Modeling & Simulation, 14(3):1063–1088, 2016.

[24] P. Chaudhari, A. Oberman, S. Osher, S. Soatto, and G. Carlier. Deep relaxation: partial differential equations for optimizing deep neural networks. arxiv, 2017.

[25] J. Melbourne, S. Talukdar, and M. Salapaka. Realizing information erasure in finite time. In Conference on Decision and Control (preprint arXiv:1809.09216[cond-mat.stat-mech]), Dec 2018.

[26] J. D. Cole. On a quasi-linear parabolic equation occurring in aerodynamics. Quarterly of Applied Mathematics, 9(3):225–236, 1951.

[27] E. Hopf. The partial differential equation ut + uux = uxx. Communications on Pure and Applied Mathematics, 3(3):201–230, 1950.

[28] W. H. Fleming and H. M. Soner. Controlled Markov processes and viscosity solutions. Applications of mathematics. Springer, New York, 2nd edition, 2006.

[29] H. J. Kappen. Linear theory for control of nonlinear stochastic systems. Phys. Rev. Lett., 95:200201, Nov 2005.

[30] W. Wiegerinck, B. van den Broek, and H. J. Kappen. Stochastic optimal control in continuous space-time multi-agent system. In UAI, 2006.

[31] E. Theodorou. Iterative Path Integral Stochastic Optimal Control: Theory and Applications to Motor Control. PhD thesis, University of Southern California, Los Angeles, CA, USA, 2011.

[32] Haomin Zhou Shui-Nee Chow, Wuchen Li. A discrete Schrödinger equation via optimal transport on graphs. arXiv:1705.07583, 2017.

[33] Y. Chen, M. Pavon, and T. Georgiou. On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. Journal of Optimal Control Theory and Applications, 169:671-691, 2016.

[34] Igor Swiecicki, Thierry Gobron, and Denis Ullmo. Schrödinger approach to mean field games. Phys. Rev. Lett., 116:128701, Mar 2016.

[35] J. Yong and X. Zhou. Stochastic Controls: Hamiltonian Systems and HJB Equations. Applications of mathematics. Springer, 1999.

[36] V. Borkar. Ergodic control of diffusions. In International Congress of Mathematicians, volume 3, pages 1299–1309, Aug 2016.

[37] G. Pavliotis. Stochastic Processes and Applications. Springer, 1st edition, 2014.

[38] H. Risken. The Fokker-Planck Equation: Methods of Solution and Applications. Number 16 in Springer Series in Synergetics. Springer-Verlag, 1984.

[39] F.A. Berezin and Schubin M.A. The Schrödinger equation. Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[40] E Todorov. Eigenfunction approximation methods for linearly-solvable optimal control problems. In IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning, pages 161–168, 2009.

[41] Tobin A Driscoll, Nicholas Hale, and Lloyd N Trefethen. Chebfun guide, 2014.