# LARGE SCALE STOCHASTIC CONTROL: ALGORITHMS, OPTIMALITY AND STABILITY

A Dissertation
Presented to
The Academic Faculty

By

Kaivalya Bakshi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Aerospace Engineering

Georgia Institute of Technology

May 2019

**LARGE SCALE STOCHASTIC CONTROL: ALGORITHMS, OPTIMALITY AND STABILITY**

Approved by:

Dr. Evangelos Theodorou, Advisor
School of Aerospace Engineering
*Georgia Institute of Technology*

Dr. Piyush Grover
Control and Dynamical Systems
*Mitsubishi Electric Research Laboratories*

Dr. Eric Feron
School of Aerospace Engineering
*Georgia Institute of Technology*

Dr. Ionel Popescu
School of Mathematics
*Georgia Institute of Technology*

Dr. Paul Bogdan
School of Electrical Engineering
*University of Southern California, Viterbi*

Dr. Yongxin Chen
School of Aerospace Engineering
*Georgia Institute of Technology*

Date Approved: November 12, 2018

To family.

# ACKNOWLEDGEMENTS

Foremost, I thank my advisor Prof. Evangelos Theodorou for providing direction to my research and putting his faith in our ideas. Specifically, I am ingratiated for his confidence in me, working closely with me in the initial period of my studies and letting me work independently in the latter. I was fortunate to work with Dr. Piyush Grover in the Control and Dynamical Systems Group at Mitsubishi Electric Research Labs, Cambridge, MA in the concluding years of my docotoral studies. I express my gratitude to Dr. Grover for introducing me to the exciting world of mean-field games, for treating me very kindly as a student at some times and as a colleague at others. It was a pleasure to work with him on exciting and new topics in dynamical systems theory.

I had several fun and insightful chats with Prof. Popescu in my PhD journey which motivated me me to think anew about research. I also thank Prof. Yongxin Chen for his informed comments and suggestions on my thesis topic. I am indebted to Prof. Eric Feron and Prof. Paul Bogdan for serving on my thesis committee.

An interesting aspect of this PhD journey was our lab and a unique set of lab members, brimming with confidence and energy. I congratulate Prof. Theodorou for his contribution in setting up this culture. I thank my collaborators and labmates who always seemed up for an interesting discussion, even if it required them to invest their valuable time and energy: David Fan, Yunpeng Pan, Harleen Brar, George Boutselis, Manan Gandhi, Kamil Saigol, Keuntaek Lee, Marcus Pereira, Andrew Aldrich, Ethan Evans, Yannis Exarchos, Grady Williams, Pat Wang, Emre Yilmaz, Gabriel Nakajima An and Oktay Arslan.

Finally, I thank my friends for all those wonderful conversations and companionship: Srinivas, Karthik, Niveditha, Nivedita, Nithanth, Rahul, Aparna, Pratibha, Raunak, Achyut, Mayank, Indraneel, Dheeraj, Sudarshan, Sandeep I am grateful for my mother's support and patience, which is like an ocean, my father's encouragement, my brother, whom I've been mooching off for a while, and my grandmother.

# TABLE OF CONTENTS

## LIST OF FIGURES

# SUMMARY


Optimal control of large-scale multi-agent networked systems which describe social networks, macro-economies, traffic and robot swarms is a topic of interest in engineering, biophysics and economics. A central issue is constructing scalable control-theoretic frameworks when the number of agents is infinite.

In this work, we exploit PDE representations of the optimality laws in order to provide a tractable approach to ensemble (open loop) and closed loop control of such systems. A centralized open loop optimal control problem of an ensemble of agents driven by jump noise is solved by a sampling algorithm based on the infinite dimensional minimum principle to solve it. The relationship between the infinite dimensional minimum principle and dynamic programming principles is established for this problem.

Mean field game (MFG) models expressed as PDE systems are used to describe emergent phenomenon in decentralized feedback optimal control models of a continuum of interacting agents with stochastic dynamics. However, stability analysis of MFG models remains a challenging problem, since they exhibit non-unique solutions in the absence of a monotonicity assumption on the cost function. This thesis addresses the key issue of stability and control design in MFGs. Specifically, we present detailed results on a models for flocking and population evolution.

An interesting connection between MFG models and the imaginary-time Schrödinger equation is used to obtain explicit stability constraints on the control design in the case of non-interacting agents. Compared to prior works on this topic which apply only to agents obeying very simple integrator dynamics, we treat nonlinear agent dynamics and also provide analytical design constraints.

# CHAPTER 1
# INTRODUCTION

## 1.1 Motivation and Prior Work

Control of continuous time nonlinear stochastic differential equations (SDEs) is at the core of nonlinear stochastic optimal control (SOC) theory. Stochastic systems obeying SDEs with Gaussian and non-Gaussian noise appear in several areas of research including economics, autonomous and biological systems and population modeling [1],[2, 3, 4, 5]. In most applications, the noise is used to model model or environmental uncertainty. Applications of stochastic control include robotics and autonomous systems such as in the control of ground and aerial vehicles, articulated mechanisms and manipulators, and humanoid robots [108110, 123, 127, 131], for modeling the control of movement in computational neuroscience [130, 132] and stock option pricing in financial engineering [102, 121]. Certain systems involve large number of more or less identical subsystems which may be manipulated by individual or identical input signals. If the number of subsystems is as large as $10^3 - 10^6$, it becomes difficult to conceive of a control framework treating each individual separately. Examples of such systems are schools of fish and neurons in bio-physics, agents in a wireless network and swarms of aerial drones.

In this thesis we study the optimal control problems and models related to non-cooperative multi-agent and possibly networked systems, in the case that the number of agents is very large. Individual agents in such systems maybe controlled by *individual state feedback* or an *identical broadcast* input signal. The overarching goals are to provide (1) scalable control-theoretic frameworks for such systems (2) control design constraints to guarantee stability and (3) numerical schemes to solve the control problem.

The contents of this thesis are expressed in two parts. The first part is devoted to development of an ensemble control algorithm for the stochastic control of jump diffusion processes and understanding the relationship between various optimality principles in ensemble control. The second part considers synthesis and stability analysis of large scale, individual feedback, non-cooperative and possibly networked multi-agent systems, also known as mean-field games, for applications in modeling flocks and control design.

### 1.1.1 Ensemble Control

The term ensemble control applies to systems consisting of of a system consisting of a large number of identical stochastic subsystems being manipulated by a single source of command signals. In this context, the collection of subsystems is called the system and an individual subsystem is called an elemental system. Examples of ensembles appear in classical thermodynamics which models collections of identical particles, weakly interacting particles appearing in quantum systems such as in nuclear magnetic resonance problem and dynamical models of neurons. Prior works on optimal open-loop or ensemble (broadcast) control consider several copies of a particular deterministic [6] or stochastic ([7], [8] system and have applications in quantum control [9] and neuroscience [10].

These applications have the common goal of controlling large-scale weakly interacting individual systems using a single or perhaps a small number of control inputs. This means that the control applied to each elemental system is identical, that is, lacks local feedback, but depends on the overall distribution of the system at each instant of time. Considering infinite copies of the finite stochastic state models of the elemental systems clearly cannot provide a scalable mathematical framework in this case. Treating the collective ensemble dynamics modeled by the Liouville or Fokker Planck (FP) PDE governing the distribution of states of elemental systems provides is a more tenable approach. The infinite dimensional minimum principle (MP) has been applied previously to solve such optimal control problems when the individual subsystems are driven by Gaussian noise. The connection between the MP and dynamic programming for ensemble control was qualitatively explored, in the case of diffusions [20] and jump diffusions [21]. On the other hand the dynamic programming principle obeyed by the value function corresponding to the infinite dimensional problems has been theoretically explored [22], [23]. The motivation of this thesis on the topic of of ensemble control is to present a complete exposition of the optimality principles applied to ensemble control by explaining the relationship between the MP and dynamic programming and to devise an algorithm for this control problem when the elemental subsystems with jump diffusion dynamics.

### 1.1.2 Mean Field Games

A standard idea in engineering, economics and biology is regulation using local feedback information and is used to model decision making in large-size populations of *rational* agents, for example in economics. Therefore 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 social networks, telecommunications, electrical micro-grids, renewable energy systems, vehicle formations, competing or cooperating mobile robots, micro-economics, finance and bio-physics such as in flocks or swarms. Optimal feedback control applications of large-size populations of *small* robots with individual state-feedback controllers have been proposed for inspection of industrial machinery [11], centralized control of hybrid automata [12] and decentralized control of robotic bee swarms for pollinating crops [13].

There has therefore been an interest in modeling, control and optimization of large-scale multi-agent stochastic dynamical systems in the mathematics and controls community. The sources of complexity in such systems are the uncertainty in individual agent dynamics or communication between agents, interaction among agents and the large number ( $10^3$ to $10^6$) of agents. Due to the large number of agents it is more prudent to develop decentralized solutions to such control problems so that individual agents take actions based on their local information and certain statistical information about the population. Consequently providing tractable control-theoretic frameworks for modeling and control of large-scale systems is a critical topic shared by several areas of research.

A viable approach which provides scalable mathematical models for such systems applies the notion of the 'mean-field', which is inspired from particle physics which focuses on quantifying the interaction among particles. Traditional physics approaches which study interactions between couples or triples of particles cannot be used in case of particle physics due to their very large numbers. The mean-field approach is the statistical idea that it is sufficient to study interactions between particles and the collection of all other particles contained in a media, which is referred to as the mean-field. An example of such behavior is air pressure which is created by microscopic motions of particles but impacts each particle in a macroscopic way. This micro-macro interaction is a salient feature of the mean-field approach. The interactions between individual agents is therefore replaced by the interaction between a single particle and the mean-field. This is the key idea which makes the mathematically tractable frameworks possible for large-scale systems. Finally, the optimal control framework allows us to model explicit interaction between agents and the mean-field by using state dependent costs which depend on the statistics of the population.

One of the earliest applications of mean field theory to large-scale systems using optimal control is seen in [14], which presents a Nash equilibrium interpretation of non-cooperative behavior of a continuum of agents. Several works on game-theoretic models of large population models have appeared in the economics literature following this paper. The idea of the general equilibrium lies at the core of modern economics. The earliest work formally applying the mean-field approach in combination with a optimal control game-theoretic interpretation of large population dynamics [14] approximates a Markov perfect equilibrium (MPE) of a dynamic game involving several firms using the idea of the oblivious equilibrium. Oblivious equilibrium describes a model in which each individual agent takes decisions based on its own state variable and the mean-field but is oblivious to the state of the overall system.

Large scale non-cooperative multi-agent systems involving coupled costs were introduced independently by Huang et. al [15] and as as mean field games (MFG) by Lasry et. al [16] in 2007. The key ideas of MFG theory are assuming an infinite number of anonymous agents with rational expectations and that individual decisions are based on statistical information about the collection of agents. This theory has therefore become a viable tool by providing a tractable framework to model self-organizing large-scale networked-systems due to its game-theoretic optimal control interpretation of emergent behavior observed in bio-physical systems [8], financial [6], traffic [5] and energy [7].

Figure 1.1: Flocking of birds (left) and (right) flow of city traffic in a multi-way intersection.

In the continuum approach, the simplest MFG models prescribe interaction between the agent and mean-field through density dependent state cost functions and are synthesized as standard stochastic optimal control problems (OCP). Quadratic MFGs refer to systems with quadratic control cost and control affine agent dynamics constitute. If the state cost function has only local density dependence and is strictly increasing, steady state solutions to the MF system can be shown to be unique [9] in many cases. In the continuum case, MFG models are synthesized as standard [3] stochastic optimal control problems (OCP). Fokker Planck (FP) and Hamilton Jacobi Bellman (HJB) equations form a fully coupled mean field (MF) optimality system governing agent density and value functions. Assumptions of quadratic control cost and control affine agent dynamics constitute quadratic MFG models [4]. However, in the absence of monotonicity of the state cost function, MFGs may exhibit non-unique solutions and related phase transitions ( [4], [8], [10]). Real-world large-scale networked systems often have several operating regimes so that non-monotonicity in the corresponding MFG models is expected to be the norm, rather than an exception. The closed-loop stability analysis of MFG models that do not satisfy this monotonicity condition usually has to be treated on a case-by-case basis. We say that a given fixed point of the MFG is called (linearly) closed-loop stable if any perturbation to the fixed-point density decays to zero under the action of the control, where both the density and control evolution are computed using the (linearized) coupled forward-backward system of FP-HJB PDEs.

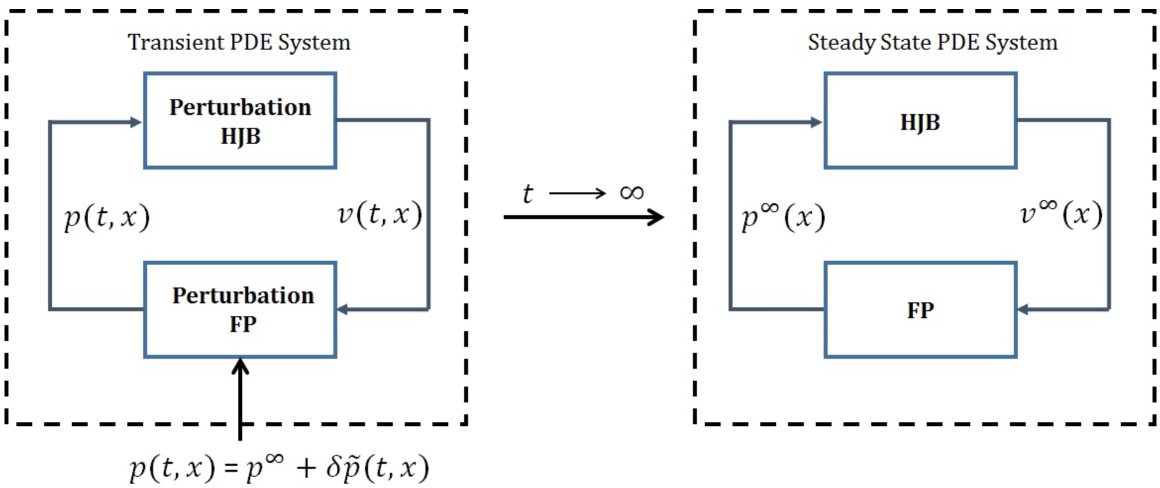



Figure 1.2: Stability of MFGs models

Stability of a MFGs was first studied by Guant [11] for a reference model with a negative log density cost. Other works on this topic include a Kuramoto oscillator model with nonlocal cost coupling by Yin et. al [10] and a mean consensus cost by Nourian et. al ( [12], [13]). These prior works are limited by the fact that they exclusively treat the case where the agents obey very simple integrator systems. On the other hand, the MF approach to large-scale networked systems with nonlinear agent dynamics have proved to be useful in modeling flocks [15], neural networks [16], crowds [14] and robotic control [17]. This motivates our work on synthesis and linear stability analysis of MFG models for applications to modeling flocks and nonlinear stability analysis, specifically in the case that agents have nonlinear mobilities.

Multi-agent populations consisting of a large number of identical and non-cooperative agents are of interest in several applications including macro-economics, robotic swarms, traffic and neuroscience. Optimal open-loop or closed-loop ensemble (broadcast) control has been used in prior works which consider several copies of a particular deterministic [6] or stochastic ([7], [8], [17]) system and have applications in quantum control [9] and neuroscience [10]. Optimal control models of collective behavior typically treat agents which are driven by individual noise and state-feedback control, and interact with each other through the coupling of their passive dynamics or utility with the overall statistics of the population. The *mean-field* approach provides a tractable framework for describing collective behavior of a continuum of agents, by approximating their individual actions [14] as the *oblivious* control [18] of a single *representative* agent and was formalized by the Mean field games (MFGs) ([16], [15]) framework.

Most works on MFGs consider explicit *interactions* between agents through the dependence of their dynamics or cost function on the population density. 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 the 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. Agents which lack explicit interaction have been studied using the mean-field approach in macro-economics [14]. In certain physical systems such as robot swarms ([12],[13]), 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. Designing optimal decentralized controllers for such systems, which guarantee closed-loop stability of the stationary density of agents under the action of their individual steady state controls is therefore a challenging problem in control theory.

The final topic of interest in this thesis related to MFGs is the connection between stochastic control and quantum mechanics. This connection is well known and is rooted in the deep relationship between PDE theory and SDEs. More recently, the related theoretical facts, in particular the path integral representation of the log transformed value function has been popular in applications. The knowledge of this log transform dates back to Schrödinger and was formalized in 1950 in the context of the heat and fluid PDEs by Cole and Hopf independently. This transform proved to be of fundamental significance in showing the relationship between noisy Newtonian systems and quantum mechanics. In 2017, the same change of variable was used along with a hermit transform to show the equivalence of the coupled FP-HJB system comprising the MFG optimality system and the Schrödinger equation pair [19], when agents obey simple integrator dynamics and interact locally. Since the Schrödinger equations are linear, leveraging this transformation is a powerful technique to analyze stability and obtain numerical schemes to solve the optimality system. Thus the theory of solitons in quantum mechanics was fashioned for understanding the mechanics of a MFG model in ([20]). However the present literature is limited to treating the simplest possible integrator agent dynamics for specific costs. We introduce a closely related but novel transformation which enables us to make the MFG-Schrödinger equation connection and present general control design constraints for stability and synthesize a computationally advantageous sampling based method to solve the optimality system.

## 1.2 Structure of the Thesis and Contributions

- **Chapter 3 *Ensemble control of Jump Diffusions*:** In chapter 3 we solve the problem of the control of a density an ensemble of stochastic ensembles driven by marked jump diffusions. We assume the input to be a broadcast controller, which is identical for each agent and uses only the macro density as the feedback. We derive the optimality system using the MP which gives the necessary conditions for the optimal control. A sampling control algorithm is derived by solving a nonlinear HJB PIDE using only forward sampling of the open loop individual element dynamics. We explain the infinite dimensional MP-DPP relationship for this problem by explicitly showing the relationship between the costate and infinite dimensional value functions.

In relation to prior work on control of jump diffusion processes, the proposed algorithm applies to the most general class of marked jump diffusions. Compared to the restricted linear-quadratic problems which can be solved by modifications of linear-quadratic-regulators, our theoretic and algorithmic framework admits nonlinear passive dynamics. A closely related sampling approach was used in [8], [12] but applies only to simple jump diffusions. Additionally, we introduce state parameterization of the controls to incorporate implicit feedback to enhance the performance of our algorithm which appears first in our work [17].

The relationship between the MP and DPP is a well explored topic in control theory for finite dimensional systems. The explanation hinges on showing the relation between the Lagrangian multiplier or costate and the value function or optimal cost-to-go. In the context of infinite dimensional systems, this topic was first broached in [], wherein the authors allude to the fact that the relationship in the finite dimensional case can be viewed as a special case of the infinite dimensional case. We explicitly show the relationship in the infinite dimensional setting when elemental systems obey marked jump diffusion dynamics by directly showing the relation between the infinite dimensional value function and the costate function.

- **Chapter 4** ***Mean Field Games for Agents with Langevin Dynamics*****:** In this chapter we analyze the linear stability of MFG models in the case that the passive agent dynamics obey nonlinear Langevin dynamics. We explain how this result can be extended to the more general case when agents obey a class of nonlinear dynamics called reversible diffusions. Explicit control design constraints required to guarantee stability are obtained for two specific models with local and non-local density dependent cost functions. Further, it is observed and verified numerically, that the fixed point static controller is also stabilizing under small density perturbations.

Our main contribution on this topic is to generalize the functional-analytic method used to analyze linear stability of MFGs for agents with integrator dynamics to the class of MFGs in which agents obey Langevin dynamics. Linear stability was introduced for a specific MFG by Guéant [21] in 2009. This was followed by works on models for a Kuramoto synchronization model [18] and a mean consensus model [22]. In [23] we show that the detailed balance property of reversible diffusions allows us to generalize the linear systems based method introduced in [21] to show stability of MFG models in which agents obey this broad class of nonlinear diffusion dynamics. Explicit stability constraints on the control design are obtained using this method for two models which have local and non-local interactions between agents.

- **Chapter 5** ***Modeling Flocks using Mean Field Games*****:** A control system mimicking homogeneous flocking is presented in chapter 5 by constructing a MFG with non-cooperative agents possessing nonlinear mobilities.

  With respect to prior works we show stability results for the flocking MFG model in which agents possess nonlinear mobilities. Prior works based on consensus models [22], [24], [18] apply exclusively to the case in which agents obey integrator dynamics. Phase transitions observed in an earlier proposed uncontrolled model [25], [26] are recovered numerically from the proposed controlled flocking model, along with some new ones, by tuning the control parameter. A contraction mapping argument is used to show stability of the proposed model. The low-rank perturbative nature of the nonlocal term in the forward-backward optimality system governing the state and control distributions is exploited to provide a closed-loop linear stability analysis demonstrating that our model exhibits bifurcations similar to those found in the empirical model.

- **Chapter 6** ***Schrödinger Approach to Large Scale Control*****:** In this chapter we consider the problem of designing state-feedback controllers which guarantee closed-loop stability and computation of the control for large-size populations of identical, non-cooperative and non-interacting stochastic agents. A novel Cole-Hopf type transform is introduced to represent the optimality system constituted by coupled forward-backward PDEs in terms of decoupled Schrödinger equations. We propose a quadrature based sampling algorithm to compute the control in the finite time horizon case.

  Prior works on change of variables to treat the HJB PDEs [27], [28], [29] use the logarithmic transform of the value function to linearize the HJB PDE. This transform was first formalized by Cole and Hopf independently. In [19] this transform was used to express the coupled FP-HJB system for certain MFG models with integrator agent dynamics, as the Schrödinger PDE pair. In this chapter we explain a closely related but novel transform introduced by us in [23] which makes it possible to obtain the same result in case agents obey nonlinear Langevin dynamics. In the context of algorithms to solve the optimality system, the sampling based algorithms proposed based on the introduced transform was introduced first by us in [23].

- **Chapter 7 Conclusions:** In this chapter we state our conclusions and give some future research directions.

# CHAPTER 2

# TECHNICAL BACKGROUND

In this chapter, we explain the notation used in this thesis and provide a brief introduction to the required technical background for this work. In section 2.1, we describe the basic building blocks of stochastic systems including probability spaces, stochastic processes and the Brownian motion. In section 2.2 we review basic results on SDEs and related existence and uniqueness results and state the standard formulation of the stochastic optimal control problem. Finally, in section 2.4 we state the stochastic control problems we study in this thesis and also state the principle of dynamic programming and corresponding optimality systems.

Note that the scope of this thesis extends beyond the fundamental literature on stochastic control elucidated this chapter. For instance, chapter 3 is concerned with control of densities affiliated with stochastic systems driven by non-Brownian noise and chapters 4, 5 and 6 consider interacting and non-interacting multi-agent control problems. The background material specific to these topics is not commonly contained in standard texts on stochastic control and will be discussed in detail in those chapters.

## 2.1 Notation

The following list summarizes frequently used notation and abbreviations.

| | |
|---|---|
| $\mathbb{R}^d$ | $d$ dimensional Euclidean space |
| $A^{\mathrm{T}}$ | transpose of matrix A |
| $\mathrm{tr}(A)$ | the trace of matrix A |
| $C^n$ | the space of functions $f : A \to \mathbb{R}$ which is $n$ times continuously differentiable on the set $A$ |
| $C^{n,m}(A \times B)$ | the space of functions $f : A \times B \to \mathbb{R}$ which are $C^n$ on the set $A$ and $C^m$ on set $B$ |
| $\nabla(\cdot), (\cdot)_x$ | the gradient operator |
| $\nabla \cdot (\cdot)$ | the divergence operator |
| $(\cdot)_{xx}$ | the Hessian operator |
| $\Delta(\cdot)$ | the Laplacian operator |
| $:=$ | defined as |
| $\approx$ | approximately equal to |
| $\mathcal{F}_s$ | the filtration at time s |
| $\mathbb{E}$ | the expectation operator |
| $\mathcal{N}(\mu, \sigma^2)$ | Gaussian (normal) distribution with mean $\mu$ and variance $\sigma^2$ |
| $w_t$ | the standard Brownian motion |
| $L^p([0,T]; \mathbb{R}^n)$ | set of $\{\mathcal{F}_t\}_{0 \le t}$ adapted, $p$ integrable, $\mathbb{R}^n$ valued processes |
| DPP | Dynamic Programming Principle |
| MP | Minimum Principle |
| HJB | Hamilton Jacobi Bellman equation |
| FP | Fokker Planck equation |

## 2.2 Stochastic Processes

In this section we provide a brief summary of the mathematical background required for this thesis. Please see references [30], [31], [28] for further details.

The basic building block for constructing a probability space is the $\sigma$ algebra which is defined below.

**Definition 2.2.1.** *($\sigma$ algebra) Let $\Omega$ be nonempty. A nonempty class $\mathcal{F} \subseteq 2^\Omega$ ($2^\Omega$) being the set of all subsets in $\Omega$ is called a $\sigma$ algebra if $\Omega \in \mathcal{F}$, $B \setminus A \in \mathcal{F}$ for every $A, B \in \mathcal{F}$ and $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ if $A_i \in \mathcal{F}$ for every $i = 1, 2, 3...$ We call $(\Omega, \mathcal{F})$ as the measurable space.*

**Definition 2.2.2.** *(Probability space) Let $\Omega$ be a nonempty set and $\mathcal{F}$ be a $\sigma$ field on $\Omega$ so that $(\sigma, \mathcal{F})$ is a measurable space. We call en element $\omega \in \Omega$ a sample. Further, any $A \in \mathcal{F}$ is called an event. A map $\mathbb{P} : \mathcal{F} \to [0, 1]$ is called a probability measure if $\mathbb{P}(\phi) = 0$, $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ if $A_i \in \mathcal{F}$ and $A_i \cap A_j = \phi$ for all $i, j = 1, 2, 3...$, $i \neq j$. We call the triple $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space. We call $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space if $A \in \mathcal{F}$ for every $A \subseteq \Omega$ with the outer measure of $\mathbb{P}$ being zero.*

**Definition 2.2.3.** *(Independence of events) let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Two events $A, B \in \mathcal{F}$ are called independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Random variables are functions from the event space to the set of real numbers.

**Definition 2.2.4.** *(Random variable) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A function $x : \Omega \to \mathbb{R}^d$ is called $\mathcal{F}$ measurable if $x^{-1}(A) = \{\omega \in \Omega | x(\omega) \in A\} \in \mathcal{F}$. A $\mathcal{F}$ measurable function $x : \Omega \to \mathbb{R}^d$ is called a random variable. We denote by $\mathbb{P}_x$ the induced probability measure of the random variable $x$ defined as $\mathbb{P}_x(A) := \mathbb{P}(x^{-1}(A))$. We define the mathematical expectation of $x$ is defined as*

$$\mathbb{E}[x] := \int_\Omega x(\omega) \mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{R}^d} y \; \mathrm{d}\mathbb{P}_x(y).$$

*Further, if $f : \mathbb{R}^d \to \mathbb{R}^n$ is measurable, then we define*

$$\mathbb{E}[f(x)] := \int_\Omega f(x(\omega)) \mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{R}^d} f(y) \; \mathrm{d}\mathbb{P}_x(y).$$

The definition of independent random variables follows from that of independence of events. If two random variables $x, y$ on $\Omega$ then $\mathbb{E}[xy] = \mathbb{E}[x]\mathbb{E}[y]$.

**Definition 2.2.5.** *(Stochastic processes) We define a stochastic process as being a set of parameterized random variables* $\{x_t\}_{t\in A}$ *on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$.

In this thesis we exclusively use the sets of positive numbers $A = [0, +\infty]$ or the interval $A = [0, T]$ where $T > 0$ as the parameter space. The notation $x_s(\omega)$ is used to denote the stochastic process consisting of random variables $x : \Omega \to \mathbb{R}^d$, with the shorthand notation $x_s$.

**Definition 2.2.6.** *(Filtration and adapted process) Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. A filtration on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is defined as the set of* $\sigma$ *algebras* $\mathcal{F}_t$*, denoted by* $\{\mathcal{F}_t\}_{t\geq 0}$*, such that* $\mathcal{F}_s \subset \mathcal{F}_t$ *for every* $0 \leq s < t$*. The stochastic process* $\{x_s\}_{s\geq 0}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called* $\mathcal{F}_t$ *adapted if* $x_s : \Omega \to \mathbb{R}^d$ *is* $\mathcal{F}_t$ *measurable for every* $t \geq 0$.

In this work, when we say that a process is adapted to a filtration, we also mean that it is progressively measurable with respect to that process.

**Definition 2.2.7.** *(Usual Condition) We say that* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t}, \mathbb{P})$ *satisfies the usual condition if* $(\Omega, \mathcal{F}, \mathbb{P})$ *is complete,* $\mathcal{F}_0$ *contains all the* $\mathbb{P}$ *null sets in* $\mathcal{F}$ *and* $\{\mathcal{F}_t\}_{0\leq t}$ *is right continuous.*

**Definition 2.2.8.** *(Square integrable stochastic process) We say that a stochastic process* $x_s$ *is square integrable if* $\mathbb{E}[\int_t^T x_s^2 \mathrm{d}s] < +\infty$ *for every* $T > t$.

Most stochastic systems in the literature typically use a specific type of model for the stochastic process governing the noise, namely, the Brownian motion. However In this thesis we work on stochastic systems which are driven by Brownian as well as non-Brownian noise, specifically the Poisson process. However we will elaborate on the class of Poisson processes we consider in detail only in chapter 3. Below we define the standard Brownian process which is widely used to model noise in control and estimation.

**Definition 2.2.9.** *(Standard Brownian Motion) Standard Brownian motion is a stochastic process denoted as* $\{w_t\}_{t\geq 0}$ *is a stochastic process on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *which satisfies the following rules:*

*(i) zero initial condition:* $w_0 = 0$

*(ii) independence of increments: the random variables* $w_{k+1} - w_k$ *are mutually independent for* $k = 0, 1, 2...$

*(iii) Gaussian increments with variance increasing linearly with time:* $w_t - w_s \sim \mathcal{N}(0, s)$

*(iv) continuous paths:* $w_t(\omega)$ *is everywhere continuous for almost every* $\omega \in \Omega$*.*

This definition can be extended easily to the case of multi-dimensional Brownian motions. Note the following properties in the one dimensional case: $\mathbb{E}[w_t - w_s] = 0$ and $\mathbb{E}[(w_t - w_s)^2] = (t - s)$ from the Gaussian increments assumption. Further, note that the ratio $\mathrm{d}w_t/\mathrm{d}t \sim \mathcal{N}(0, 1/\mathrm{d}t)$ referred to as white noise in engineering applications, technically has infinite variance as $\mathrm{d}t \to +\infty$.

We do note discuss in details the stochastic integration rule called Itô's rule and the related itô stochastic calculus for the Brownian motion. Let $w_t$ be a standard Brownian motion and $x_t$ be a measurable, square integrable, $\mathcal{F}_t$ adapted process. The Itô integral of $x_t$ over $w_t$ to time $T$ is the stochastic process denoted by

$$I_t = \int_0^t x_s \mathrm{d}w_s. \tag{2.1}$$

We refer the reader to the book [28] for a detailed explanation of the Itô calculus. In this work, we exclusively use Itô integration when we refer to a stochastic integral.

## 2.3 Stochastic Differential Equations

Consider the difference equation

$$\mathrm{d}x_t = f(t, x_t) + \sigma(t, x_t)\mathrm{d}w_t$$

with $x_0 = \xi \ \mathbb{P} - a.s$, which governs a stochastic process $x_t$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In this equation we would like to obtain the random variable $x_t$, which can be obtained by the Itô integration

$$x_t = \xi + \int_0^t f(s, x_s)\mathrm{d}s + \int_0^t \sigma(s, x_s)\mathrm{d}w_s.$$

We utilize this model of stochastic processes in this work and assume that $f : [0, +\infty) \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, +\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $w_s$ is a standard $m$ dimensional Brownian motion. We assume here that the initial state is known deterministically $\mathbb{P}$ a.s. The functions above are required to have certain continuity properties in order to ensure that the assumed model has a solution and that it is unique. Namely, we require Lipschitz continuity and almost linear growth conditions on the passive dynamics $f$ and noise matrix $\sigma$, that is there exist $C, D > 0$ such that for all $t \in [0, +\infty)$, $x, y \in \mathbb{R}^d$ such that

(E1) $|f(t, x) - f(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq C|x - y|$

(E2) $|f(t, x)| + |\sigma(t, x)| \leq D(1 + |x|)$

**Theorem 2.3.1.** *If $T > 0$, $f$ and $\sigma$ are uniformly continuous, measurable functions and assumptions (E1), (E2) are true, then the SDE* (2.3) *has a unique, square integrable and adapted solution for all* $0 \leq t < T$.

Consider the controlled SDE

$$\mathrm{d}x_t = f(t, x_t, u(t)) + \sigma(t, x_t, u(t))\mathrm{d}w_t$$

with $x_0 = \xi \ \mathbb{P} - a.s$, which governs a stochastic process $x_t$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We utilize this model of controlled stochastic processes in this work and assume that $f : [0, +\infty) \times \mathbb{R}^d \times U \to \mathbb{R}^d$, $\sigma : [0, +\infty) \times \mathbb{R}^d \times U \to \mathbb{R}^{d\times m}$ and $w_s$ is a standard $m$ dimensional Brownian motion. We assume that $U$ is a separable metric space and $T > 0$ is a fixed number.

The function $u(t)$ is referred to as the control, action, decision or policy of the decision maker or controller. We assume that the control sequence $u(\cdot) \in U[0,T] := \{u : [0,T] \to U | u(\cdot)$ is measurable$\}$ which is the class of feasible controls. The control is explicitly parameterized by the time, but at any instant of time possesses knowledge about the state of the system as specified by the information field given by the filtration $\{\mathcal{F}_t\}_{0\leq t}$. In this thesis, we assume that the controls are *not anticipative* or $u(\cdot)$ is $\{\mathcal{F}_t\}_{0\leq t}$ adapted, that is it cannot foretell the future of the system, as a result of the inherent uncertainty in the stochastic model. A simple consequence of this pertinent assumption the controller cannot execute her/his decision meant for a particular time, before that time arrives. Therefore, the control $u(\cdot) \in \mathcal{U}[0,T] := \{u : [0,T] \times \Omega \to U | u(\cdot)$ is measurable and $\{\mathcal{F}_t\}_{0\leq t}$ adapted$\}$. Notice that the domain $U$ specifies a time invariant control constraint. However, in most cases we will choose to apply control constraints implicitly through the cost function, for various problems encountered in this thesis.

In lieu with theorem 2.3.1, we introduce the following assumptions to ensure that the controlled SDE (2.3) has unique solutions given an admissible control. There exist $C, D > 0$ such that for all $t \in [0, +\infty)$, $x, y \in \mathbb{R}^d$ such that

(E3) $|f(t,x,u) - f(t,\hat{x},\hat{u})| + |\sigma(t,x,u) - \sigma(t,\hat{x},\hat{u})| \leq C(|x-\hat{x}| + |u-\hat{u}|)$

(E4) $|\sigma(t,x,u) - \sigma(t,\hat{x},\hat{u})| + |\sigma(t,x,u) - \sigma(t,\hat{x},\hat{u})| \leq C(|x-\hat{x}| + |u-\hat{u}|)$

(E5) $|f(t,x,u)| + |\sigma(t,x,u)| \leq D(1 + |x| + |u|)$.

We now introduce the cost functional

$$J(u(\cdot)) := \mathbb{E}[\int_0^T \ell(s, x_s, u(s))\,\mathrm{d}s + \Phi(T, x_T)] \tag{2.2}$$

**Definition 2.3.1.** *(admissible controls) Let* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_0 \leq t, \mathcal{P})$ *satisfy the usual condition. A control* $u(\cdot)$ *for the system* (2.3) *is called admissible and* $(x_\cdot, u(\cdot))$ *is called an admissible pair if* *(i)* $u(\cdot) \in U[0,T]$

*(ii)* $x.$ *is the unique solution to the SDE*

*(iii)* $\ell \in L^1_{\mathcal{F}}[0,T]$, $\Phi \in L^1_{\mathcal{F}_T}([0,T])$.

*The set of all admissible controls is denoted by* $\mathcal{U}_{ad}[0,T]$.

The following type of stochastic optimal control problem for a single agent is dealt with in this work: find $u(\cdot) \in \mathcal{U}_{ad}[0,T]$, if it exists, such that $J(u^*(\cdot)) = \underset{u(\cdot)\in\mathcal{U}_{ad}}{J(u(\cdot))}$. We will abuse our own notation for convenience and write $\mathcal{U}_{ad}[0,T]$ simply as $\mathcal{U}[0,T]$ in the following part of this text.

## 2.4 Dynamic Programming Principle

A standard tool in optimal control is to create a look up table of the optimal cost-to-go map from which the control maybe inferred by a form of gradient descent on the map, which is also called the value function $v(t,x) := \min_{u\in\mathcal{U}} \mathbb{E}[\int_t^T \ell(s,x_s,u(s))\ \mathrm{d}s + \Phi(T,x_T)]$ under the dynamics (2.3) with $x_t \overset{\mathbb{P}-a.s.}{=} x$. We restate the stochastic version of Bellman's [32] principle of optimality for the standard stochastic control problem above.

**Theorem 2.4.1.** *If* $\ell : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ *and* $\Phi : \mathbb{R}^d \to \mathbb{R}$ *are uniformly continuous and there exists* $D > 0$ *such that*

*(D1)* $|\ell(t,x,u) - \ell(t,\hat{x},\hat{u})| \leq D(|x-\hat{x}|$

*(D2)* $|\Phi(t,x) - \Phi(t,\hat{x})| \leq D(|x-\hat{x}|)$

*(D3)* $|\ell(t,0,u)| \leq D$

*(D4)* $|\Phi(T,x)| \leq D$

*for every* $t \in [0,T]$, $x, \hat{x} \in \mathbb{R}^d$ *and* $u, \hat{u} \in U$ *then for every* $(t,x) \in [0,T] \times \mathbb{R}^d$ *and* $0 \leq t \leq t' \leq T$ *the value function satisfies*

$$v(t,x) = \min_{u:[t,t']\in\mathcal{U}} \mathbb{E}\left[\int_t^{t'} \ell(s, x_s, u(s)) + v(t', x_{t'})\right]. \tag{2.3}$$

It is well known that under certain smoothness conditions of the value function, the DPP may be applied to obtain a PDE representation of the value function. This backward-in-time nonlinear PDE is called the HJB equation. We state the standard result on the HJB equation related to control affine systems with state multiplicative Brownian noise and quadratic control cost. We denote $|u|_R^2 = u^{\mathrm{T}} R u$ and assume that $\mathbb{R}^{m\times m} \ni R > 0$.

**Theorem 2.4.2.** *Let* $f, \sigma$ *be uniformly continuous mappings and let assumptions (E3), (E4), (E5) hold. let* $\ell, \Phi$ *be uniformly continuous and let there exist a constant* $D > 0$ *such that assumptions (D1), (D2), (D3), (D4) hold. Further, let* $f(t,x,u) = b(x) + g(x)u$, $\sigma(t,x,u) = \sigma(x)$ *and* $\ell = q(x) + \frac{1}{2}|u^2|_R$. *Let us denote the Hamiltonian by* $\mathcal{H} := \ell + v_x \mathrm{T}(b+gu) + \frac{1}{2}\mathrm{tr}(v_{xx}\sigma\sigma^{\mathrm{T}})$. *If the value function* $v \in C^{1,2}([0,T]\times\mathbb{R}^d)$, $\mathcal{H}(t,x,\cdot) \in C^1(U)$ *and there exists* $u^* \in \mathcal{U}[t,T]$ *such that* $\underset{u\in U}{\mathcal{H}} = \mathcal{H}(t,x,u^*)$ *for every* $(t,x) \in [0,T] \times \mathbb{R}^d$, *then the value function satisfies*

$$-\partial_t v = q(x) + v_x^{\mathrm{T}} f - \frac{1}{2} v_x^{\mathrm{T}} g R^{-1} g^{\mathrm{T}} v_x + \frac{1}{2}\mathrm{tr}(v_{xx}\sigma\sigma^{\mathrm{T}}) \tag{2.4}$$

*for every* $(t,x) \in [0,T] \times \mathbb{R}^d$ *with the terminal time boundary condition* $v(T,x) = \Phi(T,x)$ *and the optimal control is given by*

$$u^*(t,x) = -R^{-1} g^{\mathrm{T}} v_x(t,x). \tag{2.5}$$

The benefit of using a DPP based approach is that we treat the uncertainty in the system explicitly. However, the major disadvantage is issue of scalability in computing the value function, referred to as the *curse of dimensionality* in the literature. In this thesis we will employ sampling based methods which use the Feynman Kac lemma discussed here, to mitigate this problem to a certain extent. The sampling based approach was introduced in [33] which linearizes the HJB equation by using the logarithmic variable transform $\psi(t,x) = \exp(-v)$ along with the ad-hoc constraint $gR^{-1}g^{\mathrm{T}} = \sigma\sigma^{\mathrm{T}}$, which results in the following representation:

$$-\partial_t \psi = -q\psi + \psi_x^{\mathrm{T}} f + \frac{1}{2}\mathrm{tr}(\psi_{xx}\sigma\sigma^{\mathrm{T}}) \tag{2.6}$$

with the boundary condition $\psi(T,x) = \exp(-\Phi(T,x))$. This equation can be computed using the path integral representation:

$$\psi(t,x) = \mathbb{E}\left[\exp(-\int_t^T q(x_s)\mathrm{d}s)\psi(T,x)\right] \tag{2.7}$$

with the expectation corresponding to the uncontrolled dynamics (2.3) with $u(\cdot) = 0$ and $x_t \overset{\mathbb{P}-a.s.}{=} x$. Finally the optimal control is given by

$$u^*(t,x) = -R^{-1}g^{\mathrm{T}}\frac{\psi_x}{\Psi}(t,x). \tag{2.8}$$

This representation allows parallelizable algorithms to approximate the control at a given position by forward sampling from the stochastic passive dynamics over the time horizon. Therefore, although the related sampling based algorithms do not address the problem of the curse of dimensionality directly, they offer a scalable alternative to compute the control for high dimensional systems. A final comment on the constraint $gR^{-1}g^{\mathrm{T}} = \sigma\sigma^{\mathrm{T}}$ is that although ad-hoc, it was shown in [33] that the behavior of systems controlled by such a sampling algorithm exhibit symmetry breaking, owing to this constraint. Additionally, it clearly elucidates the adversarial relationship between control authority and noise which is required to use the path integral representation of the control. However there are recent works [34] which avoid the use of this constraint by using a nonlinear Feynman

Kac representation of the value function through the theory of forward-backward SDEs. This approach was also used to deal with systems with control multiplicative noise [35], in which case the constraint above cannot be applied.

In the following chapters we will develop sampling based algorithms related to two different problems. In chapter 3 we utilize the Feynman Kac lemma to solve an ensemble control problem wherein the underlying subsystems obey jump diffusion dynamics. In chapter 6 we will introduce a key modification of the logarithmic transform above in order to fashion a sampling algorithm which does not require sampling from nonlinear passive dynamics, the idea being that the numerical error resulting from propagating nonlinear dynamics can be mitigated, especially when the model is uncertain.

# CHAPTER 3

# ENSEMBLE CONTROL OF JUMP DIFFUSIONS

In this chapter we discuss the control of an ensemble of stochastic systems which have continuous dynamics driven by Gaussian and non-Gaussian noise and do not interact with each other. Since the number of agents is very large, the ensemble dynamics corresponds to the evolution of the probability density function (PDF). The control problem is framed as control of the forward Chapman-Kolmogorov partial integro differential equation (PIDE) governing the PDF evolution. Necessary conditions corresponding to the infinite dimensional MP for the optimal control of the forward Chapman-Kolmogorov PIDE are derived.

The relationship between infinite dimensional MP and DPP is investigated for this control problem. The relationship between infinite dimensional MP and stochastic dynamic programming (SDP) is also shown. A value function corresponding to PIDE control problem is defined and is shown to obey a DPP. We prove the precise relationship between the value function and optimal costate function satisfying the DPP optimality system and infinite dimensional MP optimality system respectively. A sampling linear Feynman-Kac formula based scheme, applicable to control of such SDEs with control dependent nonlinear drift and noise terms, is derived and demonstrated.

## 3.1 Introduction

In this chapter we discuss the stochastic control of Q-marked Markov jump diffusion (QMJD) or doubly stochastic jump diffusions processes. Such processes represented by stochastic differential equations (SDEs) are used to model ecological population, financial and manufacturing processes [36, 37]. The motivation for this problem is to present a complete exposition of the optimality principles applied to ensemble control of QMJD processes, the relationship between the optimality principles and devise an algorithm for the control problem. Considering our objectives to interpret the optimality principles as well as devise an algorithm, we will take a more pragmatic view in this work and consider QMJD processes for which a smooth probability density function (PDF) exists.

The problem of ensemble control in the case of deterministic systems with initial state uncertainty or diffusion SDEs is in general a parabolic partial differential equation (PDE) control problem [38], [39]. Mathematical formulation of the stochastic optimal control (SOC) problem in these settings is inherently infinite dimensional since the PDF dynamics have partial differential character. For the case of QMJD processes, time evolution of the corresponding PDF is described by a partial integro-differential equation (PIDE). As a result, SOC problem of QMJD processes may be formulated as deterministic control of the corresponding Chapman-Kolmogorov PIDE. To solve this infinite dimensional optimal control problem, we apply the infinite dimensional MP. The infinite dimensional MP has been studied [40], [41] and previously applied [42], [43] in stochastic problems for PDF control of diffusions. Given this prior work on infinite dimensional MP for stochastic control, the work in this chapter has the following main contributions:

- Detailed proofs of the necessary conditions for optimal control of the Chapman-Kolmogorov PIDE dynamics for the case of Q-Marked Jump Diffusions processes using the infinite dimensional MP.
- A generalized Bellman type equation satisfied by the optimal costate in order to show the relationship of the infinite dimensional MP with stochastic Dynamic Programming (SDP) applicable to a single stochastic agent.
- Proof of a Dynamic Programming Principle (DPP) obeyed by the infinite dimensional value function for the PIDE control of QMJD processes.
- Exposition of the relationship between the infinite dimensional MP and DPP for dynamical systems with PIDE dynamics. We do this by showing the explicit relationship between the optimal costate function and infinite dimensional value function.

- An algorithm to compute the infinite dimensional MP optimal control using the SDE representation of ensembles

The fundamental relationship between minimum principle (MP) and DPP in control of deterministic differential dynamics has been a topic of great interest in the control theory literature [31], [44], [45], [46]. In the control of SDE dynamics, the connection between the stochastic minimum principle (SMP) and SDP was studied by Zhou [47] for the diffusion processes. These studies were in the context of Forward Backward SDEs (FBSDEs) theory for quasilinear [48] as well as fully nonlinear backward PDEs [49]. More recently this connection was established in case of jump diffusions and marked jump diffusions [50], [51], [52] and [53]. In all these works the MP and DPP are related through the equality of MP costate function and derivatives of the value function and their corresponding derivatives.

The connection between the infinite dimensional MP and SDP was qualitatively explored recently, in the case of diffusions [54] and jump diffusions [17]. On the other hand the DPP obeyed by the value function corresponding to the infinite dimensional problems has been theoretically explored [55], [56]. However these works are not enlightening on a few aspects of the infinite dimensional MP-DPP relationship. These prior works do not explain clearly how one may compute the infinite dimensional value function defined therein. Understanding this is essential in order to compute the control. Further, they do not explain the precise relationship between the value function satisfying the DPP and the costate function satisfying infinite dimensional MP optimality, when used for the same SOC problem. There is no material on this topic, to the knowledge of the authors, in the area of control of PIDEs corresponding to QMJD processes. In this chapter, we present a generalized connection between infinite dimensional MP and DPP by explaining this relationship. We define the value function for the PIDE control problem considered and prove the DPP satisfied by it. Using the infinite dimensional MP optimality system derived in this chapter, we prove a Bellman type equation for the optimal costate and PDF. This property enables us to explain precisely the relationship between the infinite dimensional value function and infinite dimensional MP optimal costate function.

The optimality system derived in this chapter leads to forward sampling based algorithms for solving nonlinear SOC problems. It is applicable for control of SDEs with control multiplicative Gaussian [57] and marked Poisson noise and non degenerate initial distribution. Thus it expands applicability of sampling algorithms for problems which cannot be solved by the linearly solvable control framework [58, 27, 59] and the forward backward SDEs (FBSDEs) schemes [60, 61, 34] that permit semilinear PDEs. The results presented in this section were published in [17].

This chapter is structured as follows. Section 3.2 contains preliminaries and statement of the SOC problem addressed. In section 3.3 we prove necessary conditions for infinite dimensional MP applied to QMJD processes. Section 3.4 contains complete exposition of infinite dimensional MP-SDP and infinite dimensional MP-DPP relationships. A sampling based PDF control framework for QMJD processes is presented in section 3.5 along with illustrations. Finally we state our conclusions and future research directions in section 3.6.

## 3.2 Problem Formulation

In this section we provide assumptions and conditions related to existence and uniqueness of solutions to a general class of controlled QMJD processes. Two theorems describing the forward and backward Chapman-Kolmogorov PIDEs corresponding to evolution of the PDF representing QMJD processes are stated. The proofs of these theorems are given in the appendices.

### 3.2.1 Definitions

Let $(\Omega, \mathcal{B}, \{\mathcal{B}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $(x_t)_{t\geq 0}$ a process which is progressively measurable with respect to it. We follow [62] and define this process over $\mathbb{R}^{n_x}$

described by

$$
\begin{aligned}
\mathrm{d}x_t =& F(t,x_t,u(t))\mathrm{d}t + B(t,x_t,u(t))\mathrm{d}w_t + H(t,x_t,Q)\mathrm{d}P(t,x_t;t,Q) \\
=& F(t,x_t,u(t))\mathrm{d}t + B(t,x_t,u(t))\mathrm{d}w_t + \int_{D_Q} H(t,x_t,q)\mathcal{P}(t,x_t;\mathrm{d}t,\mathrm{d}q),
\end{aligned} \tag{3.1}
$$

where $x_t \in \mathbb{R}^{n_x}$, $u(t) \in D_u \subseteq \mathbb{R}^{n_u}$, $w_t \in \mathbb{R}^{n_w}$, $Q \in D_Q \subset \mathbb{R}^{n_p}$, $P \in \mathbb{R}^{n_p}$, $F : [0,T] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$, $B : [0,T] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x \times n_w}$, $H : [0,T] \times \mathbb{R}^{n_x} \times D_Q \to \mathbb{R}^{n_x \times n_p}$. The process $w_t$ is the standard Brownian motion. We denote by $H(t,x_t,q) = H(t,x_{t^-},q)$ in the above expressions so that $\Pi(t,x_t) = H(t,x_t,Q)\mathrm{d}P(t,x_t;t,Q) = \int_0^t \int_{D_Q} H(t,x_t,q)\mathcal{P}(t,x_t;\mathrm{d}t,\mathrm{d}q)$ under the zero-one law, is the doubly stochastic Poisson process [63]. Processes $x_t$, $w_t$, $P_t$ and functions on these processes are adapted to the considered filtration such that there exists a unique solution to this SDE given $x_0 = \mathbf{z} \in \mathbb{R}^{n_x}$. The conditions for the existence and uniqueness of the solutions are provided in this subsection. Writing in matrix notation, vectors $\mathcal{P} = [\mathcal{P}_j]$ and $P = [P_j]$ are such that $\{\mathcal{P}_j\}_{1 \le j \le n_p}$ are independent Poisson random measures and $Q$ is the mark vector with marks $\{Q_j\}_{1 \le j \le n_p}$ such that $Q_j \in D_{Q_j} \subset \mathbb{R}$, are independently distributed random variables independent of $P_j$. In this notation realizations of the mark random vector $Q$ in the Poisson random measure formulation are denoted by $q$. The advantage of this notation is that the mark vector has a deterministic representation. We notice that the processes $P_j$ conditioned on $x_t = x$ are Poisson distributed. We now assume that there exists the mark density function $\mathrm{p}_{Q_j}$ corresponding to the mean measure $\nu_j$ of the Poisson random measure $\mathcal{P}_j$ so that $\mathbb{E}[\mathcal{P}_{j\omega}(t,x_t;\mathrm{d}t,\mathrm{d}q_j)|x_t = x] = \nu_j(\mathrm{d}q_j) = \mathrm{p}_{Q_j}(t,q_j;t,x)\lambda_j(t,q_j;t,x)\mathrm{d}q_j\mathrm{d}t$ where $\lambda_j \in \mathbb{R}$ is called the jump rate for the doubly stochastic Poisson process $P_j$. Since $\int_{D_{Q_j}} \mathrm{p}_{Q_j}(t,q_j;t,x)\mathrm{d}q_j = 1$ it is observed, writing in matrix vector notation, that

$$
\begin{aligned}
&\mathbb{E}[\mathrm{d}P(t,Q;t,x_t)|x_t = x] = [\mathbb{E}[\mathrm{d}P_j(t,Q_j;t,x_t)]|x_t = x] = \\
&[\mathbb{E}[\int_{D_{Q_j}} \mathcal{P}_j(\mathrm{d}t,\mathrm{d}q_j;t,x_t)|x_t = x]] = [\mathbb{E}_{\mathrm{p}_{Q_j}}[\lambda_j(t,Q_j;t,x)\mathrm{d}t]].
\end{aligned}
$$

Consequently for the case of mark independent jump rates $\lambda_j$ we would have $\mathbb{E}[\mathrm{d}P(t,x_t;t,Q)|x_t = x] = [\lambda_j(t,x)\mathrm{d}t] = \lambda(t,x)\mathrm{d}t$ which recovers the same result as in the simple Markov jump diffusions process. We denote by $h_j : [0,T] \times \mathbb{R}^{n_x} \times D_{Q_j} \to \mathbb{R}^{n_x}$ the $j^{th}$ column vector of the matrix $H(t,x_t,Q) = [h_{i,j}(t,x_t,Q_j)]$ as well as $\Sigma = BB^{\mathrm{T}}$ and assume $h_{i,j}(t,x_t,Q) = h_{i,j}(t,x_t,Q_j)$. The process $x_t$ is referred to as the state variable and $\mathcal{V}[t,T] \ni u : [t,T] \to \mathbb{R}^{n_u}$ as the control variable where $\mathcal{V}[t,T] := \{u(s) \in D_u | t \leq s\}$ fora ll $t \in [0,T]$ is the class of optimal controls.We comment on this class of controls later in section 3.2.3.

Let us make the following assumptions for the above defined controlled stochastic process:

(S1) there exists a constant $C_1 < \infty$ such that for all $t \in [0,T]$, for all $x \in \mathbb{R}^{n_x}, u \in D_u$

$$||B(t,x,u)||^2 + |F(t,x,u)|^2 + \sum_{k=1}^{n_p} \int_{D_{Q_j}} |h_j(t,x)|^2 \nu_j(\mathrm{d}q_j) \leq C_1(1 + |x|^2 + |u|^2)$$

(S2) there exists a constant $C_2 < \infty$ such that for all $t \in [0,T]$, for all $x, \hat{x} \in \mathbb{R}^{n_x}$ and $u, \hat{u} \in \mathbb{R}^{n_u}$

$$||B(t,x,u) - B(t,\hat{x},\hat{u})||^2 + |F(t,x,u) - F(t,\hat{x},\hat{u})|^2 + \sum_{j=1}^{n_p} \int_{D_{Q_j}} |h_j(t,x,q_j) - h_j(t,\hat{x},q_j)|^2 \nu_j(\mathrm{d}q_j) \leq C_2(|x-\hat{x}|^2 + |u-\hat{u}|^2)$$

(S3) $F_i(t,x,u)$ is once continuously differentiable w.r.t. $x$ for all $i$

(S4) $\Sigma_{ij}(t,x,u)$ is twice continuously differentiable w.r.t. $x$ for all $i,j$

(S5) $h_j : [0,T] \times \mathbb{R}^{n_x} \times D_{Q_j} \to \mathbb{R}^{n_x}$ is a bijection from $\mathbb{R}^{n_x}$ to $\mathbb{R}^{n_x}$, for all $t \in [0,T]$, for all $q_j \in D_{Q_j}$ and $h_j(t,x,q_j) = \eta_j(t,xi_j,q_j)$, $I - \eta_{jxi_j}(t,xi_j,q_j) \neq \mathbf{0}$ for all $(t,xi_j)$ where $xi_j = x + h_j(t,x,q_j)$

(S6) $F_i(t,x,u)$, $\Sigma_{ij}(t,x,u)$ is once continuously differentiable w.r.t. $u \in D_u$ for all $i$

Under the assumptions (S1), (S2), it is well known (pp 10, theorem 1.19) [62] that (3.1) admits a unique cádlág adapted integrable solution $x_t$ for all $\in [0,T]$ given $x_0 = \mathbf{z} \in \mathbb{R}^{n_x}$. The controlled stochastic process represented by the SDE (3.1) is variously called as the marked, compound or doubly stochastic jump diffusion process. Assumptions (S3) through (S5) are typical differentiability assumptions for the existence of the forward and backward Chapman-Kolmogorov operators and corresponding PIDEs.

**Definition 3.2.1.** *Assuming (S5) we define the forward Chapman-Kolmogorov operator that corresponds to QMJD process* (3.1)*, denoted by $\mathcal{F}^{u}_{MJD}(\cdot)$ acting on a function $(\cdot) : [0,T] \times \mathbb{R}^{n_x} \to \mathbb{R}$, given the control $u$, when it exists, as*

$$\begin{aligned}
&\mathcal{F}^{u}_{MJD}(\cdot)(t,x) \\
=&\sum_{i=1}^{n_x} -\frac{\partial}{\partial x_i}((\cdot)(t,x)F_i(t,x,u)) + \frac{1}{2}\sum_{i,j=1}^{n_x}\frac{\partial^2}{\partial x_i \partial x_j}(\Sigma_{ij}(t,x,u)(\cdot)(t,x)) \\
&+\sum_{j=1}^{n_p}\int_{D_{Q_j}}\Big((\cdot)(t,x-\eta_j(t,x,q_j))|I-\eta_{jx}(t,x,q_j)| - (\cdot)(t,x)\Big)\mathrm{p}_{Q_j}\lambda_j(t,x;t,q_j)\mathrm{d}q_j.
\end{aligned} \tag{3.2}$$

**Definition 3.2.2.** *We define the backward Chapman-Kolmogorov operator that corresponds to QMJD process* (3.1)*, denoted by $\mathcal{F}^{\dagger\, u}_{QMJD}(\cdot)$ acting on a function $(\cdot) : [0,T] \times \mathbb{R}^{n_x} \to \mathbb{R}$, given the control $u$, when it exists, as*

$$\mathcal{F}^{\dagger\, u}_{MJD}(\cdot)(t,x) = \sum_{i=1}^{n_x} F_i(t,x,u)\frac{\partial(\cdot)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n_x}[\Sigma(t,x,u)]_{ij}\frac{\partial^2(\cdot)}{\partial x_i \partial x_j} + \mathrm{d}_{jump}(\cdot), \tag{3.3}$$

*where $Jump_j(\cdot)(t,x,q_j) := (\cdot)(t,x+h_{i,j}(t,x,q_j)) - (\cdot)(t,x)$ and*

$$\mathrm{d}_{jump}(\cdot)(t,x) := \sum_{j=1}^{n_p}\int_{D_{Q_j}} Jump_j(\cdot)(t,x,q_j)\mathrm{p}_{Q_j}\lambda_j(t,q_j;t,x)\mathrm{d}q_j.$$

### 3.2.2 Forward and Backward Chapman-Komogorov PIDEs

We do not discuss the existence and uniqueness of solutions to the forward Chapman-Kolmogorov PIDE here. Instead we refer the interested reader to [64] for properties of continuous solutions to the forward Chapman-Kolmogorov PIDEs. Instead we follow the approach of [65] which only derives the forward PIDE that the PDF should satisfy given that the PDF exists and has certain smoothness. Derivations of the forward Chapman-Kolmogorov PIDE is a well explored topic for the case of spontaneous and forced jumps [65], [66], [67]. An explicit derivation in case of simple jump diffusions, called the differential Chapman-Kolmogorov PIDE can be seen in (pp 51, equation 3.4.22) [66].

In this work we follow the approach of Hanson [37]. Here we give complete proof of the second part of the result stated (pp 203-204, Theorem 7.7) [37] in theorem 3.2.1 below. This is the multidimensional version of the more detailed one dimensional result (pp 199-202, Theorem 7.5) [37]. Theorem 3.2.1 states the additional smoothness conditions for the PDF besides (S1), (S2), (S3), (S4), (S5) under which the forward Chapman-Kolmogorov equation is satisfied by the PDF of multidimensional QMJD processes (3.1). Theorem 3.2.2 shows how the backward and forward Chapman-Kolmogorov operators are formal adjoints of each other under certain conditions. Proofs are presented in the appendix 3.7 for completeness. Define $\mathcal{L}^2$ inner product of functions $f_1, f_2$

$$\left\langle f_1, f_2 \right\rangle = \int f_1(x) f_2(x) \mathrm{d}x. \tag{3.4}$$

**Theorem 3.2.1.** *Consider the QMJD process in* (3.1) *such that assumptions (S1), (S2), (S3) and (S4) in section 3.2.1 are true. Let there exist* $\mathrm{p}(t, x|\tau, \mathbf{y}_\tau)$*, the transition probability density of the state* $x_t$ *for all* $t \in [0, T]$ *written in short as* $\mathrm{p}(t, x)$*. If (S5) is true,*

*(T1) there exists* $v : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ *a bounded arbitrary test function which is twice differen-*

*tiable w.r.t. $x$ such that for $u \in D_u$ the conjunct below vanishes*

$$\sum_{i,j=1}^{n_x} \int_{\mathbb{R}^{n_x}} \frac{\partial}{\partial x_i}\Bigg(F_i(t,x,u)\mathrm{p}(t,x)v(x) - \frac{1}{2}\frac{\partial}{\partial x_j}\Big(\Sigma_{ij}(t,x,u)\mathrm{p}(t,x)\Big)v(x) + \frac{1}{2}\Sigma_{ij}(t,x,u)\mathrm{p}(t,x)\frac{\partial}{\partial x_j}v(x)\Bigg)\mathrm{d}x = 0, \tag{3.5}$$

*(T2)* $\mathrm{p}(t,x)$ *is once continuously differentiable w.r.t.* $t$*,* $(\mathrm{p}F_i)(t,x,u)$ *is once continuously differentiable w.r.t.* $x$ *for all* $i$*,* $(\Sigma_{ij}\mathrm{p})(t,x,u)$ *is twice continuously differentiable w.r.t.* $x$ *for all* $i, j$*,* $u \in D_u$*,*

*then* $\mathrm{p}(t,x)$ *satisfies the forward Chapman-Kolmogorov PIDE*

$$\frac{\partial \mathrm{p}(t,x)}{\partial t} = \mathcal{F}^u_{\scriptscriptstyle MJD}\, \mathrm{p}(t,x). \tag{3.6}$$

*in the weak sense. Further, given* $x_{t_0} = x_0$*, the PDF* $\mathrm{p}(t_0,x)$ *satisfies the delta function initial condition*

$$\lim_{t \downarrow t_0} \mathrm{p}(t,x) = \delta(x - x_0). \tag{3.7}$$

**Theorem 3.2.2.** *Consider the QMJD process in (3.1) such that assumptions (S1) through (S5) in subsection 3.2.1 are true. We assume that conditions (T1) and (T2) stated in theorem 3.2.1 are true. If*

*(T3)* $\pi : [0,T] \times \mathbb{R}^{n_x} \to \mathbb{R}$ *is a bounded function which is twice differentiable w.r.t.* $\mathbf{x}$ *and*

*once continuously differentiable w.r.t. $t$ such that $\forall u \in D_u$*

$$\sum_{i,j=1}^{n_x} \int\limits_{\mathbb{R}^{n_x}} \frac{\partial}{\partial x_i}\Bigg( F_i(t,x,u)\mathrm{p}(t,x)\pi(t,x) - \frac{1}{2}\frac{\partial}{\partial x_j}\Big(\Sigma_{ij}(t,x,u)\mathrm{p}(t,x)\Big)\pi(t,x) + \frac{1}{2}\Sigma_{ij}(t,x,u)\mathrm{p}(t,x)\frac{\partial}{\partial x_j}\pi(t,x)\Bigg)\,\mathrm{d}x = 0, \tag{3.8}$$

*(T4) the left and right hand sides of equation* (3.9) *are bounded,*

*then for all $u \in D_u$ $\pi(t,x)$ satisfies the Green's identity [68] for all $u \in D_u$*

$$\Big\langle \pi(t,x), \mathcal{F}^{u}_{\scriptscriptstyle MJD}\mathrm{p}(t,x)\Big\rangle = \Big\langle \mathrm{p}(t,x), \mathcal{F}^{\dagger\, u}_{\scriptscriptstyle MJD}\pi(t,x)\Big\rangle. \tag{3.9}$$

*We then call $\pi$ the adjoint function to the PDF $\mathrm{p}$ so that $\mathcal{F}^{\dagger\, u}_{\scriptscriptstyle MJD}(\cdot)$ is the adjoint operator of $\mathcal{F}^{u}_{\scriptscriptstyle MJD}(\cdot)$.*

### 3.2.3 Problem Statement

Using the inner product definition (3.4), we may define the cost functional, when it exists, as

$$J\left(t;\mathrm{p},u\right) = \Phi(T;\mathrm{p}(T,x)) + \int\limits_{t}^{T} \mathcal{L}(s,u(s);\mathrm{p}(s,x))\,\mathrm{d}t, \tag{3.10}$$

where $\Phi(T;\mathrm{p}(T,x))$ is called the expected terminal cost functional, $\mathcal{L}(t,u(t);\mathrm{p}(t,x))$ is called the expected running cost functional and $t \in [0,T)$. Define

$$\Phi(T;\mathrm{p}(T,x)) = \Big\langle \phi(T,x), \mathrm{p}(T,x)\Big\rangle \text{ and } \mathcal{L}(s,u(s);\mathrm{p}(s,x)) = \Big\langle \ell(s,x,u(s)), \mathrm{p}(s,x)\Big\rangle, \tag{3.11}$$

wherein $\phi : [0,T] \times \mathbb{R}^{n_x} \to \mathbb{R}$ is called the terminal cost function and $\ell : [0,T] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ the running cost function. The choice of these functions as well as the class of admissible control functions $\mathcal{V}[t,T]$ is restricted such that the $\phi$ and $\ell$ are integrable at all times $t \in [0,T]$. The infinite dimensional optimal control problem for the QMJD processes (3.1) is then stated as

$$\min_{u \in \mathcal{V}[t,T]} J\left(t; \mathrm{p}, u\right), \tag{3.12}$$

subject to the dynamics

$$\frac{\partial \mathrm{p}(s,x)}{\partial s} = \mathcal{F}_{\text{MJD}}^{u(s)} \mathrm{p}(s,x), \;\; \mathrm{p}(t,x) = \mathrm{p}_0(x). \tag{3.13}$$

Henceforth the stochastic control problem (3.12) subject to the dynamics (3.13), is referred to simply as problem (3.12). Put in words, the problem undertaken is to find a *deterministic open loop* optimal control for all $s \in [t,T]$ to minimize the cost (3.10) over $[t,T]$ given the PDE dynamics for $\mathrm{p}(s,x)$. The solution is a *broadcast* controller for all the stochastic ensembles governed by (3.1) which solves the problem (3.12).

### 3.3 Infinite Dimensional Minimum Principle for Q-marked Jump Diffusions

This section contains a detailed application of the infinite dimensional MP as applied to the SOC problem 3.12. First a Hamiltonian functional for the infinite dimensional MP is defined. We then derive the Euler-Lagrange equations representing the necessary optimality conditions for the formulated problem. We conclude by defining an infinite dimensional optimal control for PIDE control of QMJD processes.

**Definition 3.3.1.** *We define the Hamiltonian functional for the infinite dimensional MP given the control $u$, when it exists, by*

$$\mathcal{H}\big(s, u; \mathrm{p}(s,x), \pi(s,x)\big) = \Big\langle \ell(s,x,u), \mathrm{p}(s,x) \Big\rangle + \Big\langle \mathcal{F}_{MJD}^{\dagger\, u} \pi(s,x), \mathrm{p}(s,x) \Big\rangle, \tag{3.14}$$

*where* $\pi : [0, T] \times \mathbb{R}^{n_x} \to \mathbb{R}$ *is called the infinite dimensional MP costate function,* $\ell$ *is the running cost function and* $\mathrm{p}$ *is the probability density function representing the MJD process* (3.1)*.*

**Theorem 3.3.1.** ***(Infinite Dimensional Minimum Principle)*** *Consider the Markov Jump Diffusion Process in* (3.1) *such that assumptions (S1), (S2), (S3), (S4), (S5), (S6) in section 3.2.1 are true. We assume that conditions (T1) and (T2) stated in the theorem 3.2.1 hold true. and that the running cost* $\ell$ *is once continuously differentiable w.r.t.* $u$*. Furthermore we assume that there exists a function* $\pi$ *called the infinite dimensional MP costate function such that the Hamiltonian functional for the infinite dimensional MP,* $\mathcal{H}\big(s; \mathrm{u}(s), \mathrm{p}(s, x), \pi(s, x)\big)$*, exists for this choice and is Frechet differentiable w.r.t.* $u$*. If the infinite dimensional MP costate function* $\pi$ *satisfies conditions (T3) and (T4) stated in theorem 3.2.2 then the necessary conditions for optimality on the domain* $[t, T]$ *for the infinite dimensional optimal control problem* (3.12) *subject to the dynamics of the forward Chapman-Kolmogorov PIDE* (3.13)*, are the Euler-Lagrange equations and terminal condition:*

$$\mathcal{H}_u(s, u(s); \mathrm{p}(s, x), \pi(s, x)) = 0 \tag{3.15}$$

$$-\frac{\partial \pi(s, x)}{\partial s} = \ell(s, x, u(s)) + \mathcal{F}^{\dagger\, u(s)}_{MJD} \pi(s, x) \tag{3.16}$$

$$\pi(T, x) = \phi(T, x). \tag{3.17}$$

*Further the Frechet derivative of the Hamiltonian functional* $\mathcal{H}$ *w.r.t.* $u$ *can be specified as*

$$\mathcal{H}_u(s, u; \mathrm{p}(s, x), \pi(s, x)) = \\ \left\langle \ell_u(s, x, u) + \sum_{i=1}^{n_x} \frac{\partial}{\partial u} F_i(s, x, u) \frac{\partial \pi(s, x)}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n_x} \frac{\partial}{\partial u} \left( \frac{\partial^2}{\partial x_i x_j} \Big( \Sigma(s, x, u)_{ij} \pi(s, x) \Big), \mathrm{p}(s, x) \right) \right\rangle. \tag{3.18}$$

*Proof.* Slight abuse of notation is employed in this proof by neglecting to write function arguments for brevity. In the spirit of applying the infinite dimensional minimum principle we append the dynamics into the cost by introducing the Lagrange multiplier or infinite dimensional MP costate. We write the auxiliary cost functional

$$J^{\star}(t;\mathrm{p},u) = \Big\langle \phi(T,x), \mathrm{p}(T,x) \Big\rangle + \int_t^T \left[ \Big\langle \ell, \mathrm{p} \Big\rangle + \Big\langle \pi, \Big( \mathcal{F}^{u}_{\text{MJD}}\mathrm{p} - \frac{\partial \mathrm{p}}{\partial s} \Big) \Big\rangle \right] \mathrm{d}s$$

$$= \Big\langle \phi(T,x), \mathrm{p}(T,x) \Big\rangle + \int_t^T \left[ \Big\langle \ell, \mathrm{p} \Big\rangle + \Big\langle \mathrm{p}, \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi \Big\rangle - \Big\langle \pi, \frac{\partial \mathrm{p}}{\partial s} \Big\rangle \right] \mathrm{d}s, \tag{3.19}$$

since $\mathcal{F}^{\dagger\, u}_{\text{MJD}}(\cdot)$ satisfies $\langle \pi, \mathcal{F}^{u}_{\text{MJD}}\mathrm{p} \rangle = \langle \mathrm{p}, \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi \rangle$ since necessary conditions for theorem 3.2.2 are assumed to be true here. We then have

$$J^{\star}(t;\mathrm{p},u) = \Big\langle \phi(T,x), \mathrm{p}(T,x) \Big\rangle + \int_t^T \left[ \Big\langle \ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi, \mathrm{p} \Big\rangle - \Big\langle \pi, \frac{\partial \mathrm{p}}{\partial s} \Big\rangle \right] \mathrm{d}s$$

$$= \Big\langle \phi(T,x), \mathrm{p}(T,x) \Big\rangle + \int_t^T \left[ \mathcal{H}(s,u;\mathrm{p},\pi) - \Big\langle \pi, \frac{\partial \mathrm{p}}{\partial s} \Big\rangle \right] \mathrm{d}s, \tag{3.20}$$

from the definition of Hamiltonian for teh infinite dimensional MP (3.14). As the Hamiltonian functional is Frechet differentiable,

$$J^{\star}(t;\mathrm{p}+\delta\mathrm{p},u+\delta u) = \Big\langle \phi(T,x), \mathrm{p}(T,x) + \delta\mathrm{p}(T,x) \Big\rangle$$
$$+ \int_t^T \left[ \Big\langle \ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi + (\ell_u + (\mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi)_u)^{\mathrm{T}}\delta u, \mathrm{p}+\delta\mathrm{p} \Big\rangle - \Big\langle \pi, \frac{\partial}{\partial s}(\mathrm{p}+\delta\mathrm{p}) \Big\rangle \right] \mathrm{d}s, \tag{3.21}$$

on neglecting higher order terms of the variations of $\mathrm{p}$ and $u$ so that

$$J^{\star}(t;\mathrm{p}+\delta\mathrm{p},u+\delta u) = J^{\star}(\mathrm{p},u) + \int_t^T \left[ \Big\langle \ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi, \delta\mathrm{p} \Big\rangle + \Big\langle \Big( \ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi \Big)_u^{\mathrm{T}} \delta u, \mathrm{p} \Big\rangle \right] \mathrm{d}s, \tag{3.22}$$

on neglecting higher order mixed variational terms of $\mathrm{p}$ and $u$. Therefore the first order variation of the auxiliary cost functional $J^\star$ with respect to the functions $\mathrm{p}(s,x)$, $u(s)$ is given by

$$\delta J^\star(t;\mathrm{p},u) = \Big\langle \phi(T,x), \delta\mathrm{p}(T,x)\Big\rangle + \int_t^T \Bigg[ \Big\langle \ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi, \delta\mathrm{p}\Big\rangle + \Big\langle \Big(\ell_u + (\mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi)_u\Big)^{\mathrm{T}} \delta u, \mathrm{p}\Big\rangle - \Big\langle \pi, \frac{\partial}{\partial s}(\delta\mathrm{p})\Big\rangle\Bigg] \,\mathrm{d}t. \tag{3.23}$$

Equations (3.14), (3.19) under assumption of Frechet differentiability of the Hamiltonian functional imply

$$\Big\langle \Big(\ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi\Big)^{\mathrm{T}}_u \delta u, \mathrm{p}\Big\rangle = \Big\langle \Big(\ell + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\Big)^{\mathrm{T}}_u, \mathrm{p}\Big\rangle \delta u(s) = \mathcal{H}^{\mathrm{T}}_u(s,u;\mathrm{p},\pi)\,\delta u(s). \tag{3.24}$$

The last term inside the integral in the RHS of (3.23) can now be integrated by parts over the time $t$. Changing the order of integration is permitted since conditions of Fubini's theorem [69], namely the relevant continuous differentiability conditions are satisfied. Therefore

$$\int_t^T \int_{\mathbb{R}^{n_x}} \pi(s,x)\frac{\partial \delta\mathrm{p}}{\partial t}(s,x)\,\mathrm{d}x\,\mathrm{d}s = \int_{\mathbb{R}^{n_x}} \int_t^T \pi(s,x)\frac{\partial \delta\mathrm{p}}{\partial s}(s,x)\,\mathrm{d}s\,\mathrm{d}x$$
$$= \Big\langle \pi(T,x), \delta\mathrm{p}(T,x)\Big\rangle - \Big\langle \pi(t,x), \delta\mathrm{p}(t,x)\Big\rangle - \int_t^T \Big\langle \frac{\partial \pi(s,x)}{\partial s}, \delta\mathrm{p}(s,x)\Big\rangle\,\mathrm{d}s, \tag{3.25}$$

where we note that $\delta\mathrm{p}(t,x)$, $\delta\mathrm{p}(T,x)$ are the values of the first variation functions of $\mathrm{p}(s,x)$ at the fixed time boundaries namely $s=t$ and $s=T$. Here $\delta\mathrm{p}(t,x) = \delta\mathrm{p}(s,x)\Big|_{s=t} - \frac{\partial \mathrm{p}}{\partial s}(s,x)\Big|_{s=t}\delta(t)$ and $\delta\mathrm{p}(T,x) = \delta\mathrm{p}(s,x)\Big|_{s=T} - \frac{\partial \mathrm{p}}{\partial s}(s,x)\Big|_{s=T}\delta(T)$, where $\delta\mathrm{p}(s,x)|_s$ denotes the variation of $\mathrm{p}(s,x)$ at time time $s$. We also know that the variation of $\mathrm{p}$ at $s=t$ is zero or $\delta\mathrm{p}|_{s=t}=0$ since the distribution $\mathrm{p}$ at the boundary $s=t$ or at the initial instant of time is known to be $\mathrm{p}_0(x)$ and that $\delta T = 0$ since we assume a fixed time boundary. Therefore equations (3.23), (3.24), (3.25) imply

$$\delta J^{\star}(t;\mathrm{p},u) = \left\langle (\phi(s,x)\big|_T - \pi(s,x)\big|_T), \delta\mathrm{p}(s,x)|_T \right\rangle$$
$$+ \int_t^T \left[ \left\langle \ell(s,x,u) + \mathcal{F}_{\text{MJD}}^{\dagger\, u(s)}\pi(s,x) + \frac{\partial}{\partial s}\pi(s,x), \delta\mathrm{p}(s,x) \right\rangle \right] \mathrm{d}s$$
$$+ \int_t^T \mathcal{H}_u^{\mathrm{T}}(s,u(s);\mathrm{p}(s,x),\pi(s,x))\delta u(s)\,\mathrm{d}s, \tag{3.26}$$

because $\delta\mathrm{p}(t,x) = 0$ as explained earlier in comments on equation (3.25). The variations $\delta\mathrm{p}(s,x)|_T$, $\delta u(s)$ and $\delta\mathrm{p}(s,x)$ which appear in the above equation are arbitrary and are non zero, so that the three terms above are independent of each other. Therefore, by using the Fundamental Lemma of the Calculus of Variations, with the usual mild conditions [70] and equation (3.26), we have that $\delta J^{\star}(t;\mathrm{p},u) = 0$ implies equations (3.15) (3.16), (3.17). These are the conditions for minimizing $J\big(t;\mathrm{p},u\big)$ using $u$ subject to the governing dynamics of the Kolmogorov Feller PDE. Explicit formulation of $\mathcal{H}_u$ can be obtained due to partial differentiatiability conditions w.r.t. $u$ as given by (3.18). recalling that $\mathrm{d}_{\text{jump}}\pi(x,t)$ does not have explicit dependence on $u$. □

**Definition 3.3.2.** *Consider the necessary conditions for the solution of the optimal control problem* (3.12)*. If there exists an admissible control* $u^*(t)$ *and a corresponding infinite dimensional MP costate function* $\pi^*(t,x)$*, such that the Euler-Lagrange equations* (3.15), (3.16), (3.17) *are satisfied at time* $t$*, then they are called an infinite dimensional optimal control policy and the corresponding optimal infinite dimensional MP costate function at time* $t$*. The PDF for* $\mathrm{p}^*(t,x)$ *denotes the the corresponding optimal PDF satisfying the forward Chapman-Kolmogorov PIDE* (3.6) *under the optimal control.*

## 3.4 Relationship with Dynamic Programming Principle

### 3.4.1 Linear Feynman-Kac lemma and the SDP connection

Under certain conditions, Dynkin's formula [37] formula for the backward costate PIDE (3.16) governing the optimal costate function gives

$$\pi^*(t,x) = \mathbb{E}\Big[\phi(T,x_T) + \int_t^T \ell(s,x_s,u^*(s))\mathrm{d}t|x_t = x\Big]. \tag{3.27}$$

Applying the iterated expectations property of conditional expectations (details in Section 3.5) we have

$$\pi^*(t,x) = \mathbb{E}\big[\ell(t,x,u_t)\mathrm{d}t + \pi^*(t+\mathrm{d}t, x_{t+\mathrm{d}t})\big|x_t = x\big]. \tag{3.28}$$

This expression of the optimal costate indicates a relationship of this function with the SDP principle [31]. We investigate this relationship in the next subsection. It is well known that SDP can be applied to the stochastic problem (3.12) only when the initial state is known with probability one. It is natural to then investigate, the relationship between a version of the DPP for non degenerate initial distributions and the infinite dimensional MP optimality system. This is the topic of subsection 3.4.3. For brevity we abuse our notation by curtailing the expression of dependent variables whenever necessary in this section.

### 3.4.2 Connection between infinite dimensional MP and SDP

*HJB theory for control of Q-marked Jump Diffusion SDEs*

We recall the Hailton Jacobi Bellman (HJB) control theory for QMJD processes here. The optimality system stated will be referred to in the next subsection to see its similarity with a form of the costate PIDE. Consider the QMJD process in (3.1). The value function is defined as the optimal cost to go at any point of time. We assume that the running and terminal cost functions are chosen such that they lead to a cost which is integrable at all instants of time. The value function may then be written as

$$v(t,x) = \min_u \mathbb{E}\left[\phi(T,x_T) + \int_t^T \ell(t,x_t,u_t)\,\mathrm{d}t \middle| x_t = x\right]. \tag{3.29}$$

**Definition 3.4.1.** *We define the HJB Hamiltonian operator for the HJB PIDE corresponding to the QMJD process* (3.1)*, when it exists, by*

$$\mathcal{H}^{\mathit{HJB}}\big(t,x,u,v(t,x)\big) := \ell(t,x,u) + v_x^{\mathrm{T}}(t,x)F(t,x,u) + \frac{1}{2}\mathrm{tr}(\Sigma v_{xx})(t,x,u) + \mathrm{d}_{\mathit{jump}}v(t,x) \tag{3.30}$$

*where* $v:[0,T]\times\mathbb{R}^{n_x}\to\mathbb{R}$ *is the value function and* $\ell$ *is the running cost function.*

Theorem 6.3 (pp 177) in [37] states that if (A1) the decomposition rules (pp 172, Rules 6.1) [37] hold, (A2) there exists a value function $v$ such that $v \in C^{1,2}([0,T]\times\mathbb{R}^{n_x})$, (A3) there exists an optimal control, called the HJB optimal control, given by

$$u^*_{\text{HJB}}(t,x) = \operatorname*{argmin}_{u\in D_u} \mathcal{H}^{\text{HJB}}(t,x,u,v) \text{ for all } (t,x)\in[0,T]\times\mathbb{R}^{n_x}, \tag{3.31}$$

then $v$ satisfies HJB equation for QMJD processes for all $(t,x)\in[0,T]\times\mathbb{R}^{n_x}$

$$-v_t(t,x) = \min_{u\in D_u} \mathcal{H}^{\text{HJB}}(t,x,u,v(t,x)) \tag{3.32}$$

$$v(T,x) = \phi(T,x). \tag{3.33}$$

*Generalized Optimal Costate equation and Relationship with HJB PDE*

Under the optimal control, Euler-Lagrange equations (3.15), (3.16), can be stated in a concise form (3.34), (3.35). This result provides the time rate of change of the optimal costate function integrated over the optimal PDF trajectory, assuming that the infinite dimensional MP optimality conditions are satisfied. We call equation (3.34) with terminal condition (3.35) as the generalized optimal costate equation. A remark at the end of this section describes the similarities and differences in the governing optimality systems obtained using the infinite dimensional MP and SDP under an additional assumption on problem (3.12).

**Lemma 3.4.1.** *Let* $u^* \in \mathcal{V}[t,T]$, $\pi^*, \mathrm{p}^* \in C_c^{1,2}([t,T] \times \mathbb{R}^{n_x})$ *be the infinite dimensional optimal control, optimal costate function and the corresponding optimal PDF for the problem* (3.12). *If*

*(L1) there exist unique* $u^* \in \mathcal{V}[t,T]$, $\pi^* \in C_c^{1,2}([t,T] \times \mathbb{R}^{n_x})$ *satisfying the Euler-Lagrange equations* (3.15), (3.16), (3.17) *for all* $s \in [t,T] \times \mathbb{R}^{n_x}$ *and* $\mathrm{p}^* \in C_c^{1,2}([t,T] \times \mathbb{R}^{n_x})$ *satisfies* (3.6) *under* $u^* : [t,T]$

*(L2)* $\mathcal{H}(s,u;\mathrm{p},\pi^*)$ *is convex and continuously differentiable w.r.t.* $u$ *on* $[t,T] \times D_u$

*then for all* $s \in [t,T]$

$$-\left\langle \frac{\partial \pi^*}{\partial s}(s), \mathrm{p}^*(s) \right\rangle = \min_{u \in D_u} \mathcal{H}(s,u;\mathrm{p}^*,\pi^*) = \min_{u \in D_u} \big\langle \ell(s,u) + \mathcal{F}_{MJD}^{\dagger\, u}\, \pi^*(s), \mathrm{p}^*(s) \big\rangle. \quad (3.34)$$

$$\text{and } \big\langle \pi^*(T), \mathrm{p}^*(T) \big\rangle = \big\langle \phi(T), \mathrm{p}^*(T) \big\rangle. \quad (3.35)$$

*Proof.* Let $u^* \in D_u$ such that $\mathcal{H}_u(s,u^*;\mathrm{p},\pi^*) = \mathbf{0}$ where $s \in [0,T]$. Conditions (L1) and (L2) then imply that

$$u^*(s) = \min_{u \in D_u} \mathcal{H}(s,u;\mathrm{p}^*,\pi^*), \quad (3.36)$$

so that we may write

$$\begin{aligned} \mathcal{H}(s,u^*;\mathrm{p}^*,\pi^*) &= \min_{u \in D_u} \mathcal{H}(s,u;\mathrm{p}^*,\pi^*) \\ &= \min_{u \in D_u} \big\langle \ell(s,u) + \mathcal{F}_{\text{MJD}}^{\dagger\, u}\, \pi^*(s), \mathrm{p}(s) \big\rangle = \big\langle \ell(s,u^*) + \mathcal{F}_{\text{MJD}}^{\dagger\, u^*}\, \pi^*(s), \mathrm{p}^*(s) \big\rangle. \end{aligned} \quad (3.37)$$

Therefore the Euler-Lagrange equations (3.16) and (3.37) give us

$$
\begin{aligned}
&-\left\langle \frac{\partial \pi^*}{\partial s}(s), \mathrm{p}^*(s) \right\rangle = \left\langle \ell(s,u^*) + \mathcal{F}^{\dagger\, u^*}_{\text{MJD}}\ \pi^*(s), \mathrm{p}^*(s) \right\rangle \\
&= \min_{u \in D_u} \mathcal{H}(s,u;\mathrm{p}^*,\pi^*) = \min_{u \in D_u} \left\langle \ell(s,u) + \mathcal{F}^{\dagger\, u}_{\text{MJD}}\ \pi^*(s), \mathrm{p}^*(s) \right\rangle. \qquad (3.38)
\end{aligned}
$$

The terminal condition (3.35) is true because the optimal costate satisfies the terminal condition (3.17). □

***Remark:*** *It is well known that application of SDP for the SOC problem* (3.12)*, is restricted to the case when the initial distribution is specified* $\mathrm{p}(t,x) = \delta(x - \mathbf{y})$ *where* $\mathbf{y} \in \mathbb{R}^{n_x}$*. Applying this condition for the generalized optimal costate equation would derive*

$$
-\frac{\partial \pi^*}{\partial s}(t,\mathbf{y}) = \min_{u \in D_u} \mathcal{H}(t,u;\delta(x-\mathbf{y}),\pi^*(t,x)). \qquad (3.39)
$$

*From the definitions of the HJB Hamiltonian operator* (3.30) *and Hamiltonian functional* (3.14)*, it is observed that if* $\mathrm{p}(t,x) = \delta(x-\mathbf{y})$ *then they are related for all* $u, v$ *at the initial time instant by*

$$
\mathcal{H}(t,u;\delta(t,x-\mathbf{y}),v) = \mathcal{H}^{\scriptscriptstyle HJB}(t,\mathbf{y},u,v(t,\mathbf{y})). \qquad (3.40)
$$

*Equations* (3.40)*,* (3.32) *then imply*

$$
-\frac{\partial v}{\partial s}(t,\mathbf{y}) = \min_{u \in D_u} \mathcal{H}(t,u;\mathrm{p}(t,x),v). \qquad (3.41)
$$

*We can therefore see that the optimal costate PIDE and HJB PIDE are idential at the initial time instant given* $\mathrm{p}(t,x) = \delta(x-\mathbf{y})$*. However, in general, the equations satisfied by the optimal costate* (3.34) *and value function* (3.32) *are distinct at all other time instants since Equation* (3.40) *is true only at the initial time instant. Observe, additionally, that the terminal conditions* (3.35)*,* (3.33) *are not equal because,* $\mathrm{p}(T)$ *is in general, not degenerate. The optimal control generated by the infinite dimensional MP,*

$u^*(s) = \text{argmin}_{u\in D_u}\mathcal{H}(s,u;\mathrm{p}^*,\pi^*)$, *is therefore distinct from the one generated by SDP,* $u^*_{HJB}(s,\mathbf{y}) = \text{argmin}_{u\in D_u}\ \mathcal{H}^{pseudo}(s,\mathbf{y},u,v(t,\mathbf{y}))$, *at all time instants including the initial instant. In addition, we recall that the infinite dimensional MP control is an open loop control which depends only implicitly on the PDF at any time instant, while the SDP control is explicitly a closed loop control.*

### 3.4.3 DPP for the Infinite Dimensional Value function and Relationship with infinite dimensional MP

Based on Lemma 3.4.1 we now quantitatively establish the relationship between the infinite dimensional MP and a DPP satisfied by the infinite dimensional value function. First we construct the DPP satisfied by such an infinite dimensional version of the value function for PIDE control of QMJD processes. Then we show how the infinite dimensional value function is related to the optimal costate function along the optimal PDF trajectory. A precise expression of this relationship has not been stated clear way in preexisting works, as far as known to the authors. Assuming sufficient smoothness of the costate function and PDF and that infinite dimensional MP optimality conditions are satisfied, the relationship between infinite dimensional MP and DPP is stated explicitly in what follows.

**Definition 3.4.2.** *We define the infinite dimensional value function at* $t \in [0,T)$ *for the problem* (3.12) *if it exists and is finite, as*

$$V(t;\mathrm{p}(t)) = \min_{u\in\mathcal{V}[t,T]} J(t;\mathrm{p},u) \tag{3.42}$$

$$\text{and } V(T;\mathrm{p}(T)) = \big\langle \phi(T),\mathrm{p}(T)\big\rangle. \tag{3.43}$$

**Theorem 3.4.2.** *Let there exist a unique finite valued value function defined in* (3.42) *for* *the problem* (3.12) *under the dynamics* (3.13). *Then for all* $s \geq t$

$$V(t;\mathrm{p}(t)) = \min_{u\in\mathcal{V}[t,T]}\left\{ \int_t^s \big\langle \ell(\tau,u),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p}(s))\right\}. \tag{3.44}$$

*Proof.* We have for all $\epsilon > 0$ there exists $u_\epsilon \in \mathcal{V}[t,T]$ such that

$$\begin{aligned} V(t;\mathrm{p}(t)) + \epsilon \geq J(t;\mathrm{p},u_\epsilon) &= \int_t^s \big\langle \ell(\tau,u_\epsilon),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + J(s;\mathrm{p},u_\epsilon) \\ &\geq \int_t^s \big\langle \ell(\tau,u_\epsilon),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p}) \end{aligned} \tag{3.45}$$

so that the following minimum on the right hand side satisfies the inequality

$$V(t;\mathrm{p}(t)) + \epsilon \geq \min_{u\in\mathcal{V}[t,T]}\left\{ \int_t^s \big\langle \ell(\tau,u),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p})\right\}. \tag{3.46}$$

Now given $\epsilon > 0$ and $u_\epsilon \in \mathcal{V}[t,T]$, we can choose $u(\tau) = u_\epsilon(\tau)$ when $\tau \in [t,s]$ such that $V(s;\mathrm{p}(s)) + \epsilon \geq J(s;\mathrm{p},u_\epsilon)$. From Equation (3.42) of Definition 6 it can be seen that for all $u_\epsilon \in \mathcal{V}[t,T]$, $\epsilon > 0$,

$$V(t;\mathrm{p}(t)) + \epsilon \leq J(t;\mathrm{p},u) = \int_t^s \big\langle \ell(\tau,u),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + J(s;u,\mathrm{p}) \leq \int_t^s \big\langle \ell(\tau,u_\epsilon),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p}(s)) + \epsilon. \tag{3.47}$$

Since this inequality is true for all $u_\epsilon \in \mathcal{V}[t,T]$, we may take the minimum over all such $u_\epsilon$

$$\begin{aligned} V(t;\mathrm{p}(t)) \leq & J(t;u,\mathrm{p}) \leq \min_{u_\epsilon\in\mathcal{V}[t,T]}\left\{ \int_t^s \big\langle \ell(\tau,u_\epsilon),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p})\right\} \\ & \leq \min_{u\in\mathcal{V}[t,T]}\left\{ \int_t^s \big\langle \ell(\tau,u),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p})\right\} + \epsilon. \end{aligned} \tag{3.48}$$

where we have replaced the symbol $u_\epsilon$ with $u$ in the last equality. Combining equations (3.46), (3.48) implies

$$V(t;\mathrm{p}(t)) - \epsilon \le \min_{u\in\mathcal{V}[t,T]}\left\{ \int_t^s \big\langle \ell(\tau,u),\mathrm{p}(\tau)\big\rangle \mathrm{d}\tau + V(s;\mathrm{p})\right\} \le V(t;\mathrm{p}(t)) + \epsilon. \quad (3.49)$$

for all $\epsilon > 0$. Under the limit $\epsilon \to 0$ the desired result is obtained. □

The following theorem quantitatively states the relationship between the infinite dimensional value function and the optimal costate function under the assumption that problem (3.12) has a unique solution.

**Theorem 3.4.3.** *Let $u^* \in \mathcal{V}[t,T]$, $\pi^*, \mathrm{p}^* \in C_c^{1,2}([t,T]\times\mathbb{R}^{n_x})$ be the unique infinite dimensional optimal control, optimal costate function and corresponding PDF for the problem* (3.12). *If (L1) and (L2) are true and*

*(T5) there exists a unique finite valued value function defined in* (3.42) *for the problem* (3.12)

*then for all $s \in [t,T]$*

$$V(s;\mathrm{p}^*(s)) = \big\langle \pi^*(s),\mathrm{p}^*(s)\big\rangle. \quad (3.50)$$

*Proof.* Dynkin's formula for the jump diffusion process (3.1) can be applied to $\pi^*(s,x) \in C_c^{1,2}([t,T]\times\mathbb{R}^{n_x})$ which satisfies (3.16), (3.17) by following Theorem 7.1, chapter 7 in [37] to obtain

$$\pi^*(s,x) = \mathbb{E}\left[\int_s^T \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \phi(T,x_T)\Big| x_s = x\right] \quad (3.51)$$

where the expectations are under the optimal PDF governed by (3.13) under the control $u^*(\tau) \in \mathcal{V}[s,T]$ is $\mathrm{p}^*(\tau) : [s,T]$. The law of iterated expectations implies for all $t \le s \le$

$\hat{s} \leq T$

$$\begin{aligned}\pi^*(s,x) &= \mathbb{E}\bigg[\,\mathbb{E}\bigg[\int\limits_s^{\hat{s}} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \int\limits_{\hat{s}}^{T} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \phi(T,x_T)\bigg| x_{\hat{s}}\bigg]\,\bigg| x_s = x\bigg]\\ &= \mathbb{E}\bigg[\int\limits_s^{\hat{s}} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \mathbb{E}\bigg[\int\limits_{\hat{s}}^{T} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \phi(T,x_T)\bigg| x_{\hat{s}}\bigg]\,\bigg| x_s = x\bigg]\\ &= \mathbb{E}\bigg[\int\limits_s^{\hat{s}} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \pi^*(\hat{s},x_{\hat{s}})\bigg| x_s = x\bigg]. \end{aligned} \tag{3.52}$$

Using the law of total expectation we have

$$\begin{aligned}&\big\langle \pi^*(s), \mathrm{p}^*(s)\big\rangle = \mathbb{E}\left[\int\limits_s^{\hat{s}} \mathbb{E}\big[\ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau + \pi^*(\hat{s},x_{\hat{s}})\big| x_s = x\big]\right]\\ &= \mathbb{E}\bigg[\int\limits_s^{\hat{s}} \ell(\tau,x_\tau,u^*(\tau))\mathrm{d}\tau\bigg] + \mathbb{E}\Big[\pi^*(\hat{s},x_{\hat{s}})\Big] = \int\limits_s^{\hat{s}} \big\langle \ell(\tau,u^*(\tau)), \mathrm{p}(\tau)\big)\big\rangle\mathrm{d}\tau + \big\langle \pi^*(\hat{s}), \mathrm{p}(\hat{s})\big\rangle. \end{aligned} \tag{3.53}$$

Due to the fact that the optimal control $u^* \in \mathcal{V}[t,T]$ solving the problem (3.12) is unique and (T5), we can write the following result from Theorem 3.4.2. Note again that we denote the optimal PDF $\mathrm{p}^*(\tau) : [s,T]$ evolving under the optimal control $u^* \in \mathcal{V}[s,T]$.

$$\begin{aligned}V(s;\mathrm{p}(s)) &= \min_{u\in\mathcal{V}[t,T]}\left\{\int\limits_s^{\hat{s}} \big\langle \ell(\tau,u), \mathrm{p}(\tau)\big\rangle\mathrm{d}\tau + V(\hat{s};\mathrm{p}(\hat{s}))\right\}\\ &= \int\limits_s^{\hat{s}} \big\langle \ell(\tau,u^*), \mathrm{p}^*(\tau)\big\rangle\mathrm{d}\tau + V(\hat{s};\mathrm{p}^*(\hat{s})) \end{aligned} \tag{3.54}$$

Recall the terminal conditions for the value function (3.43) and for optimal costate function (3.17) implies for all $\mathrm{p}(T)$ satisfying the conditions (T1) through (T4)

$$V(T;\mathrm{p}(T)) = \big\langle \pi^*(T), \mathrm{p}(T)\big\rangle = \big\langle \phi(T), \mathrm{p}(T)\big\rangle \tag{3.55}$$

which is true for $\mathrm{p}^*(T)$ as well. Let us choose $\hat{s} = T$ in equation (3.48), (3.53). Observing that equations (3.53), (3.54) are identical and the value function is unique due to (T5), we can prove easily using equation (3.55) that $V(s; \mathrm{p}^*(s)) = \langle \pi^*(s), \mathrm{p}^*(s) \rangle$ along the optimal PDF trajectory $\mathrm{p}^*(s) : [s, T]$ under the control $u^* \in \mathcal{V}[s, T]$. □

Stated in words, we have proved that the infinite dimensional value function is equal to the $\mathcal{L}^2$ product of optimal costate function with the optimal PDF along the optimal trajectory. Note that this relationship was proved using the mechanism of the linear Feynman-Kac lemma, which motivated our investigation.

### 3.5 Sampling based algorithm for PDF Control of QMJD processes

Using the Feynman-Kac formula (3.27) directly, to compute the costate or optimal costate by forward sampling would be computationally prohibitive. Direct application would require generating samples over the entire time horizon starting from each space time grid point. Instead, we use an iteratively backpropagated costate (IBC) algorithm [17], the key ingredient for which is derived below. By Dynkin's formula, Theorem 7.1, chapter 7 of [37] for QMJD process (3.1), applied to $\pi(t, x) \in C_c^{1,2}([0, T] \times \mathbb{R}^{n_x})$ which satisfies (3.16), (3.17) under arbitrary control $u(s) \in \mathcal{V}[t, T]$

$$\pi(t, x) = \mathbb{E}\left[ \int_t^T \ell(s, x_s, u(s))\mathrm{d}s + \phi(T, x_T) \Big| x_t = x \right]. \tag{3.56}$$

The law of iterated expectations implies

$$\begin{aligned}
\pi(t, x) &= \mathbb{E}\Bigg[ \mathbb{E}\Bigg[ \ell(t, x_t, u(t))\mathrm{d}t + \int_{t+\mathrm{d}t}^T \ell(s, x_s, u(s))\mathrm{d}s + \phi(T, x_T) \Bigg| x_{t+\mathrm{d}t} \Bigg] \Bigg| x_t = x \Bigg] \\
&= \mathbb{E}\Big[ \ell(t, x_t, u(t)) \Big]\mathrm{d}t + \mathbb{E}\Bigg[ \int_{t+\mathrm{d}t}^T \ell(s, x_s, u(s))\mathrm{d}s + \phi(T, x_T) \Bigg| x_{t+\mathrm{d}t} \Bigg] \Bigg| x_t = x \Bigg] \\
&= \mathbb{E}\Big[ \ell(t, x_t, u(t))\mathrm{d}t + \pi^*(t + \mathrm{d}t, x_{t+\mathrm{d}t}) \Big| x_t = x \Big]. 
\end{aligned} \tag{3.57}$$

We denote the temporal grid indexed as $[t_0, t_N] = [0, T]$. We have dropped the conditional expectation notation for brevity in the following pseudo code and pick a small number $\epsilon > 0$.

**Algorithm 2** IBC PDF control of MJD processes

**Initialize** Choose $u_t^0 : [t_0, t_N]$ arbitrarily.
**repeat**
  **Initialize** $\pi^k(t_N, x) = \phi(t_N, x)$.
  **while** $i \neq 0$ **do**
    $\pi^k(t_{i-1}, x) = \ell(t_{i-1}, x, u^k_{t_{i-1}}) + \mathbb{E}\big[\pi^k(t_i, x + \mathrm{d}x_{t_i})\big]$.
    $i = i - 1$.
  **end while**
  Compute $\mathcal{H}^k_u(t; \mathrm{p}^k, u^k_t, \mu^k_Q)$ on $[t_0, t_N]$ by (3.18).
  Update control: $u^{k+1}_t = u^k_t - \epsilon \mathcal{H}^k_u(t)$.
**until** Convergence $|\mathcal{H}_u(t)| < \epsilon$.
**return** $u^*_t : [0, T]$.

**Example Problem:** We demonstrate our algorithm for open loop control of ensembles with dynamics (3.1) with linear drift term, nonlinear diffusion coefficient and constant jump rate parameter

$$\mathrm{d}x_t = (-\alpha x_t + u(t)\mathrm{d}t + \zeta\sqrt{\frac{(\kappa - x_t)^2}{2} + u(t)^2}\mathrm{d}w_t + h(x_t, Q)\mathrm{d}P_t, \tag{3.58}$$

where $h(x_t, Q) = 0.5 \cdot Q \cdot x_t$ with constant jump rate $\lambda = 1$ and mark density of *unif*$([0, 1])$. The initial condition is assumed to be a normal distribution. The state space is specified by the constraints $x(t) \in [-3, 3]$ while the control is constrained by $u(t) \in [-3, 3]$. Further two obstructions are modeled by the state constraints $x(t) \in [-2, -1]$ at $t = 1$ and $x(t) \in [0.5, 2]$ at $t = 2$. The process is defined to terminate on reaching any of the above boundaries. The values of the constants used are $\kappa = 3$, $\alpha = 0.5$, terminal time $T = 3$ and

$\zeta = 0.1$. The task is to reach the target $x_{goal}(T) = 0$.

In this example we choose the running and terminal cost functions as $\ell(u) = \frac{R}{2}u^2$, $\phi(x) = Q_f(x - x_{goal})^2$, and a trajectory termination penalty of $\Xi - \tau$ where $\tau \in [0, T]$ is the stopping time of a trajectory colliding with a boundary or obstacle. Let $R = 2 \times 10^{-4}$, $Q_f = \frac{4}{9}$, $\zeta = 0.1$ and $\Xi = 7$. The temporal discretization $\{t_i\}_{0 \leq i \leq N}$ is chosen to satisfy $\lambda \cdot (t_i - t_{i-1}) << 1$ the zero one law [37] allowing our use of the derived form of the PIDEs. Since we generate an open loop policy, we adopt the strategy of generating an implicity feedback policy. We do this by assuming $u(t) = u_1(t) + xu_2(t)$ and treating $u(t) = [u_1(t)\ u_2(t)]^{\mathrm{T}}$ as the control we compute. This state parameterized policy results in a lower state dependent cost seen in Subfigure (1d). Trajectories are sampled in two steps in our algorithm. We sample single time step trajectories inside the costate computation loops at each spatio temporal grid point. We need full time horizon samples to compute the control gradient of the Hamiltonian at each control update iteration. Let $x_k(t)$ be the $k^{th}$ sample of trajectories at time $t$ for either case and $\mathbf{1}_k = \mathbf{1}_{\{\Xi_k < T\}}$ which indicates whether the sample was terminated by collision. Theorem 3.3.1, Eq. (3.28) imply

$$\begin{aligned}\pi(x,t) = &\frac{1}{K}\sum_{k=1}^{K}\left[\phi(x_k(T)) + \int_t^T \ell(x_k(s), u_s)\mathrm{d}s\right](1 - \mathbf{1}_k)\\ &+ \frac{1}{K}\sum_{k=1}^{K}\left[(\Xi_k - \tau) + \int_t^\tau \ell(x_k(s), u_s)\mathrm{d}s\right]\mathbf{1}_k,\end{aligned}$$

$$\mathcal{H}_u(s, u(s); \mathrm{p}, \pi) = \begin{bmatrix} Ru_1(t) + \pi_x + \xi^2 u_1(t)\pi_{xx} \\ Ru_2(t) + \pi_x + \xi^2 u_2(t)\pi_{xx} \end{bmatrix}. \tag{3.59}$$

**Results:** Optimal costate function, a set of optimally controlled trajectories and the cost per iteration depicting convergence for $\mathrm{p}_Q = \textit{unif}(0, 1)$ and $\mathrm{p}_0 = \mathcal{N}(-1, \frac{1}{2})$ are illustrated in Subfigures (1a), (1b, (1c). Converged costs are compared in Subfigure (1c) with different initial conditions. The cost is lower for initial condition $\mathcal{N}(-1, \frac{1}{2})$ since far lower number of optimally controlled trajectories end up colliding with the first obastacle depicted in the

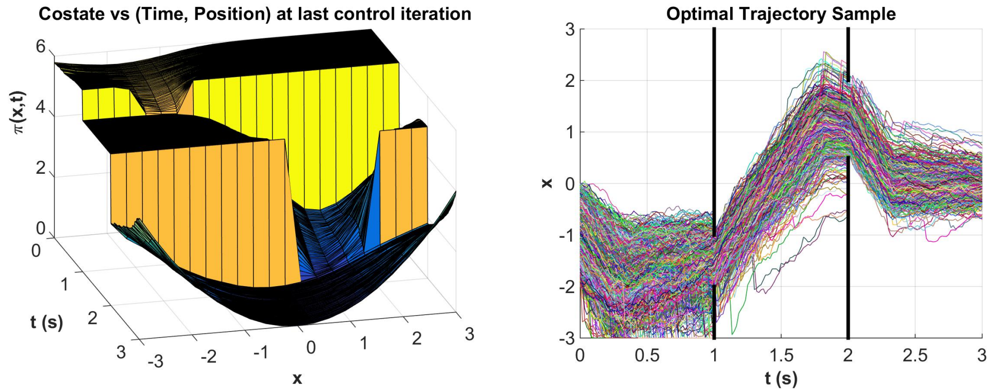



Figure 3.1: Costate at last iteration $\mathrm{p}_0 = \mathcal{N}(-1, \frac{1}{2})$ (left) and optimal trajectory samples for $\mathrm{p}_0 = \delta$

optimal trajectory sample in Subfigure (1b). We compare converged costs for simple jump diffusion $Q \sim \delta(1)$ with initial condition $\mathrm{p}_0 = \delta(0)$, when using the state parameterized policy and non parameterized control in Subfigure (1d). This shows the benefit of the implicit feedback provided by the state parameterized policy.

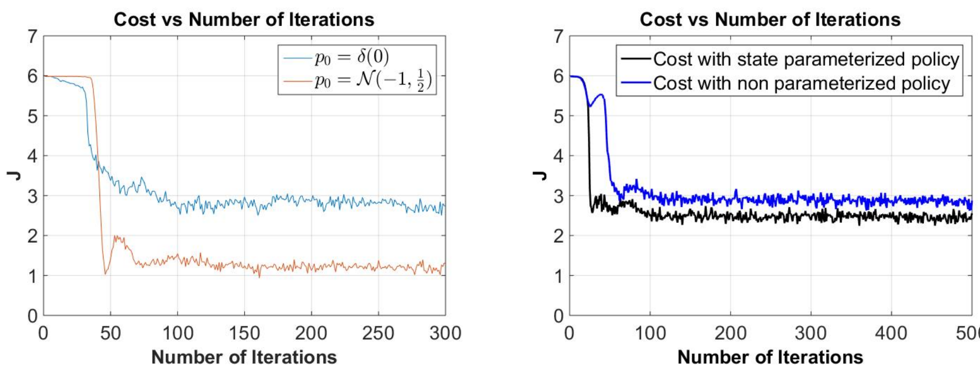



Figure 3.2: Cost vs iterations for $\mathrm{p}_0 = \mathcal{N}(-1, \frac{1}{2})$ in blue, and $\mathrm{p}_0 = \delta(0)$ in red (left) and (right) cost comparison with state parameterized policy in black, for $\mathrm{p}_0 = \delta(0)$, $Q \sim \delta(1)$

### 3.6 Conclusions

We present a complete theory for a PDE based optimal open loop control framework for marked jump diffusions in this chapter. This includes exposition of the fundamental relationship between infinite dimensional MP and DPP applied to this control framework. A sampling based algorithm is developed with good results for a nonlinear control problem with possible modified implicit feedback.

Connections between MP and DPP, including for stochastic systems, have witnessed a lot of interest in the control theory literature. The SMP-SDP connection is expressed by showing equality of the first and second costate processes with the gradient and Hessian of the value function. However the precise relationship between infinite dimensional MP and DPP for PIDE control has not been explained previously. Moreover the value function in the infinite dimensional case is not well understood. To this end we state and prove a DPP obeyed by the infinite dimensional value function. This DPP is different from the SDP, because unlike the former, SDP does not allow ensembles with non degenerate initial distributions. We precisely explain the relationship between infinite dimensional MP and DPP by showing that the $\mathcal{L}^2$ product of the optimal costate and optimal PDF equals the value function under the optimal control. Thus the infinite dimensional MP-DPP relationship shown is clearly distinct from the SMP-SDP connection, although the SOC problem considered is similar in both cases.

Appealing to the symmetry of between finite and infinite dimensional versions of the DPP, the next logical step is to compute explicit feedback control laws for PIDE dynamics as suggested in [38]. Future works on this topic will be focussed on analyzing control laws of form $u(t; \mathrm{p}(t))$ given the feedback density $\mathrm{p}(t, x)$. Potential applications of this theory would be in the field of control of elementary particle ensembles, quantum systems [7], biological systems and swarms.

### 3.7 Appendix

*Proof. of theorem 3.2.1* We permit abuse of notation in this proof by neglecting to write argument dependencies o functions for brevity when necessary in this proof. Further we denote partial derivatives as $\frac{\partial f}{\partial \mathbf{v}} = f_{\mathbf{v}}$ if needed. This proof is a stepwise analogous extension to the multidimensional case of the proof for the one dimensional case (pp 199-203, Theorem 7.5) [37]. It follows easily by differentiating the well known multidimensional Dynkin's formula (pp 203, Equation 7.32) [37] for the function $v$ and using two integration by parts steps to move the spatial derivatives operating on $v$ to $\mathrm{p}$. Differentiation of the

Dynkin's formula for $u(t,x) = \mathbb{E}[v(x_t)|x_{t_0} = x]$ yields

$$\frac{\partial}{\partial t}u(t,x) = \mathbb{E}\left[\frac{\partial}{\partial t}\int\limits_{t_0}^{t} \mathcal{F}^{\dagger\, u}_{\text{MJD}} v(x_s)\mathrm{d}\mathbf{s}|x_{t_0} = x\right] = \int\limits_{\mathbb{R}^{n_x}} \mathcal{F}^{\dagger\, u}_{\text{MJD}} v(x)\mathrm{p}(t,x)\mathrm{d}x \tag{3.60}$$

where $\mathcal{F}^{\dagger\, u}_{\text{MJD}}$ is the backward operator defined in the previous subsection 3.2.1. Note that $u(t,x) = \mathbb{E}[v(x_t)|x_{t_0} = x]$ implies that when using the PDF representation of the expectation we have

$$\frac{\partial}{\partial t}u(t,x) = \int_{\mathbb{R}^{n_x}} v(x)\frac{\partial}{\partial t}\mathrm{p}(t,x)\mathrm{d}x. \tag{3.61}$$

The Dynkin formula (3.60) and the definition of the backward operator imply

$$\frac{\partial}{\partial t}u(t,x) = \int\limits_{\mathbb{R}^{n_x}} \left(\sum_{i=1}^{n_x} F_i v_{x_i} + \frac{1}{2}\sum_{i,j=1}^{n_x} \Sigma_{ij} v_{x_j x_i} + \mathrm{d}_{\text{jump}} v\right)\mathrm{p}\mathrm{d}x. \tag{3.62}$$

Let us at first focus on the diffusion terms without the jump term in the above expression (3.62). We use integration by parts for integration w.r.t. $x_i$ in step one, and $x_j$ in step two for terms from the backward operator in (3.60) to move the spatial derivatives to the PDF using condition (T2). Using this idea along with Fubini's theorem [69] we get

$$\begin{aligned}
&\int\limits_{\mathbb{R}^{n_x}} \sum_{i=1}^{n_x}\left(F_i v_{x_i} + \frac{1}{2}\sum_{j=1}^{n_x}\Sigma_{ij} v_{x_j x_i}\right)\mathrm{p}\mathrm{d}x \\
&= \underbrace{\int\limits_{\mathbb{R}}\dots\int\limits_{\mathbb{R}}}_{(n_x-1)\text{times}} \left[\sum_{i=1}^{n_x}\left\{\int\limits_{\mathbb{R}}\left(-\frac{\partial(F_i\mathrm{p})}{\partial x_i}\,v - \frac{1}{2}\sum_{j=1}^{n_x}\frac{\partial(\Sigma_{ij}\mathrm{p})}{\partial x_i}\,v_{x_j}\right)\mathrm{d}x_i\right.\right. \\
&\qquad\qquad \left.\left. + \int\limits_{\mathbb{R}}\frac{\partial}{\partial x_i}\left(F_i\mathrm{p}v + \frac{1}{2}\sum_{j=1}^{n_x}\Sigma_{ij}\mathrm{p}v_{x_j}\right)\mathrm{d}x_i\right\}\right]\mathrm{d}x_1\dots\mathrm{d}x_{i-1}\mathrm{d}x_{i+1}\dots\mathrm{d}x_{n_x}
\end{aligned}$$

$$
\begin{aligned}
&= \int\limits_{\mathbb{R}^{n_x}} \sum_{i=1}^{n_x} -\frac{\partial (F_i \mathrm{p})}{\partial x_i}\ v \ \mathrm{d}x + \int\limits_{\mathbb{R}^{n_x}} \frac{\partial}{\partial x_i}\bigg(F_i \mathrm{p} v + \frac{1}{2}\sum_{j=1}^{n_x} \Sigma_{ij}\mathrm{p} v_{x_j}\bigg)\mathrm{d}x \\
&\quad + \underbrace{\int\limits_{\mathbb{R}} ... \int\limits_{\mathbb{R}}}_{(n_x-1)\text{times}} \Bigg[\frac{1}{2}\sum_{i,j=1}^{n_x}\int\limits_{\mathbb{R}}\bigg(\frac{\partial^2}{\partial x_i x_j}(\Sigma_{ij}\mathrm{p})v\bigg)\mathrm{d}x_j \\
&\qquad\qquad - \frac{1}{2}\sum_{i,j=1}^{n_x}\int\limits_{\mathbb{R}}\bigg(\frac{\partial}{\partial x_j}\bigg(\frac{\partial(\Sigma_{ij}\mathrm{p})}{\partial x_i} v\bigg)\bigg)\mathrm{d}x_j\Bigg]\mathrm{d}x_1...\mathrm{d}x_{j-1}\mathrm{d}x_{j+1}...\mathrm{d}x_{n_x} \\
&= \sum_{i,j=1}^{n_x}\int\limits_{\mathbb{R}^{n_x}} -\frac{\partial}{\partial x_i}(F_i\mathrm{p})\ v + \frac{1}{2}\bigg(\frac{\partial^2(\Sigma_{ij}\mathrm{p})}{\partial x_i x_j}\ v\bigg)\mathrm{d}x \\
&\quad + \sum_{i,j=1}^{n_x}\int\limits_{\mathbb{R}^{n_x}} \frac{\partial}{\partial x_i}\bigg(F_i\mathrm{p}v + \frac{1}{2}\Sigma_{ij}\mathrm{p}v_{x_j} - \frac{1}{2}\frac{\partial(\Sigma_{ij}\mathrm{p})}{\partial x_j}\ v\bigg)\mathrm{d}x,
\end{aligned} \tag{3.63}
$$

the last part of which can easily be identified as the conjunct in condition (T1). Let us now focus on the jump term in the expression (3.62) which is

$$
\begin{aligned}
&\int\limits_{\mathbb{R}^{n_x}} \mathrm{d}_{\text{jump}} v(x)\mathrm{p}(t,x)\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^{n_x}} \sum_{j=1}^{n_p} \int\limits_{D_{Q_j}} \bigg(\Big[v(x + h_j(t,x,q_j)) - v(x)\Big]\mathrm{p}_{Q_j}\lambda_j(t,q_j;t,x)\mathrm{d}q_j\bigg)\mathrm{p}(t,x)\mathrm{d}x.
\end{aligned} \tag{3.64}
$$

Consider the terms in the first summation on the right hand side of this equation. Change the variable of integration to $\xi_j$ using the Change of Variables theorem [71] where $\xi_j = x + h_j(t,x,q_j) = x + \eta_j(t,\xi_j,q_j)$ in the first step, wherein $h_j$ is assumed invertible w.r.t. $\xi_j$. This inverse mapping exists since we assumed $h_j$ is a bijection from $\mathbb{R}^{n_x}$ to $\mathbb{R}^{n_x}$. We note that transformed domain of integration is again $\mathbb{R}^{n_x}$. In the second step we change the dummy variable of integration back to $x$. We therefore have for all $j,\ 1 \leq j \leq n_p$ that

$$\begin{aligned}
&\int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \Big( v(x + h_j(t,x,q_j))(p_{Q_j}\lambda_j)(t,q_j;t,x)\mathrm{d}q_j \Big) \mathrm{p}(t,x)\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \Big( v(\xi_j)(p_{Q_j}\lambda_j)(t,q_j;t,\xi_j - \eta_j(t,\xi_j,q_j))\mathrm{p}(t,\xi_j - \eta_j(t,\xi_j,q_j)) \Big) \mathrm{d}q_j \\
&\qquad\qquad \cdot |I - \eta_{j\xi_j}(t,\xi_j,q_j)| \mathrm{d}\xi_j \\
&= \int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \Big( v(x)\mathrm{p}(t,x - \eta_j(t,x,q_j))(p_{Q_j}\lambda_j)(t,q_j;t,x - \eta_j(t,x,q_j)) \Big) \mathrm{d}q_j \\
&\qquad\qquad \cdot |I - \eta_{jx}(t,x,q_j)| \mathrm{d}x. \qquad (3.65)
\end{aligned}$$

Equations (3.62), (3.63), (3.64), (3.65) and (T1) then imply

$$\begin{aligned}
\frac{\partial}{\partial t} u(t,x) &= \int\limits_{\mathbb{R}^{n_x}} \big( \sum_{i=1}^{n_x} F_i v_{x_i} + \sum_{i,j=1}^{n_x} \Sigma_{ij} v_{x_j x_i} + \mathrm{d}_{\text{jump}} v \big) \mathrm{p}\ \mathrm{d}x \\
&= \sum_{i,j=1}^{n_x} \int\limits_{\mathbb{R}^{n_x}} \Big[ -\frac{\partial}{\partial x_i}(F_i \mathrm{p})v + \frac{1}{2}\frac{\partial^2}{\partial x_i x_j}(\Sigma_{ij}\mathrm{p})v \Big] \mathrm{d}x \\
&+ \sum_{j=1}^{n_p} \int\limits_{\mathbb{R}^{n_x}} \Big[ \int\limits_{D_{Q_j}} \Big( \mathrm{p}(t, x - \eta_j)|I - \eta_{jx}| - \mathrm{p}(t,x) \Big) \\
&\qquad \cdot (p_{Q_j}\lambda_j)(t,q_j;t,x - \eta_j(t,x,q_j)) \Big) \mathrm{d}q_j \Big] v(x)\mathrm{d}x. \qquad (3.66)
\end{aligned}$$

Equations (3.61), (3.66) and definition of forward Chapman-Kolmogorov operator in subsection 3.2.1 imply

$$\begin{aligned}
&\frac{\partial}{\partial t} u(t,x) = \int\limits_{\mathbb{R}^{n_x}} \Big[ \sum_{i,j=1}^{n_x} \Big( -\frac{\partial}{\partial x_i}(F_i \mathrm{p}) + \frac{1}{2}\frac{\partial^2}{\partial x_i x_j}(\Sigma_{ij}\mathrm{p}) \\
&+ \sum_{j=1}^{n_p} \int\limits_{D_{Q_j}} \Big( \mathrm{p}(x - \eta_j(t,x))|I - \eta_{jx}(t,x)| - \mathrm{p}(t,x) \Big)(p_{Q_j}\lambda_j)(t,q_j;t,x - \eta_j(t,x,q_j))\mathrm{d}q_j \Big] v(x)\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^{n_x}} \mathcal{F}^u_{\text{MJD}}\ \mathrm{p}(t,x)\ v(x)\mathrm{d}x = \int_{\mathbb{R}^{n_x}} \frac{\partial}{\partial t}\mathrm{p}(t,x)v(x)\mathrm{d}x. \qquad (3.67)
\end{aligned}$$

Using the calculus of variations argument (pp 201, proof of Theorem 7.5) [37], [72] since the function $v$ is any arbitrary function with assumed boundedness and smoothness properties, $\mathrm{p}(t,x)$ satisfies equation (3.6) in the weak sense. Notice that the Dynkin formula (3.60) and equation (3.67) imply

$$\frac{\partial}{\partial t}u(t,x) = \Big\langle \mathcal{F}^{\dagger\, u}_{\text{MJD}} v(x), \mathrm{p}(t,x) \Big\rangle = \Big\langle \mathcal{F}^{u}_{\text{MJD}}\, \mathrm{p}(t,x), v(x) \Big\rangle \tag{3.68}$$

which means that the backward operator $\mathcal{F}^{\dagger\, u}_{\text{MJD}}$ is the formal adjoint operator of the forward operator $\mathcal{F}^{u}_{\text{MJD}}$.

The delta initial condition to be proved is well known in the case that the jump term is absent, that is for the diffusion processes. However Poisson process $P_t$ undergoes jumps which causes $x_t$ to have discontinuous paths. However, considering that simple jump process has Poisson distribution, we see that a jump is unlikely in a small time interval $\mathrm{d}t$ from $\mathbb{P}(\mathrm{d}P_t = 0) = \exp^{-\lambda(t)\mathrm{d}t} \approxeq 1$ as $\mathrm{d}t \to 0$ proving equation (3.7). $\square$

*Proof. of theorem 3.2.2* We use the process of liberation as in [68] to liberate $\mathrm{p}$ and obtain $\mathcal{F}^{\dagger\, u}_{\text{MJD}}$. More precisely we start with the term $\pi(t,x)\mathcal{F}^{u}_{\text{MJD}}\mathrm{p}(t,x)$ and manipulate it as follows

$$\begin{aligned}
&\pi(t,x)\mathcal{F}^{u}_{\text{MJD}}\mathrm{p}(t,x)\\
&=\sum_{i=1}^{n_x}\Big(-\pi(x,t)\frac{\partial}{\partial x_i}(F_i(t,x,u)\mathrm{p}(t,x))\Big) + \frac{1}{2}\sum_{i,j=1}^{n_x}\pi(t,x)\frac{\partial^2}{\partial x_i \partial x_j}([\Sigma(t,x,u)]_{ij}\mathrm{p}(x,t))\\
&\quad + \pi(t,x)\Bigg(\sum_{j=1}^{n_p}\int_{D_{Q_j}}\Big(\mathrm{p}(x-\eta_j(t,x))\\
&\qquad\qquad \cdot |I-\eta_{jx}(t,x)| - \mathrm{p}(t,x)\Big)(p_{Q_j}\lambda_j)(t,q_j;t,x-\eta_j(t,x,q_j))\mathrm{d}q_j\Bigg)\\
&=\sum_{i=1}^{n_x}\Big[-\frac{\partial}{\partial x_i}\Big(\pi(x,t)F_i(t,x,u)\mathrm{p}(t,x)\Big) + \mathrm{p}(t,x)F_i(t,x,u(t))\frac{\partial \pi(t,x)}{\partial x_i}\Big]\\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n_x}\Big[\frac{\partial}{\partial x_i}\Big(\pi(t,x)\frac{\partial}{\partial x_j}([\Sigma(t,x,u)]_{ij}\mathrm{p}(t,x))\Big) - \frac{\partial}{\partial x_i}\Big([\Sigma(t,x,u)]_{ij}\mathrm{p}(t,x)\frac{\partial \pi(t,x)}{\partial x_j}\Big)
\end{aligned}$$

$$
+ \mathrm{p}(t,x)[\Sigma(t,x,u)]_{ij}\frac{\partial^2\pi(t,x)}{\partial x_i\partial x_j}\Bigg] - \sum_{j=1}^{n_p}\int\limits_{D_{Q_j}} \pi(t,x)\mathrm{p}(t,x)(p_{Q_j}\lambda_j)(t,q_j;t,x-\eta_j(t,x,q_j))\mathrm{d}q_j
$$

$$
+ \sum_{j=1}^{n_p}\int\limits_{D_{Q_j}} \pi(t,x)\mathrm{p}(t,x-\eta_j)|I-\eta_{jx}|(p_{Q_j}\lambda_j)(t,q_j;t,x-\eta_j(t,x,q_j))\mathrm{d}q_j. \tag{3.69}
$$

Integrating equation (3.69) over $\mathbb{R}^{n_x}$ and using (T4) gives us

$$
\begin{aligned}
&\Big\langle \pi, \mathcal{F}^u_{\text{MJD}}\mathrm{p}\Big\rangle \\
&= \int\limits_{\mathbb{R}^{n_x}} \Bigg( \sum_{i=1}^{n_x}\mathrm{p}(t,x)F_i(t,x,u)\frac{\partial\pi}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n_x}\mathrm{p}(t,x)[\Sigma(t,x,u)]_{ij}\frac{\partial^2\pi}{\partial x_i\partial x_j} \\
&+ \sum_{j=1}^{n_p}\int\limits_{D_{Q_j}} \pi\Bigg( (\mathrm{p}(t,x-\eta_j)|I-\eta_{jx}| - \mathrm{p}(t,x))(p_{Q_j}\lambda_j)(t,q_j;t,x-\eta_j(t,x,q_j))\mathrm{d}q_j\Bigg)\Bigg)\mathrm{d}x.
\end{aligned} \tag{3.70}
$$

Considering the terms in the last summation of this equation, we change the dummy variable of integration from $x$ to $\xi_j$ in the first step. Then we choose $\xi_j = x + h_j(t,x,q_j) = x + \eta_j(t,\xi_j,q_j)$ in the second step where $h_j$ is assumed to be invertible w.r.t. $\xi_j$ and change the variable of integration to $x$ using the Change of Variables Theorem [71]. This inverse mapping exists since we assume $h_j$ to be a bijection from $\mathbb{R}^{n_x}$ to $\mathbb{R}^{n_x}$. Noting that transformed domain of integration is again $\mathbb{R}^{n_x}$, we have that for all $1 \leq j \leq n_p$

$$\begin{aligned}
&\int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \pi(t,x)\mathrm{p}(t,x-\eta_j(t,x,q_j)) \\
&\qquad\qquad \cdot |I - \eta_{jx}(t,x,q_j)|(p_{Q_j}\lambda_j)(t,q_j;t,x-\eta_j(t,x,q_j))\mathrm{d}q_j\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \pi(t,\xi_j)\mathrm{p}(t,\xi_j-\eta_j(t,\xi_j,q_j)) \\
&\qquad\qquad \cdot |I - \eta_{j\xi_j}(t,\xi_j,q_j)|(p_{Q_j}\lambda_j)(t,q_j;t,\xi_j-\eta_j(t,\xi_j,q_j))\mathrm{d}q_j\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^{n_x}} \int\limits_{D_{Q_j}} \mathrm{p}(t,x)\ \pi(t,x+h_j(t,x,q_j))(p_{Q_j}\lambda_j)(t,q_j;t,x)\mathrm{d}q_j\mathrm{d}x. \qquad (3.71)
\end{aligned}$$

So that equations (3.70), (3.71) complete the proof as follows

$$\begin{aligned}
&\left\langle \pi, \mathcal{F}^u_{\text{MJD}}\mathrm{p} \right\rangle \\
&= \int\limits_{\mathbb{R}^{n_x}} \left[ \sum_{i=1}^{n_x} F_i(t,x,u)\frac{\partial \pi(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n_x} [\Sigma(t,x,u)]_{ij}\frac{\partial^2 \pi(t,x)}{\partial x_i \partial x_j} \right. \\
&\quad \left. + \sum_{j=1}^{n_p} \int\limits_{D_{Q_j}} \Big(\pi(t,x+h_j(t,x,q_j)) - \pi(t,x)\Big)(p_{Q_j}\lambda_j)(t,q_j;t,x)\mathrm{d}q_j \right] \mathrm{p}(t,x)\ \mathrm{d}x \\
&= \left\langle \mathrm{p}, \mathcal{F}^{\dagger\, u}_{\text{MJD}}\pi \right\rangle. \qquad (3.72)
\end{aligned}$$

□

# CHAPTER 4

# MEAN FIELD GAMES FOR AGENTS WITH LANGEVIN DYNAMICS

The MFG theory emerged as a viable formalism and analytical tool to understand large-scale self-organizing networked systems. The underlying mean-field approach enables a tractable framework to describe very large numbers of rational, non-cooperative and interacting agents. MFG theory provides a game-theoretic optimal control interpretation of emergent behavior of non-cooperative agents. In this chapter discuss MFG models in which individual agents obey multidimensional nonlinear Langevin dynamics, and analyze the closed-loop stability of fixed points of the corresponding coupled forward-backward PDE systems. In such MFG models, the detailed balance property of the reversible diffusions underlies the perturbation dynamics of the forward-backward system. We use our approach to analyze closed-loop stability of two specific models. Explicit control design constraints which guarantee stability are obtained for a population distribution model and a mean consensus model. It is shown that under certain conditions, that static state feedback using the steady state controller can be employed to locally stabilize a MFG equilibrium. We validate this fact numerically.

## 4.1 Introduction

Large scale non-cooperative multi-agent systems involving coupled costs were introduced as mean field games (MFG) by Huang et. al [15] and Lasry et. al [16]. Key ideas in this theory are the rational expectations hypothesis, infinitely many anonymous agents and that individual decisions are based on statistical information about the collection of agents. Subsequently, this theory has become a viable tool in the analysis of large-scale, self-organizing networked-systems, and provides a game-theoretic optimal control interpretation of the the notion of *emergent behaviour* in the non-cooperative setting. In the continuum approach, MFG models are synthesized as standard [73] stochastic optimal control problems (OCP). Fully coupled Fokker Planck (FP) and Hamilton Jacobi Bellman (HJB) equations governing agent density and value functions constitute the mean field (MF) optimality system. Assumptions of quadratic control cost and control affine agent dynamics constitute quadratic MFG models [19]. MFG models have been constructed to study several naturally occurring and engineered large-scale networked systems, including traffic [74], financial [75], energy [76], and biological systems [77].

A characteristic feature of MFGs is the ability to model interaction between networked agents by designing a suitable cost function. If the cost function has only local density dependence and is strictly increasing, steady state solutions to the MF system are unique [78] in several cases. In the absence of monotonicity, MFGs exhibit non-unique solutions and related *phase transitions* [18],[77], [19]. Since real-world large-scale networked systems soften possess several 'operating regimes', non-monotonicity in the corresponding MFG models is expected to be the norm, rather than an exception. Closed-loop stability analysis of MFG models that do not satisfy the monotonicity condition has to be done on a case-by-case basis. A given fixed point of the MFG is called (linearly) closed-loop stable if any perturbation to the fixed-point density decays to zero under the action of the control, where both the density and control evolution are computed using the (linearized) coupled forward-backward system of FP-HJB PDEs.

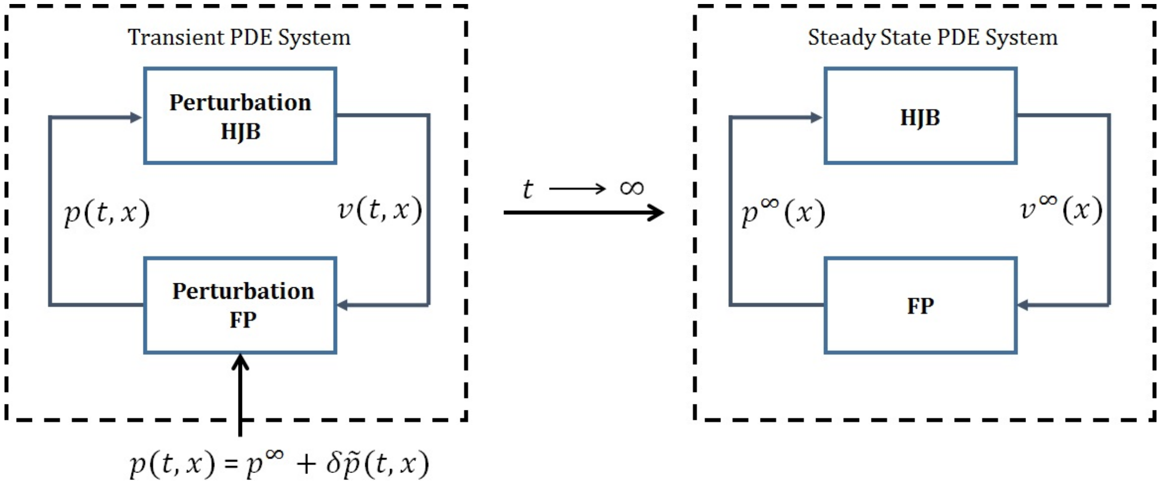



Figure 4.1: Stability of MFGs models

Guéant [21] studied the stability of an MFG model with a negative log density cost. Stability of MFGs with nonlocal cost coupling was considered for a Kuramoto oscillator model by Yin et. al [18] and a mean consensus cost by Nourian et. al ([79, 22]). A common limitation of these prior works is that the agents dynamics are assumed to be simple integrator systems. MF approach to large-scale networked systems with nonlinear agent dynamics have proved to be useful in modeling crowds [80], flocks [26], neural networks [81] and robotic control [12]. nIn our recent work [77], we analytically and numerically explored phase transitions in MFG models consisting of agents with nonlinear passive dynamics.

We expand upon the idea introduced in [77], and present rigorous closed-loop linear stability analysis for quadratic MFG models with dynamics of individual agents lying in the general class of controlled *reversible diffusions*. An example of such diffusions are the overdamped Langevin (simply Langevin for brevity) dynamics given in (4.1), while the simplest case is that of integrator systems. The key idea is that the *detailed balance* property of the generator of controlled reversible diffusions, and the resulting spectral properties of the linearized MFG system, allow for generalization of existing stability analysis techniques to this larger class of MFG systems. Furthermore, we demonstrate that static state feedback using the steady state controller can be employed to (sub-optimally) locally stabilize a MFG equilibrium.

In section 4.2, we describe the class of MFG models treated in this chapter. In section 4.3, we present the arguments detailing the main ideas for stability analysis for this class of models. Detailed analysis of closed-loop linear stability of steady states for (i) a population model with *local* cost coupling and (ii) consensus model with *nonlocal* cost coupling are presented next, which illustrate the key ideas in our approach. The population model consists of a general class of nonlinear controlled Langevin agent dynamics with a negative log density cost [21]. In section 4.4 we present technical conditions required for stability on the stationary solution and control parameters, and local stability results for this model. This analysis generalizes the stability analysis for the integrator dynamics case presented in [21]. The consensus model has flocking cost as in [22]. In section 4.5 we present stationary solutions, control design parameter constraints and linear stability results for this model in which agents obey Langevin dynamics with quadratic potential. Our results on this model generalize those of [22] concerned with integrator agent dynamics. A part of the results presented in this section will be published in [23].

Finally, in section 4.6, the action of the MF steady state controller on a population of agents in a MFG with nonlinear Langevin dynamics is considered. We show that a population of agents with perturbed (non Gaussian) initial densities will decay to the (closest) stationary density under the action of static feedback given by the corresponding steady state controller.

## 4.2 Mean Field Game Model

In this section, we first introduce some notation and then describe the MFG model treated in this chapter. $L^2(g\ \mathrm{d}x; \mathbb{R}^d)$ denotes the class of $g$-weighted 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 $||f||_{L^2(g\,\mathrm{d}x;\mathbb{R}^d)}$ and $\langle f_1, f_2 \rangle_{L^2(g\,\mathrm{d}x;\mathbb{R}^d)}$ respectively.

Let $x_s$, $u(s) \in \mathbb{R}^d$ denote the state and control inputs of a representative agent which obeys controlled Langevin dynamics in the overdamped case, given by

$$\mathrm{d}x_s = -\nabla\nu(x_s)\mathrm{d}s + u(s)\mathrm{d}s + \sigma\mathrm{d}w_s \tag{4.1}$$

for every $s \geq 0$, driven by standard $\mathbb{R}^d$ Brownian motion, with noise intensity $0 < \sigma$ on the filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$. 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 [31] containing functions $u : [t,T] \times \mathbb{R}^d \to \mathbb{R}^d$. The MFG models treated in this work can be written as the following control problem subject to (4.1)

$$\min_{u\in\mathcal{U}} J(u) := \mathbb{E}\left[\int_t^T \mathrm{e}^{-\rho s}\left(q(x_s, \mathrm{p}(s,x_s)) + \frac{R}{2}u^2(s)\right)\mathrm{d}s\right], \tag{4.2}$$

where we denote the probability density of $x_s$ by $\mathrm{p}(s,x)$ for every $s \geq 0$, with initial density being $x_t \sim \mathrm{p}(t,x)$, $q : \mathbb{R}^d \times L^1(\mathbb{R}^d) \to \mathbb{R}$ is a known deterministic function which has at most quadratic growth in $(x, \mathrm{p})$ and $R > 0$ is the control cost. We assume that the functions in the class $\mathcal{U}$ and $\nabla\nu(x), q(x,\mathrm{p})$ are measurable. The value function is defined as $v(t,x) := \min_{u\in\mathcal{U}} J(u)$ given $x_t = x$. It can be seen by standard application of *dynamic programming* [82] as in ([73], [16]), that this control problem is equivalent to the following PDE system

$$-\partial_t v = q - \rho v - \frac{(\nabla v)^2}{2R} - \nabla v \cdot \nabla\nu + \frac{\sigma^2}{2}\Delta v \tag{4.3}$$

$$\partial_t \mathrm{p} = \nabla\cdot\left((\nabla\nu + \frac{\nabla v}{R})p\right) + \frac{\sigma^2}{2}\Delta\mathrm{p} \tag{4.4}$$

with the optimal control $u^*(t,x) = -\nabla v/R$, the mass conservation constraint $\int \mathrm{p}(s,x)\mathrm{d}x = 1$ for all $s \geq 0$ and boundary constraints $\lim_{|x|\to+\infty} \mathrm{p}(t,x) = 0$, $\lim_{s\to+\infty} \mathrm{e}^{-\rho s}v(s,x_s) = 0$. These fully coupled equations identified as the HJB and FP PDEs comprise the MF optimality system. An infinite time horizon, that is $T \to +\infty$, leads to the stationary system

$$
\begin{aligned}
0 =& q(x, \mathrm{p}^{\infty}) - \rho v^{\infty} - \frac{(\nabla v^{\infty})^2}{2R} - \nabla v^{\infty} \cdot \nabla \nu \\
&+ \frac{\sigma^2}{2} \Delta v^{\infty}, \qquad (4.5)\\
0 =& \nabla \cdot \big( (\nabla \nu + \frac{\nabla v^{\infty}}{R}) \mathrm{p}^{\infty} \big) + \frac{\sigma^2}{2} \Delta \mathrm{p}^{\infty}, \qquad (4.6)
\end{aligned}
$$

governing the fixed point pair $(v^{\infty}(x), \mathrm{p}^{\infty}(x))$ of steady state value and density functions, with constraints $\int \mathrm{p}^{\infty}(x) \mathrm{d}x = 1$, and $\lim_{s \to +\infty} \mathrm{e}^{-\rho s} v^{\infty}(x_s) = 0$. The optimal control is $u^{\infty}(x) = -\nabla v^{\infty}/R$. Interesting examples of such dynamics are noisy potential wells and Kuramoto oscillator models. Note that Newtonian or second order state space dynamics cannot be modeled by these dynamics. In this case Langevin dynamics [83] are the appropriate model used. In the overdamped limit the density dynamics are precisely the Smoluchowski equation. To extend the results in this paper to Langevin dynamics we must deal with a generator operator which is the sum of an anti symmetric and a symmetric operator.

**Remark 1.** *If the MFG model has a long-time-average utility [18],*

$$
\min_{u \in \mathcal{U}} J(u) := \lim_{T \to +\infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T q(x_s, \mathrm{p}(s, x_s)) + \frac{R}{2} u^2(s) \, \mathrm{d}s \right], \qquad (4.7)
$$

*instead of the discounted version in* (4.2)*, then the corresponding stationary optimality system consists of (*(4.5), (4.6)*), on observing the limit $\rho v^{\infty} \to \lambda$ in* (4.5) *as $\rho \to 0$, where $\lambda$ is the optimal cost. Please see [84] and references therein for proof of this connection between the utilities. In this case, the time dependent, relative value function [85] obeys* (4.3) *wherein $\rho v$ is replaced by $\lambda$. Similarly, the perturbation system is obtained from* (4.17) *by setting $\rho = 0$. Thus, all the results in sections 4.3, 4.4, 4.5, 4.6 can be directly extended to the LTA utility case.*

### 4.3 Perturbation System

The FP equation for the density of an overdamped Langevin system is called the Smoluchowski PDE. From the form of the FP PDE (4.6), it can be interpreted as the Smoluchowski PDE for such a Langevin system with the restoring potential $\nu + v^{\infty}/R$. This interpretation allows us to obtain the analytical solution to the FP PDE as a Gibbs distribution, if the fixed point pair $(v^{\infty}, \mathrm{p}^{\infty})$ of the MFG (4.5, 4.6) and the Langevin potential $\nu$ satisfy certain conditions. We denote $w(x) := \nu(x) + \frac{v^{\infty}(x)}{R}$ henceforth in this chapter.

**Lemma 4.3.1.** *If $v^{\infty}(x)$, $\nu(x)$ are smooth functions satisfying $\lim\limits_{|x|\to+\infty} w(x) = +\infty$ and $\exp\left(-\frac{2}{\sigma^2} w(x)\right) \in L^1(\mathbb{R}^d)$, then the unique stationary solution to the density given by the Fokker Planck equation* (4.6) *is*

$$\mathrm{p}^{\infty}(x) := \frac{1}{Z} \exp\left(-\frac{2}{\sigma^2}\Big(w(x)\Big)\right)(x), \tag{4.8}$$

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

*Proof.* We observe that the (4.6) is the Smoluchowski equation for an overdamped Langevin system given by

$$\mathrm{d}x_s = -\nabla(\nu + v^{\infty}/R)(x_s)\,\mathrm{d}s + \sigma \mathrm{d}w_s. \tag{4.9}$$

Under the assumptions above, the proof then follows directly from proposition 4.2, pp 110 in [83]. □

Decay of an initial density of particles under uncontrolled (or open loop) overdamped Langevin dynamics to a stationary density is a classical topic [86]. We address the question of decay of a locally perturbed density of agents in a MFG to a steady state density under the closed loop time varying as well as steady state MFG optimal controls. The perturbation analysis then corresponds to a fully coupled forward-backward PDE system. The proposed approach leads to a general method to obtain stability constraints on the *control design parameters*, with explicit analytical results in certain cases.

To derive the linearization of MFG system (4.5, 4.6) around the pair $(v^\infty, \mathrm{p}^\infty)$, we write the perturbed density and value functions as $\mathrm{p}(t,x) = \mathrm{p}^\infty(x)(1+\epsilon\tilde{\mathrm{p}}(x,t))$, and $v(t,x) = v^\infty(x) + \epsilon\tilde{v}(x,t)$ respectively. The corresponding perturbed cost is $q(x;\mathrm{p}) = q(x;\mathrm{p}^\infty) + \epsilon\tilde{q}(x;\mathrm{p}^\infty,\tilde{\mathrm{p}})$ where $\epsilon > 0$. We denote $q^\infty(x) := q(x,\mathrm{p}^\infty)$, and $\tilde{q}(x) := \tilde{q}(x;\mathrm{p}^\infty(x),\tilde{\mathrm{p}}(t,x))$ for brevity.

The generator of a Langevin process is intrinsically linked to the stability properties of its density dynamics. We denote the generator of the optimally controlled agent dynamics (4.9) as $\mathcal{L}(\cdot) := -\nabla(\nu + v^\infty/R)\cdot\nabla(\cdot) + (\sigma^2/2)\Delta(\cdot)$ and its $L^2(\mathbb{R})$ adjoint $\mathcal{L}^\dagger(\cdot) := \nabla\cdot(\nabla(\nu+v^\infty/R)(\cdot)) + (\sigma^2/2)\Delta(\cdot)$.

**Theorem 4.3.2.** *If $(v^\infty(x), \mathrm{p}^\infty(x))$ are smooth steady state solutions to the MF system (4.5, 4.6) wherein $\nu$ is a smooth function such that $\lim_{|x|\to+\infty} w(x) = +\infty$ and $\exp\left(-\frac{2}{\sigma^2}w(x)\right) \in L^1(\mathbb{R}^d)$, then the linearization of the MF system (4.3, 4.4) around $(v^\infty(x), \mathrm{p}^\infty(x))$ for all $(t,x) \in [0,+\infty)\times\mathbb{R}^d$ is*

$$-\partial_t\tilde{v} = \tilde{q} - \rho\tilde{v} + \mathcal{L}\tilde{v}, \tag{4.10}$$

$$\partial_t\tilde{\mathrm{p}} = (2/\sigma^2 R)\mathcal{L}\tilde{v} + \mathcal{L}\tilde{\mathrm{p}}, \tag{4.11}$$

*where $\tilde{\mathrm{p}}(0,x)$ is given, $\int_{\mathbb{R}^d}\mathrm{p}^\infty(x)(1+\epsilon\tilde{\mathrm{p}}(t,x))\mathrm{d}x = 1$ for all $t\geq 0, \epsilon > 0$, $\lim_{|x|\to+\infty}\tilde{\mathrm{p}}(t,x) = 0$ for all $t \geq 0$ and $\lim_{t\to+\infty}\mathrm{e}^{-\rho t}\tilde{v}(t,x_t) = 0$.*

*Proof.* Substituting the perturbation density $\mathrm{p} = \mathrm{p}^\infty(1+\epsilon\tilde{\mathrm{p}})$ in (4.4), using the fixed point equation (4.6) and neglecting higher order $\epsilon$ terms we have

$$\begin{aligned}
&\partial_t(\mathrm{p}^\infty + \epsilon\mathrm{p}^\infty\tilde{\mathrm{p}})\\
&= \partial_x\Bigg(\partial_x(\nu + (v+\epsilon\tilde{v})/R)(\mathrm{p}^\infty(1+\epsilon\tilde{\mathrm{p}}))\Bigg) + \frac{\sigma^2}{2}\partial_{xx}(\mathrm{p}^\infty(1+\epsilon\tilde{\mathrm{p}}))\epsilon\partial_t(\mathrm{p}^\infty\tilde{\mathrm{p}})\\
=&\epsilon\nabla\cdot\left(\frac{\nabla v}{R}\mathrm{p}^\infty + \nabla(\nu+\frac{v}{R})\mathrm{p}^\infty\tilde{\mathrm{p}}\right) + \epsilon\frac{\sigma^2}{2}\Delta(\mathrm{p}^\infty\tilde{\mathrm{p}}),
\end{aligned} \tag{4.12}$$

so that using the operator $\mathcal{L}^\dagger$ and the fact that $\epsilon > 0$,

$$\partial_t(\mathrm{p}^\infty \tilde{\mathrm{p}}) = \nabla \cdot \left(\nabla \tilde{v}\ \mathrm{p}^\infty / R\right) + \mathcal{L}^\dagger(\mathrm{p}^\infty \tilde{\mathrm{p}}). \tag{4.13}$$

It can be verified [87] that for a smooth function $f(x)$ the generator and its adjoint operator satisfy the *detailed balance* property $\mathcal{L}^\dagger(\mathrm{p}^\infty f) = \mathrm{p}^\infty \mathcal{L} f$. Therefore from (4.13) we have

$$\mathrm{p}^\infty \partial_t \tilde{\mathrm{p}} = \frac{1}{R}\left(\mathrm{p}^\infty \Delta \tilde{v} + \nabla \tilde{v} \cdot \nabla \mathrm{p}^\infty\right) + \mathrm{p}^\infty \mathcal{L} \tilde{\mathrm{p}}. \tag{4.14}$$

From the assumed conditions on the potential and value functions, lemma 4.3.1 gives the stationary density, so that $\nabla \mathrm{p}^\infty = -\frac{2}{\sigma^2}(\nabla \nu + \frac{\nabla v^\infty}{R})\mathrm{p}^\infty$. Then the previous equation simplifies as

$$\mathrm{p}^\infty \partial_t \tilde{\mathrm{p}} = \mathrm{p}^\infty \left(\mathcal{L}\tilde{\mathrm{p}} + \frac{2}{\sigma^2 R}\mathcal{L}\tilde{v}\right), \tag{4.15}$$

giving us the density perturbation equation since $\mathrm{p}^\infty(x) > 0$. Substituting the perturbation value function $v = v^\infty + \epsilon \tilde{v}$ in (4.3), using the fixed point equation (4.5) and neglecting higher order $\epsilon$ terms gives

$$\begin{aligned}
& -\partial_t(v^\infty + \epsilon \tilde{v}) \\
=& q + \epsilon \tilde{q} - \rho(v^\infty + \epsilon \tilde{v}) - \frac{1}{2R}(\partial_x(v^\infty + \epsilon \tilde{v}))^2 - \partial_x(v^\infty + \epsilon \tilde{v})\partial_x \nu + \frac{\sigma^2}{2}\partial_{xx}(v^\infty + \epsilon \tilde{v}) \\
=& q^\infty + \epsilon \tilde{\tilde{q}} - \rho v^\infty - \epsilon \rho \tilde{v} - (\partial_x v^\infty)^2/R - \epsilon v_x^\infty \tilde{v}_x / R - v_x^\infty \nu_x \\
& \qquad - \epsilon \tilde{v}_x \nu_x + \frac{\sigma^2}{2}\partial_{xx} v^\infty + \epsilon \frac{\sigma^2}{2}\partial_{xx}\tilde{v} \\
=& -\epsilon \partial_t \tilde{v} = \epsilon \tilde{q} - \epsilon \rho \tilde{v} - \epsilon \nabla \left(\frac{v^\infty}{R} + \nu\right) \cdot \nabla \tilde{v} + \epsilon \frac{\sigma^2}{2}\Delta \tilde{v}.
\end{aligned} \tag{4.16}$$

Using the operator definition and since $\epsilon > 0$ we get the required result. The mass conservation and boundary constraints on $\tilde{v}, \tilde{p}$ follow directly from those constraints on (4.3, 4.4). □

In the following two sections we will apply the above result to obtain stability results for two MFG models. Note that the perturbation system may be written in concatenated form as

$$\partial_t \begin{bmatrix} \tilde{v} \\ \tilde{\mathrm{p}} \end{bmatrix} = \begin{bmatrix} -\mathcal{L} + \rho & 0 \\ \frac{2}{\sigma^2 R}\mathcal{L} & \mathcal{L} \end{bmatrix} \begin{bmatrix} \tilde{v} \\ \tilde{\mathrm{p}} \end{bmatrix} + \begin{bmatrix} -\tilde{q} \\ 0 \end{bmatrix}. \tag{4.17}$$

## 4.4 A Population Distribution model

We present the linear stability result for a population distribution MFG model in this section. A cost function with local density dependence is used in this model to mimic a population of agents with identical dynamics, seeking to minimize their cost functional but with a preference for imitating their peers. This model agents in an economic network [88]. A reference case for this model is [21] where the simplest case of integrator agent dynamics was treated. Note that while a strictly increasing cost function $q(p(t,x))$ models aversion among agents, a strictly decreasing one models cohesion [89]. We reiterate that, as stated in the introduction, there is no general uniqueness result for the stationary solution, in case of such a monotonically decreasing cost function. We consider a model comprised by the OCP (4.2) with the negative log density cost and agents following nonlinear Langevin dynamics (4.1).

The MF optimality system for this model consists of the coupled system (4.3, 4.4) along with the cost coupling equation

$$q(x, \mathrm{p}(t,x)) = -\ln \mathrm{p}(t,x), \tag{4.18}$$

where $\mathrm{p}(0,x) = \mathrm{p}_0(x)$ is the given initial density of agents, $\int \mathrm{p}(t,x)\mathrm{d}x = 1$ for all $t \geq 0$, $\lim_{t\to+\infty} \mathrm{p}(t,x) = 0$ and $\lim_{t\to+\infty} \mathrm{e}^{-\rho t} v(t,x_t) = 0$.

### 4.4.1 Stationary Solution

The stationary MF optimality system is given by (4.5, 4.6) and the cost coupling equation

$$q^{\infty}(x) = -\ln \mathrm{p}^{\infty}(x), \tag{4.19}$$

where $\int \mathrm{p}^{\infty}(x)\mathrm{d}x = 1$ and $\lim_{t\to+\infty} \mathrm{e}^{-\rho t} v^{\infty}(x_t) = 0$.

Calculating analytical solutions to HJB PDEs is a daunting task, examples of which are rare and mainly related to linear-quadratic regimes. The presented approach aims at being applicable to the most general class of nonlinear dynamics. We show that under certain conditions on the (unknown) stationary solution $(v^{\infty}, \mathrm{p}^{\infty})$, one may obtain sufficiency conditions required for linear stability of the population model. Conditions on the stationary solution required to guarantee stability are stated in the following assumptions. Let $w(x) := \nu(x) + \frac{v^{\infty}(x)}{R}$.

**(A1)** There exist $(v^{\infty}(x), p^{\infty}(x)) \in (C^2(\mathbb{R}^d))^2$ satisfying (4.5,4.6,4.19) such that $\lim_{|x|\to+\infty} w(x) = +\infty$ and $\exp\left(-\frac{2}{\sigma^2} w(x)\right) \in L^1(\mathbb{R}^d)$.

Due to this assumption, lemma 4.3.1 implies that the stationary density is uniquely determined by the analytical expression (4.8).

### 4.4.2 Linear Stability

Under the assumption **(A1)**, the perturbation PDEs for the value and density functions as well as the constraints follow directly from theorem 4.3.2. The only term in (4.17) specific to the cost coupling (4.18) is given by

$$\tilde{q}(x; \tilde{\mathrm{p}}(t,x)) = -\,\tilde{\mathrm{p}}(t,x), \tag{4.20}$$

using the Taylor series expansion.

We define a Hilbert space and a class perturbations in it, for which we show stability.

**Definition 4.4.1.** *Let (A1) hold. Denote the density* $\mathrm{p}^\infty(x) := \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}w(x)\right)(x)$ *with the normalizing constant* $Z$ *where* $(v^\infty, p^\infty)$ *is a pair satisfying* ***(A1)****. Denote by* $\mathcal{H}$ *the Hilbert space* $L^2(\mathrm{p}^\infty(x)\mathrm{d}x;\mathbb{R}^d)$*. The class of mass preserving density perturbations is defined as* $S_0 := \left\{q(x) \in \mathcal{H} \middle| \langle 1, q(x)\rangle_{\mathcal{H}} = 0\right\}$.

**Definition 4.4.2.** *Let us denote the set of initial perturbed densities by* $S(\epsilon) = \left\{p(0,x) = p^\infty(x)(1+\epsilon\tilde{\mathrm{p}}(0,x)) \middle| p(0,x) \geq 0, \tilde{\mathrm{p}}(0,x) \in S_0\right\}$*. We say the fixed point* $(v^\infty(x), p^\infty(x))$ *of the MF optimality system (4.3, 4.4) is linearly asymptotically stable with respect to* $S(\epsilon)$ *if there exists a solution* $(\tilde{v}(t,x), \tilde{p}(t,x))$ *to the perturbation system (4.10,4.11) such that* $\lim_{t\to+\infty} ||\tilde{\mathrm{p}}(t,x)||_{\mathcal{H}} = 0$.

Since we are concerned with stability of isolated fixed points, we *do not* assume that initial perturbations are mean preserving [21].

**Lemma 4.4.1.** *If (A1) is true then* $\mathcal{L}$ *is self adjoint in* $L^2(\mathrm{p}^\infty\mathrm{d}x;\mathbb{R}^d)$*, negative semidefinite and its kernel consists of constants.*

*Proof.* Due to **(A1)** $v^\infty(x)$ is differentiable and the operator $\mathcal{L}$ is well defined. We observe that it is the generator of an overdamped Langevin system (4.9) under a potential $\nu + v^\infty/R$ and noise intensity $\sigma$. The proof follows from proposition 4.3, pp 111 in [83]. □

We need an assumption to obtain relevant properties of the generator of the controlled process.

**(A2)** $\lim_{|x|\to+\infty}\left(\frac{|\nabla w(x)|^2}{2} - \frac{\sigma^2}{2}\Delta w(x)\right) = +\infty$
and $\nu(x) \in C^2(\mathbb{R}^d)$.

An example of an one dimensional MFG model with integrator dynamics and its corresponding stationary solution $v^\infty(x)$ satisfying this assumption was explicitly constructed in [21].

**Lemma 4.4.2.** *nLet **(A1, A2)** hold. Then* $\mathrm{p}^\infty(x)$ *satisfying **(A1)** and given by* (4.8)*, satisfies the Poincaré inequality with* $\lambda > 0$*, that is, for all* $f \in C^1(\mathbb{R}^d) \cap L^2(\mathrm{p}^\infty(x)\mathrm{d}x; \mathbb{R}^d)$ *such that* $\int f \mathrm{p}^\infty(x)\mathrm{d}x = 0$ *there exists* $\lambda > 0$ *such that*

$$\begin{aligned} &\lambda \frac{2}{\sigma^2} ||f||^2_{L^2(\mathrm{p}^\infty(x)\mathrm{d}x;\mathbb{R}^d)} \\ &\leq ||\nabla f||_{L^2(\mathrm{p}^\infty(x);\mathbb{R}^d)} = -\big\langle \mathcal{L}f, f \big\rangle_{L^2(\mathrm{p}^\infty(x);\mathbb{R}^d)}. \end{aligned} \tag{4.21}$$

*Proof.* nThe assumptions imply that $v^\infty(x) \in C^2(\mathbb{R}^d)$, and hence, $(\nu + v^\infty/R)(\cdot) \in C^2(\mathbb{R}^d)$. Observe that operator $\mathcal{L}$ is the generator of an overdamped Langevin system under a potential $\nu + v^\infty/R$ and noise intensity $\sigma$. The proof then follows from theorem 4.3, pp 112 in [83]. □

Lemma 4.4.1 implies that eigenvalues of $\mathcal{L}$ are real, negative semidefinite and its eigenfunctions are orthonormal in $L^2(\mathrm{p}^\infty(x)\mathrm{d}x; \mathbb{R}^d)$ while lemma 4.4.2 implies that the eigenvalues of $\mathcal{L}$ are discrete and its eigenfunctions are complete on $L^2(\mathrm{p}^\infty(x)\mathrm{d}x; \mathbb{R}^d)$ [87]. We denote the eigenvalues $\{\xi_n\}_{n\geq 0}$ and corresponding eigenfunctions $\{\Xi_n\}_{n\geq 0}$ of $\mathcal{L}$ which form a complete orthonormal basis of $\mathcal{H}$. Let eigenvalues $\{\xi_n\}_{n\geq 0}$ be indexed in descending order of magnitude $0 = \xi_0 > \xi_1 > ... > \xi_n > ...$ and let $\Xi_0 = 1$.

**Remark 2.** *The detailed balance* $\mathcal{L}^\dagger(\mathrm{p}^\infty f) = \mathrm{p}^\infty \mathcal{L}(f)$ *used in proof of theorem* (4.3.2) *is the key property, because of which we have distinct, real and non negative eigenvalues [87] of the generator* $\mathcal{L}$*. nThese eigen properties make the presented approach to stability analysis of MFGs possible, through the result in theorem 4.3.2.*

**(A3)** $\rho - \frac{2}{\sigma^2 R} > \xi_n$ for all $n \geq 1$.

The assumption above is the explicit control design constraint required to show stability. Denote the matrix associated with the MF system for the population distribution model $A_n := \begin{bmatrix} -\xi_n + \rho & 1 \\ \frac{2}{\sigma^2 R}\xi_n & \xi_n \end{bmatrix}$.

**Lemma 4.4.3.** *If $\xi_n \neq 0$ and $\rho - \frac{2}{\sigma^2 R} > \xi_n$ then the eigenvalues of $A_n$ are real, distinct and ordered $\lambda_n^1 < 0 < \lambda_n^2$.*

*Proof.* The characteristic equation of $A_n$ is $\lambda_n^2 - \rho\lambda_n + (\rho - \xi_n)\lambda_n + \frac{2}{\sigma^2 R}\xi_n = 0$ has the eigenvalue roots $\lambda_n^{1,2} = \frac{\rho}{2} \pm \sqrt{\left(\frac{\rho}{2}\right)^2 - (\rho - \xi_n)\xi_n + \frac{2}{\sigma^2 R}\xi_n}$ from which the result follows. □

The spectral properties of perturbation MFG system derived in this section allow us to extend the methods in [21] (applied to integrator agent dynamics) to the case of nonlinear Langevin agent dynamics. Note that the stationary solution as well as the eigenbasis are not explicitly known here, unlike in previous works which exploit the Hermite basis resulting from explicitly known quadratic-Gaussian stationary solutions.

**Theorem 4.4.4.** *Let (**A1**, **A2**, **A3**) hold, and $(v^\infty(x), \mathrm{p}^\infty(x))$ be a stationary solution to the MF system (4.3, 4.4, 4.18). If perturbation $\tilde{\mathrm{p}}(0,x) \in S_0$ and $\{v_n, p_n\}_{n\geq 0}$ is determined by $p_0(t) = 0$, and for $n \geq 0$*

$$\begin{bmatrix} \dot{v_n} \\ \dot{p_n} \end{bmatrix} = A_n \begin{bmatrix} v_n \\ p_n \end{bmatrix}, \tag{4.22}$$

$$\mathrm{p}_n(0) = \langle \tilde{p}(0,x), \Xi_n(x) \rangle_{\mathcal{H}}, \tag{4.23}$$

*then $\{\tilde{v}(t,x) = \sum_{n=0}^{+\infty} v_n(t)\Xi_n(x), \tilde{\mathrm{p}}(t,x) = \sum_{n=0}^{+\infty} p_n(t)\Xi_n(x)\}$ are unique $\mathcal{H}$ solutions to the perturbation MF system (4.10,4.11,4.20). $\mathrm{p}^\infty(x)$ is linearly asyptotically stable with respect to $S(\epsilon)$.*

*Proof.* ***Finite time solution:*** We first construct finite time solutions to the perturbation system (4.10,4.11,4.20) under initial and terminal time boundary conditions $\tilde{v}(T,x) \in \mathcal{H}$, $\tilde{p}(0,x) \in S_0$. We have the unique representations $\tilde{v}(T,x) = \sum_{n=0}^{+\infty} v_n(T)\Xi_n(x)$ and $\tilde{\mathrm{p}}(0,x) = \sum_{n=0}^{+\infty} p_n(0)\Xi_n(x)$, where

$$v_n(T)_{n\geq 0} = \langle \tilde{v}(T,x), \Xi_n(x) \rangle_{\mathcal{H}}, \tag{4.24}$$

and $p_n(0)_{n\geq 0}$ is given by (4.23).

Consider the infinite sums $\{\sum_{n=0}^{+\infty} v_n(t)\Xi_n(x),\ \sum_{n=0}^{+\infty} p_n(t)\Xi_n(x)\}$. Using the eigen property $\mathcal{L}\Xi_n(x) = \xi_n \Xi_n(x)$, and inserting the infinite sums into the perturbation system (4.10, 4.11, 4.20) yields the ODE system (4.22).

For $n = 0$, since $\tilde{p}(0, x) \in S_0$ and $\Xi_0 = 1$, we know that $p_0(0) = \langle \Xi_0, \tilde{p}(0, x)\rangle = 0$. Since $\xi_0 = 0$, from the matrix $A_n$ we have $\dot{p}_0(t) = 0$ implying $p_0(t) = 0$ for all $t \in [0, T]$. Therefore, $v_0(t) = v_0(T)\mathrm{e}^{-\rho(T-t)}$.

For $n \geq 1$, from lemma 4.4.3 we have that the eigenvalues $spec(A_n) = \lambda_n^{1,2}$ are distinct, real and are ordered $\lambda_n^1 < 0 < \lambda_n^2$. We may write

$$\begin{bmatrix} v_n(t) \\ p_n(t) \end{bmatrix} = C_1^{n,T}\mathrm{e}^{\lambda_n^1 t}\begin{bmatrix} 1 \\ e_n^1 \end{bmatrix} + C_2^{n,T}\mathrm{e}^{\lambda_n^2 t}\begin{bmatrix} 1 \\ e_n^2 \end{bmatrix}, \tag{4.25}$$

with eigenvector components $e_n^{1,2} = \xi_n - \rho + \lambda_n^{1,2}$. Boundary conditions give us $v_n(T) = C_1^{n,T}\mathrm{e}^{\lambda_n^1 T} + C_2^{n,T}\mathrm{e}^{\lambda_n^2 T}$ and $p_n(0) = C_1^{n,T}e_n^1 + C_2^{n,T}e_n^2$ implying

$$C_1^{n,T} = \frac{(e_n^2/e_n^1)v_n(T) - \mathrm{e}^{\lambda_n^2 T}(p_n(0)/e_n^1)}{(e_n^2/e_n^1)\mathrm{e}^{\lambda_n^1 T} - \mathrm{e}^{\lambda_n^2 T}}, \tag{4.26}$$

$$C_2^{n,T} = \frac{(p_n(0)/e_n^1) - v_n(T)}{(e_n^2/e_n^1)\mathrm{e}^{\lambda_n^1 T} - \mathrm{e}^{\lambda_n^2 T}}. \tag{4.27}$$

From the eigenvalues given by lemma 4.4.3 and since in the limit $\xi_n \to -\infty$, we observe that $e_n^1 \sim -2|\xi_n|$ and $e_n^2 \sim \frac{\rho}{2}$ as $n \to +\infty$ so that in the limit $C_1^{n,T} \sim \frac{p_n(0)}{e_n^1}$ and $C_2^{n,T} \sim \frac{v_n(T)}{\mathrm{e}^{\lambda_n^2 T}}$. We can therefore say that $v_n(t) = O\left(-\frac{p_n(0)}{2|\xi_n|}\mathrm{e}^{-\lambda_n^1 t}\right) + O\left(v_n(T)\mathrm{e}^{-\lambda_n^2(T-t)}\right)$ and $p_n(t) = O\left(p_n(0)\mathrm{e}^{\lambda_n^1 t}\right) + O\left(v_n(T)\mathrm{e}^{-\lambda_n^2(T-t)}\right)$. From these estimates we can say that $\sum_{n=0}^{+\infty} v_n(t)\Xi_n(x)$, $\sum_{n=0}^{+\infty} p_n(t)\Xi_n(x)$ given by the ODE system (4.22, 4.24, 4.23) are in $C^\infty([0, T] \times \mathbb{R}^d)$ and $\mathcal{H}$.

Since $\{\Xi_n\}_{n\geq 0}$ is a complete basis to $\mathcal{H}$, any solution in $\mathcal{H}$ to the system (4.10, 4.11, 4.20) must have the form $\{\tilde{v}(t, x) = \sum_{n=0}^{+\infty} v_n(t)\Xi_n(x), \tilde{p}(t, x) = \sum_{n=0}^{+\infty} p_n(t)\Xi_n(x)\}$ where $\{v_n, p_n\}_{n\geq 0}$ are finite for all $t \in [0, T]$. This concludes the proof that such a

$\{\tilde{v}(t,x), \tilde{\mathrm{p}}(t,x)\}$ governed by the ODE system (4.22, 4.23, 4.24) is a unique $\mathcal{H}$ solution to the perturbation system (4.10, 4.11, 4.20).

***Asymptotic stability:*** Now, we construct infinite time solutions by considering the limit $T \to +\infty$ of the solutions in the finite time case. As explained in the finite time solutions case, it can be shown that $p_0(t) = 0$ at all times.

The pair $\{\tilde{v}(t,x) = \sum_{n=0}^{+\infty} v_n(t)\Xi_n(x), \tilde{p}(t,x) = \sum_{n=0}^{+\infty} p_n(t)\Xi_n(x)\}$ is a unique solution specified by (4.25) given the initial and terminal coefficients $\mathrm{p}_n(0)$ and $v_n(T)$ for all $n \geq 0$. Now, if $\tilde{v}(t,x) \in \mathcal{H}$ then $\lim_{t\to+\infty} |v_n(t)| < +\infty$ for all $n \geq 0$. It is also known that $|p_n(0)| < +\infty$. Therefore, for $n = 0$, this means that $p_0(t) = 0$ and $v_0(t) = v_0(T)\mathrm{e}^{-\rho(T-t)} \xrightarrow{T\to+\infty} 0$.

From lemma 4.4.3, the eigenvalues of $A_n$ are ordered $\lambda_n^1 < 0 < \lambda_n^2$ for all $n \geq 1$ due to **(A3)**. Therefore, for all $n \geq 1$, we observe from the finite time solutions (4.26, 4.27) to the ODE system (4.22), that $C_1^{n,T} \to \frac{p_n(0)}{e_n^1}$ and $C_2^{n,T} \to v_n(T)\mathrm{e}^{-\lambda_n^2 T}$ as $T \to +\infty$. Since $\lambda_n^1 < 0 < \lambda_n^2$, for any $\alpha \in (0, \frac{1}{2})$ and as $T \to +\infty$, it can be obtained from (4.25) that

$$\begin{aligned} \sup_{t\in[\alpha T,(1-\alpha T)]} |v_n(t)| &\leq |C_1^{n,T}|\mathrm{e}^{\lambda_n^1 \alpha T} + |C_2^{n,T}|\mathrm{e}^{\lambda_n^2(1-\alpha)T} \\ &\leq \left|\frac{p_n(0)}{e_n^1}\right| \mathrm{e}^{\lambda_n^1 \alpha T} + |v_n(T)|\mathrm{e}^{-\lambda_n^2 \alpha T}, \end{aligned} \tag{4.28}$$

$$\begin{aligned} \sup_{t\in[\alpha T,(1-\alpha T)]} |p_n(t)| &\leq |C_1^{n,T}||e_1^1|\mathrm{e}^{\lambda_n^1 \alpha T} + |C_2^{n,T}||e_1^2|\mathrm{e}^{\lambda_n^2(1-\alpha)T} \\ &\leq |p_n(0)|\mathrm{e}^{\lambda_1^1 \alpha T} + |v_n(T)|\mathrm{e}^{-\lambda_1^2 \alpha T}, \end{aligned} \tag{4.29}$$

the right sides of which vanish in the limit since $|v_n(T)| < +\infty$ and $|p_n(0)| < +\infty$.

We have shown that the unique solution in $\mathcal{H}$ to the MF perturbation system has the properties $v_0(t) = p_0(t) = 0$, $\lim_{t\to+\infty} v_n(t) = 0$ and $\lim_{t\to+\infty} p_n(t) = 0$ for all $n \geq 1$. Therefore using Parseval's theorem $||\tilde{v}(t,x)||_{L^2(\mathrm{p}^\infty(x);\mathbb{R}^d)} = \left(\sum_{n=1}^{+\infty} v_n(t)\right)^{\frac{1}{2}}$, $||\tilde{p}(t,x)||_{L^2(\mathrm{p}^\infty(x);\mathbb{R}^d)} = \left(\sum_{n=1}^{+\infty} p_n^2(t)\right)^{\frac{1}{2}}$ and the Lebesgue dominated convergence theorem, we have that $\mathrm{p}^\infty(x)$ is linearly asymptotically stable with respect to perturbing densities in $S(\epsilon)$. □

## 4.5 A Mean Consensus Model

In this section we obtain stability results for a mean consensus MFG model using theorem (4.3.2). The model consists of the problem statement (4.2) with the *nonlocal* consensus cost $q(x, \mathrm{p}(t,x)) = \frac{1}{2}\left(\int (x - x')\mathrm{p}(t,x')\,\mathrm{d}x'\right)^2$ and agents following controlled one dimensional Langevin dynamics (4.1) with quadratic restoring potential $\nu = \frac{1}{2}ax^2$, $a \neq 0$. A MFG model with consensus cost has been previously studied in [22] wherein it is assumed that all agents follow integrator dynamics, that is, the case $a = 0$. Although a more general potential $\nu(x)$ can be treated using the result (4.3.2) to obtain stability results, we choose to present the generalization only to the quadratic potential. This choice allows us to obtain analytical fixed point solutions for the stationary MF system, inspired by related work in [90] where fixed points solutions were found for a different class of MFGs. The linearity in passive agent dynamics also allows for mean consensus, as discussed later in this section.

The MF optimality system for this model consists of the coupled system (4.3, 4.4) wherein $\nu = \frac{1}{2}ax^2$, along with the cost coupling equation

$$q(x, \mathrm{p}(t,x)) = \frac{1}{2}\left(\int (x - x')\mathrm{p}(t,x')\,\mathrm{d}x'\right)^2 \tag{4.30}$$

where $\mathrm{p}(0,x) = \mathrm{p}_0(x)$ is the given initial density of agents, $\int \mathrm{p}(t,x)\mathrm{d}x = 1$ for all $t \geq 0$, $\lim_{|x|\to+\infty} \mathrm{p}(t,x) = 0$ and $\lim_{t\to+\infty} \mathrm{e}^{-\rho t} v(t, x_t) = 0$.

### 4.5.1 Gaussian Stationary Solution

The stationary MF optimality system for this model consists of (4.5, 4.6) wherein $\nu = \frac{1}{2}ax^2$, along with the cost coupling equation

$$q^\infty(x) = \frac{1}{2}\left(\int (x - x')\mathrm{p}^\infty(x')\mathrm{d}x'\right)^2 \tag{4.31}$$

where $\int \mathrm{p}^\infty(x)\mathrm{d}x = 1$ and $\lim_{t\to+\infty} \mathrm{e}^{-\rho t} v^\infty(x_t) = 0$. We denote $\mu^* := \int_{\mathbb{R}} x' \mathrm{p}^\infty(x')\mathrm{d}x'$. In this subsection we will obtain solutions of the form

$$v^\infty(x) = \frac{\eta}{2}x^2 + \beta x + \omega, \tag{4.32}$$

$$\mathrm{p}^\infty(x) = \frac{1}{\sqrt{2\pi s^2}} \mathrm{e}^{-\frac{(x-\mu^*)^2}{2s^2}}, \tag{4.33}$$

to the value and density functions in the coupled optimality system (4.5, 4.6, 4.31). Parameters $\eta, \beta$ and $\omega$ can be obtained by substituting (4.32) into (4.5), using (4.31) and equating coefficients of powers of x:

$$\omega = \frac{1}{\rho}\left(\frac{1}{2}(\mu^*)^2 - \frac{\beta^2}{2R} + \frac{\sigma^2}{2}\eta\right), \tag{4.34}$$

$$\beta = \frac{-\mu^*}{\rho + \frac{\eta}{R} + a}, \tag{4.35}$$

$$\eta^2 + 2R(\rho/2 + a)\eta - R = 0. \tag{4.36}$$

These parameters must satisfy additional conditions related to the validity of the solution ansatz, namely, $s^2 > 0$ and $v^\infty(x) > 0$ for all $x \in \mathbb{R}$. The unique positive solution to the Algebraic Riccati Equation (ARE) (4.36) which permits $v^\infty(x) > 0$ for all $x \in \mathbb{R}$ is

$$\eta = -R\left(\frac{\rho}{2} + a\right) + \sqrt{R^2\left(\frac{\rho}{2} + a\right)^2 + R}. \tag{4.37}$$

Choosing this solution, it is easily verified that $\rho + \frac{\eta}{R} + a > 0$. Equating our stationary density ansatz (4.33) with the unique Gibbs distribution solution (4.8) from lemma 4.3.1 implies

$$\mu^* = \frac{-\beta}{(aR + \eta)}, \tag{4.38}$$

$$s^2 = \frac{\sigma^2}{2(a + \frac{\eta}{R})}. \tag{4.39}$$

Equations (4.35) and (4.38) are compatible only if $\mu^* = 0$ or $\frac{1}{\rho+\frac{\eta}{R}+a} = aR + \eta$. Using the ARE (4.36) it can be verified that the latter condition is equivalent to $a = -\rho$. We conclude that the Gaussian stationary solutions can be categorized into two distinct cases depending upon problem parameters: (1) if $a \neq -\rho$, there exists a unique solution with $\mu^* = 0$ and (2) if $a = -\rho$, there exist a continuum of solutions, since $\mu^* \in \mathbb{R}$ can be chosen arbitrarily. The following assumption is needed to ensure $s^2 > 0$.

**(B1)** $a + \frac{\eta}{R} > 0$ for all $a \neq 0$.

nGiven a value of $a$, we provide the range of control design parameters for which **(B1)** is true in the following lemma, which can be verified by substitution in equation (4.37).

**Lemma 4.5.1.** *Let* $a_{l,u} := \frac{-\rho}{2} \pm \sqrt{\left(\frac{\rho}{2}\right)^2 - \frac{1}{R}}$. *Then **(B1)** holds if either*

- $\rho < \frac{2}{\sqrt{R}}$

  *or*

- $\rho > \frac{2}{\sqrt{R}}$ *and* $a \in (-\infty, a_u) \cup (a_l, +\infty)$.

**Remark 3.** *MFG models with either no quadratic-Gaussian stationary solutions or a continuum of such solutions were studied in [21] and linear-quadratic models [90] in the case of an long-time-average cost functional. Our consensus model has a continuum of such solutions in the case $a = -\rho$. Since the stability analysis for restrictive case is similar to that in [22], we analyze the case $a \neq -\rho$ in what follows.*

We summarize the obtained quadratic-Gaussian solution to the stationary MF system below.

**Lemma 4.5.2.** *Let (**B1**) hold. 1) Case $a \neq -\rho$ : The unique quadratic-Gaussian solution to the stationary MF optimality system (4.5, 4.6) (with $\nu(x) = \frac{1}{2}ax^2, a \neq 0$) is the pair $\left(v^\infty(x) = \frac{\eta}{2}x^2 + \frac{\sigma^2\eta}{2\rho}, \mathrm{p}^\infty(x) = \frac{1}{\sqrt{2\pi s^2}}\mathrm{e}^{-\frac{x^2}{2s^2}}\right)$ where $(\eta, s)$ are defined by (4.37, 4.39). Furthermore, $q^\infty(x) = \frac{1}{2}x^2$.*

*2) Case $a = -\rho$ : For each $\mu^* \in \mathbb{R}$, there exists a pair $(v^\infty(x), \mathrm{p}^\infty(x))$ given by equations (4.32, 4.33) that is a solution to the stationary MF optimality system (4.5, 4.6) (with $\nu(x) = \frac{1}{2}ax^2, a \neq 0$). The parameters $(\omega, \beta, \eta, s)$ are given by equations (4.34, 4.35, 4.37, 4.39). Furthermore, $q^\infty(x) = \frac{1}{2}(x - \mu^*)^2$.*

*Proof.* In both cases, $q^\infty(x) = \frac{1}{2}(x - \mu^*)^2$ follows from equation (4.31) and assumption **(B1)** ensures that $s^2 > 0$ in the unique Gaussian Gibbs distribution (4.33) corresponding to the quadratic value function (4.32).

In case 1, the solution to the stationary value function is obtained by substituting $\mu^* = 0$ in equations (4.34), (4.35). This completes the first part of the proof. In case 2, for a given value of $\mu^* \in \mathbb{R}$, the solution to the value function maybe obtained similarly to the previous case.

From the expression for the Gibbs distribution (4.8) and equation (4.33) we have $\frac{2}{\sigma^2}v^\infty(x) = \frac{(x-\mu^*)^2}{s^2} \geq 0$, which concludes the proof. □

### 4.5.2 Linear Stability

We define a Hilbert space and a class perturbations in it, for which we show stability.

**Definition 4.5.1.** *Denote Gaussian density $\mathrm{p}_G^\infty(x) := \frac{1}{\sqrt{2\pi s^2}}\mathrm{e}^{-\frac{(x-\mu^*)^2}{2s^2}}$ with $\mu^* \in \mathbb{R}$, $s^2 > 0$. Denote by $\mathcal{H}_G$ the Hilbert space $L^2(\mathrm{p}_G^\infty(x)\mathrm{d}x; \mathbb{R})$. The class of mass preserving density perturbations is defined as $S_0 := \left\{q(x) \in \mathcal{H} \middle| \langle 1, q(x)\rangle_{\mathcal{H}_G} = 0\right\}$. The class of mass and mean preserving density perturbations is defined as $S_1 := \left\{q(x) \middle| \langle 1, q(x)\rangle_{\mathcal{H}_G} = 0, \langle x, q(x)\rangle_{\mathcal{H}_G} = 0\right\}$.*

The class of initial perturbed densities and linear asymptotic stability can be defined analogously from the previous section by replacing $\mathrm{p}^\infty(x)$ by $\mathrm{p}_G^\infty(x)$ in definition 4.4.2.

The lemma below follows from theorem 4.3.2 and Taylor expansion of $q$ in (4.30) around the fixed point.

**Lemma 4.5.3.** *Let $\nu(x) = \frac{1}{2}ax^2$, $a \neq 0$. If **(B1)** holds, and ($v^{\infty}(x)$, $\mathrm{p}^{\infty}(x)$, $q^{\infty}(x)$) given by lemma 4.5.2 is a stationary solution to the nonlinear MF system (4.5, 4.6, 4.31) then the linearization of the system around this solution for all $(t, x) \in [0, +\infty) \times \mathbb{R}$ is given by (4.10,4.11) and*

$$\tilde{q}(x; \tilde{\mathrm{p}}(t,x)) = -\left(x - \mu^*\right)\left(\int_{\mathbb{R}} x' \mathrm{p}^{\infty}(x')\tilde{\mathrm{p}}(t,x')\mathrm{d}x'\right), \tag{4.40}$$

*where $\tilde{\mathrm{p}}(0,x)$ is given, $\int_{\mathbb{R}} \mathrm{p}^{\infty}(x)(1+\epsilon\tilde{\mathrm{p}}(t,x))\mathrm{d}x = 1$ for all $t \geq 0, \epsilon > 0$, $\lim_{|x|\to+\infty} \tilde{\mathrm{p}}(t,x) = 0$ for all $t \geq 0$ and $\lim_{t\to+\infty} \mathrm{e}^{-\rho t}\tilde{v}(t, x_t) = 0$.*

We now state eigen properties of the generator ([91, 83]) of the controlled process for the consensus model. We define normalized Hermite polynomials $\{H_n(x)\}_{n\in\mathbb{W}}$ for the space $L^2(\mathrm{p}_G^{\infty}\mathrm{d}x; \mathbb{R})$ as $H_n(x) = s^n \frac{1}{\sqrt{n!}}(-1)^n \mathrm{e}^{\frac{(x-\mu^*)^2}{2s^2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-\frac{(x-\mu^*)^2}{2s^2}}$. These polynomials with $n \geq 0$, form a countable orthonormal basis of the space $\mathcal{H}_G$. $\{H_n(x)\}_{n\in\mathbb{W}}$ are eigenfunctions of the operator $\mathcal{L}$ wherein $\nu(x) = \frac{1}{2}ax^2$, with the [22] eigenproperty $\mathcal{L}H_n = -\frac{\sigma^2}{2s^2}nH_n = -(a + \frac{\eta}{R})nH_n$. The following condition is needed for stability of the consensus model.

**(B2)** $a(a+\rho) \geq 0$.

Note that this assumption is true if and only if $a \in (-\infty, -\rho] \cup (0, +\infty)$, recalling that $a \neq 0$. Denote the matrix associated with the MF system for the consensus model $B_n := \begin{bmatrix} \frac{\sigma^2 n}{2s^2} + \rho & s^2\delta(n-1) \\ \frac{-n}{s^2 R} & \frac{-\sigma^2 n}{2s^2} \end{bmatrix}$.

**Lemma 4.5.4.** *Let (**B1**, **B2**) hold. Then, for all $n \geq 2$ the eigenvalues of $B_n$, $\lambda_n^{1,2} = \frac{\rho}{2} \pm \sqrt{\left(\frac{\rho}{2}\right)^2 + \frac{\sigma^2 n}{2s^2}\left(\frac{\sigma^2 n}{2s^2} + \rho\right)} = \{-\frac{\sigma^2 n}{2s^2}, \frac{\sigma^2 n}{2s^2} + \rho\} = \{-\left(a + \frac{\eta}{R}\right)n, \left(a + \frac{\eta}{R}\right)n + \rho\}$ are real, distinct and ordered $\lambda_n^1 < 0 < \lambda_n^2$. Furthermore, the eigenvalues of $B_1$ denoted $\lambda_1^{1,2}$ are real, distinct and ordered $\lambda_1^1 < 0 < \lambda_1^2$ if $a \in (-\infty, -\rho) \cup (0, +\infty)$ and $\lambda_1^{1,2} = \{0, \rho\}$ if $a = -\rho$.*

On applying the ARE (4.36), we see that the eigenvalues $\lambda_1^{1,2} = \frac{\rho}{2} \pm \sqrt{\left(\frac{\rho}{2}\right)^2 + \frac{\sigma^2}{2s^2}\left(\frac{\sigma^2}{2s^2} + \rho - \frac{2s^2}{\sigma^2 R}\right)} = \{-a, a+\rho\}$. Choosing to denote the lower of the eigenvalues by $\lambda_1^1$, we see that $\lambda_1^1 = -a$ if $a \in (0, +\infty)$ and $\lambda_1^1 = a + \rho$ if $a \in (-\infty, -\rho)$.

Spectral properties of the perturbation MFG system obtained in this section allow us to generalize the methods in [22] (applied to integrator agent dynamics) to prove stability of fixed points for MFG with linear Langevin agent passive dynamics. In the following theorem, we show linear stability of unique zero mean stationary density ($\mu^* = 0$, corresponding to $a \neq -\rho$) with respect to mass preserving density perturbations ($\tilde{\mathrm{p}}(0,x) \in S_0$).

**Theorem 4.5.5.** *n Let $\nu(x) = \frac{1}{2}ax^2, a \notin \{\{0\}, \{-\rho\}\}$. Let (**B1, B2**) hold. Let $(v^\infty(x), \mathrm{p}^\infty(x), q^\infty(x))$ given by lemma 4.5.2 be a stationary solution to the MF system (4.5, 4.6, 4.31). If perturbation $\tilde{\mathrm{p}}(0,x) \in S_0$, and $\{v_n, p_n\}_{n\geq 0}$ are determined by*

$$\begin{bmatrix} \dot{v_n} \\ \dot{p_n} \end{bmatrix} = B_n \begin{bmatrix} v_n \\ p_n \end{bmatrix}, \quad n \geq 0, \tag{4.41}$$

*then $\tilde{v}(t,x) = \sum_{n=0}^{+\infty} v_n(t) H_n(x)$, $\tilde{\mathrm{p}}(t,x) = \sum_{n=0}^{+\infty} p_n(t) H_n(x)$ are unique $\mathcal{H}_G$ solutions to the perturbation MF system (4.10, 4.11, 4.40). Moreover, the steady state density $\mathrm{p}^\infty(x) = \mathrm{p}_G^\infty(x)$ is linearly asyptotically stable with respect to $S(\epsilon)$. Furthermore, $\tilde{\mathrm{p}}(t,x) = p_1(0)\mathrm{e}^{\lambda_1^1 t} H_1(x) + \sum_{n=2}^{+\infty} p_n(0)\mathrm{e}^{-\frac{-\sigma^2 n}{2s^2}t} H_n(x)$, $\tilde{q}(x; \tilde{\mathrm{p}}(t,x)) = -s^2 p_1(0) H_1(x) = 0$, and $\tilde{v}(t,x) = \frac{s^2 p_1(0)}{\frac{\sigma^2}{2s^2} + \rho - \lambda_1^1} \mathrm{e}^{\lambda_1^1 t} H_1(x)$ where $\lambda_1^1$ is defined in lemma 4.5.4.*

*Proof.* We construct $\mathcal{H}_G$ solutions of form $\tilde{v}(t,x) = \sum_{n=0}^{+\infty} v_n(t) H_n(x)$, $\tilde{\mathrm{p}}(t,x) = \sum_{n=0}^{+\infty} p_n(t) H_n(x)$ to the perturbation MF system (4.10, 4.11, 4.40) and show that they are unique. Since $\tilde{\mathrm{p}}(0,x) \in \mathcal{H}_G$ we have the unique representation $\tilde{\mathrm{p}}(0,x) = \sum_{n=0}^{+\infty} p_n(0) H_n(x)$.

Since $H_1(x) = \frac{x - \mu^*}{s}$, from (4.40) we note that $\tilde{q}(x; \tilde{\mathrm{p}}(t,x)) = -sH_1(x) \sum_{n=0}^{+\infty} p_n(t) \langle \tilde{\mathrm{p}}(t,x), x \rangle = -s^2 p_1(t) H_1(x)$. Substituting the selected form of the solutions into the perturbation system (4.10, 4.11, 4.40) and using the eigen property of the operator yields the ODEs (4.41).

***(i) Case*** $n=0$: Since $H_0(x)=1$, and $\tilde{\mathrm{p}}(0,x)\in S_0$, we have $p_0(0)=\left\langle \tilde{\mathrm{p}}(0,x),1\right\rangle_{\mathcal{H}_G}=0$. Therefore, from the ODE system (4.41) and matrix $B_n$, we have $\dot{p}_0=0$ and $\dot{v}_0=\rho v_0$ implying $p_0(t)=0$ and $v_0(t)=v_0(0)\mathrm{e}^{\rho t}$ for all $t>0$. So, the only solution allowing $\tilde{v}(t,x)\in\mathcal{H}_G$ is $v_0(t)=0$.

***(ii) Case*** $n=1$: In this case, from (4.39),

$$B_1=\begin{bmatrix}\left(a+\frac{\eta}{R}+\rho\right) & s^2\\ -\frac{1}{s^2R} & -\left(a+\frac{\eta}{R}\right)\end{bmatrix}. \tag{4.42}$$

The assumptions imply $a\in(-\infty,-\rho)\cup(0,+\infty)$. Hence, from lemma 4.5.4, the eigenvalues $spec(B_1)=\lambda_1^{1,2}$ are ordered $\lambda_1^1<0<\lambda_1^2$. Consider the finite time boundary conditions $p_1(0),v_1(T)$ to ODE system in this case. We may write

$$\begin{bmatrix}v_1(t)\\ p_1(t)\end{bmatrix}=C_1^{1,T}\mathrm{e}^{\lambda_1^1t}\begin{bmatrix}1\\ e_1^1\end{bmatrix}+C_2^{1,T}\mathrm{e}^{\lambda_1^2t}\begin{bmatrix}1\\ e_1^2\end{bmatrix} \tag{4.43}$$

with the eigenvector components $e_1^{1,2}=\frac{1}{s^2}\left(\frac{\sigma^2}{2s^2}+\rho-\lambda_1^{1,2}\right)$. Boundary conditions give us $C_1^{1,T}=\frac{(e_1^2/e_1^1)v_1(T)\mathrm{e}^{-\lambda_1^2T}-p_1(0)/e_1^1}{(e_1^2/e_1^1)\mathrm{e}^{(\lambda_1^1-\lambda_1^2)T}-1}$, and $C_2^{1,T}=\frac{-\mathrm{e}^{(\lambda_1^1-\lambda_1^2)T}p_1(0)/e_1^1-\mathrm{e}^{-\lambda_1^2T}v_1(T)}{(e_1^2/e_1^1)\mathrm{e}^{(\lambda_1^1-\lambda_1^2)T}-1}$. Note that if $\tilde{v}(t,x)\in\mathcal{H}$ then $\lim_{t\to+\infty}|v_n(t)|<+\infty$ for all $n\geq 0$. It is also known that $|p_n(0)|<+\infty$. Since $\lambda_1^1<0<\lambda_1^2$ we observe that $\mathrm{e}^{(\lambda_1^1-\lambda_1^2)T},\mathrm{e}^{-\lambda_1^2T}\to 0$ as $T\to+\infty$ so that in the limit, $C_1^{1,T}\to p_1(0)/e_1^1$ and $C_2^{1,T}\to 0$. Therefore we have the unique solutions $v_1(t)=(p_1(0)/e_1^1)\mathrm{e}^{\lambda_1^1t}$ and $p_1(t)=p_1(0)\mathrm{e}^{\lambda_1^1t}$. Therefore, if $\tilde{\mathrm{p}}(0,x)\in S_1$ so that $p_1(0)=\langle\tilde{\mathrm{p}}(0,x),H_1(x)\rangle=0$ then $v_1(t)=0,p_1(t)=0$ for all $t\geq 0$.

***(iii) Case*** $n\geq 2$: In this case, from the ODE system we have

$$\begin{bmatrix}\dot{v}_n\\ \dot{p}_n\end{bmatrix}=\begin{bmatrix}\frac{\sigma^2n}{2s^2}+\rho & 0\\ -\frac{n}{s^2R} & -\frac{\sigma^2n}{2s^2}\end{bmatrix}. \tag{4.44}$$

Therefore $v_n(t)=v_n(0)\mathrm{e}^{(\frac{\sigma^2n}{2s^2}+\rho)t}$, for which the unique solution allowing $\tilde{v}(t,x)\in\mathcal{H}_G$ for all $t \geq 0$ is $v_n(t) = 0$. Therefore $p_n(t) = p_n(0)\mathrm{e}^{-\frac{nt}{s^2 R}}$ is the unique solution to the ODE on $p_n$.

In the preceeding discussion we have shown that the unique $\mathcal{H}_G$ solution to the perturbation system has the properties $\{v_0(t) = 0,\ v_1(t) = \frac{s^2 p_1(0)}{\frac{\sigma^2}{2s^2}+\rho-\lambda_1^1}\mathrm{e}^{\lambda_1^1 t},\ v_n(t) = 0$ for all $n \geq 2,\}$, and $\{\ p_0(t) = 0,\ p_1(t) = p_1(0)\mathrm{e}^{\lambda_1^1 t}$ and $p_n(t) = p_n(0)\mathrm{e}^{\frac{-nt}{s^2 R}}$ for all $n \geq 2\}$. Therefore using Parseval's theorem $||\tilde{\mathrm{p}}(t,x)||_{L^2(\mathrm{p}^\infty(x)\mathrm{d}x;\mathbb{R})} = \left(p_1^2(0)\mathrm{e}^{2\lambda_1^1 t} + \sum_{n=2}^{+\infty} p_n^2(0)\mathrm{e}^{-\frac{2nt}{s^2 R}}\right)^{\frac{1}{2}}$ where $\lambda_1^1 < 0$, and the Lebesgue dominated convergence theorem, we have that $\mathrm{p}_G^\infty(x)$ is linearly asymptotically stable with respect to perturbing densities in $S(\epsilon)$. □

**Remark 4.** *For the case $a = -\rho$, there exists a continuum of stationary solutions, similar to the models considered in ([21], [22]). Stability of the mean consensus model for $a = -\rho$ with mass and mean preserving perturbations ($\tilde{\mathrm{p}}(0,x) \in S_1$) can be proved by following the approach in [22], or using contraction mapping arguments ([15], [18], [77]).*

We state a theorem regarding the mean consensus property [22] of the steady state MFG control law. Let us denote a finite set of agents $\mathcal{A} := \{x^i\}_{1\leq i\leq N}$, identified by their individual states $x^i$ with individual dynamics given by equation (4.1). The set of agents $\mathcal{A}$ is said to have the mean consensus property if $\lim_{t\to+\infty} |\mathbb{E}[x_t^i - x_t^j]| = 0$ for any two agents $x^i, x^j \in \mathcal{A}$. Let us denote a finite set of agents$\mathcal{A} = \{x_t^i\}_{1\leq i\leq N}$, identified by their individual states $x_t^i$ with individual dynamics given by equation (4.1), $\nu = \frac{1}{2}ax^2$, $a \neq 0$. We say that $\mathcal{A}$ has the initial mean consensus property if $\lim_{t\to+\infty} |\mathbb{E}[x_t^i - x_t^j]| = 0$ for any two agents $x_t^i, x_t^j \in \mathcal{A}$ under the action of the MF control law given by the MFG (4.2) with the consensus cost (4.30). The assumption below is required to prove mean consensus for a set of agents in our consensus model.

**(B3)** $\sup_{1\leq i\leq N} \mathbb{E}[|x_0^i|^2] < +\infty$ for the set $\mathcal{A}$.

**Theorem 4.5.6.** *Let (**B1**, **B2**, **B3**) hold. Let $(v^\infty, \mathrm{p}^\infty)$ be the steady state solutions to the optimality system (4.5, 4.6) given in lemma* (4.5.2). *The steady state MF control law $u^\infty(x) = -\frac{1}{R}\partial_x v^\infty(x)$ applied to a set of agents $\mathcal{A}$, in the MFG model given by equations*

*(4.1,4.2,) with* $\nu(x) = \frac{1}{2}ax^2$ *and consensus cost* (4.30) *results in a mean consensus with individual asymptotic variance* $s^2 = \frac{\sigma^2}{2(a+\frac{\eta}{R})^2}$.

*Proof.* Since $\partial v^\infty(x)/R = \eta x/R + \beta/R$ and from the properties of the solution given in lemma (4.5.2), the controlled individual dynamics for agent $x^i \in \mathcal{A}$ can be obtained from equation (4.1) as $\mathrm{d}x_t^i = \mathrm{d}(x_t^i - \mu^*) = -(a+\frac{\eta}{R})(x_t^i - \mu^*)\mathrm{d}t + \sigma \mathrm{d}w_t$. This gives the stochastic integral solution $x_t^i = \mu^* + \mathrm{e}^{-(a+\frac{\eta}{R})t}(x_0^i - \mu^*) + \sigma \int\limits_0^t \mathrm{e}^{-(a+\frac{\eta}{R})(t-s)}\mathrm{d}w_s$.

$$x_t^i = \mu^* + \mathrm{e}^{-(a+\frac{\eta}{R})t}(x_0^i - \mu^*) + \sigma \int\limits_0^t \mathrm{e}^{-(a+\frac{\eta}{R})(t-s)}\mathrm{d}w_s. \tag{4.45}$$

Using assumption **(B4)** and Jensen's inequality we have,
$\sup_{1\le i\le N}(\mathbb{E}[|x_0^i - \mu^*|]) \le \sup_{1\le i\le N}(\mathbb{E}[|x_0^i| - \mu^*]) \overset{Jensen's}{\le} \sqrt{k} - \mu^* < +\infty$. Taking expectation along with applying assumption **(B2)** and using the Itó isometry [82] to get the individual asymptotic variance we have $\lim\limits_{t\to+\infty} \mathbb{E}[x_t^i] = \mu^*$ and $\lim\limits_{t\to+\infty} \mathbb{E}[(x_t^i - \mu^*)^2] = \lim\limits_{t\to+\infty} \sigma^2 \int\limits_0^t \mathrm{e}^{-2(a+\frac{\eta}{R})(t-s)}\mathrm{d}s = \frac{\sigma^2}{2(a+\frac{\eta}{R})}$.

$$\lim_{t\to+\infty} \mathbb{E}[x_t^i] = \mu^* \tag{4.46}$$

$$\lim_{t\to+\infty} \mathbb{E}[(x_t^i - \mu^*)^2] = \lim_{t\to+\infty} \sigma^2 \int\limits_0^t \mathrm{e}^{-2(a+\frac{\eta}{R})(t-s)}\mathrm{d}s = \frac{\sigma^2}{2(a+\frac{\eta}{R})}. \tag{4.47}$$

□

A consequence of the continuum of solutions (remark 3) to the MF system for our consensus model is that the mean of the stationary density, $\mu^*$, is selected uniquely as the mean of the initial density of agents. This can be observed using the fact that $p_0(t) = p_1(t) = 0$ and $x = \mu^* H_0(x) + sH_1(x)$ as $\int x\mathrm{p}^\infty(x)(1+\tilde{\mathrm{p}}(0,x))\mathrm{d}x = \mu^* + \sum\limits_{n=0}^{+\infty} p_n(t)\langle H_n(x), x\rangle_{\mathcal{H}_G} = \mu^* + \sum\limits_{n=2}^{+\infty} p_n(t)\langle H_n(x), \mu H_0(x) + sH_1(x)\rangle_{\mathcal{H}_G} = \mu^*$.

We simulate an example of the mean consensus of agents with linear passive drift in the dynamics. The agents attain mean consensus asymtotically with good agreement between theoretical and numerical values of the asymptotic variance.

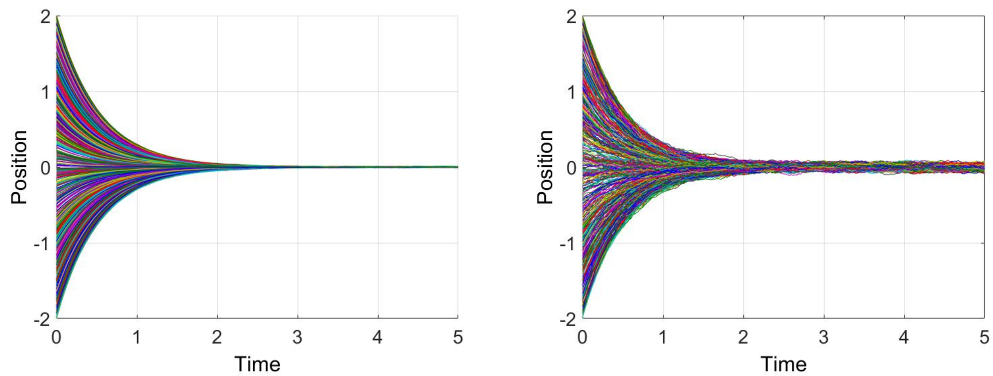



Figure 4.2: Stochastic trajectories and correspoding means of 500 agents starting from an initial uniform distribution. At large time theoretical and numerical values of variance are in close agreement

Theorems (4.4.4,4.5.5) show that in the population model (with nonlinear agent dynamics) as well as the consensus model (with linear agent dynamics , $a \neq -\rho$), the optimal MF control law $u^*(t,x) = u^\infty(x) - \partial_x \tilde{v}(t,x)/R$ is in general time-varying, and hence different from the static steady controller $u^\infty(x) = -\partial_x v^\infty(x)/R$. In the next section we study the local stabilizing property of the static steady MF controller with respect to small $S_0$ perturbations in the steady state density, for both MFG models with nonlinear Langevin agent dynamics and general cost functions.

As indicated in remark 3, a consequence of the continuum of solutions to the MF system for our consensus model is that the mean of the stationary density, $\mu^*$, is selected uniquely as the mean of the initial density of agents. This can be observed using the fact that $p_0(t) = p_1(t) = 0$ and $x = \mu^* H_0(x) + sH_1(x)$ as $\int x\mathrm{p}^\infty(x)(1+\tilde{\mathrm{p}}(0,x))\mathrm{d}x = \mu^* + \sum_{n=2}^{+\infty} p_n(t)\big\langle H_n(x), \mu H_0(x) + sH_1(x)\big\rangle_{\mathcal{H}_G} = \mu^*$. The previous theorem shows that the value function and hence the optimal control does not change under small $S_1$ density perturbations. Further it stabilizing with respect to small $S_1$ perturbations in the density. That is, since $\tilde{v}(t,x) = 0$ under small $S_1$ perturbations, the MF control law in this case is

$u^*(t,x) = u^\infty(x) = -\partial_x v^\infty(x)/R.$

### 4.6 Steady Controller: Static State Feedback

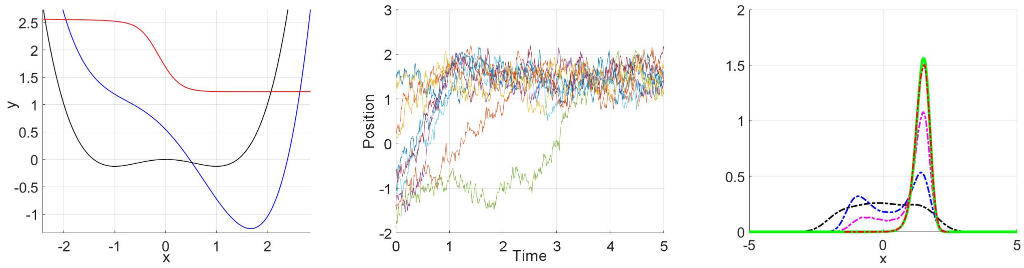
We consider the stability of a population of agents in a MFG, under the action of static state feedback provided by the steady state MFG solution. Let $(v^\infty, \mathrm{p}^\infty)$ be a fixed point for the MF system (4.5, 4.6). Consider a perturbed density of agents $\mathrm{p}^\infty(1+\epsilon\tilde{\mathrm{p}})$ as before. The static feedback MF control law $u^\infty(t,x) = -\partial_x(v^\infty(x))/R$ for agents governed by (4.1) is said to be locally *stabilizing* for a steady state density $\mathrm{p}^\infty(x)$, if the density perturbation $\tilde{\mathrm{p}}(t,x)$ governed by (4.11) with $\tilde{v}(t,x) \equiv 0$, decays to zero.

From equation (4.17), the perturbation dynamics under the static feedback are given by $\partial_t\tilde{\mathrm{p}} = \mathcal{L}\tilde{\mathrm{p}}$. Local stability therefore depends only on the eigen properties of the generator $\mathcal{L}$. Assuming (**A1**, **A2**) hold, theorems 4.4.1 and 4.4.2 imply non-negativity of spectrum of $\mathcal{L}$, which in turn yields stability w.r.t. density perturbations in $S_0$. Notice that this result is independent of the cost function $q(x,\mathrm{p})$. Therefore, the static feedback under the steady controller is locally stabilizing.

We demonstrate local linear stability property under decentralized static state feedback in two 1D numerical examples. We consider the example of a bistable Langevin potential, $\nu(x) = \alpha(\frac{x^4}{4} - \frac{x^2}{2})$, $\alpha > 0$, for both models. Open loop dynamics (4.1) under this potential would cause agents to *fall* into either one of the wells and exhibit a bimodal distribution at infinite time.

We use Chebfun [92] to solve for nsteady states of the MF system (4.5, 4.6) [77]. Monte Carlo simulations are performed for Langevin dynamics (4.1) using the nonlinear static feedback controller. Trajectories for $N = 500$ agents are simulated with $100$ stochastic realizations each. We observe an initial distribution of agents decay to the steady state density over the total simulation time $T$, in both cases.

In the population model, a combined quadratic state and log density cost $q(t,x) = \frac{1}{2}Q(x-1)^2 - \ln\mathrm{p}(t,x)$ is designed. This models a population of agents with a tendency to

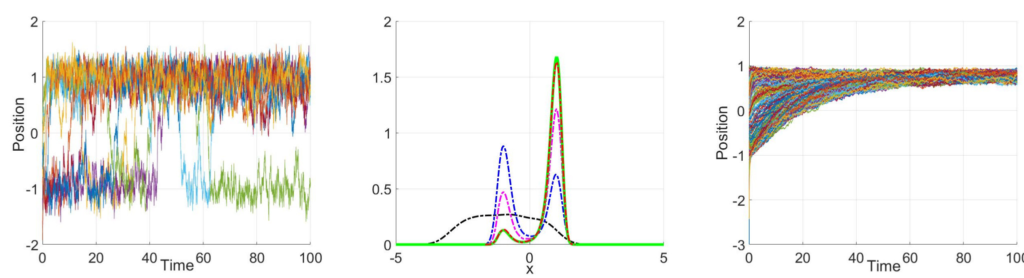


(a) (b) (c)

Figure 4.3: (a) Bistable potential (black), $v^{\infty}(x)$ for population model (blue) and the consensus cost case (red). Population model, $\alpha = 0.5$, $\sigma = 1$, $\rho = 5$, $Q = 10$ and $R = 0.5$: (b) Stochastic paths for ten agents (c) Evolution of density at various times, $t = 0$ (black), $t = T/5$ (blue), $t = 2T/5$ (pink), $t = T$ (red) to the PDE solution (green)

imitate each other while moving towards the preferred state $x = 1$. Initial states of agents are sampled from a uniform density over $[-2, 2]$. We observe that for the log density cost, in Fig. 4.3b that some agents which are initially stuck in the potential well centered at $x = -1$ are able to escape it, to the preferential well centered at $x = 1$, given sufficient time. In figure 4.3c, we see that at $t = T/5$ the dynamics are dominated by the bistable potential but as time increases $t = 2T/5, t = T$, the density becomes unimodal with a mean close to the preferred state $x = 1$. Finally the stationary density from the PDE computation is achieved by the agents at $t = T$.

In the consensus model case, the cost (4.30) is used in conjunction with the long-time-average utility (4.7). Analytical stability results in the consensus cost case with the bistable potential, were presented by the authors in [77]. However, those results pertain to local



(a) (b) (c)

Figure 4.4: Consensus cost model with long-time-average utility, $\alpha = 1.5$, $\sigma = 0.5$, and $R = 235$: (a) Stochastic paths for ten agents (b) Evolution of density at various times, $t = 0$ (black), $t = T/5$ (blue), $t = 2T/5$ (pink), $t = T$ (red) to the PDE solution (green) (c) Stochastic means of all agents

stability of the optimal (time-varying) MFG control, in contrast with the decentralized static MF control considered here. Note that there are two steady state densities, with means $\mu^* = \pm 1$. We use the control law corresponding to the right well ($\mu^* = 1$). Initial states of agents are sampled from a uniform density over $[-3, 1]$. Since the initial density has a negative mean, at $t = T/5$ we notice that there are more agents in the left well. However as time increases, we see that more agents migrate into the right well under the control. At $t = T$ the PDE solution to the stationary density which is slightly bimodal, is recovered by the Monte Carlo simulation. Although we are using the consensus cost, a high control cost causes some agents to be in the well centered at $x = -1$. Most agents are seen to escape from the left well and move into the right well in figure 4.4a. However, due to the high noise intensity combined with low control authority, some agents are seen to move in the opposite direction as well. Finally, from stochastic means in Fig. 4.4c we see that unlike the linear case where mean consensus is guaranteed (theorem 4.5.6), mean consensus is not achieved in the case with nonlinear passive dynamics.

## 4.7 Conclusions

In this chapter, we have studied MFGs for agents with multidimensional nonlinear Langevin dynamics, and provided a framework for closed-loop stability analysis of fixed points in such systems. The key idea is to use the detailed balanced property of the generator to characterize the eigenvalue spectrum of perturbation forward-backward system, hence extending existing methods that deal with integrator agent dynamics. While we demonstrate this approach in the discounted cost case, it is also applicable to MFGs using the long-time-average cost functional. Using the presented approach, conditions on the stationary solutions and explicit control design constraints have been obtained for guaranteeing stability in a population distribution and a mean consensus model. We also provide a mean consensus result for the case where the Langevin potential is quadratic, with individual asymptotic variance depending on the linear drift.

It is also shown that under certain conditions on the stationary solution, the steady MF controller providing decentralized static feedback is locally stabilizing. We illustrate this fact by Monte Carlo simulations for population and consensus cost models with non-Gaussian steady state behaviour.

The most general class of (uncontrolled) diffusions which possess the detailed balance property are *reversible diffusions* with possibly multiplicative noise. Hence, the approach presented here can be extended to provide stability results for the corresponding MFG models. Generalizing our results to second order Langevin systems will be a topic of future work. Such MFG systems must be treated separately, since the concerned closed loop generator in that case is a combination of a Liouville operator and generator $\mathcal{L}$ in this chapter.

# CHAPTER 5

# MODELING FLOCKS USING MEAN FIELD GAMES

The analysis of emergent behavior in a large population of dynamic agents is a classical topic. However the design of desired macroscopic behavior in such systems, including in bio-physics, remains a challenge. Such systems are often studied using continuum models, involving empirically derived systems of nonlinear partial differential equations that govern the distribution of agents in the phase space. The various terms in these equations represent intrinsic dynamics of the agents, mutual attraction and/or repulsion, and noise. An important class of such models concern flocking, both in nature, and engineering applications such as bio-inspired control of multi-agent robotics, traffic modeling, power-grid synchronization etc. We take a mean-field game approach to derive a control system that mimics the behavior of one such class of models in the setting of non-cooperative agents. A mean-field game is a coupled system of partial differential equations that govern the state and optimal control distributions of a representative agent in a Nash equilibrium with the population. Using a linear stability analysis, in this section, we recover phase transitions that have been observed in the corresponding empirical model, as well as find some new ones, as the control penalty is changed.

## 5.1 Introduction

Continuum models of large populations of interacting dynamic agents are popular in mathematical biology[93], and also have been employed in numerous applications such as multi-agent robotics [94], finance [95] and traffic modeling [96]. The aim of such models is to accurately represent the macroscopic dynamics of the population, and its dependence on parameters. Typically, such models are derived by starting with an empirical dynamical system for a representative agent. This system typically involves the intrinsic dynamics of the agent, a coupling function[97] describing its interaction with the population, and noise. From this single agent dynamical system, a continuum description is obtained by deriving a macroscopic equation for the distribution of agents in the phase space. We call this class of models *uncontrolled*.

An alternative way of deriving continuum models of collective behavior is via a corresponding variational principle. In this approach, the dynamical system for a representative agent includes its intrinsic dynamics, a control term and noise. The unknown control term is obtained as a solution to an optimization problem. Within this variational (or optimization) framework for large populations, there are multiple classes of modeling strategies [98]. If one takes a centralized global optimization viewpoint, the corresponding problem is that of *mean-field control*, i.e. it is assumed that each agent is being controlled by a central entity whose goal is to optimize a macroscopic cost function[99] that includes interaction among the population. In a distributed setting, there is no central entity, and the agents can either be *cooperative* or *non-cooperative*. In the former case, each agent choses its control to optimize a global sum of cost functions of the population.

On the other hand, in the *non-cooperative* mean field setting that we are interested in, each agent optimizes only its individual cost function. This cost function involves coupling with the population solely via a mean-field term. This is the setting of mean-field games (MFG)[100, 16, 15]. In this setting, a Hamilton-Jacobi-Bellman (HJB) equation (posed backward in time) characterizes the optimal feedback control for a representative agent under the assumption that the (cost) coupling function depends only on its own state, and possibly time. A Fokker-Planck (FP) equation governs the evolution of agent density in phase space. A consistency principle [15] requires that the coupling function used in the agent HJB equation is reproduced as its own average over the continuum of agents. Under fairly general conditions, solutions to MFG model can be shown to possess $\epsilon$-Nash property, i.e., unilateral benefit of any deviation from the computed control policy by a single agent vanishes rapidly as the population becomes large.

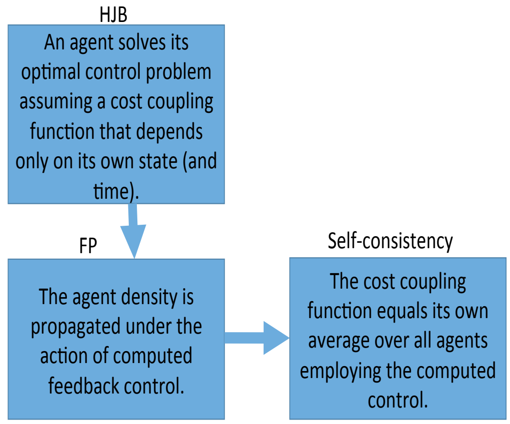



Figure 5.1: The MFG framework

The classical (uncontrolled) Cucker-Smale (CS) flocking model[101] describes a system of finite population of coupled agents with trivial intrinsic dynamics, moving solely under the influence of an alignment force, and noise. This was followed by several continuum descriptions[102, 103], and was recently generalized to a continuum model with self-propulsion effects in the homogeneous case [104] (i.e., assuming spatial homogeneity). This latter generalization results in existence of non-zero mean velocity distribution resulting from symmetry breaking, a wide range of 'disordered' states consisting of multiple flocks, and other phase transitions.

A MFG model for a continuum of coupled Kuramoto oscillators[105] was described in a seminal work [106] that influences the development in the current chapter. Building upon this work, a MFG model for the classical inhomogeneous CS was then proposed[24]; the stability analysis was partially addressed. This was followed by a homogeneous flocking MFG model for coupled agents with trivial intrinsic dynamics, along with linear and nonlinear stability analysis[22]. Also of interest is an approach [25] where agents apply a gradient descent rather than solve an HJB equation, since the Nash equilibria of the MFG are recovered under certain conditions using this approach.

The contributions of this chapter are as follows. We formulate a MFG model for homogeneous flocking of agents driven by self-propulsion and noise. In contrast to the earlier work on homogeneous MFG model with trivial intrinsic dynamics [22], this model exhibits phase transitions (bifurcations) that mimic those present in the corresponding uncontrolled model [104]. We generalize the stability analysis developed in previous MFG models [18, 22, 15, 21] to agents with gradient nonlinear dynamics, and employ a method used to study reaction-diffusion equations[107] to derive a semi-analytical stability criterion. Besides qualitatively explaining the phase transition phenomena, quantitative results useful in control design are obtained from the numerical analysis. Decreasing the control control penalty below a threshold causes the zero mean velocity steady state of the MFG model to lose stability via pitchfork bifurcation [108]. This results in a pair of stable steady states with non-zero mean velocity. If the control is made even cheaper, a new stable regime (nonexistent in the uncontrolled model) emerges for zero mean velocity steady states in the small noise case via a subcritical pitchfork bifurcation. Results of this section were published in [77].

### 5.2 Uncontrolled formulation

We briefly review here the uncontrolled formulation from Ref. [104] which provides a homogeneous model for CS flocking with self-propulsion. Consider a population of $N$ agents moving in phase space ($(q,p) \in \mathbb{R}^2$), where each agent is acted upon by a gradient self-propulsion term, a CS coupling force with localization kernel $K$ in position space that aligns the agents' velocity with the neighbors, and noise. The dynamics for $i$th agent are

$$dq_i = p_i dt,$$
$$dp_i = a(p_i)dt + F(q_i, p_i, q_{-i}, p_{-i})dt + \sigma d\omega_i,$$

where $a(p_i) = -\partial_p U(p_i)$, $U(p_i) = \alpha(\frac{p_i^4}{4} - \frac{p_i^2}{2})$, $F(q_i, p_i, q_{-i}, p_{-i}) = \dfrac{1}{N}\dfrac{\sum_{j=1}^N K(q_i, q_j)(p_j - p_i)}{\sum_{j=1}^N K(q_i, q_j)}$, $\sigma > 0$ is the noise intensity, $\alpha > 0$ defines the strength of the self-propulsion term, $K(q, q') = K(q', q) \geq 0$, and $K(q, q) = 1$, $q_{-i} = \{q_1, .., q_{i-1}, q_{i+1}..\}$, $p_{-i} = \{p_1, .., p_{i-1}, p_{i+1}..\}$.

In the continuum limit ($N \to \infty$), the agent density $f(q, p, t)$ in phase space is governed by

$$\partial_t f + \partial_q(pf) + \partial_p(a(p)f + F[f]f) = \frac{\sigma^2}{2}\partial_{pp} f,$$

where $F[f](q,p,t) = (\bar{p} - p)$ and,

$$\bar{p}(q,t) = \frac{\int\int K(q,q')\, p\, f(q',p,t)dq'dp}{\int\int K(q,q')f(q',p,t)dq'dp}.$$

We denote the action of the operator $F$ on a function $f$ by $F[f](.)$.

$$F[f](q,p,t) = (\bar{p} - p), \bar{p}(q,t) = \frac{\int\int K(q,q')\, p\, f(q',p,t)dq'dp}{\int\int K(q,q')f(q',p,t)dq'dp}.$$

Hence the explicit form of kinetic equation is

$$\partial_t f + q\partial_p f = \partial_q(\alpha(q^2-1)qf + (q - q_f)f) + \frac{\sigma^2}{2}\partial_{qq} f \tag{5.1}$$

From here onwards, we consider the homogeneous case by dropping dependence on $q$, and use $x$ to denote the velocity $p$. The uncontrolled dynamics for the velocity of agent $i$ are

$$dx_i = a(x_i)dt + \frac{1}{N}\sum_{j=1}^{N}(x_j - x_i)dt + \sigma d\omega_i, \tag{5.2}$$

with corresponding density evolution

$$\partial_t f = \partial_x(\alpha(x^2-1)xf + (x - \bar{x})f) + \frac{\sigma^2}{2}\partial_{xx} f, \tag{5.3}$$

where $\bar{x}(t) = \dfrac{\int xf(x,t)dx}{\int f(x,t)dx}$.

The gradient structure of Eq. 5.3 can be made explicit by rewriting it as

$$\partial_t f = \partial_x(f\partial_x \xi), \tag{5.4}$$

As the fixed points $f$ satisfy $\partial_x \xi = 0$.

### 5.2.1 Fixed Points and Stability Analysis

It is known [109, 104] that fixed points of Eq. (5.3) are given by

$$f_\infty(x;\mu) = \frac{1}{Z}\exp\left(\frac{-2}{\sigma^2}\left[\alpha\frac{x^4}{4} + (1-\alpha)\frac{x^2}{2} - \mu x\right]\right), \tag{5.5}$$

where $\mu \in \mathbb{R}$ is the mean of the distribution, and $Z$ is the normalization factor. For all positive values of parameters $(\sigma, \alpha)$, the zero mean velocity solution $f_\infty(\cdot, 0)$ always exists. For a range of parameters, two additional stable non-zero mean velocity solutions are created via a supercritical bifurcation, resulting in loss of stability of the zero mean solution. In Ref. [104], these stability properties were inferred numerically by a Monte-Carlo approach.

We take a different approach, and consider the spectral stability of steady state solutions of Eq. (5.3). In addition to gaining additional insight into the properties of the uncontrolled system, this also sets the stage for stability analysis of the MFG system in the next section. We consider perturbations of the form $f(x,t) = f_\infty(x)(1 + \epsilon\tilde{f}(x,t))$. Then, the linearization of Eq. (5.3) is

$$\partial_t \tilde{f}(x,t) = L[\tilde{f}](x,t) = L_{loc}[\tilde{f}](x,t) + L_{nonloc}[\tilde{f}](x,t),$$

where, $\hat{U}(x) = U(x) + x^2/2 - \mu x$,

$$L_{loc}[\tilde{f}](x,t) = -\partial_x\hat{U}(x)\partial_x\tilde{f}(x,t) + (\sigma^2/2)\partial_{xx}\tilde{f}(x,t)$$

is a local linear operator, and

$$L_{nonloc}[\tilde{f}](x,t) = \frac{2}{\sigma^2}\partial_x\hat{U}(x)\int y\tilde{f}(y,t)f_\infty(y)dy$$

is a *nonlocal* linear operator. An operator $O$ is called nonlocal if $O[f](x_1)$ depends on $f(x_2)$ (or the derivatives $\partial_x f(x_2), \partial_{xx} f(x_2)$) for some $x_2 \neq x_1$, and local otherwise. Let $q(x) \equiv \frac{2}{\sigma^2}\partial_x \hat{U}(x)$. Then, $\partial_x f_\infty(x) = -q(x) f_\infty(x)$. We define a Hilbert space $\mathbb{H} = L^2(\mathbb{R}, f_\infty dx)$ , i.e., the $f_\infty$-weighted inner-product space of square-integrable functions on the real line. Then we can write a general form of the full linearized operator as

$$L[\tilde{f}](x,t) = L_{loc}[\tilde{f}](x,t) + s_1(x)\langle g_1(\cdot), \tilde{f}(\cdot,t)\rangle, \tag{5.6}$$

where $s_1(x) = q(x), g_1(x) = x$ for our case, and the inner product is understood to be $\langle\cdot,\cdot\rangle_{\mathbb{H}}$. We note that $L_{loc}$ is a self-adjoint operator[83] on $\mathbb{H}$ which has a non-positive discrete real spectrum of the form $0 = \lambda_1 > \lambda_2 > \lambda_3 \ldots$. It has a complete set of orthogonal eigenfunctions $\{\xi_i(x)\}_{i\in\mathbb{N}}$. The first eigenfunction $\xi_1$, spanning the kernel of $L_{loc}$, is a constant function. Following the approach presented in Refs. [107, 110] for nonlocal eigenvalue problems in reaction-diffusion equations (also see Ref. [111]), we consider the following eigenvalue problem

$$\begin{aligned}
&\lambda w = L_{loc} w + s_1(x)\langle g_1, w\rangle \implies \\
&\qquad 0 = (L_{loc} - \lambda I) w + s_1(x)\langle g_1, w\rangle.
\end{aligned} \tag{5.7}$$

Note that an eigenfunction $w$ of $L$ satisfying $\langle w, g_1\rangle = 0$ is also an eigenfunction of $L_{loc}$, i.e. $w = v_i$ for some $i$ with eigenvalue $\lambda = \lambda_i$. We search for eigenfunctions such that $\langle w, g_1\rangle$ is nonzero. The corresponding eigenvalues are called 'moving' eigenvalues in Ref. [110]. Multiplying both sides of Eq. (5.7) with the resolvent $R_\lambda = (L_{loc} - \lambda I)^{-1}$,

$$0 = w + R_\lambda s_1(x)\langle g_1, w\rangle.$$

Taking the inner product of the above equation with $g_1$,

$$0 = \langle g_1, w\rangle + \langle R_\lambda s_1(x), g_1\rangle\langle g_1, w\rangle. \tag{5.8}$$

For an arbitrary function $z(x)$, $R_\lambda z = \sum_{i=1}^{\infty} \frac{\langle \xi_i, z\rangle}{\lambda_i - \lambda}\xi_i$. Evaluating the inner product in Eqs. 5.8,

$$\langle R_\lambda s_1(x), g_1\rangle = \langle R_\lambda q(x), x\rangle = \sum_{i=2}^{\infty} \frac{\langle \xi_i, q(x)\rangle}{\lambda_i - \lambda}\langle \xi_i, x\rangle. \tag{5.9}$$

Using this result in Eq. (5.8),

$$\langle w, x\rangle(1 + \sum_{i=2}^{\infty} \frac{\langle \xi_i, q(x)\rangle}{\lambda_i - \lambda}\langle \xi_i, x\rangle) = 0.$$

Hence, either $\langle w, x\rangle = 0$, or $1 + \sum_{i=2}^{\infty} \frac{\langle \xi_i, q(x)\rangle}{\lambda_i - \lambda}\langle \xi_i, x\rangle = 0$. But we are looking for moving eigenvalues, i.e. $w$ s.t. $\langle w, x\rangle \neq 0$, hence the eigenvalue equation reduces to:

$$h(\lambda) \equiv 1 + \sum_{i=2}^{\infty} \frac{\langle \xi_i, q(x)\rangle}{\lambda_i - \lambda}\langle \xi_i, x\rangle = 0. \tag{5.10}$$

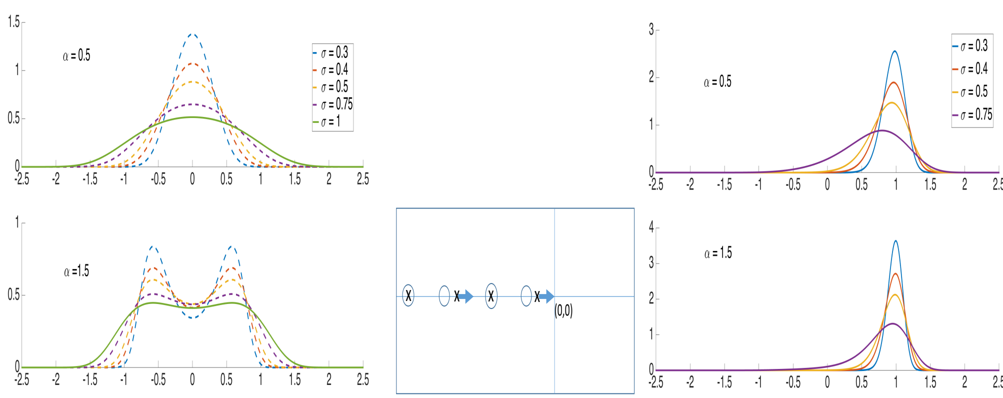



Figure 5.2: Uncontrolled system. (left) Stable (solid) and unstable (dashed) zero mean steady state solutions for $\alpha = 0.5$ (uni-modal, top) and $\alpha = 1.5$ (bi-modal, bottom). (middle) Eigenvalues of $L_{loc}$ (o) and $L$ ($\times$) for a typical zero mean case. The first eigenvalue ($= 0$) is omitted. Notice that alternating eigenvalues are same for both operators. The arrows indicate the direction of motion of the other ('moving') eigenvalues of $L$ as $\sigma$ is reduced. The rightmost eigenvalue of $L$ reaches 0 at $\sigma = \sigma_c(\alpha)$ with non-zero speed. (right) Non-zero mean solutions.

A sufficient condition for Eq. (5.10) to have only real roots is that the function $h(\lambda)$

is Herglotz, or equivalently, the product $\langle \xi_i, q(x)\rangle\langle \xi_i, x\rangle$ has the same sign for all $i$. Using integration by parts on eigenvalue equation for $L_{loc}$, one can show that $\langle \xi_i, x\rangle = -\frac{\sigma^2}{2\lambda_i}\langle \xi_i, q(x)\rangle$. Thus the Herglotz condition is satisfied since $\lambda_i < 0$ for all $i > 1$.

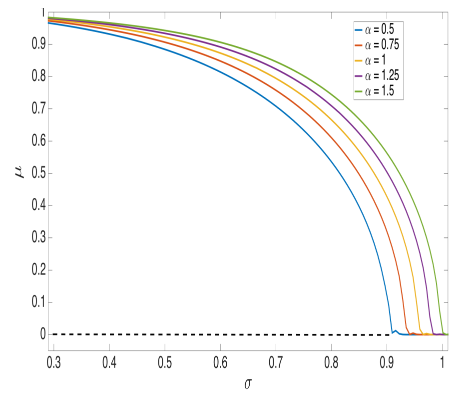



Figure 5.3: The $\mu > 0$ branch (solid) bifurcating from $\mu = 0$ solution (dashed) via a supercritical pitchfork bifurcation as $\sigma$ occurs is reduced below $\sigma_c(\alpha)$

**Numerical Results:** We use Chebfun [112] to perform all computations. The non-zero mean steady state solutions to Eq. (5.3) are computed using a simple fixed point iteration for $\mu$. The solutions are shown in Figure 5.2. The supercritical pitchfork bifurcation that occurs as $\sigma$ is reduced below critical value $\sigma_c(\alpha)$, is shown for a range of $\alpha$ values. To evaluate $h(\lambda)$ in Eq. (5.10), we compute the spectrum of $L_{loc}$ for $\mu = 0$. The odd-numbered eigenfunctions are even functions of $x$, and hence $\langle \xi_{2k+1}, g_1\rangle = 0$. Therefore, eigenvalues $\lambda_{2k+1}$ of $L_{loc}$ are also eigenvalues of $L$, and the eigenvalues $\lambda_{2k}$ are moving eigenvalues. We find that at $\sigma = \sigma_c(\alpha)$, $h(0) = 0$. Hence, as $\sigma$ is decreased below $\sigma_c(\alpha)$, the least stable eigenvalue $\lambda_2$ of $L_{loc}$ moves to the positive real axis due to the effect of the nonlocal term, resulting in instability of the zero mean solution.

**Lemma 5.2.1.** *The function $h$ defined in Eq. 5.10 is Herglotz, and hence all its roots are real for all parameter values of* $(\alpha, \sigma)$.

*Proof.* We will show that $\langle \xi_i(x;\alpha,\sigma), x\rangle = -\frac{\sigma^2}{2\lambda_i}\langle \xi_i(x;\alpha,\sigma), q(x)\rangle$, from which the result follows since $\lambda_i < 0$ for all $i > 1$. Now, consider the eigenvalue equation for $L_{loc}$

corresponding to $\lambda_i$.

$$\begin{aligned}\lambda_i \xi_i &= -\partial_x \tilde{U} \partial_x \xi_i + \frac{\sigma^2}{2} \partial_{xx} \xi_i \\ &= -\frac{\sigma^2}{2} q \partial_x \xi_i + \frac{\sigma^2}{2} \partial_{xx} \xi_i.\end{aligned} \tag{5.11}$$

Computing the inner product of Eq. 5.11 with $x$, we get

$$\lambda_i \langle \xi_i, x \rangle = \frac{\sigma^2}{2}(-\int xq\partial_x \xi_i f_\infty dx + \int x \partial_{xx} \xi_i f_\infty dx)$$
$$= \frac{\sigma^2}{2}(\int x\partial_x \xi_i \partial_x f_\infty dx - \int x\partial_x \xi_i \partial_x f_\infty dx - \int \partial_x \xi_i f_\infty dx),$$

where we have integrated by parts. Then

$$\lambda_i \langle \xi_i, x \rangle = \frac{\sigma^2}{2}(-\int \partial_x \xi_i f_\infty dx)$$
$$= \frac{\sigma^2}{2} \int \xi_i \partial_x f_\infty dx = -\frac{\sigma^2}{2} \int \xi_i q f_\infty dx$$
$$= -\frac{\sigma^2}{2} \langle \xi_i, q \rangle.$$

□

**Asymptotics:** Eq. 5.10 can now be written as

$$h(\lambda) \equiv 1 - \frac{2}{\sigma^2} \sum_{i=2}^{\infty} \lambda_i(\alpha,\sigma) \frac{\langle \xi_i(x;\alpha,\sigma), x \rangle^2}{\lambda_i(\alpha,\sigma) - \lambda} = 0 \tag{5.12}$$

In the limit $\alpha \to 0$, $(\lambda_i, \xi_i)$'s are Hermite eigenvalues and eigenfunctions. Then the above equation be evaluated analytically.

$$h(\lambda) = \sigma^2/2 + \sum_{i=2}^{\infty} \alpha \frac{\langle v_i(x;\alpha,\sigma),(x^3-x)\rangle}{\lambda_i(\alpha,\sigma)-\lambda} \langle v_i, x\rangle + \sum_{i=2}^{\infty} \frac{\langle v_i(x;\alpha,\sigma),x\rangle^2}{\lambda_i(\alpha,\sigma)-\lambda} = 0 \quad (5.13)$$

We note that for $\alpha = 0$, $v_i$ are the Hermite functions. In the limit $\alpha \to 0$, $h(\lambda) \approx \sigma^2/2 + \sum_{i=2}^{\infty} \frac{\langle v_i(x;\alpha,\sigma),x\rangle^2}{\lambda_i(\alpha,\sigma)-\lambda}$, and hence it is clear it is Herglotz.

## 5.3 MFG Formulation

In this section we describe a MFG formulation for homogeneous equation Eq. (5.3). The velocity of $i$th agent evolves via the following equation (compare with Eq (5.2))

$$dx_i(t) = a(x_i)dt + u_i(t)dt + \sigma d\omega_i(t), \quad (5.14)$$

where $u_i$ is the optimal control. Let $F[x_i, x_{-i}](t) \equiv (x_i - \frac{1}{N-1}\sum_{j\neq i} x_j)^2$, and $\beta = \frac{1}{r\sigma^2}$, where $r > 0$ is the control cost or penalty. Here $x_{-i} \equiv \{x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots\}$. Then the $i$th agent is minimizing the following long time average cost

$$J = \limsup_{T\to\infty} \frac{1}{T}\left[\int_0^T \beta F[x_i, x_{-i}](t) + \frac{1}{2\sigma^2} u_i(t)^2\right] dt,$$

that depends on states of all other agents.

To derive the MFG equations (recall Fig. 5.1), we rewrite the single-agent cost in terms of $\hat{F}(x_i, t)$, the unknown coupling function with dependence on $x_i$ only

$$J = \limsup_{T\to\infty} \frac{1}{T}\left[\int_0^T \beta \hat{F}(x_i, t) + \frac{1}{2\sigma^2} u_i(t)^2\right] dt.$$

The resulting single agent HJB equation [85] is

$$\partial_t v_i(x,t) = c - \beta \hat{F}(x,t) - a(x)\partial_x v_i(x,t) + \frac{\sigma^2}{2}(\partial_x v_i(x,t))^2 - \frac{\sigma^2}{2}\partial_{xx} v_i(x,t), \tag{5.15}$$

where $v_i(x,t)$ is the single-agent relative value function, $c$ is the minimum average cost, and $u_i(x,t) = -\sigma^2 \partial_x v_i(x_i,t)$ given in feedback form. Note that the HJB equation is well-posed backward in time. The self-consistency principle yields the expression for $\hat{F}$ in terms of agent density $f(x,t)$ (in the limit $N \to \infty$):

$$\hat{F}[f](x,t) = \left[\int (x-y) f(y,t) dy\right]^2. \tag{5.16}$$

Hence, the following set of FP-HJB MFG equations govern the density and value function evolution:

$$\partial_t f(x,t) + \partial_x \left[\left(a(x) - \sigma^2 \partial_x v(x,t)\right) f(x,t)\right] = \frac{\sigma^2}{2}\partial_{xx} f(x,t), \tag{5.17}$$

$$\partial_t v(x,t) = c - \beta \hat{F}[f](x,t) - a(x)\partial_x v(x,t) + \frac{\sigma^2}{2}(\partial_x v(x,t))^2 - \frac{\sigma^2}{2}\partial_{xx} v(x,t). \tag{5.18}$$

The unique invariant density satisfying Eq. (5.17) is

$$f_\infty(x) = \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}(U(x) + \sigma^2 v_\infty(x))\right). \tag{5.19}$$

Inserting this expression into Eq. (5.18), and using the Cole-Hopf transformation [113] $\phi(x) = \exp(-v_\infty(x))$, results in the following nonlinear nonlocal eigenvalue problem for

$\phi(x)$:

$$
\begin{aligned}
c\phi(x) = \beta\{x - \int y \exp(\frac{-2}{\sigma^2}U(y))\phi^2(y)dy\}^2\phi(x) \\
- a(x)\phi(x) - \frac{\sigma^2}{2}\partial_{xx}\phi(x),
\end{aligned} \tag{5.20}
$$

with the constraint $\int \exp(\frac{-2}{\sigma^2}U(y))\phi^2(y)dy = 1$ to ensure normalization of $f_\infty$. The ground state of this problem yields the desired steady state solutions, with corresponding eigenvalue being the minimum cost $c$.

### 5.3.1 Stability Analysis

In this section we extend the resolvent based analysis from section 5.2.1 to the MFG system, and find conditions for *closed-loop* stability of an arbitrary steady state $(f_\infty(x), v_\infty(x))$ to an initial perturbation in density. We consider mass preserving perturbations in density of the form $f(x,t) = f_\infty(x)(1+\epsilon\tilde{f}(x,t))$, i.e., the initial conditions satisfy $\int f_\infty(x)\tilde{f}(x,0)dx = 0$. The perturbed value function is taken to be of the form $v(x,t) = v_\infty(x) + \epsilon\tilde{v}(x,t)$. A given steady state is called linearly stable if any perturbation to the density decays to zero under the action of the control, where both the density and control evolution are computed using linearized MFG equations.

Linearization of MFG equations (5.17,5.18) yields the nonlocal system

$$
\begin{bmatrix} \partial_t\tilde{f}(x,t) \\ \partial_t\tilde{v}(x,t) \end{bmatrix} = L^{FB}_{loc} \begin{bmatrix} \tilde{f}(x,t) \\ \tilde{v}(x,t) \end{bmatrix}, \tag{5.21}
$$

where $L^{FB} = L^{FB}_{loc} + L^{FB}_{nonloc}$,

$$
L^{FB}_{loc} = \begin{bmatrix} L_{loc} & 2L_{loc} \\ 0 & -L_{loc} \end{bmatrix}, L^{FB}_{nonloc} = \begin{bmatrix} 0 & 0 \\ 2\beta s_1\langle g_1, .\rangle & 0 \end{bmatrix},
$$

$s_1(x) = x - \mu$, $g_1(x) = x$, $L_{loc} = -\partial_x(\hat{U})\partial_x + \frac{\sigma^2}{2}\partial_{xx}$, with eigenvalue/eigenfunction pairs denoted by $\{\lambda_i, \xi_i\}$, and $\hat{U}(x) = U(x) + \sigma^2 v^\infty(x)$ in analogy with the definition of $L_{loc}$ in Section 5.2. In addition to the Hilbert space $\mathbb{H} = L^2(\mathbb{R}, f_\infty dx)$ and $R_\lambda$ as defined earlier, we also consider a subspace $\bar{\mathbb{H}} = \{f \in \mathbb{H} | \langle f, 1 \rangle = 0\}$.

*Eigenspectrum of the linearized forward-backward operator*

We start off by noting that the characteristic equation of $L_{loc}^{FB}$ is $(L_{loc} - \lambda I)(L_{loc} + \lambda I) = 0$. Hence, its eigenvalues are $\cup_{i \in \mathbb{N}} \{\pm \lambda_i\}$. Now consider the eigenvalue problem for $L^{FB}$ with eigenvalue $\lambda$ and eigenfunction $[w_f(x)\ w_v(x)]^T$:

$$\lambda \begin{bmatrix} w_f \\ w_v \end{bmatrix} = \begin{bmatrix} L_{loc} w_f + 2 L_{loc} w_v \\ 2\beta s_1 \langle g_1, w_f \rangle - L_{loc} w_v \end{bmatrix}. \tag{5.22}$$

Assuming $\lambda \notin \cup_{i \in \mathbb{N}} \{\pm \lambda_i\}$, $R_\lambda$ and $R_{-\lambda}$ are well defined. The second equation of Eq. (5.22) gives

$$w_v = 2\beta R_{-\lambda} s_1 \langle g_1, w_f \rangle.$$

Substituting this expression in the first equation of Eq. (5.22), and re-arranging,

$$w_f = -4\beta R_\lambda L_{loc} R_{-\lambda} s_1 \langle g_1, w_f \rangle. \tag{5.23}$$

Taking the inner product of the above equation with $g_1$,

$$\langle g_1, w_f \rangle (1 + 4\beta \langle g_1, R_\lambda L_{loc} R_{-\lambda} s_1 \rangle) = 0.$$

The eigenvalue equation for the $\langle g_1, w_f \rangle \neq 0$ case for moving eigenvalues (as in Section 5.2.1) is

$$h(\lambda) \equiv 1 + 4\beta \langle g_1, R_\lambda L_{loc} R_{-\lambda} s_1 \rangle = 0. \tag{5.24}$$

Using the definition of resolvent in Eq. (5.24),

$$h(\lambda) = 1 + 4\beta \sum_{i=2}^{\infty} \lambda_i \frac{\langle g_1, \xi_i \rangle \langle s_1, \xi_i \rangle}{\lambda_i^2 - \lambda^2}, \tag{5.25}$$

and hence,

$$h(\lambda) = 1 + 4\beta \sum_{i=2}^{\infty} \lambda_i \frac{\langle x, \xi_i \rangle^2}{\lambda_i^2 - \lambda^2}. \tag{5.26}$$

Since Eq. (5.26) is Herglotz in $\lambda^2$, this implies that the eigenvalues come in pairs, either real or purely imaginary. Let $\omega \equiv h(0) = 1 + 4\beta \sum_{i=2}^{\infty} \dfrac{\langle x, \xi_i \rangle^2}{\lambda_i}$.

**Lemma 5.3.1.** *Consider the eigenvalue equation $h(\lambda) = 0$ for moving eigenvalues.*

*(i) If $\langle x, \xi_i \rangle \neq 0$ for all $i \geq 2$, then there exists a pair of real roots $\pm\delta_i$ for each $i \geq 2$, such that $\lambda_{i+1} < \delta_i < \lambda_i$.*

*(ii) Recall that $\lambda_1 = 0$. If $\langle x, \xi_2 \rangle \neq 0$ and $\omega > 0$, there exists a pair of real roots $\pm\delta_1$, such that $\lambda_2 < \delta_1 < 0$.*

*(iii) If $\langle x, \xi_2 \rangle \neq 0$ and $\omega < 0$, there exists a pair of purely imaginary roots $\pm i\gamma$.*

*Proof.* (i) Consider the interval $I_i = (\lambda_{i+1}, \lambda_i)$. As $\lambda \to \lambda_i^-$, $h(\lambda) \to \infty$, and as $\lambda \to \lambda_{i+1}^+$, $h(\lambda) \to -\infty$. It is easy to check that $h(\lambda)$ is monotonic in $I_i$. By intermediate value theorem, a root $\delta_i$ exists in $I_i$, and by the monotonicity property, it is unique. The result for $-\delta_i$ follows by symmetry.

(ii) Consider the interval $I_1 = (\lambda_2, 0)$. Note that as $\lambda \to \lambda_2^+$, $h(\lambda) \to -\infty$, and as $\lambda \to 0^-$, $h(\lambda) \to \omega$. Hence, if $w > 0$, arguments similar to those in part (i) yield the existence of a real root $\delta_1$ between $\lambda_2$ and $0$.

(iii) Consider the function $h(i\gamma)$ for real $\gamma > 0$. Clearly, $h$ is monotonic in this interval. Furthermore, as $\gamma \to \infty$, $h(i\gamma) \to 0$, and as $\gamma \to 0^+$, $h(i\gamma) \to \omega$. By arguments similar to those in part (i), $\omega < 0$ implies that there is a unique root $i\gamma$ of $h$.

□

*Contraction analysis of the linearized forward-backward operator*

Since the MFG system has a forward-backward nature, spectral information alone is insufficient to derive conclusions about the stability of steady state solutions. A contraction analysis is therefore adopted following Refs. [18, 15]. Consider the linear dynamical system given by Eq. (5.21), with initial perturbation in density $f(x,0) = f_\infty(1+\epsilon\tilde{f}(x,0))$. Assuming that $\tilde{v}(x,T) \to 0$ as $T \to \infty$, the conditions for existence of a unique solution satisfying this assumption are derived. These conditions also provide a stability criterion. Integrating the $\tilde{v}$ equation in Eq. (5.21) from $t$ to $T$,

$$\begin{aligned}\tilde{v}(x,T) = {} & e^{-L_{loc}(T-t)}\tilde{v}(x,t) \\ & + 2\beta\int_t^T e^{-L_{loc}(T-s)}s_1(x)\langle g_1,\tilde{f}(.,s)\rangle ds.\end{aligned}$$

Taking the limit $T \to \infty$,

$$\tilde{v}(x,t) = -2\beta e^{-L_{loc}t}\int_t^\infty e^{L_{loc}s}s_1(x)\langle g_1(.),\tilde{f}(.,s)\rangle ds. \tag{5.27}$$

Substituting above equation in the $\tilde{f}$ equation,

$$\begin{aligned}\partial_t\tilde{f}(x,t) = {} & L_{loc}\tilde{f}(x,t) \\ & - 4\beta L_{loc}e^{-L_{loc}t}\int_t^\infty e^{L_{loc}s}s_1(x)\langle g_1(.),\tilde{f}(.,s)\rangle ds.\end{aligned} \tag{5.28}$$

Integrating from $0$ to $t$ yields the fixed point equation,

$$\tilde{f}(x,t) = e^{L_{loc}t}\tilde{f}(x,0) + M\tilde{f}(x,t), \tag{5.29}$$

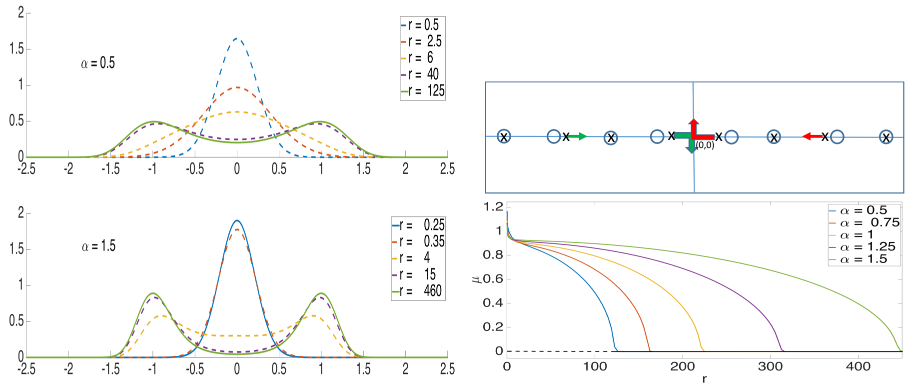



Figure 5.4: The MFG system with $\sigma = 0.5$. (left): Zero mean MFG steady state densities for $\alpha = 0.5$ (top) and $\alpha = 1.5$ (bottom) for various control penalty values. (right): (Top) Eigenvalues of $L^{FB}_{loc}$ (o) and $L^{FB}$ ($\times$) for a typical zero mean case. The twin zero eigenvalues of $L^{FB}$ are omitted. The arrows indicate the direction of motion of the 'moving' eigenvalues of $L^{FB}$ as $r$ is reduced starting from $r > r_{sup}(\alpha, \sigma)$. Note that the assumption in Lemma 5.3.1(i) is violated in this particular case due to the symmetric nature of the self-propulsion term, and hence, only alternating eigenvalues are actually 'moving'. The pair of eigenvalues of $L^{FB}$ closest to imaginary axis, $\pm\delta_1$, reaches 0 at $r = r_{sup}$, and moves up/down the imaginary axis for $r < r_{sup}$. (Bottom) The $\mu > 0$ branch (solid) bifurcating from $\mu = 0$ solution (dashed) via a supercritical pitchfork bifurcation as $r$ is reduced below $r_{sup}$.

where the operator $M$ acting on $\tilde{f}(x,t)$ is defined as

$$M\tilde{f}(x,t) = -4\beta e^{L_{loc}t}\int_{r=0}^{r=t} e^{-L_{loc}r}L_{loc}e^{-L_{loc}r}\int_{s=r}^{s=\infty} e^{L_{loc}s}s_1(x)\langle g_1(.), \tilde{f}(., s)\rangle ds dr. \tag{5.30}$$

Applying the Laplace transform in time to Eq. (5.30),

$$\hat{M}(\lambda) = -4\beta R_\lambda L_{loc} R_{-\lambda} s_1 \langle g_1, .\rangle. \tag{5.31}$$

The operator norm $\|M\|$ is given by

$$\begin{aligned} \|M\| =& \sup_{\lambda\in\mathbb{I}} \sup_{\|\tilde{f}\|=1} \|\hat{M}(\lambda)\tilde{f}\|, \\ =& 4\beta \sup_{\lambda\in\mathbb{I}} \sup_{\|\tilde{f}\|=1} \|\sum_{i=2}^{\infty} \frac{\lambda_i\langle s_1, \xi_i\rangle\langle g_1, \tilde{f}\rangle}{\lambda_i^2 - \lambda^2}\xi_i\|, \end{aligned} \tag{5.32}$$

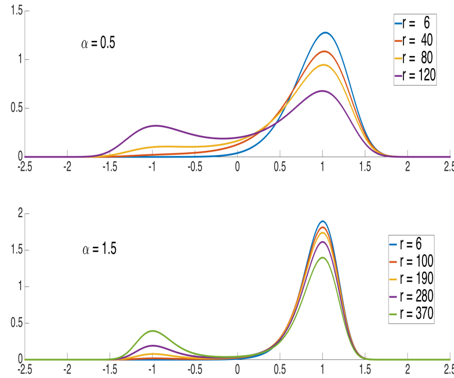



Figure 5.5: Non-zero mean MFG steady state densities on the supercritical branch.

$$
\begin{aligned}
=&4\beta \sup_{\lambda\in\mathbb{I}} \sup_{\|\tilde{f}\|=1} \sqrt{\sum_{i=2}^{\infty}\left[\frac{\lambda_i\langle s_1,\xi_i\rangle\langle g_1,\tilde{f}\rangle}{\lambda_i^2-\lambda^2}\right]^2},\\
=&4\beta\|g_1\|\sqrt{\sum_{i=2}^{\infty}\frac{\langle s_1,\xi_i\rangle^2}{\lambda_i^2}} = 4\beta\|x\|\sqrt{\sum_{i=2}^{\infty}\frac{\langle x,\xi_i\rangle^2}{\lambda_i^2}}.
\end{aligned}
\tag{5.33}
$$

Lemma 5.3.2 proved next implies that $\|M\| < 1$ is a sufficient condition for a steady state $(f_\infty(x), v_\infty(x))$ of the nonlinear MFG system Eqs. (5.17,5.18) to be linearly stable to density perturbations.

**Lemma 5.3.2.** *Consider the initial value problem for the linearized system in Eqs. 5.21, with mass-preserving initial condition* $\tilde{f}(x,0)$ *i.e.,* $\int f_\infty(x)\tilde{f}(x,0)dx = 0$*. If the operator* $M$ *is a contraction (i.e.,* $\|M\| < 1$*), then the perturbation in density,* $\tilde{f}(.,t)$*, decays to* $0$ *as* $t\to\infty$*. Moreover,* $\tilde{v}(.,t)$ *also decays to* $0$ *as* $t\to\infty$*.*

*Proof.* If $M$ is a contraction, then we can (formally) invert the Eq. (5.29), and write the unique solution

$$
\begin{aligned}
\tilde{f}(x,t) &= (\mathbf{I}-M)^{-1}e^{L_{loc}t}\tilde{f}(x,0)\\
&= (\mathbf{I}+M+M^2+\dots)e^{L_{loc}t}\tilde{f}(x,0).
\end{aligned}
\tag{5.34}
$$

We note that mass conservation property is equivalent to $\langle\tilde{f}(x,0),1\rangle = 0$, i.e. $\tilde{f}(x,0)\in\bar{\mathbb{H}}$. Recall that $L_{loc}$ restricted to $\bar{\mathbb{H}}$ is a self-adjoint operator with negative eigenvalues $\lambda_i, i =$

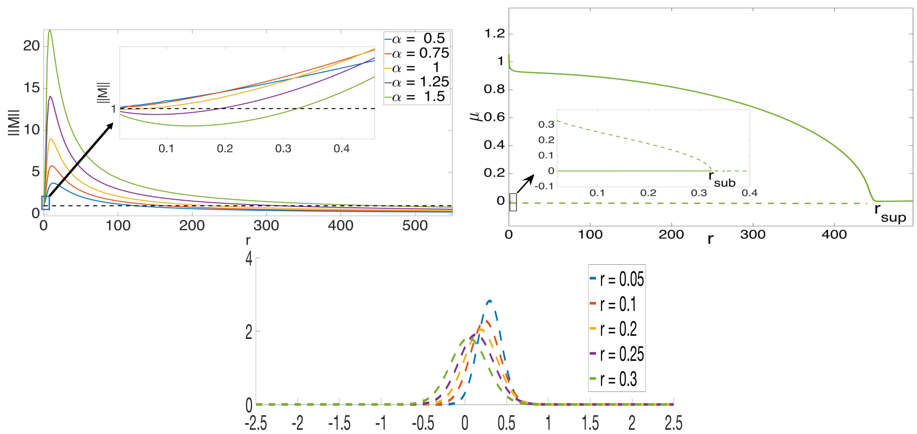



Figure 5.6: The MFG system with $\sigma = 0.5$. (left): The norm of operator $M$ for zero mean steady state as control penalty is varied, for various $\alpha$. (right): The bifurcation diagram for $\alpha = 1.5$, showing supercritical and subcritical (inset) bifurcations . Only the $\mu > 0$ non-zero mean branches are shown. (bottom): Non-zero mean MFG steady state densities on the subcritical branch for $\alpha = 1.5$.

$2, 3, \ldots$. Then, $\lim_{t\to\infty} \|e^{L_{loc}t}\|_{\bar{\mathbb{H}}} = \lim_{t\to\infty} e^{\lambda_2 t} = 0$. This proves the decay of $\tilde{f}(., t)$. The corresponding result for $\tilde{v}(., t)$ is obtained by inserting the expression for $\tilde{f}(., t)$ into Eq. (5.27). □

Now consider a case where eigenvalue equation in Eq. (5.23) has a pair of purely imaginary roots $\pm i\gamma (\neq 0)$. Then there is a eigenfunction $z_f$ s.t.

$$
\begin{aligned}
z_f &= -4\beta R_{i\gamma} L_{loc} R_{-i\gamma} s_1 \langle g_1, z_f \rangle \\
&= \hat{M}(i\gamma) z_f,
\end{aligned}
$$

by noting Eq. (5.31). But this implies that norm of $\hat{M}$ is at least 1, hence it is not a contraction. This implies that a necessary condition for $M$ to be a contraction is the absence of non-zero spectra of $L^{FB}$ on the imaginary axis.

### 5.3.2 Numerical Results

Recall that in the MFG problem described by Eqs. (5.17,5.18), the representative agent is minimizing a weighted sum of two costs: one penalizes deviation of its velocity from the mean velocity of the agent population, and the other penalizes the control action. In this section, we compute fixed points, and identify phase transitions of this system of equations as the problem parameters are varied. Rather than solving the resulting constrained nonlinear eigenvalue problem 5.20 directly, we use an iterative algorithm to compute steady state solutions of the MFG system.

We note that the coupling term $\hat{F}[f](x,t)$ evaluated at any steady state density $f_\infty$ is $\hat{F}[f_\infty](x) = (x-\mu)^2$, where $\mu = \int y f_\infty(y)dy$. Again using Cole-Hopf transformation $\phi(x) = \exp(-v(x))$ on the HJB equation leads to a **linear** eigenvalue problem in $\phi$:

$$c\phi(x) = \mathcal{L}[\phi](x), \tag{5.35}$$

where $\mathcal{L}[\phi] = \beta(x-\mu)^2\phi(x) - a(x)\partial_x\phi(x) - \frac{\sigma^2}{2}\partial_{xx}\phi(x)$. We solve Eq. (5.35) iteratively along with Eq. (5.19) to find zero mean MFG steady states $(f_\infty, v_\infty)$ for a range of $r$, keeping $\sigma$ and $\alpha$ fixed (See Fig. 5.4).These solutions are stable (i.e., $\|M\| < 1$) for large $r$, implying that when control is expensive, the agents use minimal control action. The resulting steady state distribution is bi-modal due to dominance of the self-propulsion force, and dispersion via noise.

These zero mean solutions lose stability (i.e., $\|M\| > 1$) via a supercritical bifurcation as $r$ is reduced below a critical value $r_{sup}$. The Eq. (5.26) for moving eigenvalues of $L^{FB}$ has a double zero root at $r = r_{sup}$, and a pair of purely imaginary roots emerges as $r$ is reduced below $r_{sup}$. This implies that the pair of symmetric eigenvalues of $L^{FB}$ closest to the imaginary axis reaches $0$ at the critical parameter, and then moves up/down the imaginary axis. The stable non-zero mean MFG steady state solutions on the supercritical branch are computed by combining fixed point iteration in $\mu$ with a continuation step. This bifurcation provides a MFG interpretation to the pitchfork bifurcation observed in the uncontrolled system, i.e., cheaper control makes it economical to compensate for noise. Hence, the agents apply larger control action to flock together (and reduce the cost of deviation from the population mean), resulting in symmetry breaking non-zero mean solutions.

When noise strength $\sigma$ is fixed below a critical value, the zero mean solution branch undergoes a *subcritical* bifurcation as control penalty $r$ is further reduced, i.e, at $r = r_{sub} < r_{sup}$ (See Fig. 5.6). The corresponding non-zero mean solutions were computed using bisection method. This bifurcation is not seen in the uncontrolled system. For instance, when ($\sigma = 0.5, \alpha = 1.5$), it results in creation of uni-modal stable zero mean solutions in the case of cheap control, $r < r_{sub}$, as compared to the bi-modal stable zero mean solution that exist for expensive control, $r > r_{sup}$. Hence, we conclude that for $r < r_{sub}$, the control is cheap enough to counteract the intrinsic dynamics, and make zero mean uni-modal solution stable.

## 5.4 Conclusions

We have presented a MFG formulation for homogeneous flocking of agents with gradient nonlinearity in their intrinsic dynamics. We have employed tools from theory of reaction-diffusion equations, and exploited the low rank nature of the nonlocal coupling term to study the linear stability of the MFG equations. The explicit formulae for verifying the stability of steady state solutions of the nonlocal forward-backward MFG system require relatively simple numerical computation of spectra of the local self-adjoint Fokker-Planck operators. The MFG system shows rich nonlinear behavior, such as supercritical and sub-critical pitchfork bifurcations that result in wide range of collective behaviors, some of which are not present in the uncontrolled model.

Much of the analysis in the current work can be generalized to higher dimensional state space for homogeneous flocking with self-propulsion, similar in spirit to the generalization[104] of one-dimensional *uncontrolled* flocking model. Furthermore, the abstract results presented in this work apply to models other than homogeneous flocking, e.g. nonlocally coupled agents with arbitrary first order gradient dynamics. Extension to non-homogeneous flocking would be a natural next step; the resulting second-order dynamics could require more sophisticated tools [114] for stability analysis. Implementation of the MFG control laws in an engineered large population system requires the control to be provided in a causal form. Algorithms that can learn the MFG laws can be used to convert the control laws obtained by solving the FP-HJB equations into an implementable form [115].

The use of bifurcation and singularity theory to develop bio-inspired control and decision making algorithms for multi-agent systems has been explored recently[116, 117, 118]. Our work adds to the toolbox for systematic analysis of collective behavior of non-cooperative dynamic agents via an inverse modeling approach. The qualitative and quantitative insight provided by the stability analysis can be exploited in *mechanism design*, i.e., design of penalties or incentives to drive the population to asymptotic states with desirable characteristics. We believe that a systematic study of bifurcations in MFG models can lead to progress in tackling the grand challenge of designing or manipulating collective behavior of a large population of non-cooperative dynamic agents.

# CHAPTER 6
# SCHRÖDINGER APPROACH TO LARGE SCALE CONTROL

Large-size populations consisting of a continuum of identical, non-cooperative and non-interacting agents with stochastic dynamics are useful in modeling various biological and engineered systems. This work addresses the problem of designing optimal state-feedback controllers for such systems which guarantee closed-loop stability of the stationary density of agents, in the case that individual agents have Langevin type passive dynamics. 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. We show the deep connection between the nonlinear Schrödinger equation and mean field games for agents with nonlinear Langevin dynamics. 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 control.

## 6.1 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 [6] or stochastic ([7], [8], [17]) system and have applications in quantum control [9] and neuroscience [10]. A standard idea in engineering, economics and biology is regulation using local feedback information, and is used to model decision making in large-size populations of *rational* agents with limited information. Optimal feedback control applications of large-size populations of *small* robots with individual state-feedback controllers have been proposed for inspection of industrial machinery [11], centralized control of hybrid automata [12] and decentralized control of robotic bee swarms for pollinating crops [13].

Optimal control models of collective behavior typically treat agents which are driven by individual noise and state-feedback control, and interact with each other through the coupling of their passive dynamics or utility with the overall statistics of the population. The *mean-field* approach provides a tractable framework for describing collective behavior of a continuum of agents, by approximating their individual actions [14] as the *oblivious* control [18] of a single *representative* agent. Mean field games (MFGs) ([16], [15]) utilize PDE optimality systems to model such continuum systems and are used to obtain a game-theoretic interpretation of *emergent* behaviour 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. 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 linear 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 the 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. Agents which lack explicit interaction have been studied using the mean-field approach in macro-economics [14]. In certain physical systems such as robot swarms ([12],[13]), 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 work 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. In this work, we address this question for large-size populations wherein agents obey Langevin dynamics and provide explicit control design constraints required for stability. For the finite time case, we present a 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 ([21], [106], [22]). A common limitation of prior works on this topic is that individual agent dynamics are assumed to be simple integrator systems. In the recent work [77] by some of the authors, linear (local) 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.

In this work we take a different approach. Since we assume that agent dynamics and cost functions have 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. In section 6.3.1, we introduce a novel Cole-Hopf type transform in order to obtain a decoupled representation of the coupled HJB-FP equations consisting of linear imaginary-time Schrödinger equations. Spectral properties of the corresponding Schrödinger operator underlie the stability properties of the fixed point density. In section 6.3.3 the Schrödinger potential of this operator is interpreted as the cost function of a closely related optimal control problem subject to simple integrator dynamics. This motivates a quadrature based scheme to compute the time varying control, which is explained in section 6.5. We observe 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. In section 6.4 we provide explicit stability constraints on the control design which guarantee closed-loop stability of the steady state density.

The connection between the imaginary-time Schrödinger equation and optimal control has been explored previously in the context of OMT [119], Schrödinger bridges [120] and in [20] 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 [20]. In section 6.3.2 we show that this connection is true for the broader class of MFGs in which agents obey non-linear Langevin dynamics. Our conclusions and directions for future research are presented in section 6.6.

## 6.2 Control of Large-Size Populations

We first introduce some notation and then describe the large scale stochastic control problem considered in this work. $L^2(\mathbb{R}^d)$ 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:

$$\mathrm{d}x_s^i = -\nabla\nu(x_s^i)\mathrm{d}s + u^i(s)\mathrm{d}s + \sigma\mathrm{d}w_s^i \tag{6.1}$$

where $x_s^i, u^i(s) \in \mathbb{R}^d$ are the state and control inputs and $w_s^i$ is a standard $\mathbb{R}^d$ Brownian motion. Suppose that the $i^{th}$ agent minimizes its individual performance objective given by

$$J^i(u) := \lim_{T\to+\infty}\frac{1}{T}\mathbb{E}\left[\int_0^T q(x_s^i)\mathrm{d}s + \frac{R}{2}(u^i)^2\ \mathrm{d}s\right], \tag{6.2}$$

then under certain standard conditions, the equivalent stationary HJB PDE problem is

$$0 = q - c^i - \frac{(\nabla v^i)^2}{2R} - \nabla v^i \cdot \nabla\nu + \frac{\sigma^2}{2}\Delta v^i \tag{6.3}$$

with the optimal control $u^{i,\infty}(x) = -\nabla v^i(x)/R$. Since the noise driving each agent is mutually independent and Brownian, the states of each agent $x_s^i$ are i.i.d. random variables independent of $w_s^i$. The set of the states $\{x_s^i\}_{1\leq<i<N}$ represents the population of agents. Next, we assume that the number of agents is infinite, $N \to +\infty$. We take use the mean-field approach to represent the problem as a standard OCP [73] of a *representative agent* with state $x_s \sim p(s,\cdot)$ obeying dynamics (6.4) 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 $\sup\limits_{1\leq i\leq N} \mathbb{E}[(x_0^i)^2] < k < +\infty$, the initial distribution is approximated by the emperical density $p_N(0,x) = \sum_{i=1}^N \delta(x - \mathbb{E}[x_0^i])$ 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 $\lim\limits_{N\to+\infty} \int \gamma(x)p_N(0,x) = p(0,x)$ for any bounded continuous function $\gamma(x)$ on $\mathbb{R}^d$.

### 6.2.1 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:

$$\mathrm{d}x_s = -\nabla\nu(x_s)\mathrm{d}s + u(s)\mathrm{d}s + \sigma\mathrm{d}w_s \tag{6.4}$$

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 [31] containing functions $u : [t,T] \times \mathbb{R}^d \to \mathbb{R}^d$. Consider the following optimal control problem (OCP)

$$\min_{u\in\mathcal{U}} J(u) := \lim_{T\to+\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T q(x_s)\mathrm{d}s + \frac{R}{2}u^2 \,\mathrm{d}s\right] \tag{6.5}$$

subject to (6.4), where we denote the probability density of $x_s$ by $\mathrm{p}(s, x)$ for every $s \geq 0$ which represents the density of all agents, with initial density being $x_0 \sim \mathrm{p}(0, x)$, $q : \mathbb{R}^d \rightarrow \mathbb{R}$ is a known deterministic function which has at most quadratic growth and $R > 0$ is the control cost. We assume that $\nabla\nu(x), q(x, \mathrm{p})$ and functions in the class $\mathcal{U}$ are measurable. We refer to the OCP (6.5) subject to dynamics (6.4) as problem **(P1)**.

### 6.2.2 PDE Optimality System

Standard application of *dynamic programming* [82] as in ([73], [16]), implies that under certain regularity conditions [106], problem **(P1)** is equivalent to the following HJB-FP PDE optimality system governing the value and density functions respectively:

$$q - c - \frac{(v_x^\infty)^2}{2R} - \nabla v^\infty \cdot \nabla\nu + \frac{\sigma^2}{2}\Delta v^\infty = 0 \tag{6.6}$$

$$\nabla((\nabla\nu + \frac{\nabla v^\infty}{R})p^\infty) + \frac{\sigma^2}{2}\Delta p = 0 \tag{6.7}$$

with the constraint $\int p^\infty \mathrm{d}x = 1$, where $c$ is the optimal cost. The optimal control is given by $u^\infty(x) = -\nabla v^\infty / R$. Under certain regularity conditions [106] which we assume to be true, the time-varying relative value [85] function and density corresponding to problem **(P1)** are governed by the optimality system:

$$-\partial_t v = q - c - \frac{(\nabla v)^2}{2R} - \nabla v \cdot \nabla\nu + \frac{\sigma^2}{2}\Delta v \tag{6.8}$$

$$\partial_t p = \nabla \cdot ((\nabla\nu + \frac{\nabla v}{R})p) + \frac{\sigma^2}{2}\Delta p \tag{6.9}$$

with the constraint $\int p(t, x)\mathrm{d}x = 1$ for all $t \geq 0$. In this work, we assume to be true, the additional conditions [31] which are required to show that the HJB PDEs (6.6) and (6.8) have unique solutions. Note that steady state and time varying HJB PDEs are both semilinear.

**Remark 5.** *The finite time OCP analogous to the infinite time OCP **(P1)** given by:*

$$\min_{u\in\mathcal{U}} J(u) := \mathbb{E}\left[\int_0^T q(x_s)\mathrm{d}s + \frac{R}{2}u^2 \,\mathrm{d}s\right]. \tag{6.10}$$

*subject to the dynamics* (6.4) *has the optimality system given by equations* (6.8)*,* (6.9) *with* $c = 0$*, initial density given by* $p(0,x)$ *and constraint* $\int p(t,x)\mathrm{d}x = 1$.

### 6.2.3 Stationary Solution

The FP equation governing the density of an overdamped Langevin system is called the Smoluchowski PDE. The FP PDE (6.7), 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 FP PDE can be obtained as a Gibbs distribution using this interpretation, under certain conditions on the fixed point pair $(v^\infty, \mathrm{p}^\infty)$ of the optimality system (6.6, 6.7) and the Langevin potential $\nu$. We denote $w(x) := \nu(x) + \frac{v^\infty(x)}{R}$.

**(A0)** There exist $(v^\infty(x), p^\infty(x)) \in (C^2(\mathbb{R}^d))^2$ satisfying (6.6,6.7) such that $\lim\limits_{|x|\to+\infty} w(x) = +\infty$ and $\exp\left(-\frac{2}{\sigma^2}w(x)\right) \in L^1(\mathbb{R}^d)$.

**Lemma 6.2.1.** *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* (6.7) *is*

$$\mathrm{p}^\infty(x) := \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}\Big(w(x)\Big)\right)(x), \tag{6.11}$$

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

*Proof.* We observe that the (6.7) is the Smoluchowski equation for an overdamped Langevin system given by

$$\mathrm{d}x_s = -\nabla(\nu + v^\infty/R)(x_s)\,\mathrm{d}s + \sigma \mathrm{d}w_s. \tag{6.12}$$

Under the assumptions above, the proof then follows directly from proposition 4.2, pp 110 in [83]. □

## 6.3 Schrödinger 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 [121] transform

$$\phi := \exp(-v/\sigma^2 R) \tag{6.13}$$

which was applied in stochastic control theory by Kappen [122]:

$$-\phi_t = -\frac{q\phi}{\sigma^2 R} - \phi_x \nu_x + \frac{\sigma^2}{2}\phi_{xx}. \tag{6.14}$$

The advection-diffusion equation above has a path integral solution [123] which is useful in computing the control [122, 33, 29]. 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 (6.8, 6.9) based on the spectral properties of a Schrödinger operator in section 6.4. 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 6.5.

### 6.3.1 Cole-Hopf Type transform

We introduce a Cole-Hopf type transform:

$$f(t,x) := \exp(-(v(t,x) + R\nu(x))/\sigma^2 R), \tag{6.15}$$

which leads to the following representation of equation (6.8):

$$-f_t = \frac{cf}{\sigma^2 R} - \frac{Vf}{\sigma^2 R} + \frac{\sigma^2}{2}\Delta f = \frac{cf}{\sigma^2 R} - Hf, \tag{6.16}$$

where we denote the modified cost function $V := q + (R/2)(\nabla\nu)^2 - (\sigma^2 R/2)\Delta\nu$ and the operator $H := \frac{V}{\sigma^2 R} - \frac{\sigma^2}{2}\Delta$ is a Schrödinger operator with potential $\frac{V(x)}{\sigma^2 R}$. The transformed PDE can be verified by using the calculations $\partial_t v = -\frac{\sigma^2 R \partial_t f}{f}$, $\nabla f = -\frac{f}{\sigma^2 R}\nabla(v + R\nu)$, $\Delta f = -\frac{\nabla f}{\sigma^2 R} \cdot \nabla(v + R\nu) - \frac{f}{\sigma^2 R}\Delta(v + R\nu)$ and $\frac{(\nabla v)^2}{2R} = \left(\frac{\sigma^4 R}{2}\left(\frac{\nabla f}{f}\right)^2 + \sigma^2 R\frac{\nabla f}{f} \cdot \nabla\nu + \frac{R}{2}((\nabla\nu)^2\right)$ in equation (6.8) and and recovering equation (6.16). Similarly, it can be shown that if $v(t, x)$ is a solution of equation (6.8) then $f(t, x)$ given by (6.15) is a solution to equation (6.16).

*Hermitizing* [19] the density as:

$$g := \frac{\mathrm{p}}{f}, \tag{6.17}$$

then gives the following representation of equation (6.9):

$$-g_t = -\frac{cg}{\sigma^2 R} + \frac{Vg}{\sigma^2 R} - \frac{\sigma^2}{2} g_{xx} = -\frac{cg}{\sigma^2 R} + Hg, \tag{6.18}$$

with the initial time boundary condition $g(0, x) = \frac{\mathrm{p}}{f}(0, x)$ and normalizing constraint $\int f(t, x)g(t, x)\mathrm{d}x = 1$ for all $t \geq 0$. This can be verified by using the derivatives $\partial_t p = \partial_t g f + g\partial_t f$, $\nabla p = f\nabla g + g\nabla f$, $\Delta p = f\Delta g + 2\nabla g \cdot \nabla f + g\Delta f$, $\nabla(\sigma^2 \ln f)p = \sigma^2 g f \frac{\nabla f}{f} = \sigma^2 g\nabla f$ and equation (6.16) in equation (6.9), thus recovering the equation above. Similarly, it can be shown that if $p(t, x)$ is a solution of equation (6.9) then $g(t, x) = \frac{\mathrm{p}}{f}$, with $f(t, x)$ given by (6.15), is a solution to equation (6.18). We summarize this fact in the following theorem.

**Theorem 6.3.1.** $(f(t, x), g(t, x))$ *is a solution to the linear PDE system (6.16, 6.18) such that* $\int f(t, x)g(t, x)\mathrm{d}x = 1$ *for all* $t \geq 0$ *if and only if*

$$v(t, x) = -\sigma^2 R \ln(f)(t, x) - R\nu(x) \tag{6.19}$$

$$p(t, x) = f(t, x)g(t, x) \tag{6.20}$$

*is a solution to the nonlinear optimality system (6.8, 6.9). Further, the optimal control is* *given by* $u^* = -\nabla v/R = \sigma^2 \nabla f/f$.

The introduced Cole-Hopf transform combined with hermitization of the density corresponds to a *diagonalization* of the coupled optimality system (6.8), (6.9) as follows:

$$\partial_t \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} H - \frac{c}{\sigma^2 R} & 0 \\ 0 & -H + \frac{c}{\sigma^2 R} \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}. \tag{6.21}$$

The diagonalization provides a linear representation of the FP PDE (6.9) which is not coupled with the HJB equation (6.8).

Analogously, it can be shown that the stationary value and density functions satisfying the stationary nonlinear optimality system (6.6, 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

$$He(x) = \frac{c}{\sigma^2 R} e(x) \tag{6.22}$$

subject to the normalizing constraint $\int f^\infty(x) g^\infty(x) \mathrm{d}x = 1$.

**Theorem 6.3.2.** $(f^\infty(x), g^\infty(x))$ *are both solutions to the eigenvalue problem* (6.22) *such that* $\int f^\infty(x) g^\infty(x) \mathrm{d}x = 1$ *if and only if*

$$v^\infty(x) = -\sigma^2 R \ln(f^\infty)(x) - R\nu(x) \tag{6.23}$$

$$p^\infty(x) = f^\infty(x) g^\infty(x) \tag{6.24}$$

*is a solution to the nonlinear optimality system (6.6, 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 (6.6, 6.7) it is possible to obtain explicit solutions to functions $(f^\infty, g^\infty)$ satisfying equation (6.22) such that $\int f^\infty g^\infty \mathrm{d}x =$

1. The result in theorem 6.2.1 and the introduced Cole-Hopf transform can be used to verify the following corollary to theorem 6.3.2.

**Corollary 6.3.2.1.** *Let* $p^\infty := \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}\left(w(x)\right)\right)(x)$ *with* $w(x) := \nu(x) + \frac{v^\infty(x)}{R}$ *and* $Z$ *the normalizing constant where* $(v^\infty, p^\infty)$ *is a pair satisfying* ***(A0)****. Then* $f^\infty := \sqrt{Zp^\infty}$ *and* $g^\infty := f^\infty/Z$ *both satisfy equation* (6.22) *such that* $\int f^\infty g^\infty \mathrm{d}x = 1$.

### 6.3.2 Nonlinear Schrödinger Equation and Mean Field Games for agents with Nonlinear Langevin Dynamics

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 $q = \bar{q}[p] := q(s, p(s,x))$. The mean-field time-varying optimality system [73] for this MFG is given by equations (6.8, 6.9) and $q = \bar{q}$.

In [19] 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 [20], is that this connection was shown only 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 preceeding discussion, it can be easily verified that using the transforms (6.15, 6.17), the corresponding MFG model constituted by the time-varying optimality system (6.8, 6.9) and $q = \bar{q}$ has the following NLS representation:

$$-f_t = \frac{cf}{\sigma^2 R} - \frac{\bar{V}[fg]f}{\sigma^2 R} + \frac{\sigma^2}{2}\Delta f \tag{6.25}$$

$$-g_t = -\frac{cg}{\sigma^2 R} + \frac{\bar{V}[fg]g}{\sigma^2 R} - \frac{\sigma^2}{2}g_{xx} \tag{6.26}$$

where $\bar{V}[fg] := \bar{q}[fg] + (R/2)(\nabla\nu)^2 - (\sigma^2 R/2)\Delta\nu$. Thus, we have generalized the connection between MFGs and the imaginary time NLS equation introduced in [19], to MFG models in which agent dynamics lie in the general class of nonlinear Langevin dynamics.

### 6.3.3 Interpretation

The Schrödinger potential $\frac{V(x)}{\sigma^2 R}$ 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 6.2.1:

$$\min_{\hat{u}\in\mathcal{U}} J(u) := \lim_{T\to+\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T V(\hat{x}_s)\mathrm{d}s + \frac{R}{2}\hat{u}^2 \,\mathrm{d}s\right] \tag{6.27}$$

subject to the simple integrator dynamics

$$\mathrm{d}\hat{x}_s = \hat{u}(s)\mathrm{d}s + \sigma \mathrm{d}w_s. \tag{6.28}$$

We refer to the OCP (6.27) subject to (6.28) as problem **(P2)**. The time-varying optimality system associated with problem **(P2)** is given by:

$$-\partial_t \hat{v} = V - \hat{c} - \frac{(\nabla \hat{v})^2}{2R} + \frac{\sigma^2}{2}\Delta \hat{v} \tag{6.29}$$

$$\partial_t \hat{p} = \nabla \cdot (\frac{\nabla \hat{v}}{R} p) + \frac{\sigma^2}{2}\Delta \hat{p} \tag{6.30}$$

where $\hat{c}$ is the optimal cost.

It is easily observed that if $v$ is the solution to the HJB equation (6.8), then $\hat{v} = v + R\nu$ is a solution to the HJB equation (6.29). 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 (6.30), we can see that the PDEs (6.9), (6.30) 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 6.2.1 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* (6.8)*,* (6.9) *corresponding to the OCP **(P1)** (subject to nonlinear passive dynamics) is equivalent to solving the optimality system* (6.29)*,* (6.30) *corresponding to the OCP **(P2)** (subject to simple integrator dynamics).*

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

$$V - \frac{(\nabla \hat{v}^{\infty})^2}{2R} + \frac{\sigma^2}{2}\Delta \hat{v}^{\infty} = 0 \tag{6.31}$$

$$\nabla \cdot (\frac{\nabla \hat{v}^{\infty}}{R} p^{\infty}) + \frac{\sigma^2}{2}\Delta \hat{p}^{\infty} = 0, \tag{6.32}$$

can be similarly shown to be connected to the solutions of the optimality system (6.6), (6.7) 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 $\hat{p}^{\infty}(0,x) = p^{\infty}(0,x)$. The steady state control $u^{\infty}$ of OCP (6.5), (6.4) and $\hat{u}^{\infty}$ of OCP (6.27), (6.28) are related as $\hat{u}^{\infty} = u^{\infty} - \nabla \nu$.

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 (6.4) with $u(s) = 0$. It can be proved that if $q(x) = 0$, then $\hat{u}^{\infty} = -\nabla \hat{v}^{\infty}/R = -\nabla \nu$ by verifying that $R\nu(x)$ is a solution to the stationary HJB equation (6.31). 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* (6.4) *with smooth Langevin potential* $\nu(x)$*, 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* $\hat{u}^{\infty}(x) = -\nabla \nu(x)$.

## 6.4 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 [86]. 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 (6.8, 6.9) 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 (6.9) is analyzed through evolution of the hermitized density (6.17) governed by equation (6.18). Diagonalization of the coupled PDEs constituting the optimality system as in equation (6.21) 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.

### 6.4.1 Perturbation System

Consider a controlled large-size non-interacting population expressed by problem **(P1)**, which is controlled by the optimal steady state control $u^{\infty} = -\nabla v^{\infty}/R$ corresponding to the optimality system (6.6, 6.7) with a unique fixed point $(v^{\infty}, p^{\infty})$ satisfying assumption **(A0)**. Theorem 6.3.2 implies that in this case the steady state value and density functions can be written as (6.23), (6.24), in terms of a pair of functions $(f^{\infty}, g^{\infty})$ both satisfying equation (6.22) and $\int f^{\infty} g^{\infty} \mathrm{d}x = 1$. Corollary 6.3.2.1 gives formulae for the function pair $(f^{\infty}(x), g^{\infty}(x))$ in terms of the steady state solution $(v^{\infty}, p^{\infty})$.Time varying value and density functions can be written as (6.19), (6.20) 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 (6.17) as $p(t,x) = p^{\infty}(x) + \tilde{p}(t,x) = f^{\infty}(x)g^{\infty}(t,x) + f^{\infty}(x)\tilde{g}(t,x)$. 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 $\tilde{g}(t,x)$ corresponds to a perturbation in the hermitized density given as $g(t,x) = g^{\infty}(x) + \tilde{g}(t,x)$, which obeys the time-varying PDE (6.18). In this section we study the decay of a perturbed density $p^{\infty} + \tilde{p}$ to its steady state density $p^{\infty}$. We state the following corollary to theorem 6.3.1 which provides the perturbation equation for the hermitized density $g(t,x)$.

**Corollary 6.4.0.1.** *If $g^\infty(x)$ is a solution to the stationary PDE* (6.22) *and $g(t,x) = g^\infty(x) + \tilde{g}(t,x)$ is a solution to the PDE* (6.18) *where $\tilde{g}(t,x) \in C^{1,2}([0,+\infty), \mathbb{R}^d)$, then $\tilde{g}(t,x)$ is governed by the linear PDE*

$$\tilde{g}_t = -H\tilde{g}. \tag{6.33}$$

### 6.4.2 Stability

We define the following Hilbert space and class of density perturbations for which we study stability.

**Definition 6.4.1.** *Let (A0) hold. Denote $p^\infty := \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}\left(w(x)\right)\right)(x)$ with $w(x) := \nu(x) + \frac{v^\infty(x)}{R}$ and $Z$ the normalizing constant where $(v^\infty, p^\infty)$ is the unique pair satisfying **(A0)**. We denote by $f^\infty := \sqrt{Zp^\infty}$ and $g^\infty := f^\infty/Z$ two solutions to equation* (6.22) *such that $\int f^\infty g^\infty \mathrm{d}x = 1$. We denote the Hilbert space of $L^2(\mathbb{R})$ by $\mathcal{H}$. The class of mass preserving density perturbations is defined as $S_0 := \left\{\pi(x) \in \mathcal{H} \,\middle|\, \langle \pi, f^\infty\rangle_{\mathcal{H}} = 0\right\}$.*

**Definition 6.4.2.** *We define the class of initial perturbed densities as $S := \Big\{p(0,x) = f^\infty(g^\infty(x) + \tilde{g}(0,x)$ We say that the fixed point $p^\infty(x) = f^\infty(x)g^\infty(x)$ of the nonlinear optimality system (6.6, 6.7) is asymptotically stable with respect to $S$ if there exists a solution $\tilde{g}(t,x)$ to the perturbation equation* (6.33) *such that $\lim\limits_{t\to+\infty} ||\tilde{g}(t,x)||_{\mathcal{H}} = 0$.*

**Lemma 6.4.1.** *If there exists a positive, even, continuous function $Q(x)$ on $\mathbb{R}$ which is non-decreasing for all $x \geq 0$ such that $\frac{V(x)}{\sigma^2 R} \geq -Q(x)$ for all $x \in \mathbb{R}$ and $\int \frac{\mathrm{d}x}{\sqrt{Q(2x)}}\mathrm{d}x = +\infty$ then the closure of $H$ is self adjoint.*

We omit the proof since it follows directly from theorem 1.1, pp 50 in [124]. In particular, if $\frac{V(x)}{\sigma^2 R} \geq \mathrm{k} \in \mathbb{R}$ then it follows that $H$ is self adjoint. The following assumption implies discreteness of the spectrum of $H$.

**(A1)** $\lim\limits_{|x|\to+\infty} V(x) = +\infty.$

**Lemma 6.4.2.** *If **(A1)** is true then the closure of $H$ has a discrete spectrum.*

The proof of this theorem follows from in theorem 3.1, pp 57 of [124]. This theorem implies that under assumption **(A1)**, the spectrum of $H$ denoted by $\{\lambda_n\}_{0\leq n\leq+\infty}$ has the property that $\lambda_n \to +\infty$ as $n \to +\infty$ and the corresponding eigenfunctions denoted as $\{e_n(x)\}_{0\leq n\leq+\infty}$ form a complete orthonormal system on $L^2(\mathbb{R})$. The eigenproperty is explicitly written as $He_n(x) = \lambda_n e_n(x)$. Further from proposition 3.2, pp 65 in [124] the eigenvalues have the property $\lambda_0 < \lambda_1 < \cdots < \lambda_n < \cdots$. We state the following assumption on the Schrödinger potential required to prove MF stability.

**(A2)** $V(x) \geq 0$.

**Theorem 6.4.3.** *Let **(A0, A1, A2)** be true true. Let $(v^\infty(x), p^\infty(x))$ be the unique stationary solution to the optimality system (6.6, 6.7) and denote by $(f^\infty, g^\infty)$ the two solutions to problem* (6.22) *given in corollary 6.3.2.1. If $\tilde{g}(0,x) \in S_0$ and $\{g_n\}_{0\leq n\leq+\infty}$ are determined by*

$$\dot{g_n}(t) = -\lambda_n t. \tag{6.34}$$

*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* (6.33)*. $p^\infty(x)$ is asymptotically stable with respect to $S(\epsilon)$.*

*Proof.* Since $\tilde{g}(0,x) \in \mathcal{H}$ we have the unique representation $\tilde{g}(t,x) = \sum_{n=0}^{+\infty} g_n(0)e_n(x)$ where $g_n(0) = \langle \tilde{g}(0,x), e_n(x)\rangle_{\mathcal{H}} < +\infty$ 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 (6.33) 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 (6.33) and using the eigenproperty $He_n = \lambda_n e_n$, we obtain the ODEs (6.34). Due to assumption **(A1)** the eigenproperties of the Schrödinger operator given in lemmas 6.4.1, 6.4.2 hold. Using the eigenproperty yields the ODEs (6.34) with the unique solutions $g_n(t) = g_n(0)\mathrm{e}^{-\lambda_n t}$. Therefore $\tilde{g}(t,x) = \sum_{n=0}^{+\infty} g_n(t)e_n(x)$ wherein $g_n(t) = g_n(0)\mathrm{e}^{-\lambda_n t}$, is the unique $\mathcal{H}$ solution to the perturbation equation (6.33). From the Krein-Rutman theorem [113] under the assumption that $V(x) \geq 0$ given by **(A2)**, the first eigenvalue is $\frac{c}{\sigma^2 R} = \lambda_0$ and the first eigenfunction is $0 < f^\infty(x) = e_0(x)$ corresponding to the eigenvalue problem (6.22). Further, $\tilde{g}(0,x) \in S_0$ implies that $g_0(0) = \langle \tilde{g}(0,x), e_0(x) \rangle_{\mathcal{H}} = \langle \tilde{g}(0,x), f^\infty(x) \rangle_{\mathcal{H}} = 0$ 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 $\langle He_0, e_0 \rangle_{L^2(\mathbb{R})} = \lambda_0 = \left\langle \frac{V}{\sigma^2 R} e_0, e_0 \right\rangle_{L^2(\mathbb{R})} + \frac{\sigma^2}{2} ||\partial_x e_0||^2$. 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 $||\tilde{g}(t,x)||_{L^2(\mathbb{R})} = \left( \sum_{n=0}^{+\infty} g_n(t)^2 \right)^{\frac{1}{2}}$, noting that $g_0(t) = 0$, $g_n(t)^2 = g_n(0)^2 \mathrm{e}^{-2\lambda_n t}$ 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 $S(\epsilon)$. □

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 6.1 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 (6.22) is solved using a spectral solver [112] for the parameters $\sigma = R = 1/2$ and the steady state density is obtained using equation (6.12). 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 each, using the steady state control. In the left panel we observe the density evolve over time steps $t = 0$ (black), $t = T/5$ (blue), $t = T/2$ (pink) to the final time $t = T$ (red) at which the density from the PDE computation is recovered.

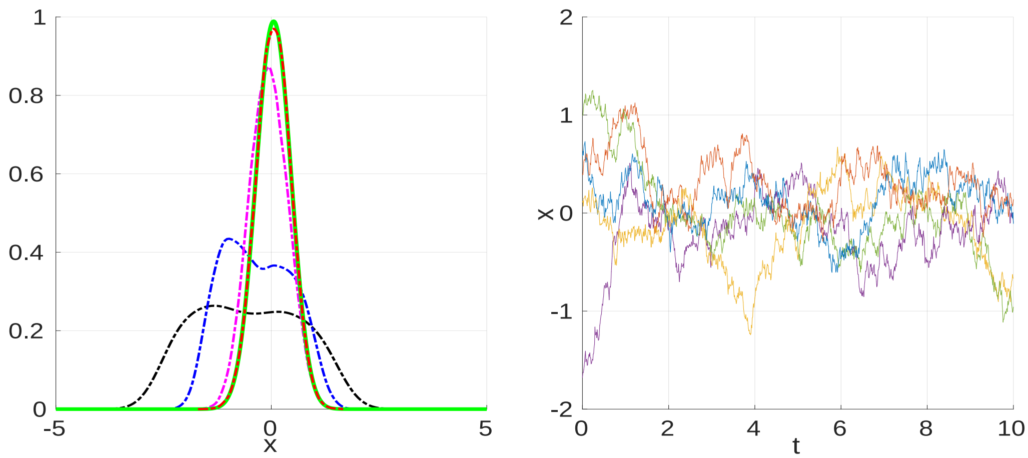



Figure 6.1: Stabilization of a density of agents with unstable passive dynamics to a fixed point density. (left) Density evolution of agents with closed loop dynamics over increasing times $t = 0$ (black), $t = T/5$ (blue), $t = T/2$ (pink), $t = T$ (red) to the steady state density (green) and (right) corresponding paths of ten agents.

## 6.5 Control Algorithm

For practical applications of the control of large-scale systems, it will be advantageous to precompute a finite time, feedback control law whose domain spans the region of state space that we are interested in. The optimal control is obtained can be obtained by solving the corresponding HJB PDE by various methods using finite difference, finite element or spectral approaches. In this work, we apply a path integral approach to solve this PDE in the finite time case and introduce an efficient quadrature method for evaluating the path integrals. Although our quadrature method could be applied to either **(P1)** or **(P2)**, the implementation becomes simpler in the case of **(P2)** due to the underlying integrator dynamics. The result is an efficient method for computing the feedback control law.

We consider the finite horizon OCP (6.10) with the HJB equation is given by (6.8), $c = 0$ as explained in remark 5. The optimal control can then be solved by treating the equivalent problem **(P2)** with HJB equation (6.16). The path integral representation of this

PDE (via Feynman-Kac) is as follows:

$$f(t,\bar{x}) = \mathbb{E}_\tau\left[\exp\Big(\int_t^T -\frac{V}{\sigma^2 R}(x_s)ds\Big)f(T,x_T)\right] \tag{6.35}$$

with the expectation over trajectories $\tau$ of brownian motions over the finite time horizon $[t, T]$, that is

$$dx_s = \sigma d\omega_s, \quad x_t = \bar{x} \tag{6.36}$$

First we approximate everything in discrete time with $N$ timesteps of duration $\delta t$, with $\delta t = (T-t)/N$, so that

$$f(t,\bar{x}) \approx \mathbb{E}_\tau\left[\exp\Big(\sum_{n=0}^{N-1} -\frac{V}{\sigma^2 R}(x_n)\delta t\Big)f(T,x_N)\right] \tag{6.37}$$

with $x_n$ governed by the discrete dynamical system:

$$x_{n+1} = x_n + \sigma\sqrt{\delta t}\epsilon, \quad x_0 = \bar{x}, \quad \epsilon \sim \boldsymbol{\mathcal{N}}(0, I) \tag{6.38}$$

with the associated transition probability $p(x_{n+1}|x_n) \sim \boldsymbol{\mathcal{N}}(x_n, \sigma^2\delta t\mathbf{I})$. Letting

$$w_n(x_n) := \exp\Big(-\frac{V}{\sigma^2 R}(x_n)\delta t\Big) \quad n = 0, \cdots, N-1 \tag{6.39}$$

$$w_N(x_N) := f(T, x_N) \tag{6.40}$$

$$w := \prod_{n=0}^{N} w_n(x_n) \tag{6.41}$$

from equation (6.37) we have

$$
\begin{aligned}
f(t,\bar{x}) &= \mathbb{E}_\tau\left[\prod_{n=0}^{N} w_n(x_n)\right] \\
&= \int\cdots\int w_N(x_N)\left[\prod_{n=2}^{N-1} w_n(x_n)p(x_{n+1}|x_n)\right]\times \\
&\left[\int w_0(\bar{x})p(x_1|x_0=\bar{x})w_1(x_1)p(x_2|x_1)dx_1\right]dx_2\cdots dx_N.
\end{aligned}
\tag{6.42}
$$

The second integral above is approximated by Gaussian quadrature with M grid points $\{\xi_1^i\}_{i=1}^M$ and weights $\alpha_1^i$ as

$$
\begin{aligned}
&\int w_0(\bar{x})p(x_1|x_0=\bar{x})w_1(x_1)p(x_2|x_1)dx_1 \approx \\
&\qquad \sum_{i=1}^{M} \underbrace{p(x_2|x_1=\xi_1^i)}_{\phi_1^i(x_2)}\underbrace{\alpha_1^i w_1(x_1=\xi_1^i)}_{\gamma_1^i}\underbrace{w_0(\bar{x})p(x_1=\xi_1^i|x_0=\bar{x})}_{\phi_0^i(\bar{x})}.
\end{aligned}
\tag{6.43}
$$

Defining the $M$ dimensional vectors $\Phi_1(x_2)$, $\gamma_1$, and $\Phi_0(\bar{x})$ to have elements $\phi_1^i(x_2)$, $\gamma_1^i$, $\phi_0^i(\bar{x})$, respectively and define $\Gamma_1 = diag(\gamma_1)$, (6.43) can be written as a set of vector products:

$$
\int w_0(\bar{x})p(x_1|x_0=\bar{x})w_1(x_1)p(x_2|x_1)dx_1 = \Phi_1(x_2)^\intercal\Gamma_1\Phi_0(\bar{x}). \tag{6.44}
$$

Recall that $p(x_1 = \xi_1^i|x_0 = \bar{x})$ is a Gaussian PDF, so each element of $\Phi_0(\bar{x})$ is Gaussian weighted by $w_0(\bar{x})$. Plugging this back into (6.42) yields:

$$
\begin{aligned}
&= \int\cdots\int w_N(x_N)\left[\prod_{n=3}^{N-1} w_n(x_n)p(x_{n+1}|x_n)\right] \\
&\quad \left[\int w_2(x_2)p(x_3|x_2)\Phi_1(x_2)^\intercal\Gamma_1\Phi_0(\bar{x})dx_2\right]dx_3\cdots dx_N.
\end{aligned}
\tag{6.45}
$$

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:

$$\int w_2(x_2)p(x_3|x_2)\Phi_1(x_2)^{\intercal}\Gamma_1\Phi_0(\bar{x})dx_2$$
$$\approx \sum_{i=1}^{M} \underbrace{p(x_3|x_2=\xi_2^i)}_{\phi^i(x_3)}\underbrace{\alpha_2^i w_2(x_2=\xi_2^i)}_{\gamma_2^i}\Phi_1(x_2=\xi_2^i)^{\intercal}\Gamma_1\Phi_0(\bar{x}) \quad (6.46)$$

Let $\tilde{\mathbf{\Phi}}_n$ be an $M \times M$ transition matrix with elements $\{\tilde{\mathbf{\Phi}}\}_{ij} = p(x_{n+1} = \xi_{n+1}^i | x_n = \xi_n^j)$. Then we can write (6.46) as:

$$\Phi_2(x_3)^{\intercal}\Gamma_2\tilde{\mathbf{\Phi}}_1\Gamma_1\Phi_0(\bar{x}) \quad (6.47)$$

Plugging this back into (6.45), 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:

$$f(t,\bar{x}) \approx \gamma_N^{\intercal}\left[\prod_{n=1}^{N-1}(\tilde{\mathbf{\Phi}}_n\Gamma_n)\right]\Phi_0(\bar{x}) \quad (6.48)$$

where we have used the definitions:

$$\gamma_n = \left[\{\alpha_i w_n(\xi_n^i)\}_{i=1}^M\right]^{\intercal} \quad (6.49)$$
$$\Gamma_n = diag(\gamma_n) \quad (6.50)$$
$$\{\tilde{\mathbf{\Phi}}_n\}_{ij} = p(x_{n+1} = \xi_{n+1}^i | x_n = \xi_n^j) \quad (6.51)$$
$$\phi_0^i(\bar{x}) = w_0(\bar{x})p(x_1 = \xi_1^i | x_0 = \bar{x}) \quad (6.52)$$
$$\Phi_0(\bar{x}) = \left[\{\phi_0^i(\bar{x})\}_{i=1}^M\right]^{\intercal} \quad (6.53)$$

Since $V(x)$ is time invariant and one chooses the same quadrature grid points at each timestep, $\gamma_n$ and $\tilde{\mathbf{\Phi}}_n$ are the same for all $n = 1, \cdots, N-1$. So (6.48) can be simplified to:

$$f(t,\bar{x}) \approx \gamma_N^{\intercal}(\tilde{\mathbf{\Phi}}\Gamma)^{N-1}\Phi_0(\bar{x}). \tag{6.54}$$

We demonstrate the resulting algorithm on a two dimensional problem where individuals obey dynamics (6.1) with $\nu(x) = 1/2\cos(x_1x_2)^2 - 1/24(x_1^4 + x_2^4)$ where $x = [x_1x_2]^{\mathrm{T}}$. In figure 6.2 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 $q(x) = \frac{1}{2}Q((x_1-1)^2 + (x_2-1)^2)((x_1+1)^2 + (x_2+1)^2)$ 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 6.3). 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.

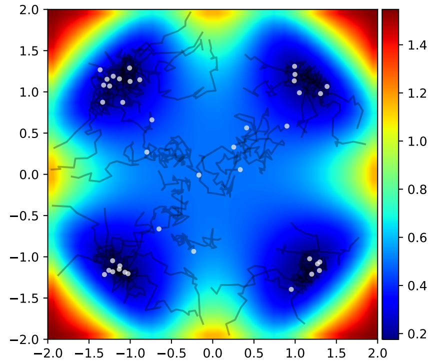



Figure 6.2: Plot of Langevin potential $\nu$ for 2 dimensional problem. x and y axes span $[-2, 2]$, and represent the state. Trajectories of 40 agents under no control (black lines) along with their final position after $T = 4.0$ seconds (white dots) are plotted. Note that the agents move into one of four potential wells.

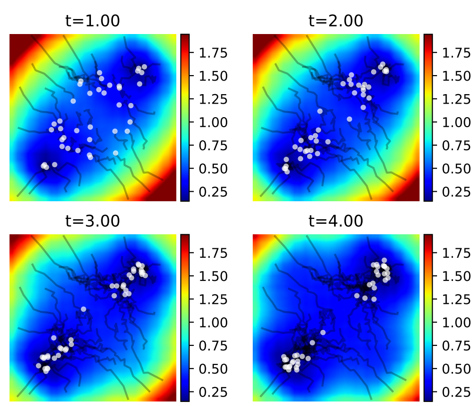



Figure 6.3: Plot of optimally controlled agents and value for 2 dimensional finite-horizon problem (T=4.0s). x and y axes span $[-2, 2]$, and represent the state. Color denotes plot of $\hat{v}(x, t) = -\sigma^2/2 \log(f(x, t))$. 4 snapshots in time are shown. Trajectories of 40 agents under the optimal control (black lines) along with their current positions (white dots) are plotted. Note that the agents move towards the regions of lowest cost at $(-1, -1)$ and $(1, 1)$ but are affected by the other potential wells at $(\pm 1, \pm 1)$. $f(x, t)$ is computed with our quadrature method on a fixed grid spanning the space.

### 6.6 Conclusions

In this chapter, we provide a framework for closed-loop stability analysis of the fixed point density of large-size populations in which agents obey multidimensional nonlinear Langevin dynamics. We utilize an imaginary time Schrödinger equation representation of the original optimality system 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. We provide explicit control design constraints which guarantee closed-loop stability of the steady state density using these spectral properties.

The corresponding Schrödinger potential is interpreted as the cost function of a related optimal control problem subject to simple integrator dynamics. This motivates a quadrature based algorithm to compute the time-varying control. 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.

In [20], the concept of solitons was used to study MFGs. The soliton theory in [20] for MFGs was based on this connection between NLS and MFGs for agents with simple integrator dynamics [20]. In section 6.3.2, 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. 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.

# CHAPTER 7

# CONCLUSIONS AND FUTURE WORK

The primary objective of this work is to develop methods and theoretical results aimed at constructing scalable control-theoretic frameworks for large-scale multi-agent systems and their control. There are several key challenges identified in the state-of-the-art controls literature related to this topic. Dynamics of individual agents can be uncertain or external disturbances, including non-Gaussian jump noise. The number of agents might be very large, to the order of $10^3 - 10^6$ agents, which render existing theoretical frameworks intractable. In the absence of a strong monotonicity assumption on the cost function, existing models exhibit non-unique solutions. Since agents are spread over a region of the state space and control algorithms must be capable of evaluating feedback controls over the entire region, this corresponds to a grid-based approach which leads to the *curse of dimensionality*. Finally, in an ensemble control setting, the relationship between the two fundamental principles of optimal control is not well understood.

We propose a PDE based approach to develop an ensemble control scheme for agents obeying a general class of marked jump diffusion dynamics. In chapter 3, a broadcast control algorithm is presented which uses a sampling scheme to compute the optimal cost-to-go on a grid. Since the control computation is computationally expensive, an iterative time-stepping strategy is used to reduce the computation time. Finally, the relationship between the dynamic programming principle and the infinite dimensional minimum principle is explained quantitatively.

In order to treat large-scale systems which use local feedback information, we take the a multi-agent systems approach in which we control a density of agents. This corresponds to the *mean-field* approach which is well known in the physics literature and enables us to synthesize and analyze several large-scale models using a non-cooperative, game-theoretic approach also referred to as mean field games (MFGs) models. The most alluring benefit of these models is that they can be used to represent large networks of rational agents. A topic of recent interest in the MFG literature is therefore, the stability properties of models which permit interaction among agents, specifically in cases with non-unique solutions. Most prior work on this topic consider agent dynamics which are very simple integrator systems. In chapter 4 we address the challenging problem of stability analysis and control design for MFG models in which agents obey nonlinear dynamics. Explicit control design constraints which guarantee stability are obtained for a consensus model and a population model. We also investigate the constraints for stability of the steady state controller, which is more pertinent to controls applications. In chapter 5 we present a MFG model for homogeneous flocking, in which agent interactions are non-local. This work recovers earlier known results on uncontrolled models of flocking as a special case. However, we observe interesting phase transitions in the synthesized controlled model which were not known previously.

A deep relationship between MFG models for agents with nonlinear Langevin dynamics and the Schrödinger equation is established in chapter 6. A novel Cole-Hopf transform is presented in order to make this connection and facilitate a sampling algorithm. When the agents do not interact explicitly through their dynamics or cost functions, a quadrature based sampling algorithm is proposed for computing the controls on a grid. Although this does not directly address the associated curse of dimensionality, it suggests a plausible approach to the challenging problem of large-scale control algorithms.

Synthesis and analysis of MFG models is presently a topic of great interest in control theory. Despite successful applications in some areas, the theory, modeling strategies and computation are far from being mature. Open problems include creation of fast, robust, and scalable numerical schemes for control computation, stability theory for models with general nonlinear and under-actuated agent dynamics, inverse MFG problems in connection with optimal transport and experimental validation. In the applied arena, data collection for accurate modeling opinion formation in social networks, finite dimensional or discrete action strategies for controlling socio-economical dynamics, say in energy consumption and election modeling, and mathematical modeling of irrational agents remain topics for future research.

# REFERENCES


[1] E. Todorov, "Classic maximum principles and estimation-control dualities for nonlinear stochastic systems," (Submitted), 2009.

[2] E. Theodorou, J. Buchli, and S. Schaal, "Reinforcement learning of motor skills in high dimensions: A path integral approach," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2010.

[3] E. Rombokas, M. Malhotra, E. Theodorou, E. Todorov, and Y. Matsuoka, "Reinforcement learning and synergistic control of the act hand," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 2, pp. 569–577, 2013.

[4] F. Stulp, E. Theodorou, and S. Schaal, "Reinforcement learning with sequences of motion primitives for robust manipulation," *IEEE Transactions on Robotics*, vol. 28, no. 6, pp. 1360–1370, 2012.

[5] K. Rawlik, M. Toussaint, and S. Vijayakumar, "On stochastic optimal control and reinforcement learning by approximate inference," *Robotics: Science and Systems*, 2012.

[6] R. Brockett, *Notes on the Control of the Liouville Equation*. Springer-Verlag Berlin Heidelberg, 2012.

[7] R. Brockett and N. Khaneja, *On the Stochastic Control of Quantum Ensembles. In: System Theory by Djaferis T.E., Schick I.C. (eds)*, ser. The Springer International Series in Engineering and Computer Science, vol 518. Springer, Boston, MA, 2000, vol. 518.

[8] A. Palmer and D. Milutinovic, "A hamiltonian approach using partial differential equations for open-loop stochastic optimal control," in *American Control Conference (ACC)*, 2011, pp. 2056–2061.

[9] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, and G. Wagner, "Optimal pulse design in quantum control: A unified computational method," *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1879–1884, 2011. eprint: `http://www.pnas.org/content/108/5/1879.full.pdf`.

[10] J. Li, I. Dasanayake, and J. Ruths, "Control and synchronization of neuron ensembles," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 1919–1930, 2013.

[11] N. Correll and A. Martinoli, "Towards multi-robot inspection of industrial machinery - from distributed coverage algorithms to experiments with miniature robotic swarms," *IEEE Robotics and Automation Magazine*, vol. 16, no. 1, pp. 103–112, 2009.

[12] D. Milutinović, *Utilizing Stochastic Processes for Computing Distributions of Large Size Robot Population Optimal Centralized Control*, ser. Springer Tracts in Advanced Robotics. Springer, Berlin, Heidelberg, 2013, vol. 83, ISBN: 978-3-642-32723-0.

[13] H. Li, C. Feng, H. Ehrhard, Y. Shen, B. Cobos, F. Zhang, K. Elamvazhuthi, S. Berman, M. Haberland, and A. Bertozzi, "Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment," in *IROS 2017 - IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2017-September, United States: Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 4341–4347.

[14] G. Y. Weintraub, C. L. Benkard, and B. V. Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," in *Advances in Neural Information Processing Systems*, 2005.

[15] M. Huang, P. E. Caines, and R. P. Malhame, "Large-population cost-coupled lqg problems with nonuniform agents: Individual-mass behavior and decentralized 949;-nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

[16] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[17] 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)*, Las Vegas, USA, 2016, pp. 1145–1152.

[18] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 920–935, 2012.

[19] U. D., S. I., and G. T., "Quadratic mean field games," *ArXiv:1708.07730*, 2017.

[20] I. Swiecicki, T. Gobron, and D. Ullmo, "Schrödinger approach to mean field games," *Phys. Rev. Lett.*, vol. 116, p. 128 701, 12 2016.

[21] O. Guéant, "A reference case for mean field games models," *Journal de Mathematiques Pures et Appliquees*, vol. 92, no. 3, pp. 276–294, 2009.

[22] 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*, vol. 59, no. 2, pp. 449–455, 2014.

[23] K. Bakshi, P. Grover, and E. A. Theodorou, "On mean field games with langevin dynamics," *Transactions on Control of Networked Systems (conditionally accepted)*, p. 061 103, 2018.

[24] M. Nourian, P. E. Caines, and R. P. Malhamé, "Mean field analysis of controlled cucker-smale type flocking: Linear analysis and perturbation equations," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 4471–4476, 2011.

[25] P. Degond, J.-G. Liu, and C. Ringhofer, "Large-scale dynamics of mean-field games driven by local nash equilibria," *Journal of Nonlinear Science*, vol. 24, no. 1, pp. 93–115, 2014.

[26] A. B. Barbaro, J. A. Cañizo, J. A. Carrillo, and P. Degond, "Phase transitions in a kinetic flocking model of Cucker–Smale type," *Multiscale Modeling & Simulation*, vol. 14, no. 3, pp. 1063–1088, 2016.

[27] H. J. Kappen, "Linear theory for control of nonlinear stochastic systems," *Phys Rev Lett*, vol. 95, p. 200 201, 2005, Journal Article United States.

[28] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, 1st, ser. Applications of mathematics. New York: Springer, 1993.

[29] 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, ISBN: 978-1-124-78907-1.

[30] K. L. Chung, *A course in probability theory*. Academic Press Inc., 1974, ISBN: 0-12-174650-X.

[31] J. Yong and X. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, 1999, ISBN: 0-387-98732-1.

[32] R. Bellman and R. Kalaba, *Selected Papers On mathematical trends in Control Theory*. Dover Publications, 1964.

[33] H. J. Kappen, "Path integrals and symmetry breaking for optimal control theory," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 11, P11011, 2005.

[34] I. Exarchos and E. Theodorou, "Learning optimal control via forward and backward differential equations," in *American Control Conference*, 2016.

[35] K. S. Bakshi, D. D. Fan, and E. A. Theodorou, “Stochastic control of systems with control multiplicative noise using second order fbsdes,” in *2017 American Control Conference (ACC)*, 2017, pp. 424–431.

[36] H. Pham, “On some recent aspects of stochastic control and their applications,” *Probability Surveys*, vol. 2, pp. 506–549, 2005.

[37] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation*. SIAM Books, 2007, pp. 4256–4261.

[38] R. Brockett, “New issues in the mathematics of control,” *Mathematics Unlimited: 2001 and Beyond*, vol. 87, 2001.

[39] ——, “On the control of a flock by a leader,” *Proceedings of the Steklov Institute of Mathematics*, vol. 268, pp. 49–57, 2010.

[40] H. Fattorini, *Infinite Dimensional Optimization and Control Theory*, ser. Encyclopaedia of Mathematics and Its Applications. Cambridge: Cambridge University Press, 1999.

[41] M. Krastanov, N. Ribarska, and T Tsachev, “A pontryagin maximum principle for infinite dimensional problems,” *SIAM Journal of Control Optimization*, vol. 49, pp. 2155–2182, 2011.

[42] T. D. Sanger, “Distributed control of uncertain systems using superpositions of linear operators.,” *Neural Computation*, vol. 23, no. 8, pp. 1911–1934, 2011.

[43] D. Milutinovic, “Utilizing stochastic processes for computing distributions of large size robot population optimal centralized control,” in *Proceedings of the 10 International Symposium on Distributed Autonomous Robotic Systems, Lausanne, Switzerland*, 2010.

[44] X. Zhou, “Maximum principle, dynamic programming and their connection in deterministic control,” *Journal of Optimization Theory and Applications*, vol. 65, pp. 363–373, 1990.

[45] R. F. Stengel, *Optimal control and estimation*, ser. Dover books on advanced mathematics. New York: Dover Publications, 1994.

[46] K. Doya, S. Ishii, and R. Pouget A. Rao, *Bayesian Brain: Probabilistic Approaches to Neural Coding*, 1st, ser. Computational Neuroscience. Cambridge: The MIT Press, 2007, ISBN: 0-262-04238-X.

[47] X. Zhou, "A unified treatment of maximum principle and dynamic programming in stochastic controls," *Journal of Optimization Theory and Applications*, vol. 65, pp. 363–373, 1990.

[48] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, ser. Forward-backward Stochastic Differential Equations and Their Applications no. 1702. Springer, 1999, ISBN: 9783540659600.

[49] H. Pham, "Feynman-Kac representation of fully nonlinear PDEs and applications," *Acta Mathematica Vietnamica*, vol. 40, pp. 255–269, 2015.

[50] Ł. Delong, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications: BSDEs with Jumps*, ser. EAA Series. Springer London, 2013, ISBN: 9781447153313.

[51] N. Framstad, B. Oksendal, and A. Sulem, "Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance," *Journal of Optimization Theory and Applications*, vol. 121, pp. 77–98, 2004.

[52] J. Shi, "Relationship between mp and dpp for the stochastic optimal control problem of jump diffusions," *Applied Mathematics and Optimization*, vol. 63, pp. 151–189, 2011.

[53] A. Deshpande, "Sufficient stochastic maximum principle for the optimal control of semi-markov modulated jump-diffusion with application to financial optimization," *Http://arxiv.org/abs/1407.3256*, vol. abs/1407.3256, 2014.

[54] M. Annunziato, A. Borzi, M. Magdziarz, and A. Weron, "A fractional fokker planck control framework for subdiffusion processes," *Optimal Control Applications and Methods*, 2015.

[55] G. Fabbri, F. Gozzi, and A. Swiech, *Stochastic Optimal Control in Infinite Dimensions: Dynamic Programming and HJB Equations*. Springer, Dec. 2016, vol. 37, pp. 290–304.

[56] M. Nisio, *Stochastic Control Theory: Dynamic Programming Principle*. Springer, Dec. 2015.

[57] E. Todorov, "Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system," *Neural Comput.*, vol. 17, no. 5, pp. 1084–1108, May 2005.

[58] E. Todorov, "Policy gradients in linearly-solvable mdps.," in *NIPS*, J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, Eds., Curran Associates, Inc., 2010, pp. 2298–2306.

[59] E. Theodorou and E. Todorov, “Stochastic optimal control for nonlinear markov jump diffusion processes,” in *The Proceedings of IEEE American Control Conference*, 2012, pp. 1633–1639.

[60] E. Pardoux and S. Peng, “Adapted solutions of backward stochastic differential equation,” *Systems and Control Letters*, vol. 14, pp. 55–61, 1990.

[61] J. Zhang, “A numerical scheme for BSDEs,” *Annals of Probability*, vol. 14, no. 1, pp. 113–147, 2001.

[62] B. Oksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, 3rd. Berlin ; New York: Springer, 2009.

[63] F. Hanson, *Stochastic processes and control for jump diffusions*, ser. Notes. IISC, 2007.

[64] M. Garroni and J. Menaldi, *Green functions for second order parabolic intego-differential problems*. Longman Scientific and Technical, 1992, ISBN: 05820215161.

[65] J. Bect, “A unifying formulation of fokker-planck-kolmogorov equation for general stochastic hybrid sytems,” *Nonlinear Analysis: Hybrid Systems*, vol. 4, pp. 357–370, 2010.

[66] C. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*. Spinger, 2004.

[67] J. Hespanha, “A model for stochastic hybrid systems with application to communication networks,” *Nonlinear Analysis*, vol. 62, pp. 1353–1383, 2005.

[68] D. Lanczos, *Linear Differential Operators*. Philadelphia, Pennsylvania: SIAM, 1961.

[69] G. B. J. Thomas and R. L. Finney, *Calculus and Analytic Geometry*. Addison-Wesley, 1996.

[70] R. Weinstock, *Calculus of Variations with Applications to Physics and Engineering*. Dover, 1952.

[71] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics*. Cambridge University Press, 1988.

[72] D. Kirk, *Optimal Control Theory: An introduction*. Mineola, NY: Prentice-Hall, 1970.

[73] A. Bensoussan, F. Jens, and P. Yam, *Mean Field Games and Mean Field Type Control Theory*. SpringerBriefs in Mathematics, 2013.

[74] P. Kachroo, S. Agarwal, and S. Sastry, “Inverse problem for non-viscous mean field control: Example from traffic,” *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3412–3421, 2016.

[75] R. Carmona, J.-P. Fouque, and L.-H. Sun, “Mean field games and systemic risk,” 2013.

[76] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, “Electrical vehicles in the smart grid: A mean field game analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1086–1096, 2012.

[77] P. Grover, K. Bakshi, and E. A. Theodorou, “A mean-field game model for homogeneous flocking,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 6, p. 061 103, 2018.

[78] D. Gomes, L. Nurbekyan, and M. Prazeres, “One-dimensional stationary mean-field games with local coupling,” *ArXiv:1611.08161*, 2016.

[79] M. Nourian, P. E. Caines, and R. P. Malhame, “Synthesis of cucker-smale type flocking via mean field stochastic control theory: Nash equilibria,” in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2010, pp. 814–819.

[80] M. Burger, M. D. Francesco, P. A. Markowich, and M.-T. Wolfram, “Mean field games with nonlinear mobilities in pedestrian dynamics,” *Discrete & Continuous Dynamical Systems - B*, vol. 19, no. 1531-3492-2014-5-1311, p. 1311, 2014.

[81] P. Chaudhari, A. Oberman, S. Osher, S. Soatto, and G. Carlier, “Deep relaxation: Partial differential equations for optimizing deep neural networks,” *Arxiv*, 2017.

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

[83] G. Pavliotis, *Stochastic Processes and Applications*, 1st. Springer, 2014.

[84] M. Bardi and I. Capuzzo Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, With Appendices by Maurizo Falcone and Pierpaolo Soravia, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston, 1997.

[85] V. Borkar, “Ergodic control of diffusions,” in *International Congress of Mathematicians*, vol. 3, 2016, pp. 1299–1309.

[86] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, ser. Springer Series in Synergetics 16. Springer-Verlag, 1984, ISBN: 3540130985.

[87] M. Cerfon, "Detailed balance and eigenfunction methods," *Applied Stochastic Analysis Lecture 12*, 2011.

[88] L. P. Guant O. Lasry JM., *Mean Field Games and Applications:Paris-Princeton Lectures on Mathematical Finance 2010*. Springer, Berlin, Heidelberg, 2011.

[89] A. Lachapelle and M. Wolfram, "On a mean field game approach modeling congestion and aversion in pedestrian crowds," *Transportation Research Part B: Methodological*, vol. 45, pp. 1572–1589, 10 2011.

[90] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Networks and Heterogeneous Media*, vol. 7, no. 1556-1801, p. 243, 2012.

[91] M. Abramoqitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graph and Mathematical Tables*. Dover, 1964.

[92] T. A. Driscoll, N. Hale, and T. L. N., *Chebfun Guide*. 2014.

[93] H. Resat, L. Petzold, and M. F. Pettigrew, "Kinetic modeling of biological systems," in, Springer, 2009, pp. 311–335.

[94] G. Beni, "From swarm intelligence to swarm robotics," in *International Workshop on Swarm Robotics*, Springer, 2004, pp. 1–9.

[95] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.

[96] G. Whitham, "On kinematic waves ii. a theory of traffic flow on long crowded roads," in *Proc. R. Soc. Lond. A*, The Royal Society, vol. 229, 1955, pp. 317–345.

[97] T. Stankovski, T. Pereira, P. V. McClintock, and A. Stefanovska, "Coupling functions: Universal insights into dynamical interaction mechanisms," *Reviews of Modern Physics*, vol. 89, no. 4, p. 045 001, 2017.

[98] M. Nourian, P. E. Caines, R. P. Malhame, and M. Huang, "Nash, social and centralized solutions to consensus problems via mean field control theory," *IEEE Transactions on Automatic Control*, vol. 58, no. 3, 2013.

[99] K. Elamvazhuthi and P. Grover, "Optimal transport over nonlinear systems via infinitesimal generators on graphs," *ArXiv preprint arXiv:1612.01193*, 2016.

[100] P. E. Caines, "Mean field games," *Encyclopedia of Systems and Control*, pp. 706–712, 2015.

[101] F. Cucker and S. Smale, “Emergent behavior in flocks,” *IEEE Transactions on automatic control*, vol. 52, no. 5, pp. 852–862, 2007.

[102] S.-Y. Ha, J.-G. Liu, *et al.*, “A simple proof of the cucker-smale flocking dynamics and mean-field limit,” *Communications in Mathematical Sciences*, vol. 7, no. 2, pp. 297–325, 2009.

[103] Y.-P. Choi, S.-Y. Ha, and Z. Li, “Emergent dynamics of the Cucker–Smale flocking model and its variants,” in *Active Particles, Volume 1*, Springer, 2017, pp. 299–331.

[104] A. B. Barbaro, J. A. Cañizo, J. A. Carrillo, and P. Degond, “Phase transitions in a kinetic flocking model of cucker–smale type,” *Multiscale Modeling & Simulation*, vol. 14, no. 3, pp. 1063–1088, 2016.

[105] Y. Kuramoto, “Self-entrainment of a population of coupled non-linear oscillators,” in *International symposium on mathematical problems in theoretical physics*, Springer, 1975, pp. 420–422.

[106] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag, “Synchronization of coupled oscillators is a game,” in *American Control Conference, 2010. ACC ’10.*, Submitted for publication, IEEE Trans. Auto. Control, 2010.

[107] T. J. Anastasio, A. K. Barreiro, and J. C. Bronski, “A geometric method for eigenvalue problems with low-rank perturbations,” *Royal Society open science*, vol. 4, no. 9, p. 170 390, 2017.

[108] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, ser. Texts in Applied Mathematics Science. Berlin: Springer-Verlag, 1990, vol. 2.

[109] J. Tugaut, “Phase transitions of McKean–Vlasov processes in double-wells landscape,” *Stochastics An International Journal of Probability and Stochastic Processes*, vol. 86, no. 2, pp. 257–284, 2014.

[110] P. Freitas, “Stability of stationary solutions for a scalar non-local reaction-diffusion equation,” *The Quarterly Journal of Mechanics and Applied Mathematics*, vol. 48, no. 4, pp. 557–582, 1995.

[111] T. Kapitula and K. Promislow, “Stability indices for constrained self-adjoint operators,” *Proceedings of the American Mathematical Society*, vol. 140, no. 3, pp. 865–880, 2012.

[112] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun guide*, 2014.

[113] E Todorov, "Eigenfunction approximation methods for linearly-solvable optimal control problems," in *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, 2009, pp. 161–168.

[114] J Alexander, R. Gardner, and C. Jones, "A topological invariant arising in the stability analysis of travelling waves," *J. reine angew. Math*, vol. 410, no. 167-212, p. 143, 1990.

[115] P. Cardaliaguet and S. Hadikhanloo, "Learning in mean field games: The fictitious play," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 23, no. 2, pp. 569–591, 2017.

[116] N. E. Leonard, "Multi-agent system dynamics: Bifurcation and behavior of animal groups," *Annual Reviews in Control*, vol. 38, no. 2, pp. 171–183, 2014.

[117] V. Srivastava and N. E. Leonard, "Bio-inspired decision-making and control: From honeybees and neurons to network design," in *American Control Conference (ACC), 2017*, IEEE, 2017, pp. 2026–2039.

[118] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Honeybee-inspired dynamics for multi-agent decision-making," *ArXiv preprint arXiv:1503.08526*, 2015.

[119] H. Z. Shui-Nee Chow Wuchen Li, "A discrete schrödinger equation via optimal transport on graphs," *ArXiv:1705.07583*, 2017.

[120] 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*, vol. 169, p. 671 691, 2016.

[121] E. Hopf, "The partial differential equation ut + uux = uxx," *Communications on Pure and Applied Mathematics*, vol. 3, no. 3, pp. 201–230, 1950. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160030302`.

[122] H. J. Kappen, "Linear theory for control of nonlinear stochastic systems," *Phys. Rev. Lett.*, vol. 95, p. 200 201, 2005.

[123] W. Fleming, "Exit probabilities and optimal stochastic control," *Applied Math. Optim*, vol. 9, pp. 329–346, 1978.

[124] F. Berezin and S. M.A., *The Schrödinger equation*, ser. Mathematics and Its Applications. Dordrecth, The Netherlands: Kluwer Academic Publishers, 1991.