Title

Connecting Mean-field Games and Generative Adversarial Networks

Author

Cao, Haoyang

Publication Date

2020

Peer reviewed|Thesis/dissertation

Connecting Mean-field Games and Generative Adversarial Networks

by

Haoyang Cao

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Industrial Engineering and Operations Research

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Xin Guo, Chair
Professor Ilan Adler
Professor Pieter Abbeel

Fall 2020

Connecting Mean-field Games and Generative Adversarial Networks

Abstract

Connecting Mean-field Games and Generative Adversarial Networks

by

Haoyang Cao

Doctor of Philosophy in Engineering – Industrial Engineering and Operations Research

University of California, Berkeley

Professor Xin Guo, Chair

The theory of mean-field games (MFGs) belongs to a branch of game theory that studies a large population of (weakly) interacting players. It serves as an analytically feasible framework to approximate stochastic differential games when the number of players is large and detailed characterization of the interactions is computationally expensive. As an effective modeling tool, the theory of MFGs attracts the attention of a variety of application fields in economics, finance and engineering. On the computation front, the development of machine learning provides abundant computational methods of solving for MFGs, which is remarkably meaningful in practice. At the same time, people may also wonder if the theory of MFGs, or stochastic analysis in general, could benefit the machine learning community.

This thesis starts with two MFG models, with singular and impulse types of controls, respectively. Theses two control types allow certain degrees of discontinuity, making them better mathematical models compared with regular controls where the interventions must be absolute continuous. However, due to the theoretical challenges brought by the discontinuous nature of the controls, these two models are less explored in existing literature compared with MFGs with regular controls. Both models are motivated by real-world problems. Explicit solutions to the MFGs are presented and shown to approximate Nash equilibria of the corresponding N-player games with an error of the order $O\left(\frac{1}{\sqrt{N}}\right)$. Further analysis of the solutions reveals the game effect from interacting with the mean-field.

Obtaining analytical solutions of MFGs is difficult in general. The thesis then turns to the computation side of MFGs and establish the connection with generative adversarial networks, a celebrated deep learning tool that enjoys tremendous empirical success since its introduction to the machine learning community. It first shows a conceptual connection between GANs and MFGs: MFGs have the structure of GANs, and GANs are MFGs under the Pareto Optimality criterion. Interpreting MFGs as GANs, on one hand, enables a GANs-based algorithm (MFGANs) to solve MFGs: one neural network (NN) for the backward HJB

equation and one NN for the forward FP equation, with the two NNs trained in an adversarial way. Viewing GANs as MFGs, on the other hand, reveals a new and probabilistic aspect of GANs. This new perspective, moreover, leads to an analytical connection between GANs and Optimal Transport (OT) problems, and sufficient conditions for the minimax games of GANs to be reformulated in the framework of OT. Numerical experiments demonstrate superior performance of this proposed algorithm, especially in higher dimensional case, when compared with existing NN approaches.

Finally, the thesis explores the possibility of enriching the theoretical understanding of the training of GANs from the perspective of stochastic analysis. It establishes approximations, with precise error bound analysis, for the training of GANs under stochastic gradient algorithms (SGAs). The approximations are in the form of coupled stochastic differential equations (SDEs). The analysis of the SDEs and the associated invariant measures yields conditions for the stability and the convergence of GANs training. Further analysis of the invariant measure for the coupled SDEs gives rise to a fluctuation-dissipation relations (FDRs) for GANs, revealing the trade-off of the loss landscape between the generator and the discriminator and providing guidance for learning rate scheduling.

To Hezhen Li and Lanzhang Zhao

Contents

List of Figures

List of Tables

Acknowledgments

I especially would like to thank my advisor, Prof. Xin Guo, for her tremendous support along the way. To me, she is such a great role model: intelligent, dedicated, and enthusiastic as a scholar and at the same time approachable, considerate and caring as an advisor. Not only she has been guiding me through my academic life, she has practically made her house my second home. For that I will always appreciate and be grateful.

I would like to thank my other committee members, Prof. Ilan Adler and Prof. Pieter Abbeel as well for constantly providing me with genuine help and insightful advice both for my research and for improving the presentation of this thesis. Besides, I deeply appreciate the kind help and guidance that Dr. Charles-Albert Lehalle from Capital Fund Management and Prof. Giorgio Ferrari from Bielefeld University have provided me during my PhD study.

I also would like to thank my collaborators for their precious contribution: Dr. Matteo Basei from EDF who was a postdoctoral fellow in our research group while we were working together, and Dr. Mathieu Laurière from Princeton University.

My life at Berkeley would not be as enjoyable as it is without my dearest friends and colleagues: Renyuan Xu, Xu Rao, Nan Yang, Xiaoli Wei, Anran Hu, Erik Bertelli, Yusuke Kikuchi, Haoyun Wu, Lidan Mu and so many more. I also would like to thank Dr. Mark Sumner, the director of UC Choral Ensembles, as well as the entire ensemble of UC Women's Chorale, with whom I have spent four happy years singing and performing.

Lastly, I would like to thank my parents for giving me their endless and unconditional love, trust and support, as they have always been doing in the past twenty-seven years of my life.

Chapter 1

Introduction

The pioneering works of Lasry and Lions[121] and Huang, Malhamè and Caines [102] lead to a rapid theoretical development of mean-field games (MFGs) studying decision making among a large population of players. There are three primary inspirations for MFGs (see [83]). The first is from particle physics, especially when the number of interacting particles in the system is so large that detailed modeling of all interactions becomes ineffective. The concept of mean-field serves as an approximation of such inter-particle interactions. The second is from game theory. MFGs are to approximate N-player games when N is large and exact characterization of an equilibrium becomes infeasible. The third is from economics. MFGs model general economic equilibrium among rational people where each individual only pays attention to her own interest and the market signals. The theory of MFGs has enjoyed tremendous growth. The existence of a solution and the relation between MFGs and the corresponding N-player games have been studied with both PDE approaches (see for instance [121, 40, 22]) and probabilistic methods (see for example [45, 46, 47, 118, 117]). Besides the development in theory, MFGs also attract the attention from various fields of applications, such as systemic risk problem [48], price impact problem [41, 4], optimal execution [103], algorithmic trading [51], portfolio management [120], growth theory [82], exhaustible resources [15, 83], power grid [10], production and inventory management [177], and so on.

While the past decade has witnessed a remarkable growth in the theory of MFGs, a majority of the theoretical results are established under regular control regime where the intervenes are absolutely continuous. For many engineering and economic problems, however, the interventions are not absolutely continuous and sometimes even discontinuous. Moreover, the derivation of an analytical form of MFG solutions largely remains a challenge beyond the linear-quadratic settings or other special cases. It calls for novel computational methods for MFGs, especially those incorporating machine learning techniques. Among the various computational tools in machine learning community, GANs are most active fields and have been enjoying great empirical success. Nonetheless, there are well recognized issues in GANs training, such as the vanishing gradient when the discriminator significantly outperforms the generator [5], the mode collapse which is believed to be linked with gradient exploding [152], and the challenge of GANs convergence [12].

This thesis will first develop the theoretical results of MFGs beyond the regular control paradigm, then propose a novel computational method for MFGs using generative adversarial networks (GANs) and discuss the conceptual connection between GANs and MFGs, and finally, analyze the convergence of GANs training via an approximation by stochastic differential equations.

In particular, Chapter 2, based on [38], will analyze a class of infinite-time-horizon MFGs with singular controls motivated from the partially reversible problem. It will first establish the existence of a solution when controls are of bounded velocity adapting the method of relaxed control from [118] to a new topological space. It then will provide an explicit solution when the controls are of finite variation, present sensitivity analysis sensitivity analysis to compare the MFG solution with that of the single-agent control problem and establish its approximation to the corresponding N-player game in the sense of ϵ-NE, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.

Chapter 3, based on [14], will discuss a general class of nonzero-sum N-player stochastic games with impulse controls, where players control the underlying dynamics with discrete interventions. It will adopt a verification approach and provide sufficient conditions for the Nash equilibria (NEs) of the game. It will then study the limit situation of $N \to \infty$, that is, an MFG with impulse controls and show that under appropriate technical conditions, the existence of unique NE solution to the MFG, which is an ϵ-NE approximation to the N-player game, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$. It will analyze in details a class of two-player stochastic games which extends the classical cash management problem to the game setting. In particular, we present numerical analysis for the cases of the single player, the two-player game, and the MFG, showing the impact of competition on the player's optimal strategy, with sensitivity analysis of the model parameters.

Chapter 4, based on [39], will first show a conceptual connection between GANs and MFGs: MFGs have the structure of GANs, and GANs are MFGs under the Pareto Optimality criterion. It will then present a novel GANs-based algorithm to compute MFG solutions based on the interpretation of MFGs as GANs: one neural network for the backward Hamilton-Jacobi-Bellman equation and one neural network for the forward Fokker-Planck equation, with the two neural networks trained in an adversarial way; numerical experiments will demonstrate superior performance of this proposed algorithm, especially in higher dimensional case, when compared with existing neural network approaches. Viewing GANs as MFGs, on the other hand, will reveal a new and probabilistic aspect of GANs. This new perspective, moreover, will lead to an analytical connection between GANs and Optimal Transport (OT) problems, and sufficient conditions for the minimax games of GANs to be reformulated in the framework of OT.

Finally, Chapter 5, based on [37], will establish approximations, with precise error bound analysis, for the training of GANs under stochastic gradient algorithms (SGAs). The approximations will be in the form of coupled stochastic differential equations (SDEs). The analysis of the SDEs and the associated invariant measures will yield conditions for the stability and the convergence of GANs training. Further analysis of the invariant measure for the coupled SDEs will give rise to a fluctuation-dissipation relations (FDRs) for GANs, revealing the trade-off of the loss landscape between the generator and the discriminator and providing guidance for learning rate scheduling.

1.1 Review on mean-field games

1.1.1 A simple example

We will use the following simple example from [83] to illustrate the essential ideas and technical components of MFGs. A meeting scheduled from time t often starts some time later than t, say T. The actual starting time T depends on the arrivals of participants. A rule is imposed saying that the meeting will start once a certain quorum is reached, i.e. the meeting starts at time t or after 90% of the participants have arrived, whichever is earlier. Assume the number of participants is so large that we consider them as a continuum of players. Players are rational, interchangeable and aware of all information provided. Let τ_i denote the time at which a representative player i decides to arrive, and $\tilde{\tau}_i$ denote time at which player i actually arrives, that is,

$$\tilde{\tau}_i = \tau_i + \sigma_i \epsilon_i$$

where $\sigma_i \epsilon_i$ is the uncertainty player i is subject to; the distribution of σ_i among population is m_0 and $\epsilon_i \sim N(0,1)$, i.i.d. across all players. Player i makes her decision upon minimizing the expectation of the total cost $\mathbb{E}[c(t,T,\tilde{\tau}_i)]$ where

$$c(t,T,\tilde{\tau}_i) = c_1(t,T,\tilde{\tau}_i) + c_2(t,T,\tilde{\tau}_i) + c_3(t,T,\tilde{\tau}_i)$$

consists of $c_1(t,T,\tilde{\tau}_i) = \alpha[\tilde{\tau}_i - t]_+$ the penalty on lateness compared to the scheduled time t, $c_2(t,T,\tilde{\tau}_i) = \beta[\tilde{\tau}_i - T]_+$ the penalty on lateness compared to the actually time T, and finally $c_3(t,T,\tilde{\tau}_i) = \gamma[T - \tilde{\tau}_i]_+$ the cost of inconvenience due to the waiting time.

The game component of this problem, in contrast to a standard optimization problem, is the interaction among players due to the actual starting time T in the decision making process: on one hand, for player i, her optimal plan τ_i depends on T through the cost function; on the other hand, once all the players make their decisions on their planned arrival times τ_i's, their actual arrival times $\tilde{\tau}_i$'s may collectively shift the actual starting time T through the rule of 90% threshold. The stochastic nature of T comes from the randomness of $\tilde{\tau}_i$ of all the players. Here, T plays the role of the *mean-field* and this problem is a mean-field game.

To solve this mean-field game, one can adopt the following fixed point approach. First, fix an arbitrary start time T and solve the optimization problem

$$\min_{\tau_i} \mathbb{E}[c(t,T,\tilde{\tau}_i)].$$

Second, update T to $\bar{T}$ according to the 90% rule as well as the optimal

$$\tau_i^* = \tau_i^*(t,T) = \arg\min_{\tau_i} \mathbb{E}[c(t,T,\tilde{\tau}_i)]$$

from the previous step. These two steps characterize the following mapping,

$$\Gamma : T \xmapsto{\text{optimize}} \{\tau_i^*(\cdot; T)\}_i \xmapsto{\text{i.i.d. noise } \epsilon_i} \{\tilde{\tau}_i^*(\cdot; T)\}_i \xmapsto{\text{90\% quantile rule}} T^*.$$

Given proper conditions on α, β and γ, *Banach fixed point theorem* cab be applied to Γ to find the fixed point of the mapping denoted by T^*. Then the pair $(\tau_i^*(\cdot, T^*), T^*)$ is the solution to the mean-field game.

This simple example characterizes three key aspects of MFGs. The first one is that the interaction lies between each individual player and the mean-field as illustrated above. The second aspect is the simplification from aggregation, highlighted during the step of updating T: instead of considering the order statistics of N random arrival times $\{\tilde{\tau}_i^*\}_i$, one can simply apply a quantile rule to one particular $\tilde{\tau}_i^*$ due to the assumption of indistinguishable players. The last aspect is the use of fixed point theorems which plays a central role in many of the prevailing methodologies of studying MFGs. Next in Section 1.1.2, some of these methodologies to establish the existence of an MFG solution will be reviewed under a more general MFG setting.

1.1.2 Existence of MFG solutions

A standard continuous-time MFG on $\mathbb{R}^d$ over a time horizon $[0, T]$ takes the following form,

$$v(s, x) = \inf_{\{\alpha_t\}_{t\in[0,T]}\in\mathcal{A}} \mathbb{E}\left[\int_s^T f(t, X_t, \mu_t, \alpha_t)dt + g(X_T, \mu_T)\middle| X_s = x\right] \tag{1.1}$$

subject to

$$dX_t = b(t, X_t, \mu_t, \alpha_t)dt + \sigma(t, X_t, \mu_t, \alpha_t)dW_t, \quad X_0 \sim \mu^0, \quad \mu_t = Law(X_t), \forall t \in [0, T].$$

Here μ^0 denotes the initial distribution of the state process with a density function $m^0(\cdot)$. $\{W_t\}_{t\geq 0}$ denotes a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. The flow of probability measures $\{\mu_t\}_t = \{Law(X_t)\}_t$ is the mean-field that denotes an aggregated status of all the players. It can appear in both the cost and the state dynamic, characterizing the interaction among players. For each individual player, her objective is to choose the optimal control $\{\alpha_t\}_t$ from a suitable admissible control set $\mathcal{A}$ to minimize the cost under the presence of the mean-field $\{\mu_t\}_t$.

Definition 1.1. *If there exists a control policy $\{\alpha_t^*\}_t$ and a flow of probability measures $\{\mu_t^*\}_t$ such that*

- *under $\{\mu_t^*\}_t$, $\{\alpha_t^*\}_t$ solves the optimal control problem*

$$v(s, x) = \inf_{\{\alpha_t\}_{t\in[0,T]}\in\mathcal{A}} \mathbb{E}\left[\int_s^T f(t, X_t, \mu_t^*, \alpha_t)dt + g(X_T, \mu_T^*)\middle| X_s = x\right]$$

subject to

$$dX_t = b(t, X_t, \mu_t^*, \alpha_t)dt + \sigma(t, X_t, \mu_t^*, \alpha_t)dW_t, \quad X_0 \sim \mu^0;$$

- *under* $\{\alpha_t^*\}_t$*, the controlled process* $\{X_t^*\}_t$ *given by*

$$dX_t^* = b(t, X_t^*, \mu_t^*, \alpha_t^*)dt + \sigma(t, X_t^*, \mu_t^*, \alpha_t^*)dW_t, \quad X_0^* \sim \mu^0$$

satisfies $\mu_t^* = Law(X_t^*)$ *for all* t*;*

then the pair $(\{\alpha_t^*\}_t, \{\mu_t^*\}_t)$ *is called a solution to the MFG* (1.1).

PDE approach

One of the central questions for MFGs is the existence of a solution. Generalizing the fixed point approach in Section 1.1.1 leads to a coupled PDE system. First, fix an arbitrary mean-field $\{\mu_t\}_t$ and solve the corresponding optimal control problem for an optimal control policy $\{\alpha_t^*\}_t$. This step gives rise to the HJB equation

$$\partial_s v(s,x) + H(s, x, \nabla_x v(s,x), \nabla_x^2 v(s,x)) = 0, \quad v(T,x) = g(x, \mu_T), \tag{1.2}$$

where the Hamiltonian is given by

$$H(s,x,y,z) = \inf_{\alpha} \left\{ b(s,x,\mu_s,\alpha) \cdot y + \frac{1}{2} Tr(\sigma\sigma^T(s,x,\mu_s,\alpha)z) + f(s,x,\mu_s,\alpha) \right\}.$$

Denote the optimal control by

$$\alpha_s^* = \arg\min_{\alpha} \left\{ b(s,x,\mu_s,\alpha) \cdot \nabla_x v(s,x) + \frac{1}{2} Tr(\sigma\sigma^T(s,x,\mu_s,\alpha)\nabla_x^2 v(s,x)) \right\}.$$

Next, the mean-field $\{\mu_t\}_t$ is updated for the controlled dynamic $\{X_t\}_t$ under the optimal policy $\{\alpha_t^*\}_t$. This gives rise to the following FP equation for the corresponding density function $m(s,x)$ of $\mu_s = Law(X_s)$,

$$\begin{aligned} &\partial_s m(s,x) + div_x\left(b(s,x,\mu_s,\alpha_s^*)m(s,x)\right) - \frac{1}{2} Tr\left[\nabla_x^2\left(\sigma\sigma^T(s,x,\mu_s,\alpha)m(s,x)\right)\right] = 0, \\ &\int_x m(s,x)dx = 1, \forall s; \quad m(0,\cdot) = m^0(\cdot). \end{aligned} \tag{1.3}$$

The coupled HJB-FP system characterizes a mapping

$$\Gamma : \{\mu_t\}_t \xmapsto{\text{HJB}} \{\alpha_t^*\}_t \xmapsto{\text{FP}} \text{updated } \{\mu_t\}_t.$$

If Γ admits a fixed point $\{\mu_t^*\}_t$, then $\{\mu_t^*\}_t$, together with its corresponding optimal control $\{\alpha_t^*\}_t$, will be a solution to the MFG in the sense of Definition 1.1. There are several fixed point theorems to be adopted under suitable circumstances. Below is one of them that is commonly used in the analysis of PDEs [74].

Proposition 1.1 (Schaefer's Fixed Point Theorem)**.** *Suppose a mapping* $A : X \to X$ *is continuous and compact, i.e. for any sequence* $\{u_k\}_{k=1}^{\infty}$ *in* X*.*

- $\lim_{k\to\infty} A(u_k) = A(u) \in X$ *if* $\lim_{k\to\infty} u_k = u \in X$*;*
- $A(u_k)_{k=1}^{\infty}$ *is precompact in* X *if* $\{u_k\}_{k=1}^{\infty}$ *is bounded in a real Banach space* X*.*

If the set

$$\{u \in X | u = \lambda A(u) \text{ for some } \lambda \in [0,1]\}$$

is bounded, then A *admits a fixed point.*

The pioneering work [121] specifies the problem setting as follows.

- the drift $b(s, x, \mu, \alpha) = -\alpha$ and the volatility $\sigma(s, x, \mu, \alpha) \equiv \sigma$ for some given σ;
- the running cost
$$f(s, x, \mu_s, \alpha) = L(x, \alpha) + \tilde{f}(s, x, m(s, x)) \tag{1.4}$$
and the terminal cost $g(x, \mu_T) = \tilde{g}(x, m(T, x))$.

Under this setting, [121] shows the existence of a MFG solution under proper technical assumptions. There, the Hamiltonian is simplified as

$$H(s, x, y, z) = \frac{1}{2} Tr(\sigma\sigma^T z) + \tilde{f}(s, x, m(s, x)) - \bar{H}(x, y) = \tilde{f}(s, x, m(s, x)) - \sup_{\alpha}\{y \cdot \alpha - L(x, \alpha)\}.$$

Consequently, the HJB becomes

$$-\partial_s v(s, x) - \frac{1}{2} Tr(\sigma\sigma^T \nabla_x^2 v(s, x)) + \bar{H}(x, \nabla_x v(s, x)) = \tilde{f}(s, x, m(s, x)), \quad v(T, x) = \tilde{g}(x, m(T, x))$$

with $\alpha_s^* = \arg\max_\alpha\{\alpha \cdot \nabla_x v(s, x) - L(x, \alpha)\}$ and the FP becomes

$$\begin{aligned} &-\partial_s m(s, x) + div_x(\alpha_s^* m(s, x)) + \frac{1}{2} Tr(\sigma\sigma^T \nabla_x^2 v(s, x)) = 0, \\ &\int_x m(s, x)dx = 1, \forall s; \quad m(0, \cdot) = m^0(\cdot). \end{aligned}$$

Theorem 1.2 (Existence [121])**.** *Let* $\mathcal{C}([0,T], \mathcal{P}^1(\mathbb{R}^d))$ *denote a set of continuous mappings from* $[0,T]$ *to the set of probability measures* $\mathcal{P}^1(\mathbb{R}^d)$ *where* $\int_{\mathbb{R}^d} |x|\mu(dx) < \infty$ *for any* $\mu \in \mathcal{P}^1(\mathbb{R}^d)$*. For any integer* $k \geq 0$ *and* $\alpha \in (0,1)$*,* $\mathcal{C}^{k,\alpha}$ *denotes the set of* $\mathcal{C}^k$ *functions with bounded and Hölder continuous of exponent* α *derivatives up to order* k*. Suppose that*

- *if the density function* $m(s, x)$ *belongs to* $\mathcal{C}([0,T], \mathcal{P}^1(\mathbb{R}^d))$*, then* $\tilde{f}(s, x, m(s, x))$ *and* $\tilde{g}(x, m(T, x))$ *are both (uniformly) bounded and continuous functions;*
- $\tilde{f}$ *is continuous with respect to the density function* $m \in \mathcal{C}([0,T], \mathcal{P}^1(\mathbb{R}^d))$*;*
- *if* $m \in \mathcal{C}^{k,\alpha}$*, then* $\tilde{f}, \tilde{g} \in \mathcal{C}^{k+1,\alpha}$*;*

- *$\bar{H}(x,y)$ is smooth on $\mathbb{R}^d \times \mathbb{R}^d$ and there exists a constant $c > 0$ such that for any (x,y),*

$$|\partial_y \bar{H}(x,y)| \le C(1+|y|), \quad |\partial_x \bar{H}(x,y)| \le c(1+|y|).$$

Then the HJB-FP system admits a pair of solutions $v^*(s,x)$ and $m^*(s,x)$ and the corresponding control-mean pair $(\{\alpha_t^*\}_t, \{\mu_t^*\}_t)$ is a solution of the MFG.

This method of HJB-FP system can also be seen in [83], [22], [46], [47] and the references within. Another PDE approach for MFGs is to use master equation, see for instance [40] and [42].

Uniqueness of a solution. [121] also provides monotonicity conditions to guarantee the uniqueness of an MFG solution. Similar conditions are also discussed in [45].

Theorem 1.3 (Uniqueness)**.** *For any $\mu, \mu' \in \mathcal{P}^2(\mathbb{R}^d)$, if for any $s \in [0,T]$, $\tilde{f}$ and g are monotone, i.e.,*

$$\int_{\mathbb{R}^d} [\tilde{f}(s,x,\mu) - \tilde{f}(s,x,\mu')](\mu-\mu')(dx) \ge 0, \quad \int_{\mathbb{R}^d} [g(x,\mu) - g(x,\mu')](\mu-\mu')(dx) \ge 0,$$

and the Hamiltonian $\bar{H}$ is strictly convex such that for any $x, y \in \mathbb{R}^d$,

$$H(x,y+z) - H(x,y) - z \cdot \partial_y H(x,y) = 0 \Rightarrow z \equiv 0;$$

or for any $s \in [0,T]$, $\tilde{f}$ and g are strictly monotone, i.e.,

$$\int_{\mathbb{R}^d} [\tilde{f}(s,x,\mu) - \tilde{f}(s,x,\mu')](\mu-\mu')(dx) \le 0 \Rightarrow \mu = \mu',$$
$$\int_{\mathbb{R}^d} [g(x,\mu) - g(x,\mu')](\mu-\mu')(dx) \le 0 \Rightarrow \mu = \mu',$$

then there exists at most one MFG solution.

Probabilistic approach

The existence of an MFG solution in [121] is subject to a problem setting where the drift of the state process is completely determined by the control process, the volatility is constant and the running cost is of a separable type as in (1.4). An immediate question is whether the existence still holds without such specifications. To allow a more general problem setup as in (1.1), [118] discusses the existence of a solution to MFGs from a probabilistic perspective, using the notion of relaxed controls from classical control theory [73, 97].

More precisely, it goes as follows. For any separable metric space (E,d), define $\mathcal{P}(E)$ the set of all probability measures on E and $\mathcal{P}^p(E) \subset \mathcal{P}(E)$ such that $\int_E d^p(x,x_0)\mu(dx) < \infty$ for any $\mu \in \mathcal{P}(E)$, equipped with a p-Wasserstein norm with

$$d_{E,p}(\mu,\nu) = \inf \left\{ \int_{E\times E} d^p(x,y)\gamma(dx,dy) | \gamma \in \mathcal{P}(E\times E) \text{ with marginal distributions } \mu,\nu \right\}^{\frac{1}{p}}.$$

Let $\mathcal{C}^d = \mathcal{C}([0,T],\mathbb{R}^d)$ denote the set of continuous process over $[0,T]$ on $\mathbb{R}^d$ equipped with a supremum norm $\|\cdot\|_T$. For $\mu \in \mathcal{P}(\mathcal{C}^d)$, let μ_t denote the image of μ under the mapping $x \in \mathcal{C}^d \mapsto x_t \in \mathbb{R}^d$. Any $\mu \in \mathcal{P}(\mathbb{R}^d)$ is equipped with a norm $\|\mu\| = (\int_{\mathbb{R}^d} \mu(dx)|x|^p)^{\frac{1}{p}}$ where $|\cdot|$ denotes Euclidean norm.

Definition 1.2 (Relaxed Control). *A relaxed control is a probability measure q on $[0,T]\times\mathcal{A}$ such that $q([s,t]\times\mathcal{A}) = t-s$ for any $0 \le s \le t \le T$ and $\int_{[0,T]\times\mathcal{A}} q(dt,da)|a|^p < \infty$ for some $p \ge 1$.*

Let $\mathcal{V}(\mathcal{A})$ be the set of all relaxed controls and $\Omega[\mathcal{A}]$ the new sample space $\mathcal{C}^d \times \mathcal{V}(\mathcal{A})$. For any $(x,q) \in \Omega[\mathcal{A}]$, denote canonical processes X and Λ on $\mathcal{C}^d$ and $\mathcal{V}(\mathcal{A})$, respectively.

Definition 1.3. *For a measure $\mu \in \mathcal{P}(\mathcal{C}^d)$, let $\mathcal{R}(\mu)$ be the set of admissible control-state joint laws $P \in \mathcal{P}(\Omega[\mathcal{A}])$ satisfying*

- $P \circ X_0^{-1} = \mu^0$;
- $\mathbb{E}^P\left[\int_0^T |\Lambda_t| dt\right] < \infty$;
- $M^{\mu,\phi} = \{M_t^{\mu,\phi}\}_{t\ge 0}$ *is a P-martingale for any smooth and compactly supported function ϕ, where*

$$M_t^{\mu,\phi}(q,x) = \phi(x_t) - \int_{[0,t]\times\mathcal{A}} q(ds,da)\left[b(s,x_s,\mu_s,a_s)\cdot\nabla_x\phi(x_s) + \frac{1}{2}Tr(\sigma\sigma^T(s,x_s,\mu_s,a_s)\nabla_x^2\phi(x_s))\right].$$

Define $\mathcal{R}^* : \mathcal{P}^p(\Omega[\mathcal{A}]) \to 2^{\mathcal{P}^p(\Omega[\mathcal{A}])}$ such that

$$\mathcal{R}^*(\mu) = \arg\min_{P\in\mathcal{R}(\mu)} \mathbb{E}^P\left[g(X_T,\mu_T) + \int_{[0,T]\times\mathcal{A}} q(dt,da) f(t,X_t,\mu_t,a_t)\right].$$

$P \in \mathcal{P}^p(\mathcal{C}^d)$ is called a relaxed MFG solution if $P \in \mathcal{R}^*(P\circ X^{-1})$; P is called a relaxed Markovian MFG solution if it is a relaxed MFG solution and $P(\Lambda = dt\tilde{q}(t,X_t)) = 1$ for some mapping $\tilde{q} : [0,T]\times\mathcal{A} \to \mathcal{P}(\mathcal{A})$; P is called a strict Markvian MFG solution if it is a relaxed MFG solution and $P(\Lambda = dt\delta_{\tilde{a}(t,X_t)}(da)) = 1$ for some mapping $\tilde{a} : [0,T]\times\mathbb{R}^d \to \mathcal{A}$.

Theorem 1.4. *Let $p_\sigma \in [0,2]$ and $p' > p \ge \max\{1,p_\sigma\}$. Suppose that*

- *$\mathcal{A}$ is a closed Euclidean space;*
- *b,σ,f,g are measurable in t and continuous in x,μ,a;*
- *b and σ are Lipschitz in x; for any $t\in[0,T]$, the growth rate of b with respect to x,μ,a is at most linear and the growth rate of σ with respect to x,μ,a is at most power of p_σ;*

- *The growth rate of g with respect to $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}^p(\mathbb{R}^d)$ is at most power of p; there exists $c_1, c_2 > 0$ such that for any $(t, x, \mu, a) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \times \mathcal{A}$,*

$$-c_2(1 + |x|^p + \|\mu\|^p + |a|^{p'}) \leq f(t, x, \mu, a) \leq c_2(1 + |x|^p + \|\mu\|^p) - c_3|a|^{p'};$$

- *$\mu^0 \in \mathcal{P}^{p'}(\mathbb{R}^d)$;*
- *for any $(t, x, \mu) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d)$,*

$$K(t, x, \mu) : - \left\{(b(t, x, \mu, a), \sigma\sigma^T)(t, x, \mu, a), z) : a \in \mathcal{A}, z \leq f(t, x, \mu, a)\right\}$$

  is a convex subset of $\mathbb{R}^d \times \mathbb{R}^{d \times d} \times \mathcal{P}^p(\mathbb{R}^d)$.

Then there exists a strict Markovian MFG solution equivalent to the one defined in Definition 1.1.

The idea of the proof starts from establishing the existence of a relaxed MFG solution, then moving on to a relaxed Markovian MFG solution following a mimicking result in [31] and finally, with a convexity condition, arriving at a strict Markovian MFG solution using a measure selection technique [97] which is equivalent to the one given by Definition 1.1. Fixed point theorems still lie at the center of this line of arguments, especially during the first step of showing the existence of a relaxed MFG solution. There, a set-valued mapping $F : \mathcal{P}^p(\mathcal{C}^d) \to 2^{\mathcal{P}^p(\mathcal{C}^d)}$ such that $F(\mu) = \{P \circ X^{-1} : P \in \mathcal{R}^*(\mu)\}$ is derived and therefore the Kakutani-Fan-Glicksberg theorem applies, see Theorem 1 in [75].

Another probabilistic approach to solve mean-field games is stochastic maximum principle, as introduced in [45]. In [45], the authors point out that in the solution to MFGs of the form of (1.1) on finite time horizon, it involves a Mckean-Vlasov type of control problem, where the dynamics of the state process is related to its own distribution, and derives a coupled forward-backward stochastic differential equation. The forward equation corresponds to the evolution of the state process; the backward equation is due to maximum principle and can be seen as the evolution of state derivative of the value function. This backward stochastic differential equation was systematically introduced in [147]. For the FBSDE approach, see also [20], [46], [22] and the references within. Note that though the controlled martingale approach can lead to the existence result under a generic problem setting, when deriving an analytical form of a MFG solution, methods of coupled PDE systems and FBSDEs are usually adopted as they lead to direct applicable forms of solutions in practice, see for instance [48, 20].

1.1.3 MFGs with singular controls

MFGs discussed in Section 1.1.2 assumes the continuity of the drift b and volatility σ with respect to the control α as well as the boundedness of $\mathcal{A}$, therefore the controlled state process is continuous. To characterize real-world problems where intervenes and state processes are mostly discontinuous, these assumptions of continuity need to be relaxed to incorporate certain level of discontinuity. It is then natural to consider MFG models with other types of controls.

The seminal work on fuel follower problem and its variants by Beneš, Shepp, and Witsenhausen [18] lays a foundation of studying singular control problems. It provides one of the principal approaches to solve singular control problems, that is, via the smooth fit principle. The simple and insightful solution structures have inspired many follow-up works in stochastic controls. See, for instance, [29], [71], [96], [107], [158], and [62]. Such problems have had a wide range of applications, including economics and finance [61], [157], [156], [104], [131] and [162], operations research [89], and queuing theory [168] and [8].

To illustrate the idea of singular control problems, consider a partially reversible investment problem in [91]. This control problem is formulated for a class of real option problems originated in the classical work of [66]. It is an optimization problem for a company whose revenue is based on the production level of a certain commodity, modeled by a geometric Brownian motion. The company can decrease its production level with a savage value and increase its production level with an investment cost, hence the term "partially reversible investment". That is, the dynamics of the production level at time t is given by

$$dX_t = X_t(\delta dt + \gamma dW_t) + d\xi_t, \quad X_{0-} \sim \mu_0,$$

where $\mu_0 \in \mathcal{P}^2(\mathbb{R})$, and the control ξ_t representing the cumulative change in the production level by time t is singular. The problem is to find an optimal investment strategy ξ_t over an appropriate control set in order to maximize its overall expected net profit

$$\mathbb{E}\left[\int_0^\infty e^{-rt}[\Pi(X_t)dt - \gamma^+ d\xi_t^+ - \gamma^- d\xi_t^-]\right]. \tag{1.5}$$

Here, the discount rate $r > 0$, $\Pi(\cdot)$ the revenue function satisfies the usual Inada condition for utility functions, γ^+ and γ^- are the unit costs of increasing and decreasing the production level respectively, subject to the assumption that $\gamma^+ + \gamma^- > 0$ for the well-posedness of the problem.

The corresponding Hamiltonian is given by

$$\begin{aligned} H(x,y) &= y \cdot \delta x + \sup_{\Delta\xi^\pm \in [0,\infty]} \{y(\Delta\xi^+ - \Delta\xi^-) - \gamma^+ \Delta\xi_t^+ - \gamma^- \Delta\xi_t^-\} \\ &= y \cdot \delta x + \sup_{\Delta\xi^\pm \in [0,\infty]} \{(y - \gamma^+)\Delta\xi_t^+ - (y + \gamma^-)\Delta\xi_t^-\}. \end{aligned}$$

Then, it is easy to see that H takes the value of either ∞ or $y \cdot \delta x$: for $y > \gamma^+$, the optimal control should be $\Delta\xi_t^+ = \infty$ and $\Delta\xi_t^- = 0$, with $H(x,y) = \infty$; for $y < -\gamma^-$, the optimal control should be $\Delta\xi_t^+ = 0$ and $\Delta\xi_t^- = \infty$, with $H(x,y) = \infty$; otherwise, the optimal control should be no invention and that leads to $H(x,y) = y\delta x$. This divergent Hamiltonian prevents the use of stochastic maximum principle. Alternatively, dynamic programming principle applies and leads to the following HJB equation with gradient constraints,

$$0 = \min\{rv(x) - \Pi(x) - \delta x v'(x) - \frac{1}{2}\gamma^2 x^2 v''(x),\, \gamma^+ - v'(x),\, v'(x) + \gamma^-\}, \tag{1.6}$$

where $v'(\cdot)$ and $v''(\cdot)$ denote the first and second order derivatives of $v(\cdot)$ respectively. In particular, the gradient constraints $\gamma^+ - v' \geq 0$ and $\gamma^- + v' \geq 0$ occur when it is optimal to make an instantaneous intervention. Considering an increase in the state by a small amount $\Delta > 0$, then by the optimality of the value function,

$$v(x) \geq v(x+\Delta) - \gamma^+\Delta.$$

Sending Δ to 0, the constraint $\gamma^+ - v' \geq 0$ naturally appears. The derivation of $\gamma^- + v' \geq 0$ is similar.

In [91], the smooth fit principle in the sense of [18] is established via regularity analysis for the value function, and the optimal control $\xi^*_\cdot = (\xi^{*,+}_\cdot, \xi^{*,-}_\cdot)$ to (1.5) is shown to be of bang-bang type characterized by a pair of threshold (x_b, x_s). It suggests that the company should spend the minimum effort to keep its production level within the interval $[x_b, x_s]$.

It is a natural extension to consider a game version of such singular control problems. In terms of MFGs with singular controls, there are extra technical challenges brought by the divergent Hamiltonian and the HJB equation with (possibly) state-dependent gradient constraints. To overcome these technical difficulties, [79] adopts the notion of relaxed controls and the techniques developed in [118] to prove the existence of solutions to MFGs with (monotone) singular controls over a finite-time horizon, with approximation analysis from MFGs with purely regular controls.

Still, very little is known on the solution structure of MFGs with singular controls, except for the recent work of [90]. They study MFGs of fuel follower problem and derive explicit solutions by exploiting symmetric structure in the cost functional. However, due to this symmetry, the optimal strategy for the MFG in [90] coincides with that for the single-agent control problem, i.e., the fuel follower problem in [18], with no demonstrated game effect.

Indeed, there are essential technical difficulties for deriving explicit solutions without certain symmetry structures in MFGs with singular controls. For instance, for a non-stationary MFG, the time-dependent mean information process leads to a parabolic HJB equation instead of an elliptic type, even in an infinite-time horizon game. This is different from classical control problems with infinite-time horizon. Moreover, the probabilistic approach of forward-backward stochastic differential equations (FBSDEs) does not work easily for the infinite-time horizon case.

In Chapter 2, the MFG counterpart of the partially reversible investment problem will be studied. This MFG with singular controls will be analyzed from two perspective: a general existence result will be established for MFGs with singular control of bounded velocity and a concrete example of MFG with singular control of finite variation will be analytically studied, where an explicit solution will be presented and analyzed. It is worth mentioning that, unlike [93], the analysis here does not rely on symmetry.

1.1.4 MFGs with impulse controls

Apart from singular controls, one can also consider control policies that solely consist of strategic jumps. This type of controls belong to impulse controls. The discontinuous nature of impulse controls lies in both the time and state spaces: interventions occur in a discrete-time fashion and each intervention shifts the state process instantaneously. Comparing to games with regular controls and singular controls, impulse control is a more natural mathematical framework for applied problems allowing for a discontinuous state space. See the examples of cash management [58], inventory controls [94, 95, 164], transaction cost in portfolio analysis [69, 136, 110, 111, 27, 146], insurance model [106, 34], liquidity risk [129], exchange rates [105, 138, 25], and real options [166, 130, 21].

One of the classical works on impulse control problems is [58] under the setting of a cash management problem. A manager possesses cash balance $X_t \in \mathbb{R}$ at time t. Without intervention, the amount of cash fluctuates along with continuous demand (withdrawal or deposit), i.e.

$$dX_t = \mu dt + \sigma dW_t, \quad X_{0-} \sim \mu^0,$$

where $\{W_t\}_{t\geq 0}$ a standard Wiener process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Denote the control policy as $\varphi = \{\tau_n, \xi_n\}_{n\geq 0}$, where $\{\tau_n\}_{n\geq 1}$ is a sequence of stopping times with respect to $\{\mathcal{F}_t\}$ such that $\tau_n \uparrow \infty$ almost surely, and the random interventions $\xi_n \in \mathcal{F}_{\tau_n}$. There $\mathcal{F}_{\tau_n}$ denotes the stopped filtration. Let $\mathcal{A}$ the set of all such controls. The objective of the manager is to minimize the total cost of managing her cash balance, i.e.,

$$v(x) = \inf_{\varphi\in\mathcal{A}} J(x,\varphi) = \inf_{\varphi\in\mathcal{A}} E\left[\int_0^\infty e^{-rt}C(X_t)dt + \sum_{n=1}^\infty e^{-r\tau_n}\phi(\xi_n)\middle| X_{0-} = x\right] \tag{1.7}$$

subject to

$$dX_t = \mu dt + \sigma dW_t + \sum_{n=1}^\infty \delta(t-\tau_n)\xi_n, \quad X_{0-} \sim \mu^0.$$

Here, the running cost is $C(x) = \max\{hx, -px\}$ where $h, p > 0$. Discount rate is $r > 0$. The cost of control is

$$\phi(x) = \begin{cases} K^+ + k^+x, & x \geq 0; \\ K^- - k^-x, & x < 0; \end{cases} \quad K^+, K^-, k^+, k^- > 0.$$

To distinguish zero control and no control, assume that $\phi(0) = K^+ > 0$. Due to the presence of the fixed cost, this cost of control possesses a crucial property, called the K-convexity. Intuitively, this property means that it is always better to shift the state in one move than to split it into two and the advantage is at least K.

The presence of discontinuity makes the analysis of impulse control problems hard and even harder for stochastic games. From a PDEs perspective, the corresponding quasi-variational inequality (QVI) contains an additional non-local operator for which most PDEs techniques are not applicable. Indeed, there was not much progress in the theory of impulse controls after Bensoussan and Lions' classical work [19], until the work of [92] where the non-local operator was found to be connected with the infinitesimal differential operator in the nonlinear PDEs via the payoffs in the action region and the waiting region. (See also [16].) [92] establishes the existence and uniqueness of the viscosity solution to the QVI with impulse control on multidimensional diffusion processes and discusses the legitimacy of applying the smooth fit principle to impulse control problems by analyzing the regularity of the value function.

The non-local operator $\mathcal{M}$ is defined as

$$\mathcal{M}v(x) = \inf_{\xi\in\mathbb{R}} v(x+\xi) + \phi(\xi),$$

and the QVI is given by

$$\min\{\frac{\sigma^2}{2}v'' - \mu v' - rv + C, \quad \mathcal{M}v - v\} = 0.$$

In [58], an optimal impulse control strategy is constructed by applying smooth fit principle on the QVI. In this case the policy is completely characterized by thresholds (d, D, U, u), meaning that if the cash balance is below d, then we raise it back to D; if it is above u then we decrease it to U. The authors also emphasize that if the proportional cost of decreasing cash balance is relatively high, then the cash management problem will degenerate to an inventory control problem. The discrete-time version of this inventory control problem was studied in [155] proposing the well-known (S, s) policy. [164] studies the continuous-time version of the inventory control problem and adopts the problem setting of the inventory control problem as in [58] and proves the uniqueness of the policy and provides the solution to the optimal S and s as the unique solution to a system of analytical equations. A similar work on continuous-time inventory control problem is done in [94].

Despite the rapid growth in recent literature on stochastic games and MFGs, the reservoir of works related to MFGs with impulse is rather limited. The technical difficulties regarding the approach HJB(QVI)-FP system are two-fold: in terms of QVI, the presence of the non-local operator brings in extra obstacle in solving for the value function under a generic problem setting; meanwhile, it still remains a challenging problem to characterize the controlled dynamic under this threshold type of impulse controls. [26] studies MFGs with impulse controls with finitely many possible choices of jumps. There, the FP equation for the controlled dynamics is established via a penalized problem.

In Chapter 3, the MFG counterpart of the cash management problem will be formulated and analyzed. Its game effect will be demonstrated through sensitivity analysis with respect to the model parameters.

1.1.5 Approximation of N-player games

One of the biggest contributions of MFGs is that they are approximations of the corresponding N-player games under the criterion of Nash equilibrium (NE). An NE of an N-player game is a set of strategies of all players from which no players has the incentive to unilaterally deviate. For more details, see Sections 2.2.5 and 3.1.

In terms of N-player stochastic differential games, some of the earliest works date back to 1970s, such as [169] and [167]. Both works consider regular type of controls. A majority of works focus on 2-player games. [23] and [78] consider 2-player stopping games, over finite and infinite time horizons, respectively. [115] studies a two-player game with singular control and optimal stopping. The authors defines strategies of a barrier type and Markovian perfect equilibrium (MPE), where the latter represents NE of Markovian type. The analysis is through verification theorem. The MPE is characterized and the conditions for the uniqueness of MPE are also provided. In [52] and [53], the authors describe a nonzero-sum N-player game with both regular and impulse control on a finite time-horizon. The timing of imposing impulse control is, however, predetermined. They use an approach combing backward stochastic differential equation and maximum principle. [2] proposes a nonzero-sum two-player impulse game where the control policy consists of both a sequence of stopping times and corresponding adjustment to make. In this paper, the authors provide a verification theorem and give an example with explicit solution. Note that in this example, two players are controlling the same one-dimensional dynamic from different directions. For 2-player impulse games, [59, 9] focus on zero-sum impulse games, [35] studies mixed impulse-stopping games, [76] analyzes nonzero-sum stochastic games involving impulse controls.

For a generic N, [48] studies an N-player game of systemic risk characterizing interaction among banks. It provides an explicit form of NE for an N-player game of systemic risk problem, under a linear-quadratic setting with regular controls. [120] also presents an explicit NE for a class of portfolio management games. The exponential and power utility functions are considered and control is again of regular type. [65] studies the a class of nonzero-sum submodular monotone follower games, an N-player game with singular controls. There the authors show the existence of an NE and also provide an approximation result from an N-player game with singular controls of bounded velocity to that of finite variation. [93] studies the N-player game version of the fuel follower problem in [18] and characterizes NE under a symmetric problem setting via analyzing its corresponding Skorokhod problem.

Solving for an explicit NE is quite challenging when N is large. Nonetheless, due to strong law of large numbers and propagation of chaos, starting from a solution to MFGs, one can construct a set of strategies for the N-player games such that the corresponding costs J^i's will be confined within a small neighborhood around the costs of an NE. This is the concept of ϵ-Nash equilibrium (ϵ-NE). For more details, see Sections 2.2.5 and 3.2.3. The error term ϵ depends on N and $\epsilon \to 0$ when $N \to \infty$. The magnitude of ϵ can be derived using strong law of large numbers and propagation of chaos according to the problem setting. For example, [45] shows that for MFGs with regular controls, the error term can be $\epsilon = O\left(N^{-\frac{1}{d+1}}\right)$; for a linear-quadratic case, [20] shows that the error term can be $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$. For MFGs with singular controls, [93] shows that $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$ under its symmetric problem setting.

For the cases where an explicit NE is available, the comparison between NE and the MFG solution con be conducted. For instance, [48] shows that the mean-field interaction create stability quantified by the systemic risk. [120] shows the impact of heterogeneity among players and the common noise in the solution structure of MFGs. In particular, without common noise or the heterogeneity, the mean-field interaction will be factored out of the optimization problem of individual players, and therefore the equilibrium strategy in MFGs solution coincides with the single-agent control problem case. [93] shows that due to the symmetric nature of its problem setup, the optimal strategy in the MFG solution coincides with that for the single-agent control problem, i.e., the fuel follower problem in [18], with no demonstrated game effect.

Chapter 2 will provide an explicit solution to the MFGs with singular control under a non-symmetric setting, analyze the difference between this MFG solution and its corresponding single-agent control problem, and study its relation with the associated N-player game. Chapter 3 will characterize the NE of an N-player game with impulse controls via a verification theorem. It will present multiple explicit NE to a 2-player game and finally, establish the approximation of the N-player games by the corresponding MFGs.

Discussion on the convergence of N-player games to MFGs. Apart from constructing approximated NE from MFG solutions, there is second angle in studying the relation between MFGs and N-player games, that is, to examine if the NE of an N-player game will eventually converge into a solution to the corresponding MFG.

[11] studies such a convergence under a linear quadratic setting. For N-player game, HJB-FP system is applied to characterize an NE where the optimal control is of linear feedback form. If assuming identical players, then the NE will eventually lead to a pair of quadratic-Gaussian solutions to the HJB-FP system that characterizes the MFG.

[40] studies the convergence of N-player games to MFGs through the master equation of the MFG. Compared with the value function as in (1.1) which is a function of time and state, a solution to the master equation also depends on the mean-field, which is the probability distribution of the state. The master equation is derived by generalizing Itô's formula, Feyman Kac formula and dynamic programming principle to functions of probability distribution. The NE of the corresponding N-player game is characterized by the solution to a system of HJB equations, called the Nash system. The convergence result relies on the uniqueness and the regularity conditions of both the solution to the master equation and the Nash system.

Following the idea of relaxed control introduced [118], [117] and [119] establish the convergence result for open- and closed-loop approximated NE towards a weak MFG solution, respectively, via a compactness argument; in addition, a Markovian projection method is adopted for the case of closed-loop equilibrium and the analysis relies on the application of Girsanov theorem. It is worth noticing that neither of the convergence results relies on uniqueness.

[142] analyzes the convergence of an N-player game of optimal stopping towards its MFG counterpart where solutions possess a transversality property and both NE and MFG solutions need not be unique. It also presents other classes of MFG solutions that cannot be limit points of the N-player game equilibria.

These existing works on the convergence problem largely focus on MFGs with regular controls. [93] studies the convergence of N-player game with singular controls towards its MFG counterpart through the action boundaries. It will be an interesting future direction to study this convergence issue under other singular control settings or impulse control setting, where the analytical tools become more restrictive.

1.1.6 Computation of MFGs

As shown in many existing works in MFGs literature, the establishment of analytical solutions still remains a challenge except for linear-quadratic structures or certain special cases. State processes in higher dimension will then bring in more difficulty. It is important to develop state-of-the-art computational methods. Most existing computational approaches for solving MFGs adopt traditional numerical schemes, such as finite differences [1] or semi-Lagrangian [44] schemes. Some exceptions are [50, 49] and [88]. [88] designs reinforcement learning algorithms with convergence and complexity analysis for learning MFGs, where the cost function of the game as well as the parameters for the underlying dynamics are unknown. [50, 49] propose deep neural networks (NNs) approaches for solving MFGs, with a particular Deep-Galerkin-Method architecture, to approximate the density and the value function by NNs separately.

Among all the computational tools from machine learning, generative adversarial networks (GANs) are particularly promising. Section 1.2 will review some basics of GANs to prepare us for the further discussions on the deeper connections between GANs, MFGs as well as other aspects of stochastic analysis in later chapters.

1.2 Review on generative adversarial networks

Since the introduction in 2014 [80], GANs have celebrated great empirical success, especially in image generation and processing. The key idea behind GANs is to interpret the process of generative modeling as a competing game between two neural networks: a generator network G and a discriminator network D. The generator network G attempts to fool the discriminator network by converting random noise into sample data, while the discriminator network D tries to identify whether the input sample is faked or true.

As minimax games, GANs provide a versatile class of generative models. Since the introduction to the machine learning community, the popularity of GANs has grown exponentially with numerous applications, including high resolution image generation [64, 149], image inpainting [181], image super-resolution [123], visual manipulation [184], text-to-image synthesis [150], video generation [173], semantic segmentation [128], and abstract reasoning diagram generation [114], and recently for simulating financial time-series data [175], [176], [182], and for asset pricing models [54].

Along with the empirical success of GANs, there is a growing emphasis on the theoretical analysis of GANs. [24] proposes a novel visualization method for the GANs training process through the gradient vector field of loss functions. In a deterministic GANs training framework, [132] demonstrates that regularization improved the convergence performance of GANs; [57]

and [68] analyze a generic zero-sum minimax game including that of GANs, and connect the mixed Nash equilibrium of the game with the invariant measure of Langevin dynamics.

This section will focus on some fundamental aspects of GANs.

1.2.1 Mathematical foundation of GANs

GANs fall into the category of generative models. The procedure of generative modeling is to approximate an unknown probability distribution $\mathbb{P}_r$ by constructing a class of suitable parametrized probability distributions $\mathbb{P}_\theta$. That is, given a latent space $\mathcal{Z}$ and a sample space $\mathcal{X}$, define a latent variable $Z \in \mathcal{Z}$ with a fixed probability distribution $\mathbb{P}_z$ and a family of functions $G_\theta : \mathcal{Z} \to \mathcal{X}$ parametrized by θ. Then $\mathbb{P}_\theta$ is defined as the probability distribution of $G_\theta(Z)$, i.e., $Law(G_\theta(Z))$.

As generative models, GANs consist of two competing neural networks: a generator network G and a discriminator network D. In GANs, the parametrized function G_θ is implemented using a neural network (NN), i.e., function approximations via specific graph structures and network architectures. Meanwhile, another neural network for the discriminator D will assign a score between 0 to 1 to the generated sample, either from the true distribution $\mathbb{P}_r$ or the approximated distribution $\mathbb{P}_\theta$; denote the parametrized D as D_ω. A higher score from the discriminator D would indicate that the sample is more likely to be from the true distribution. GANs are trained by optimizing G and D iteratively until D can no longer distinguish between samples from $\mathbb{P}_r$ or $\mathbb{P}_\theta$.

GANs as minimax games

Mathematically, GANs are minimax games as

$$\min_G \max_D \left\{\mathbb{E}_{X\sim\mathbb{P}_r}[\log D(X)] + \mathbb{E}_{Z\sim\mathbb{P}_z}[\log(1 - D(G(Z)))]\right\}. \tag{1.8}$$

Now, fixing G and optimizing for D in (1.8), the optimal discriminator would be

$$D^*_G(x) = \frac{p_r(x)}{p_r(x) + p_\theta(x)},$$

where p_r and p_θ are density functions of $\mathbb{P}_r$ and $\mathbb{P}_\theta = Law(G_\theta(Z))$ respectively. Plugging this back to Equation (1.8), we see

$$\begin{aligned}&\min_G \left\{\mathbb{E}_{X\sim\mathbb{P}_r}\left[\log \frac{p_r(X)}{p_r(X) + p_\theta(X)}\right] + \mathbb{E}_{Y\sim\mathbb{P}_\theta}\left[\log \frac{p_\theta(Y)}{p_r(Y) + p_\theta(Y)}\right]\right\}\\ &= -\log 4 + 2JS(\mathbb{P}_r, \mathbb{P}_\theta).\end{aligned}$$

That is, training of GANs with an optimal discriminator is minimizing Jensen-Shannon (JS) divergence between $\mathbb{P}_r$ and $\mathbb{P}_\theta$.

To address the instability of the vanilla GANs with JS divergence, variants of GANs with different divergences have been proposed to improve the performance of GAN training:

for instance, [141] uses f-divergence, [161] explores scaled Bregman divergence, [6] adopts Wasserstein-1 distance, [87] proposes relaxed Wasserstein divergence, and [153] and [154] utilize the Sinkhorn loss.

Equilibrium of GANs training

Under a fixed network architecture, the parametrized version of GANs training is to find

$$\begin{aligned} v_U^{GAN} &= \min_\theta \max_\omega L_{GAN}(\theta, \omega), \\ \text{where } L_{GAN}(\theta, \omega) &= \mathbb{E}_{X\sim\mathbb{P}_r}[\log D_\omega(X)] + \mathbb{E}_{Z\sim\mathbb{P}_z}[\log(1 - D_\omega(G_\theta(Z)))]. \end{aligned} \tag{1.9}$$

From a game theory viewpoint, the objective in (1.9), if attained, is in fact the upper value of the two-player zero-sum game of GANs.

Meanwhile, the lower value of the game is given by the following maximin problem,

$$v_L^{GAN} = \max_\omega \min_\theta L_{GAN}(\theta, \omega). \tag{1.10}$$

Clearly the following relation holds,

$$v_L^{GAN} \leq v_U^{GAN}. \tag{1.11}$$

Moreover, if there exists a pair of parameters (θ^*, ω^*) such that both (1.9) and (1.10) are attained, then (θ^*, ω^*) is a Nash equilibrium of this two-player zero-sum game. Indeed, if L_{GAN} is convex in θ and concave in ω, then there is no duality gap hence the equality in (1.11) holds by the minimax theorem (see [172] and [159]).

It is worth noting that conditions for such an equality in (1.11) is usually not satisfied in many common GANs models, as pointed out by [183].

SGD for GANs.

As in most NNs, stochastic gradient descent (SGD) is the standard approach for solving the optimization problem in GANs training. Accordingly, the evolution of parameters of θ and ω in (1.9) by SGD from current step t to the next step $t + 1$ is

$$\begin{aligned} \omega_{t+1} &= \omega_t + \alpha_d \nabla_\omega L_{GAN}(\theta_t, \omega_t), \\ \theta_{t+1} &= \theta_t - \alpha_g \nabla_\theta L_{GAN}(\theta_t, \omega_{t+1}). \end{aligned} \tag{1.12}$$

Here the α_d and α_g denote the step sizes of updating the discriminator and the generator, respectively.

This evolution (1.12) corresponds to the alternating updating scheme of the algorithm in [80] where at each iteration, the discriminator is updated before the generator. One of the main challenges for GANs training is the convergence of such an alternating SGD.

Later, Chapter 4 will discuss the profound connection between GANs and MFGs, as well as other aspects of stochastic analysis such as optimal transport (OT). A GANs-based algorithm for solving MFGs will be presented. Numerical experiments will demonstrates clear advantage of this variational approach, especially in terms of computational efficiency for high dimensional MFGs. Note that the idea of developing NN-based algorithm with incorporation of adversarial training is promising for more general dynamic systems with variational structures. In return, Chapter 5 will analyze the convergence of GANs training through establishing the approximation by stochastic differential equations.

Chapter 2

MFGs with singular controls

In this chapter, we analyze a class of infinite-time-horizon MFGs with singular controls, without symmetric cost structures. As in the classical singular control theory, we consider two types of singular controls: singular controls of bounded velocity and single controls of finite variation. It is generally believed that singular controls of bounded velocity are similar to regular controls and are easier to analyze. To our surprise, while we manage to establish the existence of MFGs solutions, we find it hard to derive them explicitly. In contrast, we are able to explicitly solve MFGs with finite-variation controls with a class of non-symmetric cost functional. The analytical solutions allow us to analyze in details the game effect with respect to model parameters.

In particular, we take the partially reversible investment model in [91], formulate its MFG counterpart, provide an analytical solution to the MFG and study the difference between this MFG with its corresponding single-agent control problem, as well as its relation with the associated N-player game.

In the MFG framework, instead of one company, we consider a continuum of infinitely many indistinguishable companies reacting to the market. We assume that the revenue function f is affected by the aggregated production level made by all the companies on the market, i.e., the game interaction among companies is through the revenue function f. We analyze this MFG from two aspects. First, we establish the existence of a solution to the MFG when controls are of bounded velocity. The approach is to adapt the technique by [118] to the infinite-time-horizon setting with an appropriate modification of topological spaces. Next, we analyze explicitly when the controls are of finite variation. We analyze explicitly this MFG, and compare it in details with the single-agent control problem when the revenue function is of the Cobb-Douglas type. In particular, we show that model parameters in the MFG impact both the optimal strategies (as in the single-agent case), and the equilibrium price. We then formulate the corresponding N-player game, and establish that this MFG solution is an approximation to the N-player game in the ϵ-NE sense, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.

2.1 MFGs with singular control of bounded velocity

2.1.1 Problem formulation and assumptions

In this section, we present the general mathematical framework of MFGs with singular controls of bounded velocity for the partially reversible investment problem introduced in Section 1.1.3.

Let $(\Omega, \mathcal{F}_t, P)$ be some filtered probability space supporting an $\mathcal{F}_t$-adapted standard Brownian motion W. Let $\mathcal{P}(\mathbb{R})$ be the set of probability measures over $\mathbb{R}$. For $p > 0$, denote $\mathcal{P}^p(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |x|^p \mu(dx) < \infty\}$. Fix $\theta > 0$. Let $A = [0, \theta] \times [0, \theta]$ be the control space, a compact subset of $\mathbb{R}^2$ with Euclidean norm $|\cdot|$. Then $[0, \infty) \times A$ contains the trajectories of the processes $\{(\dot{\xi}_t^+, \dot{\xi}_t^-)\}_{t \geq 0}$ such that we can define $\{\xi_t\}_{t \geq 0}$,

$$d\xi_t = \dot{\xi}_t dt = (\dot{\xi}_t^+ - \dot{\xi}_t^-)dt,$$

where the velocity of ξ stays within $[-\theta, \theta]$. That is, $\{\xi_t\}_{t \geq 0}$ is singular control of bounded velocity.

Given an initial distribution $\lambda \in \mathcal{P}^2(\mathbb{R})$, MFGs with singular controls of bounded velocity are defined as follows:

$$\begin{aligned} &\sup_{(\dot{\xi}^+, \dot{\xi}^-) \in \mathcal{U}_\theta} \mathbb{E}_\lambda \left[\int_0^\infty e^{-rt} \left[f(X_t, \mu_t) - \gamma^+ \dot{\xi}_t^+ - \gamma^- \dot{\xi}_t^- \right] dt \right] \\ &\text{subject to } dX_t = \left[b(X_t, \mu_t) + \dot{\xi}_t^+ - \dot{\xi}_t^- \right] dt + \sigma dW_t, \quad X_{0-} \sim \lambda, \\ &\qquad \mu_t = \mathbb{P}_{X_t}, \quad \forall t > 0. \end{aligned} \tag{MFG-BV}$$

Here, $\mathbb{P}_{X_t}$ is the probability distribution of X_t for any $t > 0$. The function $b : \mathbb{R} \times \mathcal{P}^1(\mathbb{R}) \to \mathbb{R}$ is the drift function and $\sigma > 0$ be the volatility coefficient. For the cost structure, let $f : \mathbb{R} \times \mathcal{P}^1(\mathbb{R}) \to \mathbb{R}$ be the running revenue function, γ^+ and γ^- be the proportional costs of per unit increase and decrease, respectively, with $\gamma^+ + \gamma^- > 0$, and finally $r > 0$ be the rate of discount.

The set of admissible controls $\mathcal{U}_\theta$ is given by

$$\begin{aligned} \mathcal{U}_\theta = \Big\{ &(\dot{\xi}^+, \dot{\xi}^-) : \dot{\xi}^+, \dot{\xi}^- \ \mathcal{F}_t\text{-adapted processes}, \\ &\dot{\xi}_t^+, \dot{\xi}_t^- \in [0, \theta], \forall t \geq 0, \mathbb{E} \left[\int_0^\infty e^{-rt} (|\dot{\xi}_t^+| + |\dot{\xi}_t^-|) dt \right] < \infty \Big\}. \end{aligned}$$

2.1.2 Assumptions and notation

We adapt the approach in [118] for finite-time horizon to our problem setting with an infinite-time horizon. The fundamental difference is that, due to the shift from finite to infinite time horizon, we will need different topologies and therefore many of the statements and proofs must be adjusted.

Throughout the discussion on the MFGs with singular controls of bounded velocity, the following assumptions hold.

Assumptions.

(A1) $b(x, \mu)$ and $f(x, \mu)$ are continuous, with b bounded and f nonnegative, concave and nonlinear in x under any fixed μ.

(A2) There exists some constant $c > 0$ such that for any $x, y \in \mathbb{R}$ and $\mu \in \mathcal{P}^1(\mathbb{R})$,

$$|b(x, \mu) - |b(y, \mu)| < c|x - y|,$$
$$|b(x, \mu)| \leq c(1 + |x| + \int_{\mathbb{R}} |z|\mu(dz)) = c(1 + |x| + |\mu|).$$

(A3) There exists some constant $c' > 0$ such that for any $x \in \mathbb{R}$ and $\mu \in \mathcal{P}^1(\mathbb{R})$,

$$|f(x, \mu)| \leq c'(1 + |x| + |\mu|).$$

With slight abuse of notation, for any $t \in [0, \infty)$, $x \in \mathbb{R}$, $\mu \in \mathcal{P}^1(\mathbb{R})$ and $a \in A$, define

$$\begin{aligned}
&b(x, \mu, a) = b(x, \mu) + a_1 - a_2, \\
&f(x, \mu, a) = f(x, \mu) - \gamma^+ a_1 - \gamma^- a_2; \\
&K(t, x, \mu) \equiv K(x, \mu) = \left\{ \left(b(x, \mu, a), \sigma^2, z\right) : \, a \in A, \, z \leq f(x, \mu, a) \right\}.
\end{aligned}$$

Remark 2.1. *It is easy to check the assumptions above, plus the facts that the volatility σ is a positive constant and there is no terminal cost, ensure Assumptions (A), (B) and (Convex) of [118] hold:*

Assumption (A): *the Lipschitz and growth rate conditions on b, σ and f;*

Assumption (Convex): *the convexity of the $K(t, x, \mu)$ of all (t, x, μ);*

Assumption (B): *the additional boundedness of b and σ as well as compactness of A.*

To facilitate our discussion on (MFG-BV) which is over an infinite-time horizon, we will introduce the following topological spaces.

Notations. To start, let $\mathcal{C} = \mathcal{C}\left([0, \infty), \mathbb{R}\right)$ denote the set of continuous functions from infinite time horizon $[0, \infty)$ to $\mathbb{R}$. For any $T \in (0, \infty)$, let $\mathcal{C}_T = \mathcal{C}\left([0, T], \mathbb{R}\right)$ denote the set of continuous functions from finite time horizon $[0, T]$ to $\mathbb{R}$, with a supremum norm $\| \cdot \|_T$,

$$\|y\|_T = \sup_{t \in [0,T]} |y_t|, \quad \forall y \in \mathcal{C}_T.$$

Define $\varphi_T : \mathcal{C} \to \mathcal{C}_T$ such that for any $x \in \mathcal{C}$, $y = \varphi_T(x) \in \mathcal{C}_T$ such that $y_t = x_t$ for all $t \in [0, T]$. That is, φ_T truncates any continuous trajectory on infinite-time horizon to that up to time T. Denote

$$d(x, x') = \sum_{T=1}^{\infty} \frac{1}{2^T} \frac{\|\varphi_T(x) - \varphi_T(x')\|_T}{1 + \|\varphi_T(x) - \varphi_T(x')\|_T}, \quad \forall x, x' \in \mathcal{C}. \tag{2.1}$$

Then $(\mathcal{C}, d)$ is a Polish space, see Section 1.3 in [163]. Denote $\mathcal{P}(\mathcal{C})$ the set of probability measures over $\mathcal{C}$ with topology of weak convergence. Then for any $\mu \in \mathcal{P}(\mathcal{C})$, μ_t is the image of μ under $x \mapsto x_t$. On top of $\mathcal{P}(\mathcal{C})$, define $\mathcal{P}^1(\mathcal{C}) \subset \mathcal{P}(\mathcal{C})$, where for $\mu \in \mathcal{P}^1(\mathcal{C})$, $\int_{\mathcal{C}} d(x, x^o)\mu(dx) < \infty$ for some $x^o \in \mathcal{C}$. Let $d_{\mathcal{C},1}(\mu, \nu)$ be the 1-Wasserstein distance between any $\mu, \nu \in \mathcal{P}^1(\mathcal{C})$, where

$$d_{\mathcal{C},1}(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{C}\times\mathcal{C}} \pi(dx, dx')d(x, x'), \tag{2.2}$$

with $\Pi(\mu, \nu)$ the set of all possible coupling of μ and ν. Note that the way in which the metric space $(\mathcal{P}^1(\mathcal{C}), d_{\mathcal{C},1})$ is defined can be generalized to other Polish space (E, ρ). For a fixed $T > 0$, given $p \geq 0$, write

$$\|\mu\|_T^p = \int_{\mathcal{C}} \|\varphi_T(x)\|_T^p \mu(dx).$$

2.1.3 Existence of solutions to (MFG-BV)

The key idea of analyzing (MFG-BV) is to consider a controlled martingale problem, as in [118]. The first step is to show the existence of a *relaxed* solution and then construct a *strict Markovian* solution based on the relaxed one.

MFGs with relaxed control

We first introduce the notion of relaxed controls.

Relaxed controls. Let $\mathcal{V}$ denote the collection of measures q on $[0, \infty) \times A$, with

$$q(dt, da) = dt[q_t](da)$$

for some $q_t \in \mathcal{P}(A)$, such that

$$\int_{[0,\infty)\times A} e^{-rt} q(dt, da)|a| = \int_{[0,\infty)\times A} e^{-rt} dt q_t(da)|a| < \infty. \tag{2.3}$$

Each $q \in \mathcal{V}$ is a *relaxed control*. From a game theory point of view, this can be interpreted as a mixed strategy in the sense that at any time point $t \geq 0$, the player chooses her action among A according to the probability distribution q_t. Given Equation (2.2), define the following metric for $\mathcal{V}$,

$$d_{\mathcal{V}}(q^1, q^2) = d_{[0,\infty)\times A,1}(re^{-rt}q^1, re^{-rt}q^2), \quad \forall q^1, q^2 \in \mathcal{V}. \tag{2.4}$$

Notice $re^{-rt}q(dt, da)$ for any $q \in \mathcal{V}$ is a probability measure on $[0, \infty) \times A$. Then $(\mathcal{V}, d_{\mathcal{V}})$ is a complete and separable metric space.

Remark 2.2. *Equation* (2.3) *implies* $\int_{[0,T]\times A} q(dt, da)|a| < \infty$ *for any* $T \in [0, \infty)$.

Remark 2.3. *To define $d_{\mathcal{V}}$ as in* (2.4)*, $\frac{q}{T}$ is used for finite-time horizon $[0,T]$ in [118]. That is, the time marginal is a uniform distribution. It is consistent with the definition of the objective functional on $[0,T]$. Since this chapter deals with discounted total payoff over an infinite-time horizon with a discount rate r, an exponential distribution is chosen instead.*

Remark 2.4. *If there exists some measurable $\alpha : [0,\infty) \to A$ such that $q_t = \delta_{\alpha_t}$ for all t, then such a q is called a* strict control. *It corresponds to a* pure strategy *in game theory. Note that if q is indeed a strict control for some measurable α, then Equation* (2.3) *becomes*

$$\int_0^\infty e^{-rt}|\alpha_t|dt < \infty,$$

restricting the control trajectory α to be of finite variation.

Denote $\Omega = \mathcal{V} \times \mathcal{C}$ and let $\mathcal{B}(\Omega)$ be its Borel σ-field. Define $\Lambda : \mathcal{V} \to \mathcal{V}$ and $X : \mathcal{C} \to \mathcal{C}$ such that $\Lambda(q) = q$ and $X(x) = x$ for any $(q,x) \in \Omega$. Denote $\mathcal{F}_t = \mathcal{F}_t^\Lambda \bigotimes \mathcal{F}_t^X$, where $\mathcal{F}_t^X = \sigma(X_s : s \in [0,t])$ on $\mathcal{C}$ and

$$\mathcal{F}_t^\Lambda = \sigma\left(\mathbb{1}_{[0,t]}\Lambda\right) = \sigma\left(\Lambda(C) : C \in \mathcal{B}([0,t] \times A)\right)$$

on $\mathcal{V}$. With a natural extension, X is said to be process on both $\mathcal{C}$ and $\mathcal{V}$. To define the canonical process on $\mathcal{V}$, we can extend Lemma 3.2 in [118] to an infinite-time horizon.

Lemma 2.5. *There exists a $\mathcal{F}_t^\Lambda$-predictable process $\overline{\Lambda} : [0,\infty) \to \mathcal{P}(A)$ such that $\forall q \in \mathcal{V}$, $\overline{\Lambda}(t,q) = q_t$ for almost all $t \geq 0$. Specifically, $q = dt\overline{\Lambda}(t,q)(da)$ for all $q \in \mathcal{V}$.*

For simplicity of notation, write $\Lambda_t = \overline{\Lambda}(t,\cdot)$ as the canonical process on $\mathcal{V}$. That is, $\Lambda(dt,da) = dt\Lambda_t(da)$.

Now we are ready to connect (MFG-BV) to controlled martingale problems.

Controlled martingale problems. A controlled process is defined by its infinitesimal generator,

$$L\phi(x,\mu,a) = b(x,\mu,a)\phi'(x) + \frac{\sigma^2}{2}\phi''(x), \quad \forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}),$$

for any $x \in \mathbb{R}$, $\mu \in \mathcal{P}^1(\mathbb{R})$ and $a \in A$. Here $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$ denotes any compactly supported smooth function on $\mathbb{R}$. For any $\mu \in \mathcal{P}^1(\mathcal{C})$ and $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$, define $M_t^{\mu,\phi}$ such that

$$M_t^{\mu,\phi}(q,x) = \phi(x_t) - \int_{[0,t]\times A} q(ds,da)L\phi(x_s,\mu_s,\alpha)$$

Now we are ready to discuss the law of the joint control-state pair.

Definition 2.6. *For a measure $\mu \in \mathcal{P}^1(\mathcal{C})$, define $\mathcal{R}(\mu) \subset \mathcal{P}(\Omega)$ to be a set of probability measures over the joint control-state process such that any $P \in \mathcal{R}(\mu)$ satisfies the following:*

1. $P \circ X_0^{-1} = \lambda$;
2. $\mathbb{E}^P \left[\int_0^\infty e^{-rt}|\Lambda_t|dt\right] < \infty$;
3. $M^{\mu,\phi} = \{M_t^{\mu,\phi}\}_{t\geq 0}$ *is a* P*-martingale for any* $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$.

The following proposition illustrates the connection between $\mathcal{R}(\mu)$ in Definition 2.6 and the controlled processes in (MFG-BV) defined as solutions to stochastic differential equations. It follows from Theorem IV-2 in [72].

Proposition 2.7. *For* $\mu \in \mathcal{P}^1(\mathcal{C})$, $\mathcal{R}(\mu)$ *is the set of laws* $P' \circ (\Lambda, X)^{-1}$ *as follows.*

1. $(\Omega', \mathcal{F}_t', P')$ *is a filtered probability space supporting a one-dimensional* $\mathcal{F}_t'$*-adapted process* X *as well as a standard Brownian motion* W;
2. $P' \circ (X_0)^{-1} = \lambda$;
3. $\mathbb{E}^{P'} \left[\int_0^\infty e^{-rt}|\Lambda_t|dt\right] < \infty$;
4. *The following state equation holds,*

$$dX_t = \int_A b(X_t, \mu, a)\Lambda_t(da)dt + \sigma dW_t. \tag{2.5}$$

Note that Assumption (A2) ensures that Equation (2.5) admits a unique strong solution.

Now we can define MFGs using controlled martingale problem. For $x \in \mathcal{C}$, $\mu \in \mathcal{P}^1(\mathcal{C})$ and $q \in \mathcal{V}$, define the objective functional

$$\Gamma^\mu(q, x) = \int_{[0,\infty)\times A} e^{-rt} q(dt, da) f(x_t, \mu_t, a), \tag{2.6}$$

define the objective function $J : \mathcal{P}^1(\mathcal{C}) \times \mathcal{P}^1(\Omega) \to \mathbb{R} \cup \{-\infty\}$, and define optimal control(s) in the form of set-valued function $\mathcal{R}^* : \mathcal{P}^1(\mathcal{C}) \to 2^{\mathcal{P}^1(\Omega)}$ as follows:

$$\begin{aligned} J(\mu, P) &:= \int_\Omega \Gamma^\mu dP, \\ \mathcal{R}^*(\mu) &:= \arg\max_{P\in\mathcal{R}(\mu)} J(\mu, P). \end{aligned}$$

Assumptions (A1)–(A3) ensure that both J and $\mathcal{R}^*(\mu)$ are well-defined. Now we introduce the notion of relaxed MFG solutions:

Definition 2.8 (Relaxed MFG solution)**.** $P \in \mathcal{P}^1(\Omega)$ *is called a relaxed MFG solution if*

$$\begin{aligned} J(P \circ X^{-1}, P') &= \int_\Omega \Gamma^{P\circ X^{-1}} dP', \\ P \in \mathcal{R}^*(P \circ X^{-1}) &= \arg\max_{P'\in\mathcal{R}(P\circ X^{-1})} J(P \circ X^{-1}, P'). \end{aligned} \tag{MFG}$$

$P \circ X^{-1} \in \mathcal{P}^1(\mathcal{C})$ *may also be interpreted as a relaxed MFG solution to* (MFG).

Definition 2.9 (Relaxed Markovian MFG solution)**.** *For $P \in \mathcal{P}^1(\Omega)$ satisfying* (MFG)*, if $P(\Lambda = dt[\hat{q}(t, X_t)](da)) = 1$ for some measurable function $\hat{q} : [0, \infty) \times \mathbb{R} \to \mathcal{P}(A)$, then P is called a relaxed Markovian solution to* (MFG)*.*

Definition 2.10 (Strict Markovian solution)**.** *For $P \in \mathcal{P}^1(\Omega)$ satisfying* (MFG)*, if $P(\Lambda = dt[\delta_{\hat{\alpha}(t,X_t)}](da)) = 1$ for some measurable function $\hat{\alpha} : [0, \infty) \times \mathbb{R} \to A$, then P is called a strict Markovian solution to* (MFG)*.*

Remark 2.11. *If $P \in \mathcal{P}^1(\Omega)$ is a* strick Markovian *solution to* (MFG) *in the sense of Definition 2.10, that is,*

$$P(\Lambda = dt\delta_{\hat{\alpha}(t,X_t)}(da)) = 1$$

for some measurable function $\hat{\alpha} : [0, \infty) \times \mathbb{R} \to A$, then $(\dot{\xi}^+, \dot{\xi}^-) = \hat{\alpha}$ together with $\mu = P \circ X^{-1}$ is the solution to (MFG-BV)*.*

The following theorem concerns the existence of a relaxed MFG solution to (MFG).

Theorem 2.12. *Under Assumptions (A1)–(A3), there exists a relaxed MFG solution to* (MFG)*.*

Theorem 2.12 is established by applying the Kakutani-Fan-Glicksberg theorem, see Theorem 1 in [75], on the following set-value mapping F defined as

$$F : \mathcal{P}^1(\mathcal{C}) \ni \mu \mapsto \{P \circ X^{-1} : P \in \mathcal{R}^*(\mu)\} \in 2^{\mathcal{P}^1(\mathcal{C})}. \tag{2.7}$$

Its continuity is defined through the notion of upper and lower hemicontinuity. In general, for any fixed metric spaces E and F, a set-value function $h : E \to 2^F$ is *lower hemicontinuous* if for any converging sequence $\{x_n\}$ in E with limit $x \in E$ and $y \in h(x)$, there exists a subsequence $\{x_{n_k}\}$ such that $\exists y_{n_k} \in h(x_{n_k})$ converging to y. Provided that $h(x)$ is closed in F for all x, h is *upper hemicontinuous* if for any converging $\{x_n\}$ in E to $x \in E$ and $y_n \in h(x_n)$, $\{y_n\}$ has a limit point $y \in h(x)$. It is continuous if it is both upper and lower hemicontinuous.

To verify the conditions of Kakutani-Fan-Glicksberg theorem, we need Lemmas 2.16–2.21 specified in Section 2.1.4. To summarize, Lemmas 2.16–2.19 ensure Lemmas 2.20 and 2.21 which in turns ensure that F is upper hemicontinuous and has non-empty compact convex values. The techniques are similar with [118]. The outline of the proof is provided in Section 2.1.4 with key steps and major differences from [118] highlighted.

From relaxed to strict Markovian MFG solution

Theorem 2.12 ensures the existence of a relaxed MFG solution P. Write $\mu = P \circ X^{-1}$. Now we are going to adopt the similar technique of constructing a *strict Markovian control* as in Theorem 3.7 of [118].

Theorem 2.13. *Under Assumptions (A1)–(A3), there exists a strict Markovian MFG solution to* (MFG)*.*

Proof. First, we can construct a *Markovian control* by conditioning. Notice that in our setting, σ is uncontrolled. Therefore, considering $\mathbb{E}^P[\Lambda_t|X_t]$ would suffice for finding a measurable function $\hat{q} : [0,\infty) \times \mathbb{R} \to \mathcal{P}(A)$ such that

$$\int_A \phi(t, X_t, a)\hat{q}(t, X_t)(da) = \mathbb{E}^P\left[\int_X \phi(t, X_t, a)\Lambda_t(da)|X_t\right], \quad P - a.s.,\ a.e.\, t \geq 0, \tag{2.8}$$

for any bounded measurable function $\phi : [0,\infty) \times \mathbb{R} \times A \to \mathbb{R}$.

This idea of construction can also be found in Proposition 5.1 of [31]. Define probability measure η on $[0,\infty) \times \mathbb{R} \times A$ such that for any $C \in \mathcal{B}([0,\infty) \times \mathbb{R} \times A)$,

$$\eta(C) := \mathbb{E}^P\left[\int_0^\infty re^{-rt}\int_A \mathbb{1}_C(t, X_t, a)\Lambda_t(da)dt\right],$$

with $\eta(dt, dx, da) = \eta_{1,2}(dt, dx)[\hat{q}(t, x)](da)$. Then for any bounded measurable function $h : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ and for any $T > 0$,

$$\begin{aligned}
&\mathbb{E}^P\left[\int_0^T h(t, X_t)\int_A \phi(t, X_t, a)\hat{q}(t, X_t)(da)dt\right] \\
&= \mathbb{E}^P\left[\int_0^\infty \mathbb{1}_{[0,T]}(t)h(t, X_t)\int_A \phi(t, X_t, a)\hat{q}(t, X_t)(da)dt\right] \\
&= \int_{[0,\infty)\times\mathbb{R}} 1_{[0,T]}(t)\frac{e^{rt}}{r}h(t, x)\int_A \phi(t, x, a)\hat{q}(t, x)(da)\eta_{1,2}(dt, dx) \\
&= \int_{[0,\infty)\times\mathbb{R}\times A} 1_{[0,T]}(t)\frac{e^{rt}}{r}h(t, x)\phi(t, x, a)\eta(dt, dx, da) \\
&= \mathbb{E}^P\left[\int_0^T h(t, X_t)\int_A \phi(t, X_t, a)\Lambda_t(da)dt\right]
\end{aligned} \tag{2.9}$$

By Lemma 5.2 of [31], Equation (2.9) implies Equation (2.8).

Corollary 3.7 of [31] implies that there exists $(\Omega', \mathcal{F}'_t, P')$ supporting a standard $\mathcal{F}'_t$-Brownian motion W' and a $\mathcal{F}'_t$-adapted process X' such that

$$dX'_t = \int_A b(X'_t, \mu_t, a)\hat{q}(t, X'_t)(da)dt + \sigma dW'_t; \quad P' \circ X'^{-1} = \mu_t,\ \forall t \geq 0.$$

Itô's formula implies $P_0 := P' \circ (dt\hat{q}(t, X'_t)(da), X')^{-1} \in \mathcal{R}(\mu)$. By Fubini's theorem and towering property of conditional expectation, it is easy to verify that $J(\mu, P) = J(\mu, P_0)$ and hence $P_0 \in \mathcal{R}^*(\mu)$. Define $\mu_0 = P_0 \circ X'^{-1}$. Since

$$(P_0 \circ X'^{-1})_t = (P' \circ X'^{-1})_t = \mu_t, \quad \forall t \geq 0,$$

then $\mathcal{R}(\mu_0) = \mathcal{R}(\mu)$, $J(\mu_0, \cdot) = J(\mu, \cdot)$ and hence, $\mathcal{R}^*(\mu_0) = \mathcal{R}^*(\mu)$. Therefore P_0 is a relaxed Markovian MFG solution.

P_0 can be made into *strict*, combining the arguments in [118] with results in [97] on infinite-time horizon, therefore the exact same technique can be used with minimal adjustment. □

Now the existence of a strict Markovian solution to (MFG) immediately follows.

Theorem 2.14. *[Existence of strict Markovian MFG solution.] Assume Assumptions (A1)-(A3). Then there exists $\mu \in \mathcal{P}^1(\mathcal{C})$ and a measurable function $\dot{\xi} : [0,\infty) \times \mathbb{R} \to A$ satisfying the following:*

(1) There exists a probability space $(\Omega, \mathcal{F}_t, P)$ supporting a standard $\mathcal{F}_t$-Brownian Motion W and $\mathcal{F}_t$-adapted process X such that

$$\begin{cases} dX_t = b(X_t, \mu_t)dt + \sigma dW_t + \dot{\xi}_1(t, X_t)dt - \dot{\xi}_2(t, X_t)dt, \quad P \circ X_0^{-1} = \lambda; \\ P \circ X^{-1} = \mu. \end{cases}$$

(2) For any other $(\Omega', \mathcal{F}'_t, P')$, W', X' and $\dot{\xi}'$,

$$\mathbb{E}^P \left[\int_0^\infty e^{-rt} [f(X_t, \mu_t) - \gamma^+ \dot{\xi}_1(t, X_t) - \gamma^- \dot{\xi}_2(t, X_t)] dt \right] \geq \mathbb{E}^{P'} \left[\int_0^\infty e^{-rt} [f(X'_t, \mu_t) - \gamma^+ \dot{\xi}'_{1,t} - \gamma^- \dot{\xi}'_{2,t}] dt \right]$$

with $dX't = b(X'_t, \mu_t)dt + \sigma dW'_t + \dot{\xi}'_{1,t}dt - \dot{\xi}'_{2,t}dt$, $P \circ X_0'^{-1} = \lambda$.

Remark 2.15. *Using the language of controlled martingale problems, condition (1) in Theorem 2.14 means $P \in \mathcal{R}(P \circ X^{-1})$ while condition (2) means $P \in \mathcal{R}^*(P \circ X^{-1})$. Let $\Lambda = dt\delta_{\dot{\xi}_t}(da)$, then P corresponds to a strict Markovian MFG solution.*

2.1.4 Proof of Theorem 2.12

The proof is established through a series of lemmas.

Lemma 2.16. *Given Assumptions (A1)–(A3) with c the Lipschitz constant appeared in Assumption (A2), fix any $\gamma \in [1, 2]$ and any $T > 0$, then there exists a constant $c''(T) = c''_{\gamma,\lambda,\theta,c}(T) > 0$ such that for all $\mu \in \mathcal{P}^1(\mathcal{C})$ and $P \in \mathcal{R}(\mu)$,*

$$\mathbb{E}^P \|X\|_T^\gamma \leq c''(T).$$

In particular, $P \in \mathcal{P}^1(\Omega)$. Moreover, if $P \circ X^{-1} = \mu$, then

$$\|\mu\|_T^\gamma \leq c''(T).$$

It follows directly from the definition of $\mathcal{R}(\mu)$ and the boundedness of b and the control set A. In fact, Lemma 2.16 is a stronger version of Lemma 4.3 in [118] due to our problem setting.

Lemma 2.17. *Suppose $K \subset \mathcal{P}(\Omega)$ satisfies $\{P \circ X^{-1} : P \in K\}$ is tight in $\mathcal{P}(\mathcal{C})$, and*

$$\sup_{P \in K} \mathbb{E}^P \|X\|_T < \infty$$

for any $T > 0$. Then K is relatively compact in $\mathcal{P}^1(\Omega)$.

Proof. The proof is by Theorems 16.8 and 16.10 in [28]. Since $\sup_{P \in K} \mathbb{E}^P \|X\|_T < \infty$ for any $T > 0$ and $\{P \circ X^{-1} : P \in K\}$ is tight, we see that $\{P \circ X^{-1} : P \in K\}$ is relatively compact in $\mathcal{P}^1(\mathcal{C})$. The rest of the proof fis similar to the proof of Proposition B.3 in [118]. □

Lemma 2.18. *Fix any $\mu \in \mathcal{P}^1(\mathcal{C})$. Let $\mathcal{Q} \subset \mathcal{P}(\Omega)$ be the set of laws $P \circ (\Lambda, X)^{-1}$ of Ω-valued random variables (Λ, X) defined on some filtered probability space $(\Omega, \mathcal{F}_t, P)$ such that*

(1) $dX_t = \int_A b(X_t, \mu_t, a)\Lambda_t(da)dt + \sigma dW_t$.

(2) W is a standard $\mathcal{F}_t$-Brownian motion.

(3) X_0 has law λ and $\mathcal{F}_0$-measurable.

Then $\mathcal{Q}$ is relatively compact in $\mathcal{P}^1(\Omega)$.

Proof. By Lemmas 2.16 and 2.17, it suffices to verify the Aldous criterion for tightness for any $T > 0$. Now the same technique in the proof of Proposition B.4 of [118] can be applied. □

Lemma 2.19. *Let $\phi : [0, \infty) \times \mathbb{R} \times A \to \mathbb{R}$ be measurable, with $\phi(t, \cdot, \cdot)$ jointly continuous for any $t \geq 0$. Suppose for some $k > 0$ and $x_0 \in \mathbb{R}$ such that $|\phi(t, x, a)| \leq k(1 + |x - x_0| + |a|)$ for all $(t, x, a) \in [0, \infty) \times \mathbb{R} \times A$, then*

$$\mathcal{C} \times \mathcal{V} \ni (x, q) \mapsto \int e^{-rt} q(dt, da)\phi(t, x_t, a)$$

is a continuous mapping.

Proof. Adapt the proof for Corollary A.5 in [118] by considering the following jointly continuous mapping

$$\mathcal{C} \times \mathcal{V} \ni (x, q) \mapsto re^{-rt} q(dt, da)\delta_{x_t}(de) \in \mathcal{P}^1([0, \infty) \times A \times \mathbb{R}).$$

□

Lemma 2.20. *Under Assumptions (A1)–(A3), J is continuous.*

Lemma 2.20 is a direct consequence of Lemma 2.19.

Lemma 2.21. *Given Assumptions (A1)–(A3), the range $\mathcal{R}(\mathcal{C}) := \{P \in \mathcal{R}(\mu) : \mu \in \mathcal{P}^1(\mathcal{C})\}$ is relatively compact in $\mathcal{P}^1(\Omega)$, and the set-valued function $\mathcal{R}$ is continuous.*

Proof. Having established Lemmas 2.18 and 2.19, same techniques in proving Lemma 4.4 in [118] can be applied to show $\mathcal{R}(\mathcal{C})$ is relatively compact in $\mathcal{P}^1(\Omega)$ and $\mathcal{R}$ is upper hemicontinuous.

Next, we show $\mathcal{R}$ is lower hemicontinuous. Let $\mu^n \to \mu$ in $\mathcal{P}^1(\mathcal{C})$ and $P \in \mathcal{R}(\mu)$. By Lemma 2.7, there exists $(\Omega', \mathcal{F}'_t, P')$ supporting a standard $\mathcal{F}'_t$-Brownian motion such that $P' \circ (\Lambda, X)^{-1} = P$. Assumption (A2) ensures the existence of a strong solution to the SDE

$$dX_t^n = \int_A b(X_t^n, \mu_t^n, a)\Lambda_t(da)dt + \sigma dW_t, \quad X_0^n = X_0.$$

For any fixed $T > 0$, similar as Lemma 4.4 in [118], we have $\mathbb{E}^{P'}\|X^n - X\|_T \to 0$ as $n \to \infty$. Define $P^n := P' \circ (\Lambda, X^n)^{-1}$. By Theorem 16.7 in [28], $P^n \to P$ in $\mathcal{P}^1(\Omega)$. Itô's formula implies $P^n \in \mathcal{R}(\mu^n)$. Therefore, $\mathcal{R}$ is lower hemicontinuous. □

Proof of Theorem 2.12. First, Lemma 2.21 and 2.20 imply that the set-valued function F defined in Equation (2.7) is upper hemicontinuous and has non-empty compact convex and for any $\mu \in \mathcal{P}^1(\mathcal{C})$, $\mathcal{R}^*(\mu)$ is convex. Define

$$F(\mathcal{P}^1(\mathcal{C})) = \{P \circ X^{-1} : \mu \in \mathcal{C},\ P \in \mathcal{R}^*(\mu)\} \subset \mathcal{P}^1(\mathcal{C}).$$

Then, Lemma 2.16 ensures that for any $T > 0$, $M_T := \sup\{\|\mu\|_T^2 : \mu \in F(\mathcal{P}^1(\mathcal{C}))\} < \infty$. Boundedness of b implies that for any $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$, there exists $C_\phi \in (0, \infty)$ such that for any (x, μ, a),

$$|L\phi(x, \mu, a)| < C_\phi.$$

Define $\mathcal{Q} \subset \mathcal{P}(\mathcal{C})$ such that for any $P \in \mathcal{Q}$,

(1) $P \circ X_0^{-1} = \lambda$.

(2) $\mathbb{E}^P\|X\|_T \leq M_T$ for all $T > 0$.

(3) $\{\phi(X_t) + C_\phi t\}_{t\geq 0}$ is a P-submartingale for nonnegative $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$.

Then Theorem 4.1 in [118] ensures $\mathcal{Q}$ is compact under the weak convergence topology $\mathcal{M}(\mathcal{C})$ of bounded signed measures of $\mathcal{C}$. Clearly $F(\mathcal{P}^1(\mathcal{C})) \subset \mathcal{Q}$. That is, F is an upper hemicontinuous maps $\mathcal{Q}$ back into itself that has nonempty, compact and convex values. Kakutani-Fan-Glicksberg theorem implies the existence of a fixed point of F, which is a relaxed MFG solution.

2.2 MFGs with singular control of finite variation

In the previous section, we have shown the existence of a strict Markovian MFG solution with singular control of bounded velocity. In this section, we will analyze a particular MFG with singular control of finite variation inspired by the partially reversible investment problem.

2.2.1 Preliminary on partially reversible investment problem

The basic idea of the partially reversible investment problem goes as follows. A company profits from producing and selling a commodity. The revenue function depends on the production level with fluctuations according to, for instance, the market demand. The company has the flexibility to adjust its production level at any time, with the expansion incurring a cost and the contraction bringing a smaller salvage value. The objective of the company is to choose an optimal investment strategy in terms of its production level to maximize the overall expected net profits.

In [91], this partially reversible investment problem is formulated as follows. Take a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ supporting a standard Brownian motion $W = \{W_t\}_{t\geq 0}$. Assume that $\mathbb{F}^W$ is the augmented filtration generated by W that satisfies the usual condition. The production level of a company x_t at time t is characterized by a geometric Brownian motion with an initial distribution $\mu_0 \in \mathcal{P}^2(\mathbb{R})$ such that

$$dX_t = X_t(\delta dt + \gamma dW_t), \quad X_{0-} \sim \mu_0,$$

where $\delta, \gamma > 0$ are drift and volatility coefficients, representing respectively the average and fluctuation in market demand. The production level can be adjusted at any time t, and possibly in a discontinuous fashion such that

$$dX_t = X_t(\delta dt + \gamma dW_t) + d\xi_t, \quad X_{0-} \sim \mu_0, \quad \xi_{0-} = 0. \tag{2.10}$$

Here, $\xi_t = \xi_t^+ - \xi_t^-$, $\xi_{0-}^{\pm} = 0$ with $\xi_\cdot^+$ and $\xi_\cdot^-$ adapted and nondecreasing cádlág processes representing the accumulated increased and decreased production level by time t respectively. (Note that when the control is of finite variation, such decomposition of $\xi_\cdot$ by $\xi_\cdot^+$ and $\xi_\cdot^-$ is unique).

The objective of the company is to adjust its production level x_t according to a policy $\xi_\cdot = (\xi_\cdot^+, \xi_\cdot^-)$ chosen from an appropriate admissible control set $\mathcal{U}$, in order to maximize its discounted expected total profit over an infinite-time horizon. That is to find

$$v(x) = \sup_{(\xi_\cdot^+, \xi_\cdot^-)\in\mathcal{U}} \mathbb{E}\left[\int_0^\infty e^{-rt}[\Pi(X_t)dt - \gamma^+ d\xi_t^+ - \gamma^- d\xi_t^-]\middle| X_{0-} = x\right], \quad \forall x > 0. \tag{2.11}$$

Here the discount rate $r > 0$, $\Pi(\cdot)$ the revenue function satisfies the standard Inada condition for utility functions, $\gamma^+ = p > 0$ is the unit investment cost to increase production level, and $-\gamma^- = p(1-\lambda)$ is the unit gain for reducing production level, with $\lambda \in (0,1)$ to ensure no-arbitrage.

Finally, the admissible control set $\mathcal{U}$ is

$$\begin{aligned} \mathcal{U} = \Big\{ & (\xi_\cdot^+, \xi_\cdot^-) : \xi_\cdot^+,\ \xi_\cdot^- \text{ nondecreasing càdlàg processes adapted to } \mathbb{F}^W, \\ & \xi_{0-}^+ = \xi_{0-}^- = 0,\ \mathbb{E}\left[\int_0^\infty e^{-rt} d\xi_t^+\right] < \infty, X_t \geq 0. \Big\} \end{aligned} \tag{2.12}$$

In [91], the smooth fit principle in the sense of [18] is established via regularity analysis for the value function, and the optimal control $\xi^*_\cdot = (\xi^{*,+}_\cdot, \xi^{*,-}_\cdot)$ to (2.11) is shown to be of bang-bang type. Moreover, the value function is shown to be the unique classical $\mathcal{C}^2$ solution to the following Hamilton-Jacobian-Bellman (HJB) equation,

$$0 = \min\{rv(x) - \Pi(x) - \delta x v'(x) - \frac{1}{2}\gamma^2 x^2 v''(x),\, p - v'(x),\, v'(x) - p(1-\lambda)\}, \tag{2.13}$$

where $v'(\cdot)$ and $v''(\cdot)$ denote the first and second order derivatives of $v(\cdot)$ respectively.

When the revenue function $\Pi(x)$ is of the Cobb-Douglas type, i.e., $\Pi(x) = c\rho x^\alpha$ with constants $\rho > 0$, $c > 0$ and $\alpha \in (0,1)$, then the optimal control is characterized by two thresholds $0 < x_b < x_s < \infty$, which are explicitly given by

$$\begin{cases} x_b = \left\{ \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} \right\}^{\frac{1}{1-\alpha}} \rho^{\frac{1}{1-\alpha}}, \\ x_s = \left\{ \frac{2c\alpha y_0^{1-\alpha}(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} \right\}^{\frac{1}{1-\alpha}} \rho^{\frac{1}{1-\alpha}}, \end{cases}$$

where

$$m = -\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right) - \sqrt{\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\gamma^2}},\ n = -\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right) + \sqrt{\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\gamma^2}},$$

and $y_0 > 1$ is a root of the following equation

$$1 - \lambda = \frac{(n-1)(\alpha-m)y^{m-1}(y^\alpha - y^n) + (1-m)(n-\alpha)y^{n-1}(y^m - y^\alpha)}{(n-1)(\alpha-m)(y^\alpha - y^n) + (1-m)(n-\alpha)(y^m - y^\alpha)}.$$

The value function is then derived by solving the following QVI via the smooth fit principle,

$$\begin{cases} p - v' = 0, & x \le x < x_b, \\ rv - c\rho x^\alpha - \delta x v' - \frac{1}{2}\gamma^2 x^2 v'' = 0, & x_b \le x \le x_s, \\ v' - p(1-\lambda) = 0, & x > x_s. \end{cases}$$

This bang-bang type of control, that is, the optimal control $\xi^*_\cdot$ characterized by a pair of threshold (x_b, x_s), suggests that the company should spend the minimum effort to keep its production level within the interval $[x_b, x_s]$.

2.2.2 Formulation of MFGs

Having reviewed a classical partially investment problem of [91], it is natural to consider the game version of this partially reversible investment problem. We will first consider an MFG in which there are infinite number of rational and indistinguishable companies, and derive an explicit solution to this MFG. We will then compare this (much simpler) MFG with the single-agent problem (in Section 2.2.4), and study its relation with the corresponding N-player game (in Section 2.2.5).

Let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ be a filtered probability space supporting a standard Brownian motion $W = \{W_t\}_{t\geq 0}$. Assume that $\mathbb{F}^W$ is the augmented filtration generated by W that satisfies the usual condition. As in the single-agent control problem, in the MFG a representative company adjusts its production level x_t according to a policy chosen from the admissible control set $\mathcal{U}$ defined as

$$\begin{aligned}\mathcal{U} = \Big\{(\xi_\cdot^+, \xi_\cdot^-) : \xi_\cdot^+, \, \xi_\cdot^- \text{ nondecreasing càdlàg processes adapted to } \mathbb{F}^W, \\ \xi_{0-}^+ = \xi_{0-}^- = 0, \, \mathbb{E}\left[\int_0^\infty e^{-rt} d\xi_t^+\right] < \infty, x_t \geq 0.\Big\}\end{aligned} \tag{2.14}$$

to maximize its discounted total profit over an infinite-time horizon,

$$\sup_{(\xi_\cdot^+, \xi_\cdot^-)\in\mathcal{U}} \mathbb{E}\left[\int_0^\infty e^{-rt}[f(X_t, \mu)dt - pd\xi_t^+ + p(1-\lambda)d\xi_t^-]\Big| x_{0-} = x\right], \quad \forall x > 0, \tag{MFG-FV}$$

subject to

$$dX_t = X_t(\delta dt + \gamma dW_t) + d\xi_t^+ - d\xi_t^-, \quad X_{0-} \sim \mu_0. \tag{2.15}$$

Unlike the single agent problem, the revenue function for a representative company in this game (MFG-FV) depends on *both* its own production level x and the aggregation of all other companies, denoted by a probability distribution μ. More precisely, $f(x, \mu)$ the revenue function of a Cobb-Douglas type takes the form of $f(x, \mu) = F(\mu)x^\alpha$ for some $\alpha \in (0, 1)$, with μ being the distribution of the production level in the long run, i.e., $\mu = Law(x_\infty)$. If we consider the inverse demand function, then the price will be given by

$$\rho = \rho(\mu) = \mathbb{E}_{X\sim\mu}[\tilde{\rho}(X)] = \int (a_0 - a_1 y^{1-\alpha})\mu(dy),$$

and $F(\mu) = c\rho(\mu)$. Effectively one can write

$$f(x, \mu) = c\rho x^\alpha.$$

Note that in this MFG companies interact through the revenue function f. It is also worth noting that, unlike the revenue function for the single-agent control problem in Section 1.1.3 where the unit price ρ is exogenously given and fixed, here in the game (MFG-FV) ρ is endogenously determined.

We will look for a solution to the (MFG-FV) in the following sense.

Definition 2.22. *If there exists a control $\xi_\cdot^* = (\xi_\cdot^{+,*}, \xi_\cdot^{-,*}) \in \mathcal{U}$ and $\rho^* > 0$ such that*

1. *Under ρ^*, $\xi_\cdot^*$ is an optimal control for*

$$\tilde{v}(x) = \sup_{(\xi_\cdot^+, \xi_\cdot^-)\in\mathcal{U}} \mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}[c\rho^* X_t^\alpha dt - pd\xi_t^+ + p(1-\lambda)d\xi_t^-]\Big| X_{0-} = x\right], \forall x > 0, \tag{2.16}$$

   subject to (2.15).

2. Under $\xi^*_\cdot$ the controlled process $x^* = \{x^*_t\}_{t\geq 0}$ given by

$$dX^*_t = X^*_t(\delta dt + \gamma dW_t) + d\xi^{+,*}_t - d\xi^{-,*}_t, \quad X^*_{0-} \sim \mu_0 \tag{2.17}$$

admits a limiting distribution $\mathbb{P}_{x^*_\infty}$ and $\rho^* = \int (a_0 - a_1 y^{1-\alpha})\mathbb{P}_{x^*_\infty}(dy)$.

then the control-mean pair $(\xi^*_\cdot, \rho^*)$ is said to be an NE solution to the game (MFG-FV)*.*

To ensure the well-posedness of (MFG-FV), we assume $2\delta + \gamma^2 < r$ and $\frac{2\delta}{\gamma^2} \notin \{\alpha, 1\}$.

Remark 2.23. *There is an alternative and equivalent definition of the solution to the game* (MFG-FV) *. That is, for any fixed $\rho \in \mathbb{R}$, we may define*

$$\tilde{v}(\mu_0) = \sup_{(\xi^+,\xi^-)\in\mathcal{U}} \mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}\left[c\rho X^\alpha_t dt - pd\xi^+_t + p(1-\lambda)d\xi^-_t\right]\right]$$

subject to (2.15)*. Then these two solutions are equivalent in the sense that $\tilde{v}(\mu_0) = \mathbb{E}_{\mu_0}[\tilde{v}(X_{0-})]$.*

2.2.3 Explicit solution to MFG

Solution to the (MFG-FV)

We shall now solve the game (MFG-FV), with the fixed-point approach as in [121].

Step 1. Control problem under fixed mean information. Fix a $\rho > 0$, then the game (MFG-FV) is a singular control problem,

$$\tilde{v}(x) = \sup_{(\xi^+_\cdot,\xi^-_\cdot)\in\mathcal{U}} \mathbb{E}\left[\int_0^\infty e^{-rt}[c\rho X^\alpha_t dt - pd\xi^+_t + p(1-\lambda)d\xi^-_t]\middle| X_{0-} = x\right], \; x > 0, \tag{Control-S}$$

subject to (2.15). The dynamic programming principle leads to the following HJB equation associated with the problem (Control-S) under the fixed ρ,

$$0 = \min\{r\tilde{v}(x) - cx^\alpha\rho - \delta x\partial_x\tilde{v}(x) - \frac{1}{2}\gamma^2x^2\partial_{xx}\tilde{v}(x), p - \partial_x\tilde{v}(x), \partial_x\tilde{v}(x) - p(1-\lambda)\}. \tag{2.18}$$

Similar to the argument in [91], we see that the optimal policy is a bang-bang type and is characterized by an expansion threshold $\tilde{x}_b$ and a contraction threshold $\tilde{x}_s$ so that $x_t \in [\tilde{x}_b, \tilde{x}_s]$ almost surely.

More precisely, at time $t = 0$, if $x \in (0, \tilde{x}_b)$, then $\xi^+_0 = \tilde{x}_b - x$ and $\xi^-_0 = 0$; if $x \in (\tilde{x}_s, \infty)$, then $\xi^+_0 = 0$ and $\xi^-_0 = x - \tilde{x}_s$. Note that $x_0 = x_{0-} + \xi^+_0 - \xi^-_0 \in [\tilde{x}_b, \tilde{x}_s]$. For $t > 0$, it is optimal to impose a minimum amount of adjustment so that $x_t \in [\tilde{x}_b, \tilde{x}_s]$.

Accordingly, the solution $\tilde{v}$ is of the form

$$\tilde{v}(x) = \begin{cases} px + C_1, & 0 \le x \le \tilde{x}_b, \\ Ax^m + Bx^n + Hx^\alpha, & \tilde{x}_b < x < \tilde{x}_s, \\ p(1-\lambda)x + C_2, & \tilde{x}_s \le x, \end{cases}$$

where $\tilde{x}_b = \inf\{x : \partial_x \tilde{v}(x) = p\}$, $\tilde{x}_s = \sup\{x : \partial_x \tilde{v}(x) = p(1-\lambda)\}$ with $0 < \tilde{x}_b \le \tilde{x}_s$ (see Lemma 4.4 in [91]), and since it is assumed that $2\delta + \gamma^2 < r$ and hence $\delta < r$,

$$m = -\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right) - \sqrt{\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\gamma^2}} < 0,\ n = -\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right) + \sqrt{\left(\frac{\delta}{\gamma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\gamma^2}} > 1,$$
$$H = \frac{2c\rho}{\gamma^2(n-\alpha)(\alpha-m)}.$$

Moreover, by the smooth-fit principle, we have

$$\begin{cases} A\tilde{x}_b^m + B\tilde{x}_b^n + H\tilde{x}_b^\alpha = p\tilde{x}_b + C_1, \\ mA\tilde{x}_b^{m-1} + nB\tilde{x}_b^{n-1} + \alpha H\tilde{x}_b^{\alpha-1} = p, \\ m(m-1)A\tilde{x}_b^{m-2} + n(n-1)B\tilde{x}_b^{n-2} + \alpha(\alpha-1)H\tilde{x}_b^{\alpha-2} = 0, \\ A\tilde{x}_s^m + B\tilde{x}_s^n + H\tilde{x}_s^\alpha = p(1-\lambda)\tilde{x}_s + C_2, \\ mA\tilde{x}_s^{m-1} + nB\tilde{x}_s^{n-1} + \alpha H\tilde{x}_s^{\alpha-1} = p(1-\lambda), \\ m(m-1)A\tilde{x}_s^{m-2} + n(n-1)B\tilde{x}_s^{n-2} + \alpha(\alpha-1)H\tilde{x}_s^{\alpha-2} = 0. \end{cases} \tag{2.19}$$

Some algebraic manipulations yield

$$A = \frac{p(n-1)\tilde{x}_b - \alpha(n-\alpha)H\tilde{x}_b^\alpha}{m(n-m)\tilde{x}_b^m} = \frac{p(1-\lambda)(n-1)\tilde{x}_s - \alpha(n-\alpha)H\tilde{x}_s^\alpha}{m(n-m)\tilde{x}_s^m}; \tag{2.20}$$

and

$$B = \frac{p(m-1)\tilde{x}_b - \alpha(m-\alpha)H\tilde{x}_b^\alpha}{n(m-n)\tilde{x}_b^n} = \frac{p(1-\lambda)(m-1)\tilde{x}_s - \alpha(m-\alpha)H\tilde{x}_s^\alpha}{n(m-n)\tilde{x}_s^n}. \tag{2.21}$$

Furthermore, denote $y_0 = \frac{\tilde{x}_s}{\tilde{x}_b}$ and $y_0 \ge 1$. By (2.20) and (2.21), we have

$$\begin{cases} p(n-1)\left[(1-\lambda)y_0 - y_0^m\right] = \alpha(n-\alpha)H\tilde{x}_b^{\alpha-1}\left(y_0^\alpha - y_0^m\right), & (2.22) \\ p(m-1)\left[(1-\lambda)y_0 - y_0^n\right] = \alpha(m-\alpha)H\tilde{x}_b^{\alpha-1}\left(y_0^\alpha - y_0^n\right), & (2.23) \end{cases}$$

and

$$\frac{(n-1)(\alpha-m)y_0^{m-1}(y_0^\alpha - y_0^n) + (1-m)(n-\alpha)y_0^{n-1}(y_0^m - y_0^\alpha)}{(n-1)(\alpha-m)(y_0^\alpha - y_0^n) + (1-m)(n-\alpha)(y_0^m - y_0^\alpha)} = 1-\lambda. \tag{2.24}$$

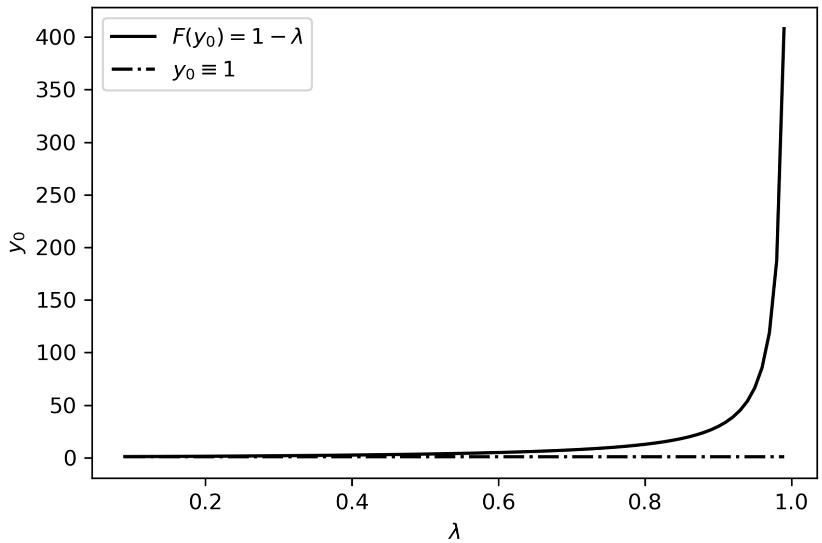


Figure 2.1: y_0 increases along with λ.

Now, to show that there exists a y_0 for (2.24), define $F(y)$ for $y > 1$ that

$$F(y) = \frac{(n-1)(\alpha-m)y^{m-1}(y^\alpha - y^n) + (1-m)(n-\alpha)y^{n-1}(y^m - y^\alpha)}{(n-1)(\alpha-m)(y^\alpha - y^n) + (1-m)(n-\alpha)(y^m - y^\alpha)}.$$

Since $\lim_{y\to 1^+} F(y) = 1$, $\lim_{y\to\infty} F(y) = 0$, and F is continuous, there exists a $y_0 > 1$ satisfying $F(y_0) = 1-\lambda \in (0,1)$ (see also Figure 2.1). Note that the function F does not depend on ρ, therefore y_0 is independent of ρ. From (2.23), we can conclude that

$$\tilde{x}_b = \left\{ \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} \right\}^{\frac{1}{1-\alpha}} \rho^{\frac{1}{1-\alpha}}. \tag{2.25}$$

where $\left\{ \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} \right\}^{\frac{1}{1-\alpha}}$ does not depend on ρ, and

$$\tilde{x}_s = \tilde{x}_b y_0 = \left\{ \frac{2c\alpha y_0^{1-\alpha}(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} \right\}^{\frac{1}{1-\alpha}} \rho^{\frac{1}{1-\alpha}}. \tag{2.26}$$

After plugging in (2.25) and (2.26), A and B are given by (2.20) and (2.21), respectively, and

$$C_1 = A\tilde{x}_b^m + B\tilde{x}_b^n + H\tilde{x}_b^\alpha - p\tilde{x}_b, \quad C_2 = A\tilde{x}_s^m + B\tilde{x}_s^n + H\tilde{x}_s^\alpha - p(1-\lambda)\tilde{x}_s.$$

To justify that the above analytical solution is indeed the solution to the problem (Control-S), one way is via the verification theorem, see for instance [93]. Alternatively, one can first show that the value function is the unique viscosity solution to the corresponding HJB and then establish the uniqueness of a classical $\mathcal{C}^2$ solution to the HJB, see for instance [91]. Here we adopt the second approach and claim that under any given $\rho > 0$, $\tilde{v}$ derived above is the value function of problem (Control-S). The proof is similar to that for the single-agent case in [91], therefore omitted here.

Step 2. Updating the price ρ and the locating the fixed point. Under any fixed $\rho > 0$, the optimal controlled process x_t is a geometric reflected Brownian motion within the interval $[\tilde{x}_b, \tilde{x}_s]$. By [30], for any $x \in [\tilde{x}_b, \tilde{x}_s]$, the scale density is given by

$$s(x) = \exp\left\{-\int_\theta^x \frac{2\delta}{\gamma^2 y} dy\right\} = \theta^{\frac{2\delta}{\gamma^2}} x^{-\frac{2\delta}{\gamma^2}}, \quad \forall \theta \in (\tilde{x}_b, \tilde{x}_s),$$

the speed density is

$$m(x) = \frac{2}{\gamma^2 x^2 s(x)} = \frac{2}{\gamma^2 \theta^{\frac{2\delta}{\gamma^2}}} x^{\frac{2\delta}{\gamma^2}-2},$$

and finally

$$M(x) = \int_{\tilde{x}_b}^x m(y)dy = \frac{2}{\gamma^2 \theta^{\frac{2\delta}{\gamma^2}}} \frac{x^{\frac{2\delta}{\gamma^2}-1} - \tilde{x}_b^{\frac{2\delta}{\gamma^2}-1}}{\frac{2\delta}{\gamma^2}-1}.$$

The density function of P_{x_∞}, the limiting distribution of x_t, is thus

$$f(x) = \frac{m(x)}{M(\tilde{x}_s)} = \frac{\frac{2\delta}{\gamma^2}-1}{\tilde{x}_s^{\frac{2\delta}{\gamma^2}-1} - \tilde{x}_b^{\frac{2\delta}{\gamma^2}-1}} x^{\frac{2\delta}{\gamma^2}-2}, \quad \forall x \in [\tilde{x}_b, \tilde{x}_s].$$

The updated price $\bar{\rho}$ under the limiting distribution $\bar{\mu} = Law(x_\infty)$ is

$$\begin{aligned} \bar{\rho} = \Gamma(\rho) &= a_0 - a_1 \int_{\tilde{x}_b}^{\tilde{x}_s} x^{1-\alpha} f(x)dx = a_0 - a_1 \frac{2\delta-\gamma^2}{2\delta-\alpha\gamma^2} \frac{\tilde{x}_s^{\frac{2\delta}{\gamma^2}-\alpha} - \tilde{x}_b^{\frac{2\delta}{\gamma^2}-\alpha}}{\tilde{x}_s^{\frac{2\delta}{\gamma^2}-1} - \tilde{x}_b^{\frac{2\delta}{\gamma^2}-1}} \\ &= a_0 - \rho \cdot a_1 \frac{2\delta-\gamma^2}{2\delta-\alpha\gamma^2} \frac{y_0^{\frac{2\delta}{\gamma^2}-\alpha} - 1}{y_0^{\frac{2\delta}{\gamma^2}-1} - 1} \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]}, \end{aligned} \tag{2.27}$$

where the coefficient $a_1 \frac{2\delta-\gamma^2}{2\delta-\alpha\gamma^2} \frac{y_0^{\frac{2\delta}{\gamma^2}-\alpha} - 1}{y_0^{\frac{2\delta}{\gamma^2}-1} - 1} \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]}$ does not rely on ρ. Clearly, for a_1 such that

$$a_1 > 0, \quad a_1 \frac{2\delta-\gamma^2}{2\delta-\alpha\gamma^2} \frac{y_0^{\frac{2\delta}{\gamma^2}-\alpha} - 1}{y_0^{\frac{2\delta}{\gamma^2}-1} - 1} \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n - (1-\lambda)y_0\right]} < 1, \tag{2.28}$$

the mapping Γ is a contraction and therefore admits a unique fixed point

$$\rho^* = \frac{a_0}{1 + a_1 \frac{2\delta-\gamma^2}{2\delta-\alpha\gamma^2} \frac{y_0^{\frac{2\delta}{\gamma^2}-\alpha}-1}{y_0^{\frac{2\delta}{\gamma^2}-1}-1} \frac{2c\alpha(y_0^n - y_0^\alpha)}{\gamma^2 p(1-m)(n-\alpha)\left[y_0^n-(1-\lambda)y_0\right]}}. \tag{2.29}$$

Substitute ρ^* of (2.29) into (2.25) and (2.26), we can derive optimal action boundaries $\tilde{x}_b^*$ and $\tilde{x}_s^*$. Denote the singular control characterized by $(\tilde{x}_b^*, \tilde{x}_s^*)$ as $\xi_\cdot^*$. Under Definition 2.22, $(\xi_\cdot^*, \rho^*)$ is a solution to the game (MFG-FV).

Remark 2.24. *Note that under the assumption $2\delta + \gamma^2 < r$, the uncontrolled process $X = \{X_t\}_{t\geq 0}$ satisfies $\mathbb{E}\left[\int_0^\infty e^{-rt} X_t^2 dt\right] < \infty$, and this property is preserved for the controlled process $X^* = \{X_t^*\}_{t\geq 0}$ under $\xi_\cdot^*$, as it is restricted to a bounded region.*

2.2.4 Sensitivity analysis and comparison with single-agent control problem

As seen from (2.27), the iterations do not stop after the first round, indicating that the game (MFG-FV) demonstrates a genuine game effect from the weak interactions among the players. Moreover, we can see that in the game (MFG-FV), model parameters λ, δ, γ, r and α affect both the optimal strategy of as in the single-agent control problem (2.11) and the equilibrium price ρ^*.

To illustrate, consider the following case where $\delta = 1$, $\gamma = 2$, $r = 3$, $\alpha = 0.6$, $\lambda = 0.6$, $p = 0.5$, $c = 1$, $a_0 = 1$ and $a_1 = 0.1$. Suppose the iterative process starts from a fixed value $\rho = 1$. In the single-agent setting (2.11) where the price $\rho = 1$ is seen as exogenously given and fixed, the optimal thresholds are given by $x_b = 0.053$ and $x_s = 0.264$. Figure 2.2 shows that both x_b and x_s increase along with the value of ρ and the non-action region $[x_b, x_s]$ expands. In the game (MFG-FV), in contrast, the equilibrium price is $\rho^* = 0.96$ under which the optimal thresholds are $\tilde{x}_b^* = 0.048$ and $\tilde{x}_s^* = 0.239$. Figure 2.3 shows the difference in the thresholds of intervention between the single-agent control problem (2.11) and the game (MFG-FV).

Impact of λ. $\lambda \in (0, 1)$ measures the irreversibility of the investment: the closer λ to 1, the more irreversible the investment. For the single-agent control problem (2.11) (Figures 2.4a and 2.4b), the expansion threshold x_b stays relatively insensitive with respect to an increasing λ, the contraction threshold x_s however increases dramatically along with λ. This means that for an individual company, if the investment is more irreversible, it becomes less profitable to frequently decrease the production level; consequently, the contraction threshold is raised to a higher level. Under the game (MFG-FV) setting, the irreversibility does not have an immediate impact on the optimal strategies (Figure 2.4c); instead, it drives down the equilibrium price (Figure 2.4d). This suggests that as it becomes less profitable to reduce production when λ approaches 1, companies in the game (MFG-FV) tend to keep a higher

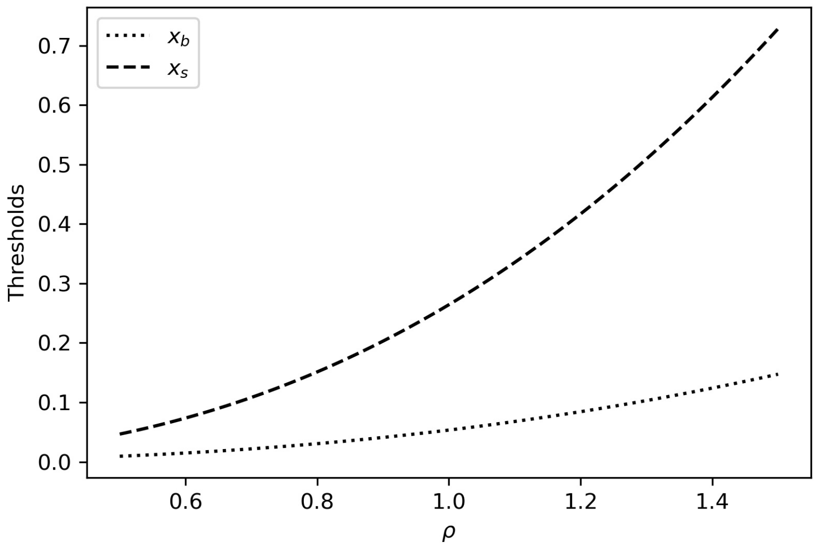


Figure 2.2: Thresholds under different values of ρ.

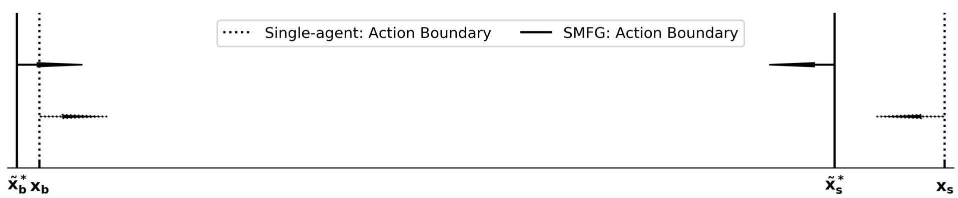
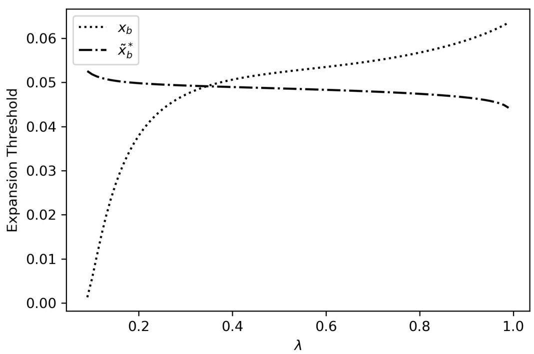


Figure 2.3: Single-agent v.s. MFG

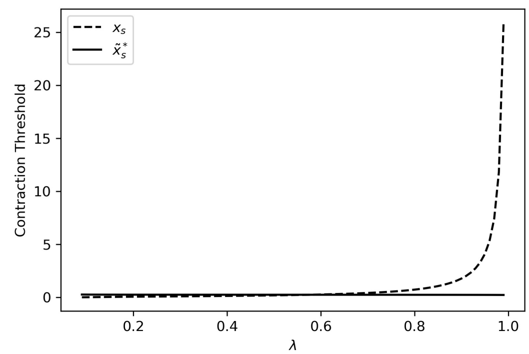
production level and this tendency collectively reduces the price due to the risk-aversion implied by the Cobb-Douglas function.

Impact of δ and γ. The drift coefficient δ represents the expected growth rate of the production and γ measures the volatility of the growth. The decision of whether or not to adjust the production level is the trade-off between the running payoff $c\rho x_t^\alpha$ and the profit from direct intervention $p(1-\lambda)d\xi_t^- - pd\xi_t^+$, with $\alpha \in (0,1)$. Without any intervention within the time interval $[t, t+\Delta t]$, $x_{t+\Delta t}^\alpha$ is given by

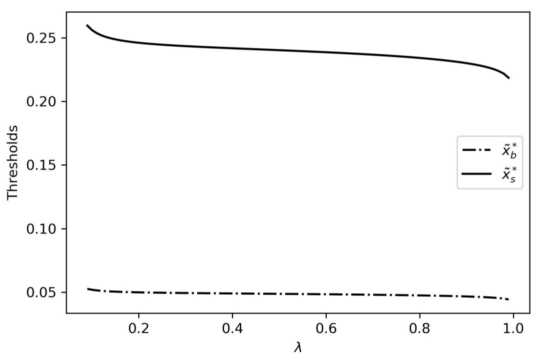
$$x_t^\alpha \exp\left\{[\alpha\delta - \frac{\gamma^2}{2}\alpha(1-\alpha)]\Delta t\right\} \exp\left\{\alpha\gamma(W_{t+\Delta t} - W_t) - \frac{\alpha^2\gamma^2}{2}\Delta t\right\}, \tag{2.30}$$



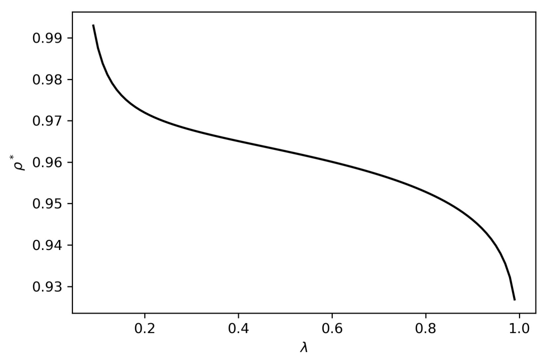
(a) Expansion threshold under different values of λ: single-agent v.s. MFG



(b) Contraction threshold under different values of λ: single-agent v.s. MFG



(c) MFG optimal thresholds versus λ

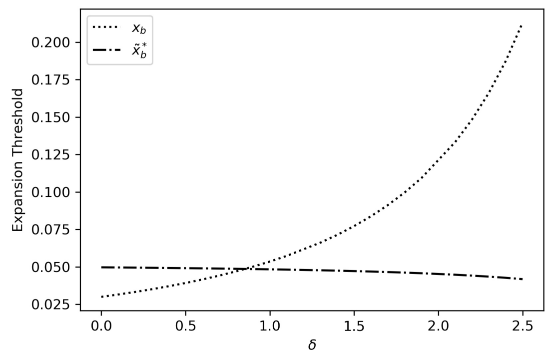


(d) Equilibrium price versus λ

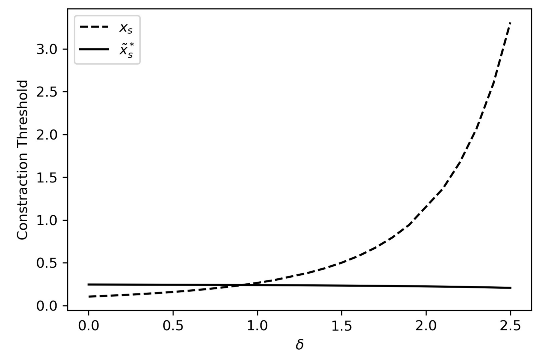
Figure 2.4: Impact of λ.

therefore $\alpha\delta - \frac{\gamma^2}{2}\alpha(1-\alpha)$ represents the expected growth rate of x_t^α. Under the single-agent setting (2.11), when δ increases, the revenue function grows faster, leading to higher expansion and contraction thresholds, as shown in Figures 2.5a and 2.5b. Moreover, the growth in δ has larger impact on the contraction threshold x_s compared to the the expansion threshold x_b. It also implies that each company tends to maintain a higher production level as δ grows. Under the game (MFG-FV), this tendency on the individual level is aggregated, driving down the equilibrium price ρ^*, as shown in Figure 2.5d.

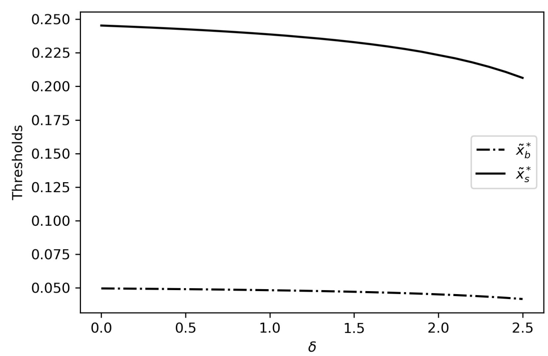
The impact of an increasing γ on both the single-agent control problem and the MFG can be seen from the following two perspectives. As γ increases, the growth rate of the revenue function $\alpha\delta - \frac{\gamma^2}{2}\alpha(1-\alpha)$ decreases, potentially causing lower expansion and contraction thresholds. An increasing γ indicates a larger volatility in the growth rate of the production level and the company can take advantage of the high volatility and reduce the frequency of intervention, potentially decreasing the expansion threshold and increasing the contraction



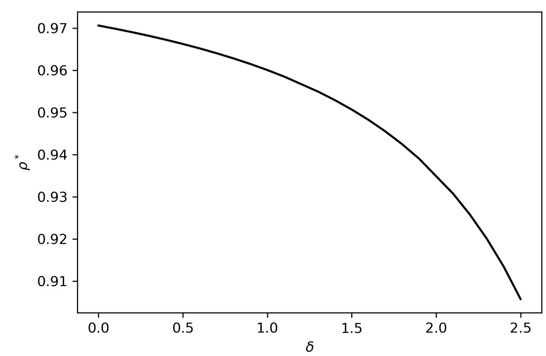
(a) Expansion threshold under different values of δ: single-agent v.s. MFG



(b) Contraction threshold under different values of δ: single-agent v.s. MFG



(c) MFG optimal thresholds versus δ



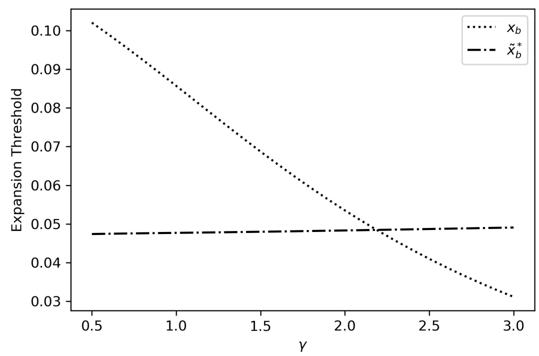
(d) Equilibrium price versus δ

Figure 2.5: Impact of δ.

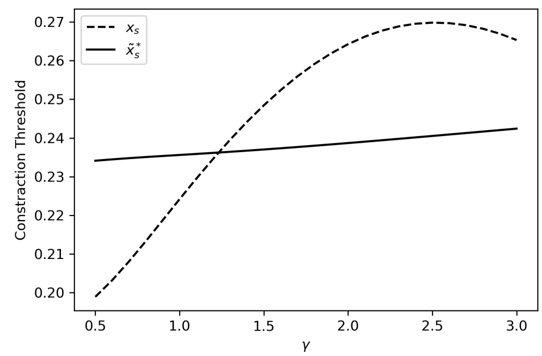
threshold.

Under both perspectives, the expansion threshold is expected to decrease when γ increases. But an increase in γ potentially has opposite effects on the contraction threshold. In the single-agent control problem (2.11), the expansion threshold x_b decreases as expected (Figure 2.6a); the contraction threshold x_s first increases and then decreases (Figure 2.6b). In the game (MFG-FV), the prevailing impact of a decreasing growth rate of x_t^α leads to higher the equilibrium price ρ^*, as shown in Figure 2.6d.

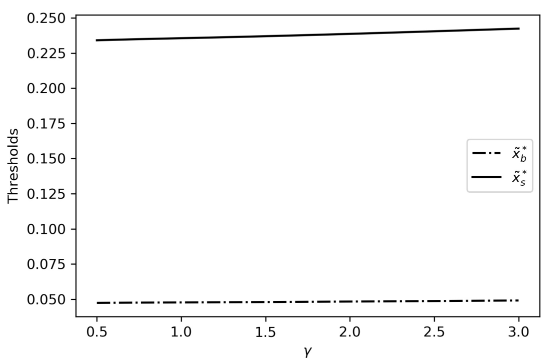
Impact of r. In the single-agent control problem (2.11), as the discount rate r increases, the revenue decays faster as time goes by, thus it becomes more beneficial to decrease the production level for profit and consequently, a significant drop in the contraction threshold in Figure 2.7b. In the game (MFG-FV), the tendency of decreasing production for each company ultimately drives up the equilibrium price, as shown in Figure 2.7d.



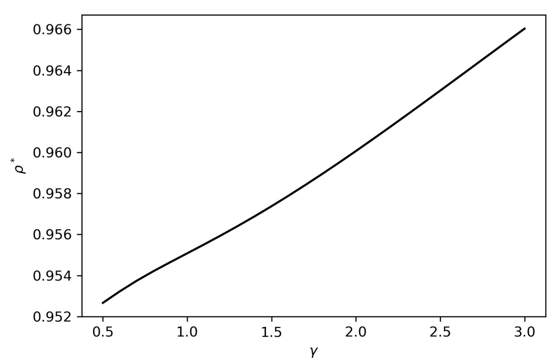
(a) Expansion threshold under different values of γ: single-agent v.s. MFG



(b) Contraction threshold under different values of γ: single-agent v.s. MFG



(c) MFG optimal thresholds versus γ



(d) Equilibrium price versus γ

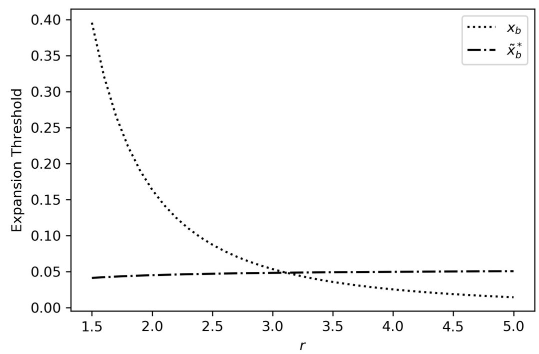
Figure 2.6: Impact of γ.

Impact of α. $\alpha \in (0, 1)$ measures the elasticity of the profit with respect to the production. Under the single-agent setting (2.11), both thresholds first increase and then decrease as α approaches 1. In the game (MFG-FV), the more sensitive the revenue with respect to production, the lower the equilibrium price, as shown in Figure 2.8d.

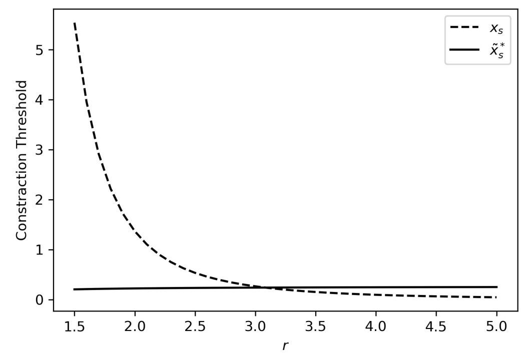
2.2.5 Approximation of N-player game

In Section 2.2.4, we compare the solution to game (MFG-FV) with the solution to the single-agent control problem (2.11), and demonstrate the game effect by analyzing the impact of the model parameters. In this section, we will show that the game (MFG-FV) is an approximation of its associated N-player game, in the sense of ϵ-NE.

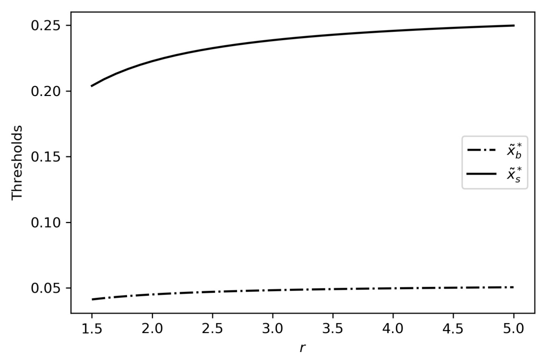
Take the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ that supports a standard Brownian motion $W = \{W_t\}_{t\geq 0}$. Take N identical copies of the Brownian motion W,



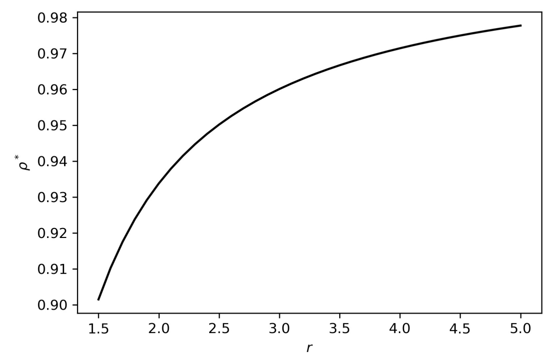
(a) Expansion threshold under different values of r: single-agent v.s. MFG



(b) Contraction threshold under different values of r: single-agent v.s. MFG



(c) MFG optimal thresholds versus r



(d) Equilibrium price versus r

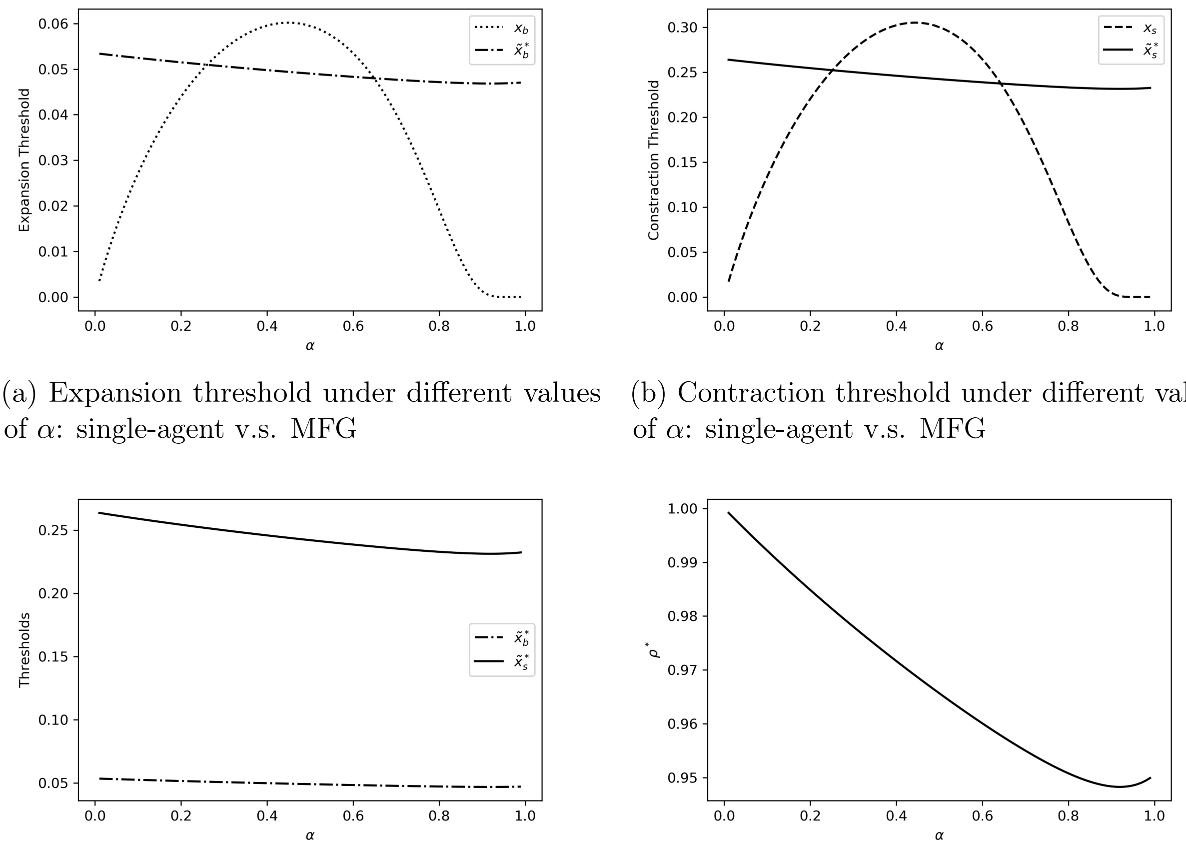
Figure 2.7: Impact of r.

$W^i = \{W_t^i\}_{t\geq 0}$ with $i = 1, \ldots, N$, such that W^i's are i.i.d. and independent of W.

Suppose there are N companies participating in the game of partially reversible investment. For each company i, denote $X^i = \{X_t^i\}_{t\geq 0}$ as its production level on $\mathbb{R}$, with initial states $X_{0-}^i \overset{i.i.d.}{\sim} \mu_0 \in \mathcal{P}^2(\mathbb{R})$. Similar to (2.14) and considering Remark 2.24, define the set of admissible controls $\mathcal{U}^N$ for each company,

$$\begin{aligned}\mathcal{U}^N = \Big\{ (\xi_\cdot^+, \xi_\cdot^-) : \xi_\cdot^\pm \text{ adapted to } \mathbb{F}^{(W^1,\ldots,W^N)}, \text{ nondecreasing, càdlàg,} \\ \xi_{0-}^+ = \xi_{0-}^- = 0,\ \mathbb{E}\left[\int_0^\infty e^{-rt} d\xi_t^+\right] < \infty, \\ \text{controlled process } X_t \geq 0,\ \forall t \geq 0,\ \mathbb{E}\left[\int_0^\infty e^{-rt} X_t^2 dt\right] < \infty \Big\},\end{aligned} \tag{2.31}$$

where $\mathbb{F}^{(W^1,\ldots,W^N)} = \{\mathcal{F}_t^{(W^1,\ldots,W^N)}\}_{t\geq 0}$ is the filtration generated by $(W^1, \ldots, W^N)$. For any



(a) Expansion threshold under different values of α: single-agent v.s. MFG

(b) Contraction threshold under different values of α: single-agent v.s. MFG

(c) MFG optimal thresholds versus α

(d) Equilibrium price versus α

Figure 2.8: Impact of α.

$\xi^i_\cdot = (\xi^{i,+}_\cdot, \xi^{i,-}_\cdot) \in \mathcal{U}^N$, assume that the process $x^i = \{x^i_t\}_{t\geq 0}$ is driven by

$$dX^i_t = X^i_t(\delta dt + \gamma dW^i_t) + d\xi^{i,+}_t - d\xi^{i,-}_t, \quad X^i_{0-} \sim \mu_0. \tag{2.32}$$

Here we consider a similar payoff function for each individual company as in problem (2.11). However, unlike (2.11) where the price in the revenue function is exogenously given, here in the N-player game ρ^i the price for company i is assumed to depend on the average of all its opponents' limiting product levels $\frac{\sum_{j-1} X^j_\infty}{N-1}$, and the price is assumes to be determined by the inverse demand function

$$\tilde{\rho}(x) = a_0 - a_1 x^{1-\alpha},$$

where a_0, a_1 are some positive constants with a_1 satisfying (2.28).

Under a given set of other companies' controls, $\boldsymbol{\xi}_\cdot^{-i} = (\xi_\cdot^1, \ldots, \xi_\cdot^{i-1}, \xi_\cdot^{i+1}, \ldots, \xi_\cdot^N)$, for any $\mathbf{x} \in \mathbb{R}^N$, the payoff function for company i, is given by

$$J^i(\mathbf{x}, \xi_\cdot^i; \boldsymbol{\xi}_\cdot^{-i}) = \mathbb{E}\left[\int_0^\infty e^{-rt}\left[\frac{c(x_t^i)^\alpha}{N-1}\sum_{j\neq i}\tilde{\rho}(x_\infty^j)dt - pd\xi_t^{i,+} + p(1-\lambda)d\xi_t^{i,-}\right]\middle|\mathbf{X}_{0-} = \mathbf{x}\right], \tag{2.33}$$

where $\mathbf{X}_{0-} = (X_{0-}^1, \ldots, X_{0-}^N)$. The objective of company i is to choose the best control policy $\xi^{*,i} \in \mathcal{U}^N$ to maximize the above payoff. That is,

$$\sup_{\xi_\cdot^{i,+}, \xi_\cdot^{i,-} \in \mathcal{U}^N} J^i(\mathbf{x}, \xi_\cdot^i; \boldsymbol{\xi}_\cdot^{-i}) \tag{N-player-S}$$

subject to (2.32).

There are various solution criteria for an N-player game. In this section, we focus on the notion of the Nash equilibrium (NE). An NE of an N-player game is a set of strategies of all agents from which no players has the incentive to unilaterally deviate. More specifically,

Definition 2.25 (NE)**.** $\boldsymbol{\xi}_\cdot^* = (\xi_\cdot^{*,1}, \ldots, \xi_\cdot^{*,N})$ *is called an NE to the game* (N-player-S) *if for any* $i = 1, \ldots, N$,

$$\mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \xi_\cdot^{*,i}; \boldsymbol{\xi}_\cdot^{*,-i})\right] \geq \mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \xi_\cdot^i; \boldsymbol{\xi}_\cdot^{*,-i})\right], \quad \forall \xi_\cdot^i \in \mathcal{U}^N,$$

where $x_{0-}^k \overset{i.i.d.}{\sim} \mu_0$, $k = 1, \ldots, N$.

Solving for such an NE analytically is challenging especially when N is large. We will show that the solution for the game (MFG-FV) in Section (2.2.3) provides an approximation of the game (N-player-S) in the following sense.

Definition 2.26 (ϵ-NE)**.** *For some* $\epsilon > 0$, $\boldsymbol{\xi}_\cdot^* = (\xi_\cdot^{*,1}, \ldots, \xi_\cdot^{*,N})$ *is called an* ϵ*-NE to the game* (N-player-S) *if for any* $i = 1, \ldots, N$,

$$\mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \xi_\cdot^{*,i}; \boldsymbol{\xi}_\cdot^{*,-i})\right] \geq \mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \xi_\cdot^i; \boldsymbol{\xi}_\cdot^{*,-i})\right] - \epsilon, \quad \forall \xi_\cdot^i \in \mathcal{U}^N,$$

where $x_{0-}^k \overset{i.i.d.}{\sim} \mu_0$, $k = 1, \ldots, N$.

To see the approximation, first recall the definition of a solution $(\xi_\cdot^*, \rho^*)$ to the (MFG-FV) given by Definition 2.22 and its explicit form given in Section 2.2.3 characterized by the pair of reflection boundaries and mean information $(\tilde{x}_b^*, \tilde{x}_s^*, \rho^*)$.

Now, for any company $k = 1, \ldots, N$, consider the following admissible control policy $\bar{\xi}_\cdot^k$ characterized by the reflection boundaries $(\tilde{x}_b^*, \tilde{x}_s^*)$ such that for the controlled process $\bar{X}^k = \{\bar{X}_t^k\}_{t\geq 0}$, $\bar{X}_t^k \in [\tilde{x}_b^*, \tilde{x}_s^*]$ for almost all $t \geq 0$. Fix a representative company i. Suppose that for any $j \neq i$, company j decides to take the policy $\bar{\xi}_\cdot^j \in \mathcal{U}^N$. Then $\bar{X}^j$'s are i.i.d.; moreover, according to Definition 2.22, the consistency condition of the solution to the game (MFG-FV) guarantees that $\rho^* = \mathbb{E}[\tilde{\rho}(\bar{x}_\infty^j)]$.

Now denote the set of strategies consisting of $\bar{\xi}^k_\cdot$'s by the following vector

$$\bar{\boldsymbol{\xi}}_\cdot = (\bar{\xi}^1_\cdot, \dots, \bar{\xi}^N_\cdot). \tag{2.34}$$

Then we have

Theorem 2.27. *For any fixed N, $\bar{\boldsymbol{\xi}}_\cdot$ given in* (2.34) *is an ϵ-NE to the game* (N-player-S) *where $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.*

Proof. It suffice to show that

$$\mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \bar{\xi}^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot)\right] \geq \sup_{\xi^i \in \mathcal{U}^N} \mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-}, \xi^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot)\right] - O\left(\frac{1}{\sqrt{N}}\right).$$

Note that the strategies of other companies are fixed as $\bar{\xi}^j_\cdot$, where $j \neq i$. By the continuity of $\tilde{\rho}(\cdot)$ and boundedness of $\bar{X}^j$'s, $\bar{\rho} := \frac{\sum_{j\neq i} \tilde{\rho}(\bar{X}^j_\infty)}{N-1}$ is bounded by a sufficiently large number $R > 0$. Therefore for any $\xi^i_\cdot \in \mathcal{U}^N$,

$$\begin{aligned} J^i(\mathbf{x}, \xi^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot) &\leq J^{u,i}(\mathbf{x}, \xi^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot) \\ &:= \mathbb{E}\left[\int_0^\infty e^{-rt}\left[cR(x^i_t)^\alpha dt - pd\xi^{i,+}_t + p(1-\lambda)d\xi^{i,-}_t\right] \middle| \mathbf{X}_{0-} = \mathbf{x}\right]. \end{aligned}$$

From [91], clearly $\sup_{\xi^i \in \mathcal{U}^N} J^{u,i}(\mathbf{x}, \xi^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot)$ is finite and

$$U := \sup_{\xi^i \in \mathcal{U}^N} J^{u,i}(\mathbf{x}, \xi^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot) < \infty.$$

For some $d > 0$ such that $U - \frac{d}{\sqrt{N}} > 0$, take $\hat{\xi}^i_\cdot \in \mathcal{U}^N$ such that $J^i(\mathbf{x}, \hat{\xi}^i_\cdot; \bar{\boldsymbol{\xi}}^{-i}_\cdot) \geq U - \frac{d}{\sqrt{N}}$ and denote the production level under policy $\hat{\xi}^i_\cdot$ by $\hat{x}^i = \{\hat{x}^i_t\}_{t\geq 0}$. According to (2.31), there exists $L > 0$ $\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(\hat{X}^i_t)^2 dt\right] < L$. Consider $\xi^i_\cdot \in \mathcal{U}^N$ such that the corresponding controlled process x^i satisfies

$$\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(X^i_t)^2 dt\right] < L. \tag{2.35}$$

Take such a control policy $\xi^i_\cdot$.

$$\frac{c(X^i_t)^\alpha}{N-1}\sum_{j\neq i}\tilde{\rho}(\bar{X}^j_\infty) = c(X^i_t)^\alpha\left\{\rho^* + \frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}^j_\infty) - \rho^*\right]}{N-1}\right\}.$$

Since $\bar{x}^j_\cdot \in [\tilde{x}^*_b, \tilde{x}^*_s]$ almost surely, $\tilde{\rho}(\bar{x}^i_\infty)$ is also bounded almost surely. For $j \neq i$, $\bar{x}^j$'s are i.i.d., then

$$\mathbb{E}_{\mu_0}\left|\frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}^j_\infty) - \rho^*\right]}{N-1}\right| \leq \mathbb{E}_{\mu_0}\left[\left|\frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}^j_\infty) - \rho^*\right]}{N-1}\right|^2\right]^{\frac{1}{2}} = O\left(\frac{1}{\sqrt{N}}\right).$$

Therefore,

$$\begin{aligned}
&\left|\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}\frac{c(X_t^i)^\alpha}{N-1}\sum_{j\neq i}\tilde{\rho}(\bar{X}_\infty^j)dt\right]-\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}c\rho^*(X_t^i)^\alpha dt\right]\right|\\
&\leq \mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}c\left|\frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}_\infty^j)-\rho^*\right]}{N-1}\right|(X_t^i)^\alpha dt\right]\\
&\leq \mathbb{E}_{\mu_0}\left[\left|\frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}_\infty^j)-\rho^*\right]}{N-1}\right|^2\right]^{\frac{1}{2}}\mathbb{E}_{\mu_0}\left[\left(\int_0^\infty e^{-rt}(X_t^i)^\alpha dt\right)^2\right]^{\frac{1}{2}}\\
&\leq \mathbb{E}_{\mu_0}\left[\left|\frac{\sum_{j\neq i}\left[\tilde{\rho}(\bar{X}_\infty^j)-\rho^*\right]}{N-1}\right|^2\right]^{\frac{1}{2}}\mathbb{E}_{\mu_0}\left[\frac{1}{r}\int_0^\infty e^{-rt}(X_t^i)^{2\alpha} dt\right]^{\frac{1}{2}},
\end{aligned}$$

where the last inequality is by the Jensen inequality. By (2.35), we have

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(X_t^i)^{2\alpha}dt\right]&=\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(X_t^i)^{2\alpha}\mathbb{1}\{X_t^i\leq 1\}dt\right]+\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(X_t^i)^{2\alpha}\mathbb{1}\{X_t^i> 1\}dt\right]\\
&\leq r+\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}(X_t^i)^2dt\right]\\
&\leq r+L.
\end{aligned}$$

Therefore,

$$\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}\frac{c(X_t^i)^\alpha}{N-1}\sum_{j\neq i}\tilde{\rho}(\bar{X}_\infty^j)dt\right]=\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}c\rho^*(X_t^i)^\alpha dt\right]+O\left(\frac{1}{\sqrt{N}}\right).$$

In particular,

$$\begin{aligned}
&\sup_{\xi_\cdot^i\in\mathcal{U}^N}\mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-},\xi_\cdot^i;\bar{\boldsymbol{\xi}}_\cdot^{-i})\right]-O\left(\frac{1}{\sqrt{N}}\right)\leq \mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-},\hat{\xi}_\cdot^i;\bar{\boldsymbol{\xi}}_\cdot^{-i})\right]\\
&=\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}\left[\frac{c(\hat{X}_t^i)^\alpha}{N-1}\sum_{j\neq i}\tilde{\rho}(X_\infty^j)dt-\gamma^+d\xi_t^{i,+}-\gamma^-d\xi_t^{i,-}\right]\right]\\
&=\mathbb{E}_{\mu_0}\left[\int_0^\infty e^{-rt}\left[c\rho^*(\hat{X}_t^i)^\alpha dt-\gamma^+d\xi_t^{i,+}-\gamma^-d\xi_t^{i,-}\right]\right]+O\left(\frac{1}{\sqrt{N}}\right)\\
&\leq \mathbb{E}_{\mu_0}\left[J^i(\mathbf{X}_{0-},\bar{\xi}_\cdot^i;\bar{\boldsymbol{\xi}}_\cdot^{-i})\right]+O\left(\frac{1}{\sqrt{N}}\right),
\end{aligned}$$

where the last inequality is due to the optimality of $\bar{\xi}_\cdot^i$ according to Step 1 of Section 2.2.3. □

2.3 Conclusion and remarks

This chapter analyzes a class of MFGs with singular controls motivated from the partially reversible problem.

It establishes the existence of a solution to the MFG when controls are of bounded velocity where the approach is adapted from the technique in [118] to the infinite-time-horizon setting with an appropriate modification of topological spaces. It provides an explicit solution to the MFG when the singular controls are of finite variation, presents sensitivity analysis to compare the solution to the MFG with that of the single-agent control problem, and establishes its approximation to the corresponding N-player game in the sense of ϵ-NE, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.

The natural next step is to study the problem of convergence of the N-player game to the associated MFG. Note that this problem has been studied for regular controls in [119, 43, 142]. It will be interesting to explore the case when controls are possibly discontinuous.

Another class of stochastic games with possibly discontinuous controls is impulse control games. Recently there are progresses in this direction, including [3] and [36] for explicit solutions of two-player games and [14] showing solutions of impulse MFGs being ϵ-NE for their corresponding N-player impulse games, with $\epsilon = O(\frac{1}{\sqrt{N}})$. Similar to MFGs with singular controls, it is challenging to establish general NE structures for impulse games, except for some special cases. The main challenge comes from the non-local operator associated with impulse controls, even with one-dimensional state processes.

Finally, it is well known that under proper technical conditions, singular controls of finite variation can be approximated by singular controls of bounded velocity. See for instance [98]. More recently, under the finite-time horizon setting, [90] establishes an ϵ-Nash Equilibrium (ϵ-NE) approximation of N-player games with singular controls of finite variation by MFGs with singular controls of bounded velocity; [65] studies in an N-player game setting the convergence of singular control of bounded velocity to that of finite variation, assuming sub-modularity of the cost function and via the notion of *weak* NE. An immediate question is whether the convergence relation holds in a MFG framework, and if so, under what form of equilibrium. This is an intriguing question to be answered in future works.

Chapter 3

Nonzero-sum stochastic games and mean-field games with impulse controls

Motivated by the classical cash management problem [58] reviewed in Section 1.1.4, this chapter will focus on the analysis of stochastic games with impulse controls, for both N players and its corresponding MFGs.

Consider N players, each managing a flow of cash balance. For player $i \in \{1, \ldots, N\}$, the uncontrolled cash balance is driven by

$$dX_t^i = b_i(X_{t-}^i)dt + \sigma_i(X_{t-}^i)dW_t^i, \qquad X_{0-}^i = x_i,$$

where W^i are independent real Brownian motions. Each player, say player i, chooses a sequence of random (stopping) times $(\tau_{i,1}, \tau_{i,2}, \cdots, \tau_{i,k}, \cdots)$ to intervene and exercise her control. At each $\tau_{i,k}$, the time of this player's k-th intervention, her control is denoted as $\tilde{\xi}_{i,k}$. Given the sequence $\{(\tau_{i,k}, \tilde{\xi}_{i,k})\}_{k\geq 1}$ for player i, the dynamic of X^i becomes

$$dX_t^i = b_i(X_{t-}^i)dt + \sigma_i(X_{t-}^i)dW_t^i + \sum_{\tau_{i,k}\leq t} \delta(t-\tau_{i,k})\tilde{\xi}_{i,k}, \qquad X_{0-}^i = x_i,$$

with $\delta(\cdot)$ the Dirac function. The payoff for player i is

$$\mathbb{E}_x \left[\int_0^\infty e^{-rt} f_i(X_t)dt + \sum_{k=1}^\infty e^{-r\tau_{i,k}} \phi_i(\tilde{\xi}_{i,k}) + \sum_{j\neq i} \sum_{k=1}^\infty e^{-r\tau_{j,k}} \psi_{i,j}(\tilde{\xi}_{j,k}) \right]. \qquad \text{(N-player-I)}$$

Here $X_t = (X_t^1, \cdots, X_t^N)$ with $x = (x_1, \cdots, x_N)$ the starting state, $r > 0$ the discount rate, f_i the running cost, ϕ_i the cost of control for player i, and $\psi_{i,j}$ the cost for player i incurred from player j's control, subject to appropriate conditions to be specified in Section 3.1.3. The goal of each player i is to find the best policy to minimize her cost among a set of admissible game strategies.

For this N-player game, it establishes a general form of system of QVIs and provides the sufficient conditions for the Nash equilibria (NEs) of the game, via the verification theorem approach. For the corresponding MFG, it presents sufficient conditions for the existence of NEs and shows that the solution of the MFG is an ϵ-NE approximation to the N-player game, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$. Through sensitivity analysis and comparisons among $N = 1, 2$ and $N = \infty$ (i.e., MFG), it analyzes the cash management game problem and the effect of competition in games and the collapse of MFG to the single-player game. In particular, it shows that in a game setting players have to take the opponents' strategies into consideration due to competition. Consequently, it is optimal (in the NE sense) that players choose to intervene less frequently; but once set to intervene, players will exert larger amount of controls. In some sense, competition induces more efficient control strategies from players.

3.1 N-Player stochastic games with impulse controls.

3.1.1 Problem formulation.

In this section, we provide the mathematical definition for the N-player stochastic games with impulse controls. The idea is clear and intuitive: N players intervening on a stochastic process by discrete-time intervention. However, the precise mathematical definition presents some non-trivial technicalities with the presence of discontinuous multi-dimensional controlled process.

Domain and underlying process. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space and let $\{W_t\}_{t\geq 0}$ be an M-dimensional Brownian motion with natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Let S be a fixed non-empty subset of $\mathbb{R}^d$, representing the set where the game takes place, in the sense that the game ends when the controlled process exits from S. For example, in portfolio optimization problems the game ends in case of bankruptcy, which may be modelled by choosing $S = (0, \infty)$.

For $t \geq 0$ and $\zeta \in L^2(\mathcal{F}_t)$, we denote by $Y^{t,\zeta} = \{Y^{t,\zeta}_s\}_{s\geq t}$ a solution to the stochastic differential equation

$$\begin{cases} dY^{t,\zeta}_s = b(Y^{t,\zeta}_s)ds + \sigma(Y^{t,\zeta}_s)dW_s, \quad s \geq t, \\ Y^{t,\zeta}_t = \zeta. \end{cases} \tag{3.1}$$

Here, $b : S \to \mathbb{R}^d$ and $\sigma : S \to \mathbb{R}^{d\times M}$ are given Lipschitz-continuous functions, i.e., there exists a constant $K > 0$ such that for all $y_1, y_2 \in S$,

$$|b(y_1) - b(y_2)| + |\sigma(y_1) - \sigma(y_2)| \leq K|y_1 - y_2|.$$

The equation in (3.1) models the underlying process when none of the players intervenes. Since we are going to stop the process as soon as it exits from S, in our framework it is enough to have the functions b and σ defined only on S.

Interventions of the players and impulse controls. N players, indexed by $i \in \{1, \dots, N\}$, can intervene on the process in (3.1) by means of discrete-time interventions. Namely, if player i intervenes with impulse $\delta \in Z_i$, where Z_i is a fixed subset of $\mathbb{R}^{l_i}$, the process is shifted from state x to state $\Gamma^i(x, \delta)$, where $\Gamma^i : S \times Z_i \to S$ is a given function. In most applied settings, the process shifts with a simple translation, i.e., $\Gamma^i(x, \delta) = x + \delta$.

The interventions of player i are described by the sequence $\{(\tau_{i,k}, \xi_{i,k})\}_{k \geq 1}$ (impulse control), where $\{\tau_{i,k}\}_{k \geq 1}$ represent the intervention times and $\{\xi_{i,k}\}_{k \geq 1}$ the corresponding amount of adjustment. Mathematically, $\tau_{i,k}$ is a stopping time with respect to a suitable filtration $\{\widetilde{\mathcal{F}}_t\}_{t \geq 0}$ (see Remark 3.1 below for details), with $\tau_{i,k+1} \geq \tau_{i,k}$, and $\xi_{i,k}$ is a $\widetilde{\mathcal{F}}_{\tau_{i,k}}$-measurable variable, for each $k \geq 1$ and $i \in \{1, \dots, N\}$.

Intervening has a cost or a gain, both for the acting player and for all her opponents. Namely, if x is the current state and player i intervenes with an impulse δ, her cost is $\phi_i(x, \delta)$, whereas the cost for player $j \neq i$ is $\psi_{j,i}(x, \delta)$, for given functions $\phi_i, \psi_{j,i} : S \times Z_i \to \mathbb{R}$. For the game to be well defined, it is necessary to have $\phi_i > 0$. That is, intervening corresponds to a cost, otherwise the game degenerates and the players could improve their payoff by continuously intervening.

Action regions, impulse functions, strategies. As seen, players' interventions on the underlying process are modelled by impulse controls. In the model we propose here, impulse controls originate from a precise strategy that each player preliminarily fixes.

Definition 3.1. *A strategy for player $i \in \{1, \dots, N\}$ is a pair $\varphi_i = (A_i, \xi_i)$, where A_i is a fixed closed subset of $\mathbb{R}^d$ (action region) and $\xi_i : S \to Z_i$ is a continuous function (impulse function). We denote by Φ_i the set of strategies for player i.*

Strategies determine the behaviour of the players, as follows. Fix a starting point $x \in S$ and an N-tuple of strategies $\varphi = (\varphi_1, \dots, \varphi_N)$, where $\varphi_i = (A_i, \xi_i) \in \Phi_i$ is the strategy of player i and the sets A_i are pairwise disjoint, that is, $A_i \cap A_j = \emptyset$ for $i \neq j$. Then, N impulse controls $\{(\tau_{i,k}^{x;\varphi}, \xi_{i,k}^{x;\varphi})\}_{k \geq 1}$ (the players' interventions), a right-continuous process $X^{x;\varphi}$ (the controlled process), a stopping time $\tau_S^{x;\varphi}$ (the end of the game) are uniquely defined by the following two rules.

1. Player i intervenes if and only if the process enters the set A_i, in which case the impulse is given by $\xi_i(y)$, where y is the current state. Recall that choosing $\xi_i(y)$ as the intervention impulse means that player i shifts the process from state y to state $\Gamma^i(y, \xi_i(y))$, as introduced earlier.

2. The game ends when the process exits from S.

More precisely, $\{(\tau_{i,k}^{x;\varphi}, \xi_{i,k}^{x;\varphi})\}_{k \geq 1}$, $X^{x;\varphi}$, $\tau_S^{x;\varphi}$ are defined in the following Definition 3.2, where we use the conventions $\inf \emptyset = \infty$ and $[\infty, \infty) = \emptyset$.

Definition 3.2. *Let $x \in S$ and $\varphi = (\varphi_1, \dots, \varphi_N)$, where $\varphi_i = (A_i, \xi_i) \in \Phi_i$ is a strategy for player $i \in \{1, \dots, N\}$. Assume that $A_i \cap A_j = \emptyset$, for $i \neq j$. For $k \in \{0, \dots, \bar{k}\}$, where*

$\bar{k} = \sup\{k \in \mathbb{N} \cup \{0\} : \widetilde{\tau}_k < \alpha_k^S\}$, *we define by induction* $\widetilde{\tau}_0 = 0$, $x_0 = x$, $\widetilde{X}^0 = Y^{\widetilde{\tau}_0, x_0}$, $\alpha_0^S = \infty$, *and*

$$
\begin{aligned}
&\alpha_k^O = \inf\{s > \widetilde{\tau}_{k-1} : \widetilde{X}_s^{k-1} \notin O\}, && \text{[exit time from } O \subseteq S\text{]}\\
&\widetilde{\tau}_k = \min\{\alpha_k^{A_1}, \dots, \alpha_k^{A_N}\}, && \text{[intervention time]}\\
&m_k = \mathbb{1}_{\{\widetilde{\tau}_k = \alpha_k^{A_1}\}} + \dots + N\mathbb{1}_{\{\widetilde{\tau}_k = \alpha_k^{A_N}\}}, && \text{[index of the player interv. at } \widetilde{\tau}_k\text{]}\\
&\widetilde{\xi}_k = \xi_{m_k}\big(\widetilde{X}_{\widetilde{\tau}_k}^{k-1}\big), && \text{[impulse]}\\
&x_k = \Gamma^{m_k}\big(\widetilde{X}_{\widetilde{\tau}_k}^{k-1}, \widetilde{\xi}_k\big), && \text{[starting point for the next step]}\\
&\widetilde{X}^k = \widetilde{X}^{k-1}\mathbb{1}_{[0,\widetilde{\tau}_k[} + Y^{\widetilde{\tau}_k, x_k}\mathbb{1}_{[\widetilde{\tau}_k,\infty[}. && \text{[contr. process up to the k-th interv.]}
\end{aligned}
$$

Let $\bar{k}_i$ *be the number of interventions by player* $i \in \{1, \dots, N\}$ *before the end of the game, and, in the case where* $\bar{k}_i \neq 0$, *let* $\eta(i,k)$ *be the index of her* k*-th intervention* $(1 \leq k \leq \bar{k}_i)$*:*

$$\bar{k}_i = \sum\nolimits_{1 \leq h \leq \bar{k}} \mathbb{1}_{\{m_h = i\}}, \qquad \eta(i,k) = \min\Big\{l \in \mathbb{N} : \sum\nolimits_{1 \leq h \leq l} \mathbb{1}_{\{m_h = i\}} = k\Big\}.$$

Assume now that the times $\{\widetilde{\tau}_k\}_{0 \leq k \leq \bar{k}}$ *never accumulate strictly before* $\alpha_{\bar{k}}^S$. *That is, we assume that* $\lim_{k \to \bar{k}} \widetilde{\tau}_k = \alpha_{\bar{k}}^S$ *in the event* $\{\bar{k} = +\infty\}$, *with the convention* $\alpha_\infty^S = \sup_k \alpha_k^S$. *The controlled process* $X^{x;\varphi}$ *and the exit time* $\tau_S^{x;\varphi}$ *are defined by*

$$X^{x;\varphi} := \widetilde{X}^{\bar{k}}, \qquad \tau_S^{x;\varphi} := \alpha_{\bar{k}}^S = \inf\{s \geq 0 : X_s^{x;\varphi} \notin S\},$$

with the convention $\widetilde{X}^\infty = \lim_{k \to +\infty} \widetilde{X}^k$. *Finally, the impulse controls* $\{(\tau_{i,k}^{x;\varphi}, \xi_{i,k}^{x;\varphi})\}_{k \geq 1}$, *with* $i \in \{1, \dots, N\}$, *are defined by*

$$\tau_{i,k}^{x;\varphi} := \begin{cases} \widetilde{\tau}_{\eta(i,k)}, & k \leq \bar{k}_i, \\ \tau_S^{x;\varphi}, & k > \bar{k}_i, \end{cases} \qquad \xi_{i,k}^{x;\varphi} := \begin{cases} \widetilde{\xi}_{\eta(i,k)}, & k \leq \bar{k}_i, \\ 0, & k > \bar{k}_i. \end{cases} \tag{3.2}$$

Notice that, if player i intervenes a finite number of times, i.e., $\bar{k}_i = \bar{k}_i(\omega)$ is finite, then the tail of the control is conventionally set to $(\tau_{i,k}, \xi_{i,k}) = (\tau_S, 0)$ for $k > \bar{k}_i$. The following lemma characterizes precisely the controlled process $X^{x;\varphi}$.

Lemma 3.3. *Let* $x \in S$ *and* $\varphi = (\varphi_1, \dots, \varphi_N)$, *where* $\varphi_i = (A_i, \xi_i) \in \Phi_i$ *is a strategy for player* $i \in \{1, \dots, N\}$. *Let* $X = X^{x;\varphi}$, $\tau_S = \tau_S^{x;\varphi}$, $\tau_{i,k} = \tau_{i,k}^{x;\varphi}$, $\xi_{i,k} = \xi_{i,k}^{x;\varphi}$ *be as in Definition 3.2, for* $i \in \{1, \dots, N\}$ *and* $k \geq 1$. *Then,*

- *X admits the following representation, with* $\widetilde{\tau}_k$, x_k *as in Definition 3.2 and* Y *as in* (3.1)*:*

$$X_s = \sum_{k=0}^{\bar{k}-1} Y_s^{\widetilde{\tau}_k, x_k}\mathbb{1}_{[\widetilde{\tau}_k, \widetilde{\tau}_{k+1}[}(s) + Y_s^{\widetilde{\tau}_{\bar{k}}, x_{\bar{k}}}\mathbb{1}_{[\widetilde{\tau}_{\bar{k}}, \infty[}(s). \tag{3.3}$$

- *X is right-continuous. More precisely, X is continuous in $[0,\infty) \setminus \{\tau_{i,k} : \tau_{i,k} < \tau_S\}$ and discontinuous in $\{\tau_{i,k} : \tau_{i,k} < \tau_S\}$, where*

$$X_{\tau_{i,k}} = \Gamma^i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big), \qquad \xi_{i,k} = \xi_i\big(X_{(\tau_{i,k})^-}\big), \qquad X_{(\tau_{i,k})^-} \in \partial A_i. \tag{3.4}$$

- *X never exits from the set $(A_1 \cup \cdots \cup A_N)^c$.*

Proof. We just prove the first property in (3.4), the other ones being immediate. Let $i \in \{1,\dots,N\}$, $k \geq 1$ with $\tau_{i,k} < \tau_S$ and set $\sigma = \eta(i,k)$, with η as in Definition 3.2. By (3.2), (3.3) and Definition 3.2, we have

$$\begin{aligned} X_{\tau_{i,k}} = X_{\widetilde{\tau}_\sigma} = Y^{\widetilde{\tau}_\sigma, x_\sigma}_{\widetilde{\tau}_\sigma} = x_\sigma = \Gamma^i\big(\widetilde{X}^{\sigma-1}_{\widetilde{\tau}_\sigma}, \widetilde{\xi}_\sigma\big) \\ = \Gamma^i\big(\widetilde{X}^{\sigma-1}_{(\widetilde{\tau}_\sigma)^-}, \widetilde{\xi}_\sigma\big) = \Gamma^i\big(X_{(\widetilde{\tau}_\sigma)^-}, \widetilde{\xi}_\sigma\big) = \Gamma^i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big), \end{aligned}$$

where the fifth equality is by the continuity of the process $\widetilde{X}^{\sigma-1}$ in $[\widetilde{\tau}_{\sigma-1}, \infty)$ and the next-to-last equality follows from $\widetilde{X}^{\sigma-1} \equiv X$ in $[0, \widetilde{\tau}_\sigma)$. □

Remark 3.1. *For $x \in S$ and $\varphi \in \Phi_x$, let $\{\mathcal{F}^X_t\}_{t\geq 0}$ denote the natural filtration of the process $X = X^{x;\varphi}$. Then, by construction, $\tau_{i,k}$ is a stopping time with respect to the filtration $\{\mathcal{F}^X_{t^-}\}_{t\geq 0}$ and $\xi_{i,k}$ is a $\mathcal{F}^X_{\tau_{i,k}}$-measurable random variable, for $i \in \{1,\dots,N\}$ and $k \in \mathbb{N}$.*

Remark 3.2. *In single-player impulse control problems (e.g., [146]), the optimal intervention times are recursively defined by*

$$\tau_{k+1} = \inf\{s \geq \tau_k : X^k_t \in A\}, \tag{3.5}$$

for a suitable set A, where X^k represents the controlled process after the k-th intervention. Notice that this procedure cannot be directly extended to N-player impulse games: In a game setting, the intervention times of player i also depend on her opponents' past interventions, so that (3.5) would not be well defined in this case. To overcome this technical difficulty and provide a rigorous framework, we have introduced the definition of strategy.

Objective functions. Each player aims at minimizing her objective function, made up of four terms: a continuous-time running cost in $[0,\tau_S]$, the discrete-time costs associated to her own interventions, the discrete-time costs associated to her opponents' interventions, a terminal cost if the game ends.

More precisely, let $f_i, h_i : S \to \mathbb{R}^d$ be given functions, and let $\rho_i > 0$ be strictly positive constants, for $i \in \{1,\dots,N\}$. For more technical details on the existence and uniqueness of the solution to impulse control problems, see [92]. The functional that player i aims at minimizing is defined as follows.

Definition 3.4. *Let $x \in S$ and $\varphi = (\varphi_1, \dots, \varphi_N)$ be a N-tuple of strategies. For $i \in \{1, \dots, N\}$, provided that the right-hand side exists and is finite, we set*

$$J^i(x;\varphi) := \mathbb{E}_x\bigg[\int_0^{\tau_S} e^{-\rho_i s}\, f_i(X_s)ds + \sum_{\substack{k\in\mathbb{N}\\ \tau_{i,k}<\tau_S}} e^{-\rho_i\tau_{i,k}}\,\phi_i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big)$$
$$+ \sum_{\substack{1\leq j\leq N\\ j\neq i}} \sum_{\substack{k\in\mathbb{N}\\ \tau_{i,k}<\tau_S}} e^{-\rho_i\tau_{j,k}}\,\psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big) + e^{-\rho_i\tau_S}\, h_i(X_{\tau_S})\mathbb{1}_{\{\tau_S<+\infty\}}\bigg], \quad (3.6)$$

with $X = X^{x;\varphi}$, $\tau_S = \tau_S^{x;\varphi}$, $\{(\tau_{i,k}, \xi_{i,k})\}_{k\geq 1} = \{(\tau_{i,k}^{x;\varphi}, \xi_{i,k}^{x;\varphi})\}_{k\geq 1}$ as in Definition 3.2.

The subscript in the expectation denotes, as in control theory, conditioning with respect to starting point $X_t^{x;\varphi} = x$. To shorten the notations, we will often omit the initial state and write $\mathbb{E}$. Also, notice that in the summations we only consider times strictly smaller than τ_S: indeed, since the game ends in $(\tau_S)^-$, interventions in the form $\tau_{i,k} = \tau_S$ are meaningless for the game.

Admissible strategies and Nash equilibria. Before defining a Nash equilibrium (NE) for the game, we define, for each starting point $x \in S$, the set Φ_x of admissible strategies, i.e., strategies as in Definition 3.1 with additional properties assuring that the game is well defined.

Definition 3.5. *Let $x \in S$ and $\varphi_i = (A_i, \xi_i)$ be a strategy for player $i \in \{1, \dots, N\}$. We say that the N-tuple $\varphi = (\varphi_1, \dots, \varphi_N)$ is x-admissible, written as $\varphi \in \Phi_x$, if:*

1. *the sets $A_1, \dots, A_N$ are pairwise disjoint, that is, $A_i \cap A_j = \emptyset$ for $i \neq j$;*
2. *for $i, j \in \{1, \dots, N\}$ with $i \neq j$, the following random variables are in $L^1(\Omega)$:*

$$\int_0^{\tau_S} e^{-\rho_i s}|f_i|(X_s)ds, \qquad e^{-\rho_i\tau_S}|h_i|(X_{\tau_S}),$$
$$\sum_{\tau_{i,k}<\tau_S} e^{-\rho_i\tau_{i,k}}|\phi_i|(X_{(\tau_{i,k})^-}, \xi_{i,k}), \qquad \sum_{\tau_{i,k}<\tau_S} e^{-\rho_i\tau_{i,k}}|\psi_{i,j}|(X_{(\tau_{j,k})^-}, \xi_{j,k}); \quad (3.7)$$

3. *for each $i \in \{1, \dots, N\}$ and $p \in \mathbb{N}$, the random variable $\|X\|_\infty = \sup_{t\geq 0}|X_t|$ is in $L^p(\Omega)$:*
$$\mathbb{E}[\|X\|_\infty^p] < \infty; \quad (3.8)$$
4. *for $i \in \{1, \dots, N\}$, we have*
$$\lim_{k\to+\infty} \tau_{i,k} = \tau_S. \quad (3.9)$$

The first condition in Definition 3.5 ensures that there are no conflicts due to two or more players willing to intervene at same time (see Remarks 3.3 and 3.4 below for further comments on this condition). The second condition assures that the functionals $J^i(x;\varphi)$ in (3.6) are well-defined, for each $i \in \{1,\dots,N\}$. The third condition will be used in the proof of the verification theorem where sufficient conditions for the NEs are specified. Finally, the fourth condition prevents the players from accumulating the interventions before the end of the game.

We now provide the definition of NE and payoffs. Given a tuple of strategies $\varphi = (\varphi_1,\dots,\varphi_N)$, an index $i \in \{1,\dots,N\}$ and a strategy $\bar{\varphi} \in \Phi_i$, we denote by $(\varphi^{-i},\bar{\varphi})$ the N-tuple we get when substituting the i-th component of φ by $\bar{\varphi}$, that is

$$(\varphi^{-i},\bar{\varphi}) := (\varphi_1,\dots,\varphi_{i-1},\bar{\varphi},\varphi_{i+1},\dots,\varphi_N).$$

Definition 3.6. *Given $x \in S$, we say that the admissible N-tuple of strategies $\varphi^* \in \Phi_x$ is a NE of the game if*

$$J^i(x;\varphi^*) \le J^i(x;(\varphi^{*,-i},\varphi_i)),$$

for each $i \in \{1,\dots,N\}$ and each $\varphi_i \in \Phi_i$ such that $(\varphi^{*,-i},\varphi_i) \in \Phi_x$. Finally, if $x \in S$ and $\varphi^* \in \Phi_x$ is a NE, then the payoff associated with the equilibrium φ^* for player $i \in \{1,\dots,N\}$ is

$$V_i(x) := J^i(x;\varphi^*).$$

Remark 3.3. *If the action regions are not pairwise disjoint (i.e., two or more players would like to intervene at the same time), one sets specific rules deciding which player has the priority. For example, player i may have priority over player j whenever $i > j$; otherwise, priority may be given to the player who is the farthest away from the state he would shift the process to.*

Several formulation are possible to handle priorities among players. Our formulation is based on the idea that, since priority rules practically partition the conflict regions, it is not restrictive to assume that the action regions are pairwise disjoint, as we now detail.

Let $\tilde{A}_1,\dots,\tilde{A}_N$ denote the action regions before any priority rules is set, so that the sets $\tilde{A}_i$ are possibly non-disjoint. Let C denote the region where two or more players would like to simultaneously intervene, $C := \cup_{i\neq j}(\tilde{A}_i \cap \tilde{A}_j)$. Deciding which player has the priority corresponds to choosing a partition $C_1,\dots,C_N$ of C, with the additional property that $C_i \subseteq \tilde{A}_i$: namely, if a point belongs to the conflict region, $x \in C$, then it is player i who intervenes, where i is the only index such that $x \in C_i$. The actual action regions A_i are then defined by $A_i = (\tilde{A}_i \setminus C) \cup C_i$, clearly pairwise disjoint. Hence, whatever the priority rule is, we get a N-uple of pairwise disjoint action regions, so that condition 1 in Definition 3.5 is not restrictive.

The advantage of this formulation is twofold. On one hand, no specific priority rules are embedded in the model, which provides more flexibility with respect to, e.g., the approach in [2] (if $N = 2$, choosing $A_1 = \tilde{A}_1$ and $A_2 = \tilde{A}_2 \setminus \tilde{A}_1$ in our setting retrieves the priority rules in [2]). On the other hand, this helps relieving the notational burden, since we do not have to deal with multiple intersections of the action regions when defining the controlled process.

Remark 3.4. *We remark that the present setting allows two or more players to intervene, one right after the other, at a same instant $t \geq 0$. For example, if the present state is $X_{t-} = x \in A_{i_1}$, then player i_1 intervenes and move it to x'. If x' happens to be in A_{i_2}, for some i_2, then player i_2 immediately intervenes, moving again the state to x''. Overall, the state jumped from $X_{t-} = x$ to $X_t = x''$. In order to have a well-defined process, only finitely many players can intervene in t, which is guaranteed by condition* (3.9) *above.*

3.1.2 Verification theorem.

In this section we establish a verification theorem for the games defined in Section 3.1, providing sufficient conditions to determine the payoffs and an NE. This verification theorem links the impulse games with a suitable system of quasi-variational inequalities (QVI). Note that a special case of this verification theorem for $N = 2$ was presented in [2].

In Section 3.1.2 we heuristically introduce the system of QVIs, providing the intuition behind each equation involved. These arguments are made rigorous in Section 3.1.2, with the precise statement and proof of the verification theorem.

The quasi-variational inequalities.

We start by heuristically guessing an expression for a NE $\varphi^* = (\varphi_1^*, \dots, \varphi_N^*)$ and for the corresponding payoffs V_i of the game.

Consider a game as in Section 3.1. Assume for a moment that the payoffs V_i, $i \in \{1, \dots, N\}$ are known. Moreover, assume that for every i there exists a (unique) function $\xi_i : S \to Z_i$ such that

$$\{\xi_i(x)\} = \arg\min_{\delta \in Z_i} \big\{ V_i(\Gamma^i(x, \delta)) + \phi_i(x, \delta) \big\}, \tag{3.10}$$

for each $x \in S$. We define the intervention operators by

$$\begin{aligned} \mathcal{M}_i V_i(x) &= V_i\big(\Gamma^i(x, \xi_i(x))\big) + \phi_i\big(x, \xi_i(x)\big), \\ \mathcal{H}_{i,j} V_i(x) &= V_i\big(\Gamma^j(x, \xi_j(x))\big) + \psi_{i,j}\big(x, \xi_j(x)\big), \end{aligned} \tag{3.11}$$

for $x \in S$ and $i, j \in \{1, \dots, N\}$, with $i \neq j$.

The functions in (3.10) and (3.11) are intuitive. If x is the current state of the process, and player i (resp. player j) intervenes with impulse δ, the payoff for player i can be represented as $V_i(\Gamma^i(x, \delta)) + \phi_i(x, \delta)$ (resp. $V_i(\Gamma^j(x, \delta)) + \psi_{i,j}(x, \delta)$), that is, as the sum of the intervention cost and the payoff in the new state. As a consequence, $\xi_i(x)$ in (3.10) is the impulse that player i would use in case she decides to intervene. Similarly, $\mathcal{M}_i V_i(x)$ (resp. $\mathcal{H}_{i,j} V_i(x)$) represents the payoff for player i when player i (resp. player $j \neq i$) takes the best immediate action and behaves optimally afterwards.

Notice that it is not always optimal to intervene, so $\mathcal{M}_i V_i(x) \geq V_i(x)$, for each $x \in S$, and that player i should intervene (with impulse $\xi_i(x)$) only if $\mathcal{M}_i V_i(x) = V_i(x)$. As a consequence, provided that an explicit expression for V_i is available, an NE is heuristically given by $\varphi^* = (\varphi_1^*, \dots, \varphi_N^*)$, where $\varphi_i^* = (A_i^*, \xi_i^*)$ is given, for each $i \in \{1, \dots, N\}$, by

$$A_i^* = \{\mathcal{M}_i V_i - V_i = 0\}, \qquad \xi_i^* = \xi_i.$$

Practically, this means that player i intervenes when the process enters the region $\{\mathcal{M}_i V_i - V_i = 0\}$, i.e, when $\mathcal{M}_i V_i(x) = V_i(x)$. When this happens, her impulse is $\xi_i(x)$, where x is the current state. The verification theorem in the next section will give a rigorous proof to this heuristic argument.

To complete the argument, we need to determine the payoffs V_i. Assume that V_i are smooth enough so that we can define

$$\mathcal{L}V_i = b \cdot \nabla V_i + \frac{1}{2}\text{tr}\left(\sigma\sigma^t D^2 V_i\right), \tag{3.12}$$

where b, σ are as in (3.1), σ^t denotes the transpose of σ and $\nabla V_i, D^2 V_i$ are the gradient and the Hessian matrix of V_i, respectively. Then V_i should satisfy the following quasi-variational inequalities (QVIs), where $i, j \in \{1, \dots, N\}$:

$$\begin{cases} V_i = h_i, & \text{in } \partial S, & \text{(3.13a)} \\ \mathcal{M}_j V_j - V_j \geq 0, & \text{in } S, & \text{(3.13b)} \\ \mathcal{H}_{i,j} V_i - V_i = 0, & \text{in } \bigcup_{j \neq i}\{\mathcal{M}_j V_j - V_j = 0\}, & \text{(3.13c)} \\ \min\left\{\mathcal{L}V_i - \rho_i V_i + f_i, \mathcal{M}_i V_i - V_i\right\} = 0, & \text{in } \bigcap_{j \neq i}\{\mathcal{M}_j V_j - V_j > 0\}. & \text{(3.13d)} \end{cases}$$

Notice that there is a small abuse of notation in (3.13a), as V_i is not defined in ∂S, so that (3.13a) means $\lim_{y \to x} V_i(y) = h_i(x)$, for each $x \in \partial S$.

Intuition behind each in (3.13): the terminal condition (3.13a) is obvious, and (3.13b), already stated above, is a standard condition in impulse control theory. As for (3.13c), if player j intervenes (i.e., $\mathcal{M}_j V_j - V_j = 0$), by the definition on NE, we expect no losses for player $i \neq j$, that is $\mathcal{H}_{i,j} V_i - V_i = 0$. Meanwhile, if all the players except i are not intervening (hence, $\mathcal{M}_j V_j - V_j > 0$ for all $j \neq i$), then player i faces a classical one-player impulse control problem, so that V_i satisfies the corresponding QVI of $\min\left\{\mathcal{L}V_i - \rho_i V_i + f_i, \mathcal{M}_i V_i - V_i\right\} = 0$, which is (3.13d). In short, the latter condition says that $\mathcal{L}V_i - \rho_i V_i + f_i = 0$ when she does not intervene, whereas $\mathcal{L}V_i - \rho_i V_i + f_i \geq 0$ when she intervenes.

Remark 3.5. *For any player i, the region where she chooses not to intervene, as in* (3.13d) *when* $\min\{\mathcal{L}V_i - \rho_i V_i + f_i, \mathcal{M}_i V_i - V_i\} = \mathcal{L}V_i - \rho_i V_i + f_i = 0$*, is decided by not just player i but all N players; it is indeed the common non-action region C. On C, it is necessary to have $\mathcal{L}V_i - \rho_i V_i + f_i = 0$ for all $i \in \{1, \dots, N\}$. The condition that $\mathcal{M}_i V_i - V_i \geq 0$, however, needs an extra verifying step: it is not entirely player i's decision to wait, yet this choice has to be the best one she can make at a NE. This marks the subtlety of the NE and one crucial difference between the single-player control problem and the multi-player game.*

Statement and proof.

We now provide a rigorous proof of the results heuristically introduced in the previous section. Notations and assumptions from Section 3.1 are adopted from now on.

Theorem 3.7 (Verification Theorem). *Let $V_1, \dots, V_N$ be functions from S to $\mathbb{R}$, assume that*

$$\{\xi_i(x)\} = \arg\min_{\delta \in Z_i} \big\{ V_i(\Gamma^i(x,\delta)) + \phi_i(x,\delta) \big\}$$

holds and set $\mathcal{D}_i := \{\mathcal{M}_i V_i - V_i > 0\}$. Moreover, for $i \in \{1,\dots,N\}$ assume that:

(i) *V_i is a solution to* (3.13a)-(3.13d)*;*

(ii) *$V_i \in C^2(\cap_{j\neq i}\mathcal{D}_j \setminus \partial\mathcal{D}_i) \cap C^1(\cap_{j\neq i}\mathcal{D}_j) \cap C(\overline{\cap_{j\neq i}\mathcal{D}_j})$ and it has polynomial growth;*

(iii) *$\partial\mathcal{D}_i$ is a Lipschitz surface, and V_i has locally bounded derivatives up to the second order in some neighbourhood of $\partial\mathcal{D}_i$.*

Finally, let $x \in S$ and define $\varphi^* = (\varphi_1^*, \dots, \varphi_N^*)$, with

$$\varphi_i^* := (A_i^*, \xi_i^*), \qquad A_i^* := \{\mathcal{M}_i V_i - V_i = 0\}, \qquad \xi_i^* := \xi_i,$$

where $i \in \{1,\dots,N\}$ and the function ξ_i is as in (3.10)*. Then, provided that $\varphi^* \in \Phi_x$,*

$$\varphi^* \text{ is an NE and } V_i(x) = J^i(x;\varphi^*) \text{ for } i \in \{1,\dots,N\}.$$

Proof. Let $x \in S$, $i \in \{1,\dots,N\}$ and $\varphi_i \in \Phi_i$ such that $(\varphi^{*,-i}, \varphi_i) \in \Phi_x$. Notice that $(\varphi^{*,-i}, \varphi_i)$ corresponds to the case where all the players except player i behave optimally. By Definition 3.6, we have to prove that

$$V_i(x) = J^i(x;\varphi^*), \qquad V_i(x) \le J^i(x;(\varphi^{*,-i}, \varphi_i)).$$

Step 1: $V_i(x) \le J^i(x;(\varphi^{*,-i}, \varphi_i))$. To simplify the notations, we omit the dependence on i, x, φ and write

$$X = X^{x;(\varphi^{*,-i},\varphi_i)}, \qquad \tau_{j,k} = \tau_{j,k}^{x;(\varphi^{*,-i},\varphi_i)}, \qquad \xi_{j,k} = \xi_{j,k}^{x;(\varphi^{*,-i},\varphi_i)}. \tag{3.14}$$

The properties in Lemma 3.3 imply that, for $j \neq i$, $s \ge 0$, $\tau_{j,k} < \infty$,

$$(\mathcal{M}_j V_j - V_j)\big(X_s\big) > 0, \tag{3.15a}$$
$$(\mathcal{M}_j V_j - V_j)\big(X_{(\tau_{j,k})^-}\big) = 0, \tag{3.15b}$$
$$\xi_{j,k} = \xi_j\big(X_{(\tau_{j,k})^-}\big). \tag{3.15c}$$

We first approximate V_i with regular functions. Since (ii) and (iii) hold, by [144, proof of Thm. 10.4.1 and App. D] there exists a sequence of functions $\{V_{i,m}\}_{m\in\mathbb{N}}$ such that:

(a) $V_{i,m} \in C^2(\cap_{j\neq i}\mathcal{D}_j) \cap C^0(\overline{\cap_{j\neq i}\mathcal{D}_j})$ for each $m \in \mathbb{N}$ (in particular, the function $\mathcal{L}V_{i,m}$ is well-defined in $\cap_{j\neq i}\mathcal{D}_j$);

(b) $V_{i,m} \to V_i$ as $m \to \infty$, uniformly on the compact subsets of $\overline{\cap_{j\neq i}\mathcal{D}_j}$;

(c) $\{\mathcal{L}V_{i,m}\}_{m\in\mathbb{N}}$ is locally bounded in $\cap_{j\neq i}\mathcal{D}_j$ and $\mathcal{L}V_{i,m} \to \mathcal{L}V_i$ as $m \to \infty$, uniformly on the compact subsets of $\cap_{j\neq i}\mathcal{D}_j \setminus \partial\mathcal{D}_i$.

For each $r > 0$ and $\ell \in \mathbb{N}$, we set

$$\tau_{r,\ell} = \tau_r \wedge \tau_{1,\ell} \wedge \cdots \wedge \tau_{N,\ell}, \tag{3.16}$$

where $\tau_r = \inf\{s > 0 : X_s \notin B(0,r)\}$ is the exit time from the ball with radius r. By (3.15a) we have that $X_s \in \cap_{j\neq i}\mathcal{D}_j$ for each $s > 0$. Since $V_{i,m} \in C^2(\cap_{j\neq i}\mathcal{D}_j)$ by (a), for each $m \in \mathbb{N}$ we can apply Itô's formula to the process $e^{-\rho_i t}V_{i,m}(X_t)$ over the interval $[0,\tau_{r,\ell})$. Taking the conditional expectations, we get

$$\begin{aligned} V_{i,m}(x) = \mathbb{E}_x\bigg[&-\int_0^{\tau_{r,\ell}} e^{-\rho_i s}(\mathcal{L}V_{i,m}-\rho_i V_{i,m})(X_s)ds - \sum_{\tau_{i,k}<\tau_{r,\ell}} e^{-\rho_i\tau_{i,k}}\Big(V_{i,m}\big(X_{\tau_{i,k}}\big)-V_{i,m}\big(X_{(\tau_{i,k})^-}\big)\Big) \\ &- \sum_{j\neq i}\sum_{\tau_{j,k}<\tau_{r,\ell}} e^{-\rho_i\tau_{j,k}}\Big(V_{i,m}\big(X_{\tau_{j,k}}\big) - V_{i,m}\big(X_{(\tau_{j,k})^-}\big)\Big) + e^{-\rho_i\tau_{r,\ell}}V_{i,m}(X_{(\tau_{r,\ell})^-})\bigg]. \end{aligned} \tag{3.17}$$

Notice that (3.17) is well defined: since $\tau_{r,\ell} \leq \tau_r$, X belongs to the compact set $\overline{B(0,r)}$, where the continuous function $V_{i,m}$ is bounded; moreover, the two summations consist in a finite number of terms since $\tau_{r,\ell} \leq \tau_{i,\ell}$ for each $i \in \{1,\dots,n\}$. Also, notice that in (3.17) we need to write $V_{i,m}(X_{(\tau_{r,\ell})^-})$, as we have a jump at time $\tau_{r,\ell}$. We now pass to the limit in (3.17) as $m \to \infty$: since X belongs to the compact set $\overline{B(0,r)}$, by the uniform convergence in (b) and (c) we get

$$\begin{aligned} V_i(x) = \mathbb{E}_x\bigg[&-\int_0^{\tau_{r,\ell}} e^{-\rho_i s}(\mathcal{L}V_i - \rho_i V_i)(X_s)ds - \sum_{\tau_{i,k}<\tau_{r,\ell}} e^{-\rho_i\tau_{i,k}}\Big(V_i\big(X_{\tau_{i,k}}\big) - V_i\big(X_{(\tau_{i,k})^-}\big)\Big) \\ &- \sum_{j\neq i}\sum_{\tau_{j,k}<\tau_{r,\ell}} e^{-\rho_i\tau_{j,k}}\Big(V_i\big(X_{\tau_{j,k}}\big) - V_i\big(X_{(\tau_{j,k})^-}\big)\Big) + e^{-\rho_i\tau_{r,\ell}}V_i(X_{(\tau_{r,\ell})^-})\bigg]. \end{aligned} \tag{3.18}$$

We now estimate each term in the right-hand side of (3.18). As for the first term, since $(\mathcal{M}_jV_j - V_j)(X_s) > 0$ for each $j \neq i$ by (3.15a), from (3.13d) it follows that

$$(\mathcal{L}V_i - \rho_i V_i)(X_s) \geq -f_i(X_s), \tag{3.19}$$

for all $s \in [0, \tau_S]$. Let us now consider the second term: by (3.13b) and the definition of $\mathcal{M}_i V_i$ in (3.11), for every stopping time $\tau_{i,k} < \tau_S$ we have

$$\begin{aligned} V_i\big(X_{(\tau_{i,k})^-}\big) &\leq \mathcal{M}_i V_i\big(X_{(\tau_{i,k})^-}\big) \\ &= \inf_{\delta \in Z_i} \big\{ V_i\big(\Gamma^i\big(X_{(\tau_{i,k})^-}, \delta\big)\big) + \phi_i\big(X_{(\tau_{i,k})^-}, \delta\big) \big\} \\ &\leq V_i\big(\Gamma^i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big)\big) + \phi_i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big) \\ &= V_i\big(X_{\tau_{i,k}}\big) + \phi_i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big). \end{aligned} \tag{3.20}$$

As for the third term, let us consider any stopping time $\tau_{j,k} < \tau_S$, with $j \neq i$. By (3.15a) we have $(\mathcal{M}_j V_j - V_j)\big(X_{(\tau_{j,k})^-}\big) = 0$; hence, the condition in (3.13c), the definition of $\mathcal{H}_{i,j} V_i$ in (3.11) and the expression of $\xi_{j,k}$ in (3.15c) imply that

$$\begin{aligned} V_i\big(X_{(\tau_{j,k})^-}\big) &= \mathcal{H}_{i,j} V_i\big(X_{(\tau_{j,k})^-}\big) \\ &= V_i\big(\Gamma^j\big(X_{(\tau_{j,k})^-}, \xi_j\big(X_{(\tau_{j,k})^-}\big)\big)\big) + \psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_j\big(X_{(\tau_{j,k})^-}\big)\big) \\ &= V_i\big(\Gamma^j\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big)\big) + \psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big) \\ &= V_i\big(X_{\tau_{j,k}}\big) + \psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big). \end{aligned} \tag{3.21}$$

By (3.18) and the estimates in (3.19)-(3.21) it follows that

$$\begin{aligned} V_i(x) \leq \mathbb{E}_x\bigg[\int_0^{\tau_{r,\ell}} e^{-\rho_i s} f_i(X_s) ds + \sum_{\tau_{i,k} < \tau_{r,\ell}} e^{-\rho_i \tau_{i,k}} \phi_i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big) \\ + \sum_{j \neq i} \sum_{\tau_{j,k} < \tau_{r,\ell}} e^{-\rho_i \tau_{j,k}} \psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big) + e^{-\rho_i \tau_{r,\ell}} V_i(X_{\tau_{r,\ell}}) \bigg]. \end{aligned}$$

Thanks to the conditions in (3.7) and (3.8) together with the polynomial growth of V_i in (ii), we now use the dominated convergence theorem and pass to the limit, first as $r \to \infty$ and then as $\ell \to \infty$, so that the stopping times $\tau_{r,\ell}$ converge to τ_S by (3.9). In particular, for the fourth term we notice that by (ii) and (3.8) we have

$$V_i(X_{(\tau_{r,\ell})^-}) \leq C(1 + |X_{(\tau_{r,\ell})^-}|^p) \leq C(1 + \|X\|_\infty^p) \in L^1(\Omega), \tag{3.22}$$

for suitable constants $C > 0$ and $p \in \mathbb{N}$. Therefore, the corresponding limit for the fourth term immediately follows by the continuity of V_i in the case $\tau_S < \infty$ and by (3.22) itself in the case $\tau_S = \infty$ (as a direct consequence of (3.8), we have $\|X\|_\infty^p < \infty$ a.s.). Hence,

$$\begin{aligned} V_i(x) \leq \mathbb{E}_x\bigg[\int_0^{\tau_S} e^{-\rho_i s} f_i(X_s) ds + \sum_{\tau_{i,k} < \tau_S} e^{-\rho_i \tau_{i,k}} \phi_i\big(X_{(\tau_{i,k})^-}, \xi_{i,k}\big) \\ + \sum_{j \neq i} \sum_{\tau_{j,k} < \tau_S} e^{-\rho_i \tau_{j,k}} \psi_{i,j}\big(X_{(\tau_{j,k})^-}, \xi_{j,k}\big) + e^{-\rho_i \tau_S} h_i(X_{(\tau_S)^-}) \mathbb{1}_{\{\tau_S < +\infty\}} \bigg] = J^i(x; (\varphi^{*,-i}, \varphi_i)). \end{aligned}$$

Step 2: $V_i(x) = J^i(x; \varphi^*)$. Similar as in Step 1, except that all the inequalities are equalities by the properties of φ^*. □

3.1.3 Example: two-player cash management game.

Now let us revisit the cash management game in Section 1.1.4, using the notations introduced in Section 3.1. We here consider the two-player game: $N = 2$, $b_i = 0$, $\sigma_i = \sigma > 0$. The uncontrolled cash level of the two players $X_t = (X_t^1, X_t^2)$ is

$$dX_t^i = \sigma dW_t^i, \qquad X_{0-}^i = x_i,$$

where W is a two-dimensional standard Brownian motion and $x \in \mathbb{R}^2$. Let

$$\varphi = (\varphi_1, \varphi_2), \qquad \varphi_1 = (A_1, \xi_1), \qquad \varphi_2 = (A_2, \xi_2),$$

denote the strategies of the players for this game, as in Definition 3.1. Since player $i \in \{1, 2\}$ intervenes by shifting her own component X^i of the cash level, we have

$$\Gamma^i(x, \delta) = x + \delta, \qquad Z_1 = \{(\delta_1, 0) : \delta_1 \in \mathbb{R}\}, \qquad Z_2 = \{(0, \delta_2) : \delta_2 \in \mathbb{R}\}.$$

This means that player 1 (resp. player 2) intervenes by moving the process from state (x_1, x_2) to state

$$\xi_1(x_1, x_2) = \big(x_1 + \tilde{\xi}_1(x_1, x_2),\ x_2\big) \qquad \Big(\text{resp. } \xi_2(x_1, x_2) = \big(x_1,\ x_2 + \tilde{\xi}_2(x_1, x_2)\big)\Big),$$

for suitable functions $\tilde{\xi}_i$. Notice that, as a consequence, the controlled process $X_t = (X_t^1, X_t^2)$ satisfies

$$dX_t^i = \sigma dW_t^i + \sum_{\tau_{i,k} \le t} \delta(t - \tau_{i,k}) \tilde{\xi}_{i,k}, \qquad X_{0-}^i = x_i,$$

where

$$\tilde{\xi}_{i,k} = \tilde{\xi}_i\Big(X^1_{(\tau_{i,k})^-},\, X^2_{(\tau_{i,k})^-}\Big).$$

Let now $i, j \in \{1, 2\}$ with $j \neq i$. The cost function for player i under the control policy $\varphi = (\varphi_1, \varphi_2)$ is given by

$$J^i(x; \varphi) = \mathbb{E}_x \left[\int_0^\infty e^{-rt} f_i(X_t) dt + \sum_{k \ge 1} e^{-r\tau_{i,k}} \phi_i(\xi_{i,k}) + \sum_{k \ge 1} e^{-r\tau_{j,k}} \psi_{i,j}(\xi_{j,k}) \right],$$

where

$$\begin{cases} f_i(x) = h \left| x_i - \dfrac{1}{N} \displaystyle\sum_{j=1}^N x_j \right|, \quad x \in \mathbb{R}^2, N = 2, \\ \phi_i(\xi) = K + k|\xi|, \quad \xi \in \mathbb{R}, \\ \psi_{i,j}(\xi) = c, \quad \xi \in \mathbb{R}, \end{cases}$$

for positive constants h, K, k, c. The goal of player i is to minimize the cost J^i: we are interested in finding $\varphi^* = (\varphi_1^*, \varphi_2^*)$ such that Definition 3.6 holds.

By the symmetry of the problem structure, we seek for an NE where the action regions take the form of

$$A_1 = \{x : x_1 - x_2 \geq u\}, \qquad A_2 = \{x : x_2 - x_1 \geq u\}$$

for some $u > 0$, with appropriate impulse functions such that

$$\xi_1(x) = (U - x_1 + x_2,\, 0), \qquad \xi_2(x) = (0,\, U - x_2 + x_1)$$

for some $U < u$. Recall that this means that player 1 (resp. player 2) intervenes when $X_t^1 - X_t^2 \geq u$ (resp. $X_t^2 - X_t^1 \geq u$) and shifts her component so as to have $X_t^1 - X_t^2 = U$ (resp. $X_t^2 - X_t^1 = U$). Note that $A_1 \cap A_2 = \emptyset$ and that $C = \{x : -u < x_1 - x_2 < u\}$ is the common waiting region, i.e., where no player intervenes.

By the same symmetry argument, we look for payoffs in the form of

$$V_i(x_1, x_2) = w_i(x_i - x_j), \quad i, j \in \{1, 2\}, \quad i \neq j,$$

for some functions w_i. In this case, player 1 and player 2 are indistinguishable, therefore it suffices to study the payoff of player 1. Now the waiting region for player 1 is $D_1 = \{x : x_1 - x_2 < u\}$, and define $D_{-1} = \{x : x_2 - x_1 < u\} = \{x : x_1 - x_2 > -u\}$. By the corresponding QVI and the regularity requirement in the Verification Theorem 3.7, the function $w_1 : \mathbb{R} \to \mathbb{R}$ need to satisfy the following system of equations and inequalities:

$$\begin{cases} w_1(s) = \begin{cases} w_1(u) + k(s-u), & s \geq u; \\ \frac{h_2}{r} s + c_1 e^{\lambda_2 s} + c_2 e^{-\lambda_2 s}, & 0 \leq s \leq u; \\ -\frac{h_2}{r} s + \left(c_1 + \frac{h_2}{r\lambda_2}\right) e^{\lambda_2 s} + \left(c_2 - \frac{h_2}{r\lambda_2}\right) e^{-\lambda_2 s}, & -u \leq s \leq 0; \\ w_1(-u), & s \leq -u; \end{cases} & \text{(3.23a)} \\ \dot{w}_1(u) = \dot{w}_1(U) = k; & \text{(3.23b)} \\ w_1(u) = w_1(U) + K + k(u - U); & \text{(3.23c)} \\ w_1(-u) = w_1(-U) + c; & \text{(3.23d)} \\ \mathcal{M} w_1(s) - w_1(s) \geq 0, \quad \forall s \in D_{-1}; & \text{(3.23e)} \end{cases}$$

where $h_2 = \frac{h}{2}$, $\sigma_2 = \sqrt{2}\sigma$, $\lambda_2 = \frac{\sqrt{2r}}{\sigma_2}$ and (c_1, c_2, u, U) remain to be determined. Accordingly, $w_2(s) = w_1(-s)$ for any $s \in \mathbb{R}$.

Now, similar argument as in [58] shows that when $h_2 - rk > 0$, $c > 0$, there exists a solution w_1 to Equations (3.23a) to (3.23d) satisfying $c_1 < 0$, $c_2 > 0$, $0 < U < u$. Moreover, if such solution w_1 as above satisfies (3.23e), then an NE to the cash management problem $\varphi^* = (\varphi_1^*, \varphi_2^*)$ is characterized by

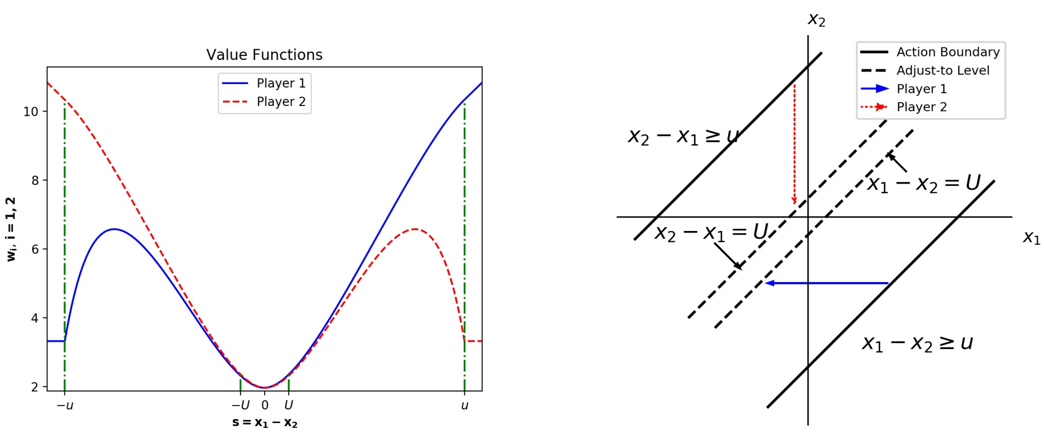
$$\begin{cases} A_1^* = \{x : x_1 - x_2 \geq u\}, \quad A_2^* = \{x : x_2 - x_1 \geq u\}, \\ \xi_1^*(x) = (U - x_1 + x_2,\, 0), \quad \xi_2^*(x) = (0,\, U - x_2 + x_1). \end{cases} \tag{NE-1}$$

The corresponding payoffs are given by

$$V_1(x_1, x_2) = w_1(x_1 - x_2), \quad V_2(x_1, x_2) = w_1(x_2 - x_1).$$

The optimality of (NE-1) can be easily verified by checking the conditions in the Verification Theorem 3.7.

Figure 3.1a shows the equilibrium payoffs for both players if they adopt the control policy specified in (NE-1). Figure 3.1b illustrates the control policy, with $h = 2$, $K = 3$, $k = 1$, $r = 0.5$, $\sigma = \frac{\sqrt{2}}{2}$ and $c = 1$, where the thresholds can be solved as $U = 0.686$ and $u = 5.658$, with $c_1 = -0.003$ and $c_2 = 1.972$.



(a) Payoffs after change of variables.

(b) Control policy.

Figure 3.1: A Nash equilibrium and the payoffs

Multiple NEs. In general, NE for nonzero-sum games may not be unique. In this example, an alternative NE can be derived by switching action regions between the two players. For instance, if player 1 is to dictate the game whereas player 2 is a complete follower, the action region for player 1 can be characterized by $A_1 = \{x \in \mathbb{R}^2 : |x_1 - x_2| > u\}$, and $A_2 = \emptyset$. That is, let $s = x_1 - x_2$ then $V_1(x) = w_1(x_1 - x_2)$ and $V_2(x) = w_2(x_1 - x_2)$, where

$$\begin{cases} w_1(s) = \begin{cases} w_1(u) + k(s-u), & s \geq u; \\ \frac{h_2}{r} s + c_1 e^{\lambda_2 s} + \left(c_1 + \frac{h_2}{r\lambda_2}\right) e^{-\lambda_2 s}, & 0 \leq s \leq u; \\ w_1(-s), & s \leq 0; \end{cases} & (3.24a) \\ \dot{w}_1(u) = \dot{w}_1(U) = k; & (3.24b) \\ w_1(u) = w_1(U) + K + k(u-U); & (3.24c) \\ \mathcal{M} w_1(s) - w_1(s) \geq 0, \quad s \in \mathbb{R}; & (3.24d) \end{cases}$$

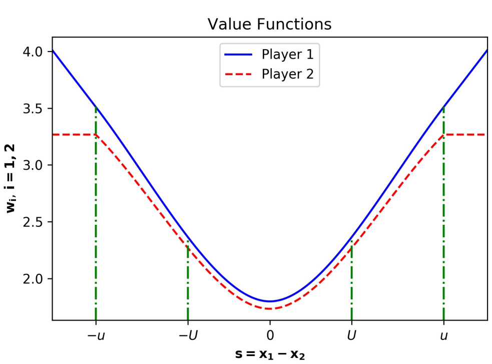
$$\begin{cases} w_2(s) = \begin{cases} w_2(u), & s \geq u; \\ \frac{h_2}{r}s + c_2 e^{\lambda_2 s} + \left(c_2 + \frac{h_2}{r\lambda_2}\right) e^{-\lambda_2 s}, & 0 \leq s \leq u; \\ w_2(-s), & s \leq 0; \end{cases} & \text{(3.25a)} \\ w_2(u) = w_2(U) + c; & \text{(3.25b)} \\ \mathcal{M}w_2(s) - w_2(s) \geq 0, \quad -u \leq s \leq u. & \text{(3.25c)} \end{cases}$$

Now, assume that $h_2 - rk > 0$, $c > 0$, then again one can show that there exists a solution w_1 satisfying Equations (3.24a) to (3.24d) with $c_1 \in (-\frac{h_2}{r\lambda_2}, 0)$ as well as $0 < U < u$, and w_2 satisfying Equations (3.25a) to (3.25b). Moreover, if such solution w_2 satisfies (3.25c), then an NE to the cash management problem in Section 3.1.3 $\varphi^* = (\varphi_1^*, \varphi_2^*)$ is characterized by

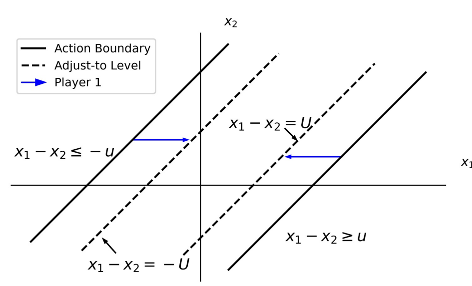
$$\begin{cases} A_1^* = \{x : |x_1 - x_2| \geq u\}, \quad A_2^* = \emptyset; \\ \xi_1^*(x) = \begin{cases} (U - x_1 + x_2,\, 0), & \text{if} \quad x_1 - x_2 \geq u, \\ (-U - x_1 + x_2,\, 0), & \text{if} \quad x_1 - x_2 \leq -u. \end{cases} \end{cases} \tag{NE-2}$$

Notice that we do not need to define ξ_2^*, as player 2 never intervenes.

Figure 3.2a shows the payoffs and Figure 3.2b demonstrates the NE, under the same values of h, K, k, r and σ, with thresholds $U = 0.993$ and $u = 1.999$, and $c_1 = -0.101$ and $c_2 = -0.133$.



(a) Player 1 is the sole controller.



(b) Control policy.

Figure 3.2: Alternative Nash equilibrium and the payoffs

3.2 MFGs with impulse controls.

As seen from the previous section, it is difficult to solve analytically the general N-player impulse control game. We will now introduce an MFG framework for the impulse control game and show that this MFG provides a reasonable approximation to the N-player game. More precisely, we show that under appropriate technical conditions, the existence of unique NE solution to the MFG, which is an ϵ-NE approximation to the N-player game, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.

3.2.1 Formulation of MFGs with impulse controls.

Given the N-player stochastic game formulation (N-player-I), its natural MFG formulation goes as follows. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space supporting an $\{\mathcal{F}_t\}_{t\geq 0}$-adapted standard Brownian motion W. Consider an infinite number of rational and indistinguishable players who interact through the cost structure consisting of a running cost f and the cost of control ϕ. For each player, her uncontrolled state process is given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_{0-} \sim \mu.$$

Each player seeks for the optimal impulse control policy φ^* among the set of admissible impulse controls $\mathcal{A}$ to minimize the total discounted cost. Controls $\varphi \in \mathcal{A}$ are here represented by $\varphi = (A, \xi)$, where A, a closed subset of $\mathbb{R}$, is called the action region and $\xi : \mathbb{R} \to \mathbb{R}$ is a measurable function. Under the control policy φ, the state process becomes

$$dX_t = b(X_{t-})dt + \sigma(X_{t-})dW_t + \sum_{n\geq 1} \delta(t-\tau_n)\xi_n, \quad X_{0-} \sim \mu,$$

where μ denotes the initial distribution of the state, and

$$\begin{aligned}
&\tau_1 = \inf\{t \geq 0 : X_{t-} \in A\}, \quad \tau_n = \inf\{t > \tau_{n-1} : X_{t-} \in A\},\ n \geq 2;\\
&\xi_n = \xi(X_{\tau_n -}) \in \mathcal{F}_{\tau_n},\ n \geq 1.
\end{aligned}$$

Therefore, the control policy φ can also be characterized by a sequence of stopping times and the associated random variables, $\varphi = \{(\tau_n, \xi_n)\}_{n\geq 1}$. The optimization problem faced by individual player is given by,

$$\begin{aligned}
V(x|m) &= \inf_{\varphi\in\mathcal{A}} J_\infty(x, \varphi|m),\\
J_\infty(x, \varphi|m) &= \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f(X_t, m)dt + \sum_{n=1}^{\infty} e^{-r\tau_n}\phi(\xi_n) \middle| X_{0-} = x \right],\\
dX_t &= b(X_{t-})dt + \sigma(X_{t-})dW_t + \sum_{\tau_n\leq t} \delta(t-\tau_n)\xi_n, \qquad X_{0-} \sim \mu,
\end{aligned} \tag{MFG-I}$$

where m denotes the mean information:

$$m = \limsup_{t\to\infty} \mathbb{E}_\mu[X_t].$$

Finally, for $x, m \in \mathbb{R}$, we define the function $\mathcal{M}V(x|m)$ in the usual way:

$$\mathcal{M}V(x|m) = \inf_{\delta \in \mathbb{R}} \big\{ V(x+\delta|m) + \phi(\delta) \big\}.$$

Compared to (N-player-I) for N-player games, individual players in an MFG now lose sight of individual opponents, hence the term $\psi_{i,j}$ has disappeared in the MFG formulation. As for the intervention costs, we chose them in the form $\phi(\xi_n)$ in order to have simpler notations in this section; however, one may also consider intervention costs in the form $\phi(\xi_n, X_{(\tau_n)^-}, m)$.

Definition 3.8. *A pair of control policy and mean information* ($\varphi^* = (A^*, \xi^*), m^*$)*, with* $\varphi^* \in \mathcal{A}$ *and* $m^* \in \mathbb{R}$*, is said to be a solution to the* (MFG-I) *if*

- $V(x|m^*) = J_\infty(x, \varphi^*|m^*)$,
- $m^* = \limsup_{t\to\infty} \mathbb{E}_\mu [X_t^*]$*, where*

$$dX_t^* = b(X_{t-}^*)dt + \sigma(X_{t-}^*)dW_t + \sum_{\tau_n^* \le t} \delta(t-\tau_n^*)\xi_n^*, \quad X_{0-}^* \sim \mu,$$

such that

$$\tau_1^* = \inf\{t \ge 0 : X_{t-}^* \in A^*\}, \quad \tau_n^* = \inf\{t > \tau_{n-1}^* : X_{t-}^* \in A^*\},\ n \ge 2;$$
$$\xi_n^* = \xi^*(X_{(\tau_n^*)-}^*) \in \mathcal{F}_{\tau_n^*},\ n \ge 1.$$

3.2.2 Solution to symmetric MFGs with impulse controls.

We first analyze the existence of the solution to (MFG-I), under some technical assumptions. In particular, we impose symmetry on the dynamics and cost functions.

(A1) The uncontrolled dynamics is symmetric in the sense that

$$dX_t = \sigma dW_t, \quad X_{0-} \sim \mu,$$

where $\sigma > 0$ is a constant, and μ, with $\int_{\mathbb{R}} |x|\mu(dx) < \infty$, is symmetric around its mean.

(A2) The cost of control satisfies

$$\begin{cases} K := \inf_{\xi\in\mathbb{R}} \phi(\xi) > 0, & (3.26)\\ \phi \in \mathcal{C}(\mathbb{R}\setminus\{0\}), & (3.27)\\ \lim_{|\xi|\to\infty} \phi(\xi) = +\infty, \quad \sup_{\xi\in\mathbb{R}} \dfrac{\phi(\xi)}{1+|\xi|} \le k \text{ for some } k > 0, & (3.28)\\ \phi(\xi_1) + \phi(\xi_2) \ge \phi(\xi_1+\xi_2) + K, \quad \forall \xi_1, \xi_2 \in \mathbb{R}, & (3.29)\\ \phi(-\xi) = \phi(\xi), \quad \forall \xi \in \mathbb{R}. & (3.30) \end{cases}$$

(A3) The running cost $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, jointly continuous with $f(x, m) \geq 0$, satisfies the following properties, for any fixed $m \in \mathbb{R}$.

(A3-1) There exists $C_f = C_f(m) > 0$ such that, for each $x, y \in \mathbb{R}$,

$$|f(x, m) - f(y, m)| < C_f |x - y|;$$

(A3-2) For each $\delta \in \mathbb{R}$,

$$f(m + \delta, m) = f(m - \delta, m).$$

(A4) For each $m \in \mathbb{R}$ fixed, the strategy $\tilde{\varphi}(m) = (A_m, \xi_m)$ defined by

$$A_m = \{x \in \mathbb{R} : \mathcal{M}V(x|m) - V(x|m) = 0\}, \qquad \xi_m(x) = \arg\min_{\delta \in \mathbb{R}} \{V(x + \delta|m) + \phi(\delta)\}, \tag{3.31}$$

is admissible and is the unique strategy φ such that $V(x|m) = J_\infty(x, \varphi|m)$.

Remark 3.6. *We remark that, when $m \in \mathbb{R}$ is fixed, problem* (MFG-I) *becomes a standard single-player impulse control problem, whose typical solution has the form* (3.31) *in Assumption (A4): see [145] for an introduction to single-player impulse problems and e.g. [13], [32], [33], [109], [134] and [137] for some applications having solutions in the form* (3.31)*.*

Theorem 3.9. *Under Assumptions (A1)–(A4),* (MFG-I) *admits a solution in the sense of Definition 3.8.*

Proof. Theorem 3.9 is proved using a three-step approach (solution for generic m, mean information update, fixed point of the composite function).

First, if the mean information m is given, we solve the corresponding optimal control problem. By (A4), the unique optimal strategy is given by $\tilde{\varphi}(m) = (A_m, \xi_m) \in \mathcal{A}$. Define a mapping from space of mean information $\mathbb{R}$ to the admissible control set $\mathcal{A}$, as

$$\Gamma_1 : \mathbb{R} \to \mathcal{A}, \qquad \Gamma_1(m) = \tilde{\varphi}(m) = (A_m, \xi_m).$$

Next, we update the mean information m. Given any $\varphi = (A, \xi) \in \mathcal{A}$, under Assumption (A1), define the corresponding controlled process

$$dX_t^\varphi = \sigma dW_t + \sum_{n \geq 1} \delta(t - \tau_n) \xi_n,$$

where

$$\tau_1 = \inf\{t \geq 0 : X_{t-}^\varphi \in A\}, \quad \tau_n = \inf\{t > \tau_{n-1} : X_{t-}^\varphi \in A\},\ n \geq 2;$$
$$\xi_n = \xi(X_{\tau_n -}^\varphi),\ n \geq 1.$$

Define a mapping from the admissible control set $\mathcal{A}$ to the extended real line $[-\infty, \infty]$, as

$$\Gamma_2 : \mathcal{A} \to [-\infty, \infty], \qquad \Gamma_2(\varphi) = \limsup_{t \to \infty} \mathbb{E}_\mu \left[X_t^\varphi\right].$$

Finally, define the composite mapping $\Gamma : \mathbb{R} \to [-\infty, \infty]$, as

$$\Gamma : m \overset{\Gamma_1}{\mapsto} \tilde{\varphi}(m) \overset{\Gamma_2}{\mapsto} \limsup_{t\to\infty} \mathbb{E}_\mu \left[X_t^{\tilde{\varphi}(m)} \right].$$

This is where we utilize the symmetric cost structures. By Lemma 3.10 below, the waiting region $D_m = \mathbb{R} \setminus A_m$ and the optimal control function ξ_m are symmetric with respect to m. Given the symmetry and L^1 condition of μ in Assumption (A1), let $m^* = \mathbb{E}_\mu[X_{0-}] \in \mathbb{R}$, then by symmetry, $\Gamma(m^*) = \limsup_{t\to\infty} \mathbb{E}_\mu \left[X_t^{\tilde{\varphi}(m^*)} \right] = m^*$. Therefore, $(\varphi^* = \tilde{\varphi}(m^*), m^*)$ is a solution to (MFG-I) in the sense of Definition 3.8. □

Lemma 3.10. *Let (A1)-(A4) hold and let $m \in \mathbb{R}$ be fixed. Then:*

(i) the function $V(\cdot|m)$ in (MFG-I) is symmetric with respect to m, i.e., $V(m+x|m) = V(m-x|m)$ for each $x \in \mathbb{R}$;

(ii) the continuation region $D_m = \{x \in \mathbb{R} : \mathcal{M}V(x|m) - V(x|m) > 0\}$ is symmetric with respect to m, i.e., $m - x \in D_m$ if and only if $m + x \in D_m$, for each $x \in \mathbb{R}$;

(iii) the optimal impulse function $\xi_m = \arg\min_{\delta\in\mathbb{R}}\{V(\,\cdot\, + \delta) + \phi(\delta)\}$ satisfies $\xi_m(m+x) = -\xi_m(m-x)$, for each $x \in \mathbb{R}$. As a consequence, if x, x' are symmetric with respect to m, then $x + \xi_m(x), x' + \xi_m(x')$ are symmetric as well (i.e., optimal interventions preserve symmetry).

Proof. (i) By [92], the function $V(\cdot|m)$ is the only viscosity solution to the QVI

$$\min\{\mathcal{L}V(x|m) - rV(x|m) + f(x,m),\ \mathcal{M}V(x|m) - V(x|m)\} = 0, \tag{3.32}$$

with $\mathcal{L} = \frac{\sigma^2}{2}\frac{d^2}{dx^2}$. We will prove that $x \mapsto V(2m - x|m)$ is a viscosity solution to (3.32), so that by uniqueness we get $V(x|m) = V(2m-x|m)$ for each $x \in \mathbb{R}$, i.e., the claim in (i). To prove that $\tilde{V}(x) := V(2m-x|m)$ is a viscosity subsolution to (3.32) (the supersolution argument is similar), let us consider $x_0 \in \mathbb{R}$ and $\tilde{\varphi} \in C^2(\mathbb{R})$ such that $\tilde{V} - \tilde{\varphi}$ has a local maximum at x_0 and $(\tilde{V} - \tilde{\varphi})(x_0) = 0$. If we set $x_1 = 2m - x_0$ and $\varphi(x) := \tilde{\varphi}(2m - x)$, we see that $V(\cdot|m) - \varphi$ has a local maximum at x_1 and $(V(\cdot|m) - \varphi)(x_1) = 0$. Since $V(\cdot|m)$ is viscosity subsolution to (3.32), we have that

$$\begin{aligned} 0 &\geq \min\left\{\mathcal{L}\varphi(x_1) - rV(x_1|m) + f(x_1,m),\ \mathcal{M}V(x_1|m) - V(x_1|m)\right\} \\ &= \min\left\{\mathcal{L}\tilde{\varphi}(x_0) - r\tilde{V}(x_0) + f(x_0,m),\ \mathcal{M}\tilde{V}(x_0) - \tilde{V}(x_0)\right\}, \end{aligned}$$

where in the last step we have used the definitions of $\tilde{V}, \tilde{\varphi}, x_1$ and the symmetry properties in (3.30) and (A3-2).

(ii) For each $x \in \mathbb{R}$, we have

$$\begin{aligned} \mathcal{M}V(m-x|m) &= \inf_{\delta\in\mathbb{R}}\left\{V(m-x+\delta|m) + \phi(\delta)\right\} = \inf_{\tilde{\delta}\in\mathbb{R}}\left\{V(m-(x+\tilde{\delta})|m) + \phi(-\tilde{\delta})\right\} \\ &= \inf_{\tilde{\delta}\in\mathbb{R}}\left\{V(m+(x+\tilde{\delta})|m) + \phi(\tilde{\delta})\right\} = \mathcal{M}V(m+x|m) \end{aligned} \tag{3.33}$$

where in the second equality we have used the change of variable $\tilde{\delta} = -\delta$, while in the second-to-last equality we have used (i) and (3.30). The claim in (ii) immediately follows from (3.33) and (i).

(iii) By arguing as in (ii), for every $x \in \mathbb{R}$ we have

$$\begin{aligned}\xi_m(m+x) &= \arg\min_{\delta\in\mathbb{R}}\{V(m+x+\delta|m)+\phi(\delta)\}\\ &= -\arg\min_{\tilde{\delta}\in\mathbb{R}}\{V(m-x+\tilde{\delta}|m)+\phi(\tilde{\delta})\} = -\xi_m(m-x).\end{aligned}$$

□

3.2.3 MFGs vs N-player games.

Next, we will demonstrate that the solution to (MFG-I) is an approximation to the corresponding N-player game with identical players under the symmetric setting. Here the state process of the N-player game on $\mathbb{R}^N$ is denoted by $\{\mathbf{X}_t\}_{t\geq 0} = \{(X_t^1,\ldots,X_t^N)\}_{t\geq 0}$, with player i only controlling her own state process X^i, $i=1,\ldots,N$.

Recall the notations in Section 2: for player i, with $i=1,\ldots,N$, $f_i:\mathbb{R}^N\to\mathbb{R}$ denotes the running cost, ϕ_i denotes the individual cost of control and $\psi_{i,j}$ denotes the cost of control done by another player j. To be consistent with the MFGs setting, we assume

(A1′) Player i's uncontrolled state process is given by

$$dX_t^i = \sigma dW_t^i, \quad X_{0-}^i \sim \mu,$$

where $(W^1,\ldots,W^N)$ denotes the N-dimensional standard Brownian motion, $\sigma>0$ is a constant and μ is symmetric around its mean.

(A2′) The cost of individual control ϕ_i equals ϕ that satisfies Assumption (A2) and the costs of other players' control, $\psi_{i,j}$ are identical for all i and j.

(A3′) The cost function f_i takes the form of

$$\begin{aligned}f_i(\mathbf{x}) &= g\left(x_i - \frac{1}{N}\sum_{j=1}^N x_j\right), \forall \mathbf{x}\in\mathbb{R}^N,\\ f(x,m) &= g(x-m), \forall x\in\mathbb{R},\end{aligned}$$

for some function $g:\mathbb{R}\to\mathbb{R}^+$, $g(x)=g(-x)$, with

$$|g(x)-g(y)| \leq L|x-y|, \forall x,y\in\mathbb{R},$$

where $L>0$ is the Lipschitz constant.

Note that the running cost of (MFG-I) under Assumption (A3′) also satisfies Assumption (A3).

Definition 3.11 (**ϵ-Nash equilibrium**). *A strategy $\varphi^* = (\varphi_1^*, \ldots, \varphi_N^*)$ is a called an ϵ-Nash equilibrium to the N-player game introduced in Section 3.1 if*

$$\mathbb{E}_\mu \left[J^i(\mathbf{X}_{0-}; \varphi^*)\right] \leq \mathbb{E}_\mu \left[J^i(\mathbf{X}_{0-}; (\varphi^{*,-i}, \varphi_i))\right] + \epsilon, \quad \forall \varphi_i \in \Phi_i(\mathbf{X}_{0-}) \ s.t. \ (\varphi^{*,-i}, \varphi_i) \in \Phi(\mathbf{X}_{0-}),$$

where $X_{0-}^i \overset{i.i.d.}{\sim} \mu$, $i = 1, \ldots, N$.

Let $(\tilde{\varphi}^*, m^*)$ be a solution to the (MFG-I) under Assumptions (A1)-(A4), where m^* is the expectation of initial distribution μ, the control policy $\tilde{\varphi}^*$ is characterized by action region $\tilde{A}^*$ and impulse function $\tilde{\xi}^* : \mathbb{R} \to \mathbb{R}$; denote the corresponding waiting region as $\tilde{D}^*$ and the state process on $\mathbb{R}$ under $\tilde{\varphi}^*$ as $\tilde{X}$. As illustrated in Theorem 3.9, the waiting region $\tilde{D}^*$ as well as the impulse function $\tilde{\xi}^*$ will be symmetric around m^*. Define the following priority sets

$$P_i = \{\mathbf{x} \in \mathbb{R}^N : |x_i - m^*| > |x_j - m^*|, \, \forall j > i; \; |x_i - m^*| \geq |x_k - m^*|, \, \forall k > i\}, \; \forall i \in \{1, \ldots, N\}.$$

Then for $i = 1, \ldots, N$, define the action region for Player i as $A_i^* = \{\mathbf{x} \in \mathbb{R}^N : x_i \in \tilde{A}^*\} \cap P_i$ and her impulse function $\xi_i^* : \mathbb{R}^N \to \mathbb{R}$ such that

$$\xi_i^*(\mathbf{x}) = \tilde{\xi}^*(x_i), \, \forall \mathbf{x} \in \mathbb{R}^N.$$

Denote $\varphi_i^* = (A_i^*, \xi_i^*)$ and $\varphi^* = (\varphi_1^*, \ldots, \varphi_N^*)$.

Theorem 3.12. *Let Assumptions (A1-A3) and (A1′-A3′) hold. Suppose that, under $\tilde{\varphi}^*$, we have $\tilde{D}^* \subset [m^* - u^*, m^* + u^*]$ for some positive constant u^* and that $\tilde{X}_t^* \in \tilde{D}^*$ almost surely for all $t \geq 0$. Then φ^* is an ϵ-NE for the N-player cash management game introduced in Section 3.1.3 for generic N, with $\epsilon = O\left(\frac{1}{\sqrt{N}}\right)$.*

Proof. Fix $i \in \{1, \ldots, N\}$. Consider $\bar{\varphi} = (\varphi^{*,-i}, \varphi_i)$ such that $\varphi_i \in \Phi_i(\mathbf{x})$ and $\bar{\varphi} \in \Phi(\mathbf{x})$. For $j \neq i$, $\bar{\varphi}_j = \varphi_j^*$ whose action region independent from the strategy of player i.

We first look at the running cost.

$$X_t^i - \frac{1}{N}\sum_{j=1}^N X_t^j = \left(X_t^i - m^*\right) - \frac{1}{N}\left(X_t^i - m^*\right) - \frac{\sum_{j\neq i}(X_t^j - m^*)}{N},$$

so that

$$\begin{aligned}\left|f_i(\mathbf{X}_t) - f(X_t^i, m^*)\right| &= \left|g\left(X_t^i - \frac{1}{N}\sum_{j=1}^N X_t^j\right) - g\left(X_t^i - m^*\right)\right| \\ &\leq L\left(\frac{1}{N}|X_t^i - m^*| + \frac{\sum_{j\neq i}|X_t^j - m^*|}{N-1}\right).\end{aligned}$$

Note that

$$\left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right| \leq u^*$$

and by the i.i.d. assumption,

$$\mathbb{E}_\mu \left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right| \leq \left(\mathbb{E}_\mu \left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right|^2 \right)^{\frac{1}{2}} \Rightarrow \mathbb{E}_\mu \left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right| = O\left(\frac{1}{\sqrt{N}} \right).$$

Without loss of generality, let us consider φ_i such that

$$\mathbb{E}_\mu \left[\int_0^\infty e^{-rt} L |X_t^i - m^*| dt \right] < M,$$

for some sufficiently large $M > 0$. Then

$$\begin{aligned}
&\left| \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f_i(\mathbf{X}_t) dt \right] - \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f(X_t^i, m^*) dt \right] \right| \\
&\qquad \leq \frac{1}{N} \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} L |X_t^i - m^*| dt \right] + \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} L \left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right| dt \right] \\
&\qquad = \frac{1}{N} \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} L |X_t^i - m^*| dt \right] + \int_0^\infty e^{-rt} L \mathbb{E}_\mu \left[\left| \frac{\sum_{j \neq i} (X_t^j - m^*)}{N-1} \right| \right] dt \quad \text{(Fubini's)} \\
&\qquad = O\left(\frac{1}{N} \right) + O\left(\frac{1}{\sqrt{N}} \right) \\
&\Rightarrow \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f_i(\mathbf{X}_t) dt \right] = \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f(X_t^i, m^*) dt \right] + O\left(\frac{1}{\sqrt{N}} \right)
\end{aligned}$$

and therefore

$$\begin{aligned}
&\mathbb{E}_\mu \left[J^i(\mathbf{X}_{0-}; \bar{\varphi}) \right] \\
&= \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f_i(\mathbf{X}_t) dt + \sum_{n \geq 1} e^{-r\tau_{i,n}} \phi(\xi_{i,n}) + \sum_{j \neq i} \sum_{n \geq 1} e^{-r\tau_{j,n}} \psi_{i,j}(\xi_{j,n}^*) \right] \\
&= \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f(X_t^i, m^*) dt + \sum_{n \geq 1} e^{-r\tau_{i,n}} \phi(\xi_{i,n}) \right] + \mathbb{E}_\mu \left[\sum_{j \neq i} \sum_{n \geq 1} e^{-r\tau_{j,n}} \psi_{i,j}(\xi_{j,n}^*) \right] + O\left(\frac{1}{\sqrt{N}} \right) \\
&\geq V(\mu) + \mathbb{E}_\mu \left[\sum_{j \neq i} \sum_{n \geq 1} e^{-r\tau_{j,n}} \psi_{i,j}(\xi_{j,n}^*) \right] + O\left(\frac{1}{\sqrt{N}} \right) \\
&= \mathbb{E}_\mu \left[J^i(\mathbf{X}_{0-}; \varphi^*) + O\left(\frac{1}{\sqrt{N}} \right) \right],
\end{aligned}$$

where we have denoted by $V(\mu)$ the payoff of the impulse MFG with initial distribution μ. □

Remark 3.7. *Among the assumption of Theorem 3.12, we ask that $\tilde{D}^* \subset [m^* - u^*, m^* + u^*]$ is bounded. Heuristically, if $f(x, m) \to +\infty$ as $x \to \infty$ and diverges at a greater rate than the intervention costs, we expect that, for $|x|$ big enough, intervening is cheaper than keeping the state as it is (in other words, we expect that the continuation region is bounded). For examples of bounded continuation regions in impulse control theory, see e.g. the references in Remark 3.6, i.e. [13], [32], [33], [109], [134], [137]. For a two-player stochastic impulse game with bounded continuation regions, see [76]. In Section 3.2.4 below, we will provide an example of impulse MFG where this condition is satisfied.*

3.2.4 Explicit solutions: MFGs for cash management problems.

In this section, we explicitly solve the MFGs cash management problems, i.e., a (slightly more general) mean-field counterpart to the two-player game in Section 3.1.3.

(A1″) The uncontrolled dynamics follow

$$dX_t = \sigma dW_t, \quad X_{0-} \sim \mu, \tag{3.34}$$

where $\sigma > 0$ is a constant.

(A2″) The cost of control satisfies

$$\phi(\xi) = \begin{cases} K^+ + k^+\xi, & \xi \geq 0, \\ K^- - k^-\xi, & \xi < 0, \end{cases} \tag{3.35}$$

where $K^\pm, k^\pm > 0$.

(A3″) The running cost takes the form

$$f(x, m) = C\left(x - \alpha(m)\right),$$

where $C : \mathbb{R} \to \mathbb{R}$ is defined by

$$C(x) = \max\{hx, -px\}, \qquad x \in \mathbb{R}, \tag{3.36}$$

with parameters $h, p > 0$ satisfying

$$h - k^+ r > 0, \quad p - k^- r > 0,$$

and where $\alpha : \mathbb{R} \to \mathbb{R}$ is a contraction, i.e.,

$$|\alpha(x) - \alpha(y)| \leq k|x - y|, \qquad 0 < k < 1, \qquad x, y \in \mathbb{R}.$$

Here the function α can be interpreted as target level depending on the mean information m. The assumption that α is a contraction mapping is for analytical tractability.

Theorem 3.13. *Under Assumptions (A1–A4),* (MFG-I) *admits a unique analytical solution in the sense of Definition 3.8.*

Proof. We will in fact explicitly derive the solution. Let us fix $m \in \mathbb{R}$. Then the corresponding QVI for the control problem is

$$\min\left\{\mathcal{L}V - rV + C(x-\alpha(m)), \mathcal{M}V - V\right\} = 0. \tag{3.37}$$

Similar to [58], one can find an optimal policy characterized by the vector (d, D, U, u) with $d < D < 0 < U < u$, by smooth-fit principle. The payoff corresponding to 3.13 has to satisfy (we set $\lambda = \frac{\sqrt{2r\sigma^2}}{\sigma^2}$)

$$\begin{cases} V(x) = \begin{cases} V\left(u+\alpha(m)\right) - k^-\left(u - x + \alpha(m)\right), & x-\alpha(m) \geq u, \\ \frac{h}{r}\left(x-\alpha(m)\right) + c_1 \exp\left\{\lambda\left(x-\alpha(m)\right)\right\} \\ \qquad + c_2 \exp\left\{-\lambda\left(x-\alpha(m)\right)\right\}, & 0 \leq x-\alpha(m) \leq u, \\ -\frac{p}{r}\left(x-\alpha(m)\right) + \left(c_1 + \frac{h+p}{2r\lambda}\right)\exp\left\{\lambda\left(x-\alpha(m)\right)\right\} \\ \qquad + \left(c_2 - \frac{h+p}{2r\lambda}\right)\exp\left\{-\lambda\left(x-\alpha(m)\right)\right\}, & d \leq x-\alpha(m) \leq 0; \\ V\left(d+\alpha(m)\right) + k^+\left(d - x + \alpha(m)\right), & x-\alpha(m) \leq d; \end{cases} & (3.38) \\ \dot{V}\left(U+\alpha(m)\right) = \dot{V}\left(u+\alpha(m)\right) = k^-, \quad \dot{V}\left(D+\alpha(m)\right) = \dot{V}\left(d+\alpha(m)\right) = -k^-; & (3.39) \\ V\left(u+\alpha(m)\right) = K^- + k^-(u-U) + V\left(U+\alpha(m)\right), \\ V\left(d+\alpha(m)\right) = K^+ + k^+(D-d) + V\left(D+\alpha(m)\right). & (3.40) \end{cases}$$

Recall that $K^\pm, k^\pm > 0$ and $h - rk^-, p - rk^+ > 0$. By [58], there exists a 6-tuple (c_1, c_2, d, D, U, u) satisfying (3.38), (3.39) and (3.40) such that $d < D < 0 < U < u$ and

$$\begin{cases} c_1 = \frac{h+p}{r\lambda}\frac{(e^{-\lambda u} - e^{-\lambda U})[\cosh(\lambda d) - \cosh(\lambda D)]}{(e^{\lambda u} - e^{\lambda U})(e^{-\lambda d} - e^{-\lambda D}) - (e^{-\lambda u} - e^{-\lambda U})(e^{\lambda d} - e^{\lambda D})} \in (-\frac{h+p}{2r\lambda}, 0), & (3.41) \\ c_2 = c_1 \frac{e^{\lambda u} - e^{\lambda U}}{e^{-\lambda u} - e^{-\lambda U}} \in \left(0, \frac{h+p}{2r\lambda}\right), & (3.42) \\ K^- - \left(\frac{h}{r} - k^-\right)(u-U) - 2c_1(e^{\lambda u} - e^{\lambda U}) = 0, & (3.43) \\ \lambda\left(c_1 e^{\lambda u} - c_2 e^{-\lambda u}\right) + \left(\frac{h}{r} - k^-\right) = 0, & (3.44) \\ K^+ - \left(\frac{p}{r} - k^+\right)(D-d) - 2\left(c_1 + \frac{h+p}{2r\lambda}\right)(e^{\lambda u} - e^{\lambda U}) = 0, & (3.45) \\ \lambda\left[\left(c_1 + \frac{h+p}{2r\lambda}\right)e^{\lambda u} - \left(c_2 - \frac{h+p}{2r\lambda}\right)e^{-\lambda u}\right] - \left(\frac{p}{r} - k^+\right) = 0, & (3.46) \end{cases}$$

where the thresholds d, D, u, U only depend on $K^{\pm}, k^{\pm}, h, p, r, \sigma$. The optimal simple control policy $\varphi^* = ((-\infty, \alpha(m)+d] \cup [\alpha(m)+u, +\infty), \xi^*) = \{(\tau_n^*, \xi_n^*)\}_{n\geq 1}$ is given by

$$\begin{aligned} \xi^*(x) &= \begin{cases} U - x + \alpha(m), & \text{if } x - \alpha(m) \geq \alpha(m) + u, \\ D - x + \alpha(m), & \text{if } x - \alpha(m) \leq \alpha(m) + d, \\ 0, & \text{otherwise.} \end{cases} \\ \tau_1^* &= \inf\{t \geq 0 : |X_{t-} - \alpha(m)| \notin (\alpha(m)+d, \alpha(m)+u)\}, \\ \tau_n^* &= \inf\{t > \tau_{n-1}^* : |X_{t-} - \alpha(m)| \notin (\alpha(m)+d, \alpha(m)+u)\}, \quad n \geq 2; \\ \xi_n^* &= \xi^*(X_{\tau_n -}) \end{aligned} \tag{3.47}$$

Assume that the initial position X_{0-} follows any given distribution μ. Recall from (MFG-I) that

$$\begin{aligned} V(x) &= \inf_{\varphi} \mathbb{E}_x \left[\int_0^\infty e^{-rt} f(X_t, m) dt + \sum_{n\geq 1} e^{-r\tau_n} \phi(\xi_n) \right] \\ &= \inf_{\varphi} \mathbb{E}_\mu \left[\int_0^\infty e^{-rt} f(X_t, m) dt + \sum_{n\geq 1} e^{-r\tau_n} \phi(\xi_n) \middle| X_{0-} = x \right]. \end{aligned}$$

Denote the updated mean information as $\bar{m} = \limsup_{t\to\infty} \mathbb{E}_x[X_t]$. We will show that this $\bar{m}$ is well-defined and invariant with respect to x.

Notice that $\bar{m} = \limsup_{t\to\infty} \mathbb{E}_x[X_t] = \limsup_{n\to\infty} \mathbb{E}_x[X_{\tau_n^*}]$ by symmetry and a Fubini argument. Note that $\mathbb{E}_x[X_{\tau_n^*}] = \alpha(m) + U\mathbb{P}\{X_{\tau_n^*} = \alpha(m)+U\} + D\left[1 - \mathbb{P}\{X_{\tau_n^*} = \alpha(m)+U\}\right]$. For simplification, denote $\mathbb{P}\{X_{\tau_n^*} = \alpha(m) + U | X_{0-} = x\}$ as $p_n(x)$. Then, by the strong Markovian property of $\bar{X}_t$

$$\begin{aligned} q_1 &\equiv \mathbb{P}\{X_{\tau_{n+1}^*} = \alpha(m) + U | X_{\tau_n^*} = \alpha(m) + U\} = \frac{U-d}{u-d}, \forall n \in \mathbb{N}, \\ q_2 &\equiv \mathbb{P}\{X_{\tau_{n+1}^*} = \alpha(m) + U | X_{\tau_n^*} = \alpha(m) + D\} = \frac{D-d}{u-d}, \forall n \in \mathbb{N}, \\ p_{n+1}(x) &= p_n(x) q_1 + [1 - p_n(x)] q_2. \end{aligned}$$

Therefore, we have

$$p_{n+1}(x) = q_2 + (q_1 - q_2) p_n(x) \Rightarrow p_n(x) - \frac{q_2}{1 - q_1 + q_2} = (q_1 - q_2)^{n-1} \left[p_1(x) - \frac{q_2}{1 - q_1 + q_2} \right].$$

Hence, $\lim_{n\to\infty} p_n(x) = \frac{q_2}{1-q_1+q_2}$ and this is independent of the initial position x. We then have $\bar{m} = \alpha(m) + \frac{uD - dU}{u - U + D - d}$. Define the update of mean information $\Gamma : m \mapsto \bar{m}$ as $\Gamma(m) = \alpha(m) + \frac{uD-dU}{u-U+D-d}$. Since α is assumed a contraction mapping, so is Γ. Denote the fixed point of Γ as m^* and let $\varphi^* = \varphi(m^*)$ be be as in (3.47). Then (φ^*, m^*) is a solution to the (MFG-I) in the sense of Definition 3.8. □

Remark 3.8. *Notice that in the example of Section 3.1.3, the cash management setting is under a symmetric cost structure with $h = p$, $K^{\pm} = K$, $k^{\pm} = k$. Its mean-field counterpart, by considering a similar derivation as in the proof of Theorem 3.13, has a symmetric solution structure with $d = -u$ and $D = -U$. Therefore $\frac{uD-dU}{u-U+D-d} = 0$, and $m^* = \mathbb{E}_\mu[X_{0-}]$ is a solution.*

3.3 Sensitivity analysis.

In this section, we come back to the *symmetric* cash management problem, and compare the solutions in the cases $N = 1$ ('monopoly', described in [58] and here recalled in Section 1) and $N = 2$ ('duopoly', introduced in Section 3.1.3 as the multi-player extension of the problem in [58]).

Namely, we want to study see how parameters h, K, k, r and σ influence the control policies and the thresholds d, D, U, u, we conduct a series of sensitivity analysis. We start with $h = 2$, $K = 3$, $k = 1$, $r = 0.5$, $\sigma = \frac{\sqrt{2}}{2}$ and $c = 1$.

We shall see similar behaviors for both the monopoly and the duopoly cases in terms of the thresholds and policy changes with respect to the underlying parameter changes. One distinction is that the thresholds and policy changes are more sensitive to parameter changes in the duopoly case due to competition.

Finally, we will study the sensitivity analysis for the MFG counterpart.

3.3.1 Duopoly vs monopoly.

Putting the thresholds for the duopoly and those of the monopoly together in Figure 3.3, one can see that due to competition in a game setting, players take the opponents' strategies into consideration.

We notice that the continuation region gets bigger in the duopoly case. Equivalently, interventions on the underlying process are less frequent in the duopoly case than in the monopoly case. Also, notice that the intervention size when the process reaches the lower or upper threshold are equal, due to the symmetric structure of the problem. In summary, the player of the monopoly makes frequent but cautious interventions whereas each player in the duopoly intervenes less often but each time with bolder moves.

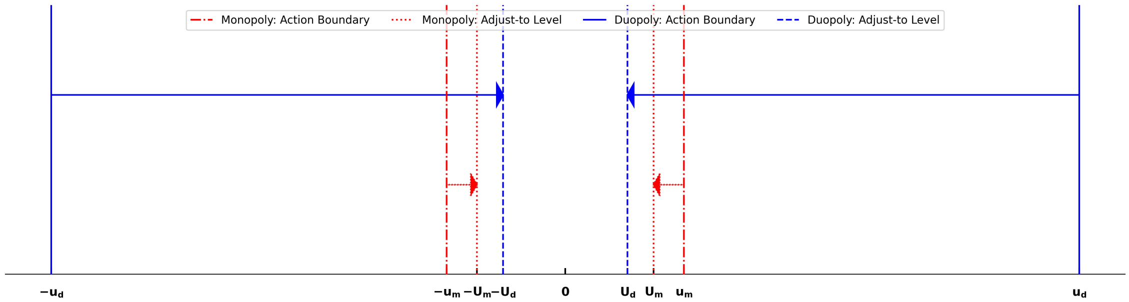
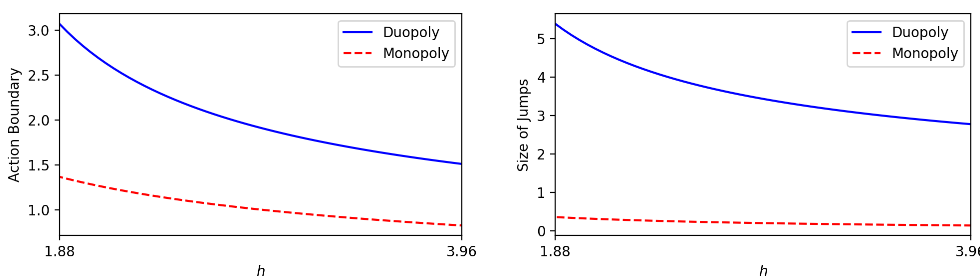


Figure 3.3: Thresholds: Duopoly vs. Monopoly

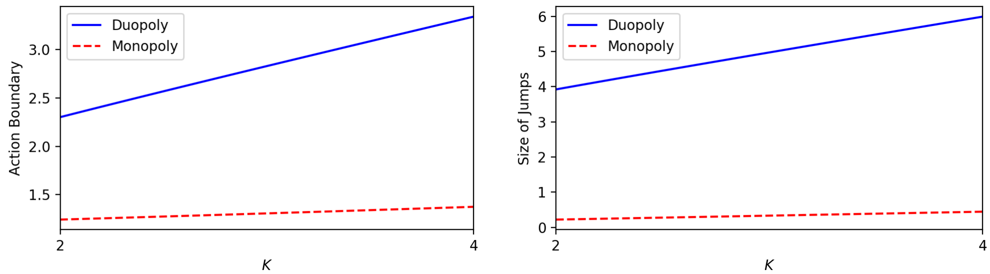
Running Cost h. When the running cost h increases, players have the incentive to intervene more frequently to prevent controlled process from deviating too far away from the target level. See Figure 3.4a. On the other hand, the presence of the cost of control makes players more cautious when exercising controls. Thus, an increased running cost encourages the players to intervene more frequently but with smaller amount of adjustment. See Figure 3.4b.



(a) Action Boundary of Monopoly and Duopoly (b) Amount of Adjustment of Monopoly and Duopoly

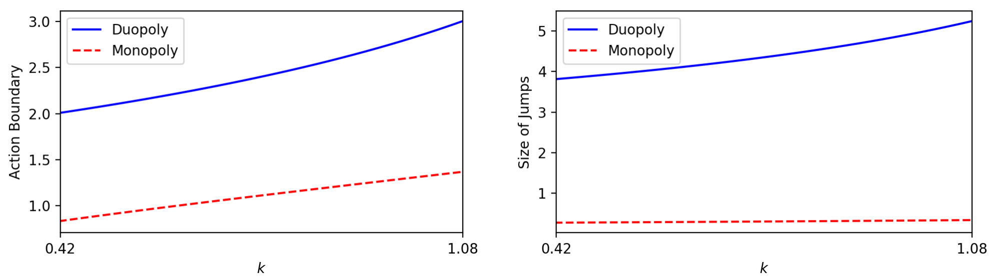
Figure 3.4: Sensitivity w.r.t. h

Cost of Control K and k. The parameter K is the fixed cost when players choose to intervene. High fixed cost K discourages the player from intervening too frequently. Therefore players have the incentive to tolerate a larger deviation from the target; and once a player chooses to intervene, the size of control needs to be bigger to compensate for less frequent controls. Meanwhile, a decreasing frequency of intervention leads to an increasing action boundary u. See Figure 3.5a and Figure 3.5b. For the per unit control cost k, similar results are shown in Figure 3.6.



(a) Action Boundary of Monopoly and Duopoly (b) Amount of Adjustment of Monopoly and Duopoly

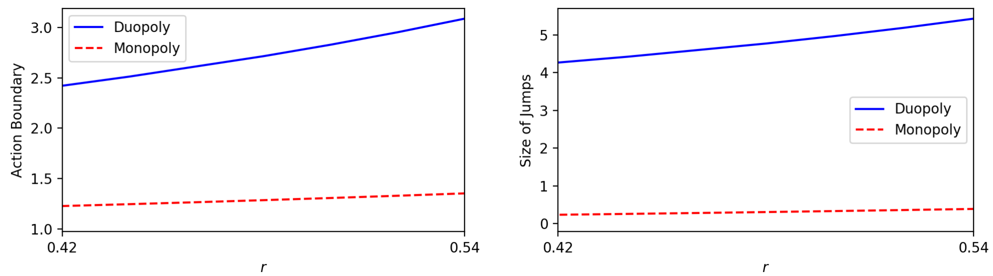
Figure 3.5: Sensitivity w.r.t. K



(a) Action Boundary of Monopoly and Duopoly (b) Amount of Adjustment of Monopoly and Duopoly

Figure 3.6: Sensitivity w.r.t. k

Discount Rate r. When the discount rate r increases, players are more tolerant with a larger deviation from the target level as the penalty is discounted by a larger factor. That is, a higher discount rate r effectively reduces both the running and control cost, hence resulting in a decreased intervention frequency with an increased size of controls, as shown in Figure 3.7.

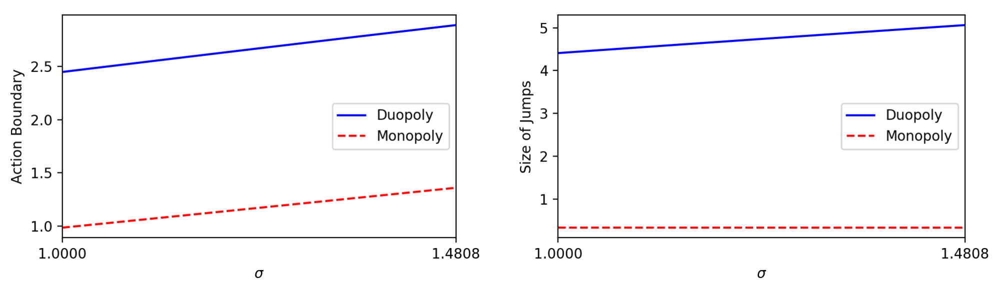


(a) Action Boundary of Monopoly and Duopoly

(b) Amount of Adjustment of Monopoly and Duopoly

Figure 3.7: Sensitivity w.r.t. r

Volatility σ. When the volatility σ is bigger, players tend to intervene less as the controlled process is more likely to move closer to the target level with a higher volatility. Therefore, a higher volatility allows players to intervene less frequently with a larger amount of adjustment, as shown in Figure 3.8.



(a) Action Boundary of Monopoly and Duopoly

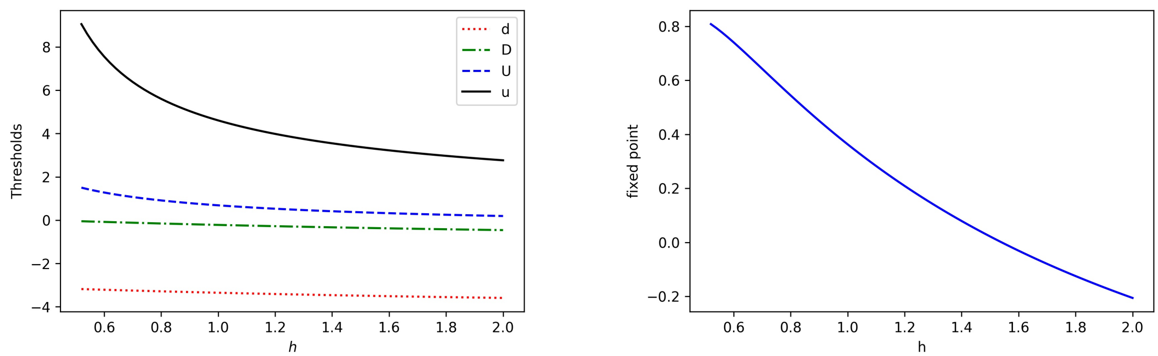
(b) Amount of Adjustment of Monopoly and Duopoly

Figure 3.8: Sensitivity w.r.t. σ

3.3.2 Sensitivity analysis of the MFG.

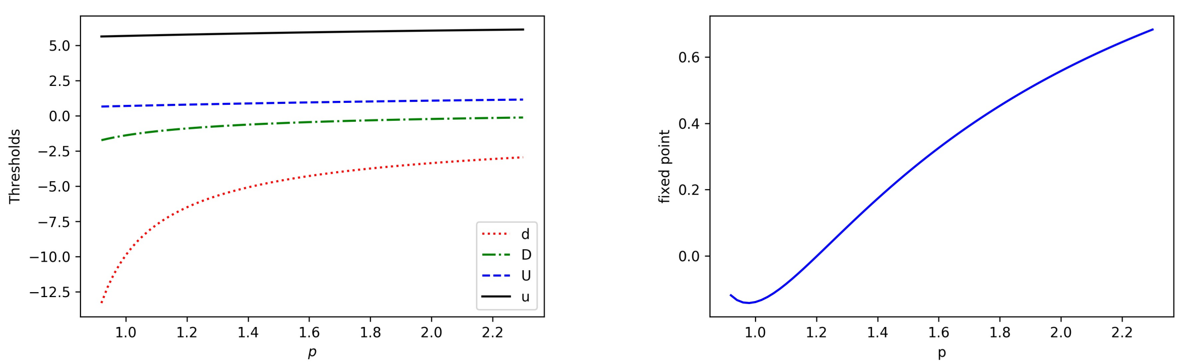
Here we present the sensitivity analysis of the MFG solution with respect the model parameters. In particular, we look into the impact of the list of parameters, namely h, p, $K^{\pm}$, $k^{\pm}$, r and σ, on the optimal impulse control policy and on the value of the mean information in the solution. Here we take one particular form of the α function $\alpha(m) = \alpha m$ with $0 < \alpha < 1$.

Running cost h and p. Parameters relevant to running costs are holding cost h and penalty cost p. When running cost increases, players have the incentive to intervene more frequently to prevent the state from deviating too far away from the target level α; at the meantime, players pay extra precaution for each intervention so the jump size reduces. Figure 3.9 shows the impact of increasing holding cost h with $p = 2$, $K^- = 3$, $k^- = 0.5$, $K^+ = 3.25$, $k^+ = 1.5$, $r = 0.5$, $\sigma = 1$. It primarily affects the upper action boundary u, causing it to decreases. Meanwhile the gap between the u and U decreases as well. As for the mean information, eventually, increasing h causes the mean information to decrease. Figure 3.10 illustrates the effect of increasing p with $h = 1$, $K^- = 3$, $k^- = 1$, $K^+ = 3.25$, $k^+ = 1.5$, $r = 0.5$, $\sigma = 1$. As opposed to increasing h, increasing p primarily impacts d, causing to increase, while the gap between d and D decreases. Mostly it leads the mean information to increase.



(a) All thresholds (b) Fixed point

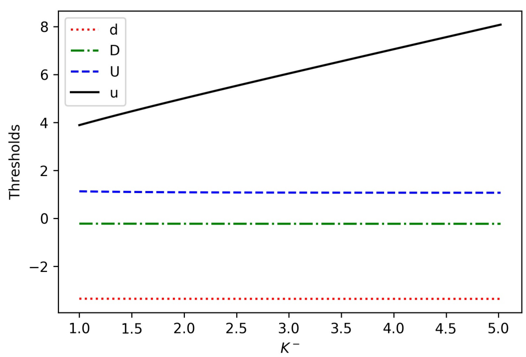
Figure 3.9: Sensitivity w.r.t. h



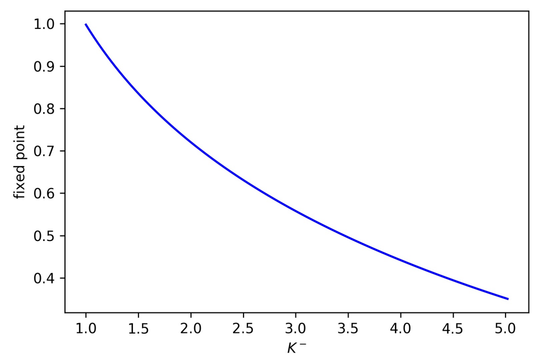
(a) All thresholds (b) Fixed point

Figure 3.10: Sensitivity w.r.t. p

Cost of instant decrease K^- and k^-. The parameter K^- is the amount of fixed cost to pay every time a player chooses to decrease the state. A high fixed cost K^- discourages the player from deceasing the state too often. Therefore the players have the incentive to tolerate a high state value compared with the reference point α and once a player chooses to decrease its state, the amount is adjustment will be larger. The more obvious impact is that the distance between α and the upper bound of the non-action region, i.e. u, increases; the impact on U is trickier, as the per unit cost of decreasing the state stays unchanged. The impact of the increasing K^- on the rest of the thresholds becomes the trade-of between unchanged holding cost, penalty cost, cost of increasing the state, and the costs associated with potentially more frequent but larger in scales, instantaneous increases in the state variables.



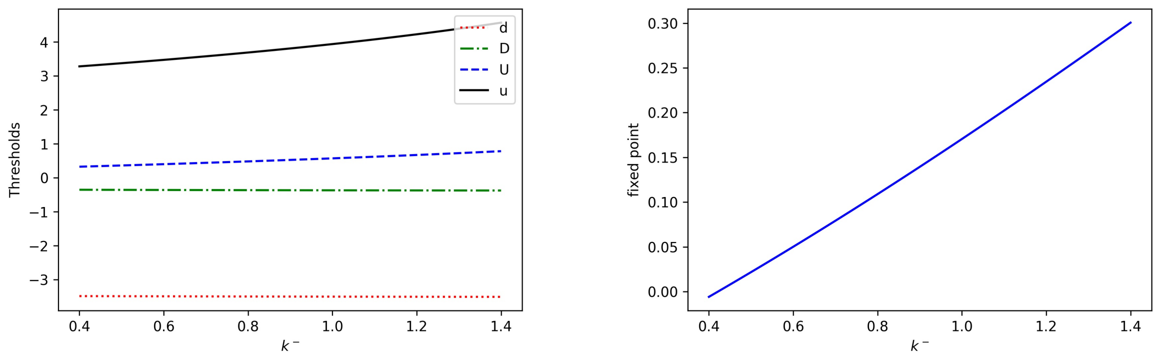
(a) All thresholds



(b) Fixed point

Figure 3.11: Sensitivity w.r.t. K^-

Figure 3.11 illustrates the effect of varying K^- from 1 to 5 under fixed values of the rest of the parameters, where $h = 1$, $p = 2$, $k^- = 1$, $K^+ = 3.25$, $k^+ = 1.5$, $r = 0.5$, $\sigma = 1$. We can see from Figure 3.11a that as K^- grows, u is affected the most and it increases significantly indicating a decreased frequency of intervention, with a larger jumper size. Similar impact on control policy can be observed with an increasing k^-, shown in Figure 3.12. It pushes the equilibrium mean information to a lower level, as shown in Figure 3.11b.



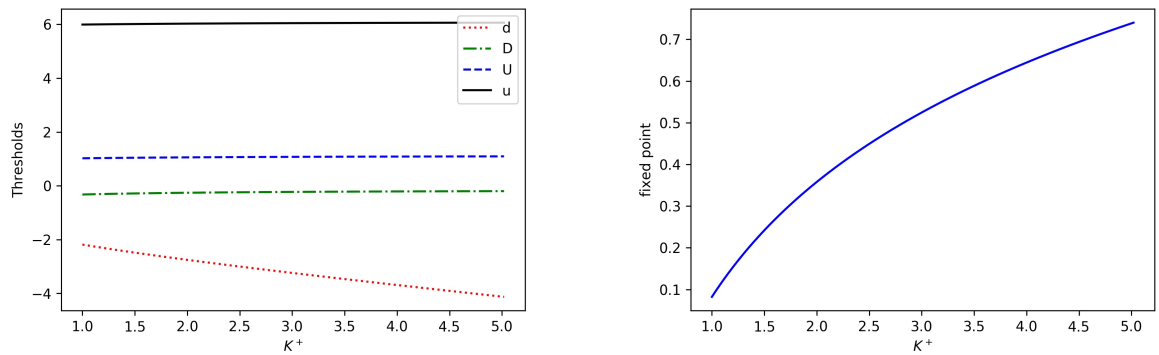
(a) All thresholds

(b) Fixed point

Figure 3.12: Sensitivity of the thresholds w.r.t. k^-

A lower frequency of decreasing can lead the state to move right while larger jump size may push the state to the left. While, as shown in Figure 3.11b, increasing K^- results in a decreasing mean information, increasing k^- causes mean information to increasing, see Figure 3.12b.

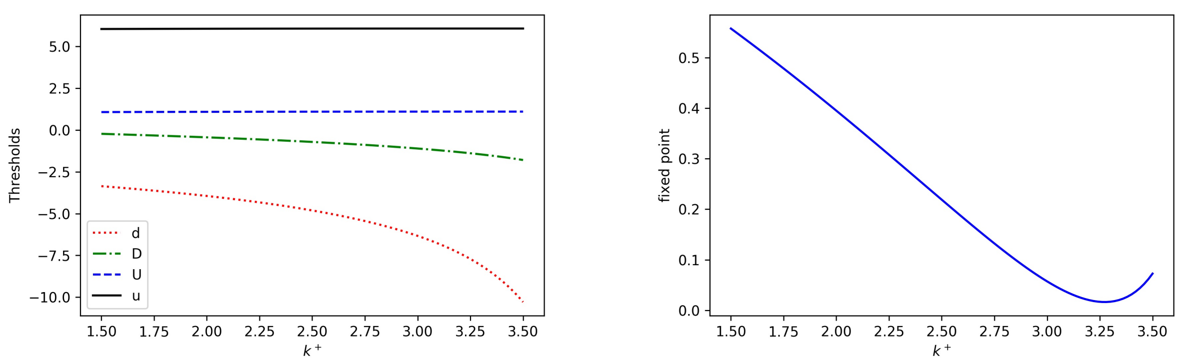
Cost of instant increase K^+ and k^+. The parameter K^+ is the amount of fixed cost to pay every time a player chooses to increase the state. A high fixed cost K^+ discourages the player from increasing the state too often. Therefore the players have the incentive to tolerate a low state value compared with the reference point α and once a player chooses to increase its state, the amount is adjustment will be larger. The more obvious impact is that the distance between α and the lower bound of the non-action region, i.e. d, decreases; the impact on D is trickier, as the per unit cost of increasing the state stays unchanged. The impact of the increasing K^+ on the rest of the thresholds becomes the trade-of between unchanged holding cost, penalty cost, cost of decreasing the state, and the costs associated with potentially more frequent but larger in scales, instantaneous decreases in the state variables.



(a) All thresholds (b) Fixed point

Figure 3.13: Sensitivity w.r.t. K^+

Figure 3.13 illustrates the effect of varying K^+ from 1 to 5 under fixed values of the rest of the parameters, where $h = 1$, $p = 2$, $K^- = 3$, $k^- = 1$, $k^+ = 1.5$, $r = 0.5$, $\sigma = 1$. We can see from Figure 3.13a that as K^+ grows, d gets the most influence and decreases significantly, indicating a lower frequency of intervention with larger jump size. Similar impact on control policy can be seen if k^+ increases.

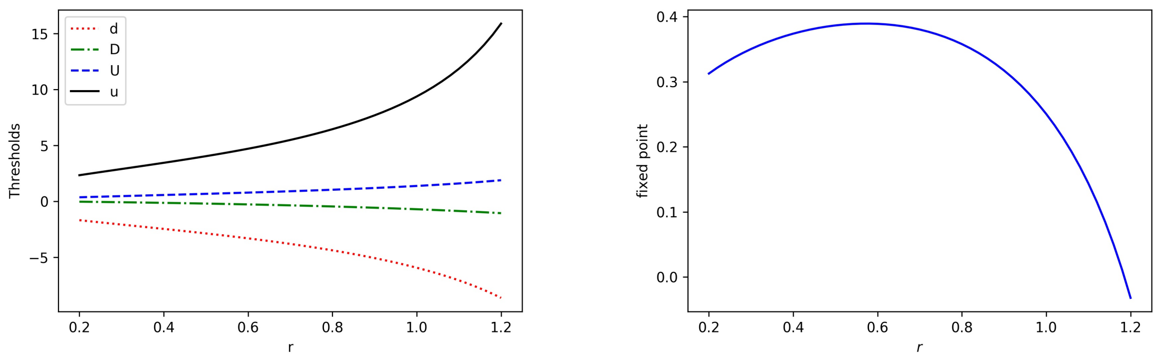


(a) All thresholds (b) Fixed point

Figure 3.14: Sensitivity w.r.t. k^+

A lower frequency of increase motivates the state to go left while larger jump size can push the state to the right. The impacts on the mean information by varying K^+ and k^+ are different as shown in Figure 3.13b and Figure 3.14b.

Discount rate r. Figure 3.15 illustrates the effect of varying r from 0.2 to 1.2 under fixed values of the rest of the parameters, where $h = 1.5$, $p = 2.5$, $K^- = 3$, $k^- = 1$, $K^+ = 3.25$, $k^+ = 1.5$, $\sigma = 1$. We can see in Figure 3.15a that as r grows, d and u get influenced the



(a) All thresholds

(b) Fixed point

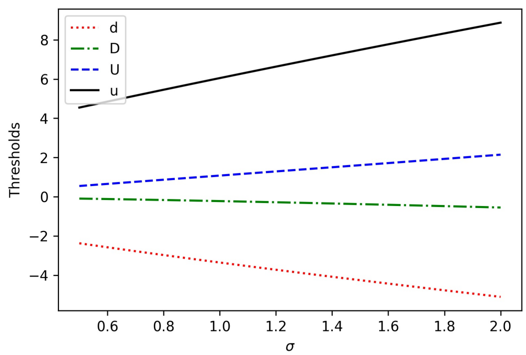
Figure 3.15: Sensitivity w.r.t. r

most indicating a decreased intervention in both directions, with a larger jump size. A lower decreasing frequency and a higher intensity of instant increases can both push the state the right while a lower increasing frequency and a higher intensity of instant decreases can both cause the state to go left. The influence on the mean information is shown in Figure 3.15b, that the mean first increases and then decreases as r grows.

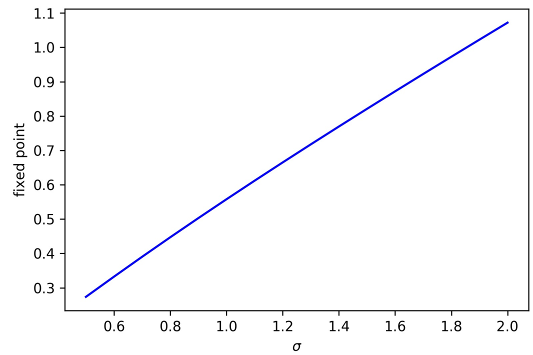
Individual volatility σ. Figure 3.16 illustrates the effect of varying σ from 0.5 to 2 under fixed values of the rest of the parameters, where $h = 1$, $p = 2$, $K^- = 3$, $k^- = 1$, $K^+ = 3.25$, $k^+ = 1.5$, $r = 0.5$. We can see from Figure 3.16a that as σ grows, d and u get influenced the most indicating a decreased intervention in both directions, with a larger jump size. The influence on the mean information is shown in Figure 3.16b, that the mean increases as σ grows.

3.4 Conclusion and remarks

In this chapter, we study the nonzero-sum stochastic differential games with impulse controls. In particular, we characterize NEs for a generic N-player games with impulse controls via a verification theorem and provide explicit forms of multiple NEs for a 2-player cash management game. We then formulate the MFG counterpart of the cash management problem and provide an explicit solution. We establish the ϵ-NE approximation from the MFG solution to its N-player NE. Finally, we compare the original control problem with the 2-player game via



(a) All thresholds



(b) Fixed point

Figure 3.16: Sensitivity w.r.t. σ

a sensitivity analysis with respect to the model parameters to demonstrate the impact of competition; the sensitivity analysis for the MFG counterpart is also presented.

MFGs with impulse controls under a generic problem setting is still an under explored topic. On one hand, for impulse control problems with generic cost structures, there is little known about characterization of the optimal impulse control policy other than the existence, uniqueness and regularity of the solution to QVI. On the other hand, it is remained to generalize the work of [26] on the FP equation for the controlled dynamics into a general framework. These two directions are worth further exploring. Alternative ways of studying general MFGs with impulse controls include the use of relaxed controls as well as analyzing a corresponding stopping problem as in [92].

Chapter 4

Connecting GANs, MFGs and OT

Up to date, theories of GANs, MFGs, and OT have been developed independently. In this chapter, we will show that they are intriguingly connected. In particular, GANs can be understood and analyzed from the perspective of MFGs and OT. More precisely,

- We first show a conceptual connection between GANs and MFGs: MFGs have the structure of GANs, and GANs are MFGs under the Pareto Optimality criterion. This intrinsic connection is transparent for a class of MFGs for which there is a minimax game representation.
- We then interpret MFGs as GANs and propose a GANs-based algorithm (MFGANs) to solve MFGs: one neural network (NN) for the backward HJB equation and one NN for the forward FP equation, with the two NNs trained in an adversarial way. Our numerical experiments demonstrate superior performance of this proposed algorithm when compared with existing approaches, especially in higher dimensional case.
- Finally, by viewing GANs as MFGs, we present new and probabilistic characteristics of GANs. This new perspective leads to an analytical connection between a broad class of GANs and Optimal Transport (OT) problems. This representation is explicit for Wasserstein GANs, by virtue of Kantorovich-Rubinstein duality theorem. Moreover, we identify sufficient conditions under which the minimax game of GANs can be redefined in the framework of OT: any form of divergence in GANs can be represented as the OT cost between the generated and the true data, provided that the OT admits a dual formulation. In this case, the discriminator corresponds to the price functions in the dual problem.

Notations. Throughout this chapter, the following notations will be adopted, unless otherwise specified.

- $\mathcal{X}$ denotes a Polish space with metric l.
- $\mathcal{P}(\mathcal{X})$ denotes the set of all probability distributions over the space $\mathcal{X}$.

- $\mathcal{P}(\mathbb{R}^d)$ denotes the set of probability distributions on $\mathbb{R}^d$ that admit corresponding density functions. That is, $\mu \in \mathcal{P}(\mathbb{R}^d)$ if there exists a mapping $m : \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} \mu(dx) = \int_{\mathbb{R}^d} m(x)dx = 1$.

- For $p > 0$, $\mathcal{P}^p(\mathbb{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \middle| \int_{\mathbb{R}^d} \|x\|_p^p \mu(dx) < \infty \right\}$, with $\| \cdot \|_p$ the p-norm on $\mathbb{R}^d$.

- For $p \geq 1$, $L^p(\mathcal{X}) := \left\{ \mu \in \mathcal{P}(\mathcal{X}) \middle| \int_{\mathcal{X}} d(x, x_0)^p \mu(dx) < \infty \right\}$ for some fixed $x_0 \in \mathcal{X}$.

- For any $\mu \in \mathcal{P}(\mathcal{X})$, $L^1(\mu) := \left\{ \psi : \mathcal{X} \to \mathbb{R} \middle| \int_{\mathcal{X}} |\psi(x)| \mu(dx) < \infty \right\}$.

4.1 Preliminaries

4.1.1 PDE system of MFGs.

As introduced in Chapter 1.1, MFGs have been introduced to tackle Nash equilibria in games with many players by considering a game with infinitely many agents. To customize into the discussion of this chapter, let us consider a class of MFGs as follows. Take $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ as a filtered probability space where $\{\mathcal{F}_t\}_{t\geq 0}$ supports a standard d-dimensional Brownian motion $W = \{W_t\}_{t\geq 0}$. Let W^i be i. i. d. copies of W. In MFGs, take any representative player i from infinitely many rational and indistinguishable players, her objective is to choose the optimal control over an admissible control set $\mathcal{A} = \{\{\alpha_t\}_{t\geq 0} : \alpha_t \in \mathbb{R}^d, \forall t \geq 0\}$ for the following minimization problem for any $s \in [0, T]$ and $x \in \mathbb{R}^d$:

$$u(s, x) = u(s, x; \{\mu_t\}_{t\in[0,T]}) = \inf_{\{\alpha_t\}_{t\geq 0} \in \mathcal{A}} E\left[\int_s^T f(t, X_t^i, \mu_t, \alpha_t)dt \,\middle|\, X_s^i = x \right] \tag{MFG}$$

subject to the state dynamics

$$\begin{aligned} dX_t^i &= b(t, X_t^i, \mu_t, \alpha_t)dt + \sigma dW_t^i, \; X_0^i \sim \mu^0, \\ \int_{\mathbb{R}^d} \mu^0(dx) &= \int_{\mathbb{R}^d} m^0(x)dx = 1. \end{aligned}$$

Here, the mean-field information is characterized by a flow of probability measures $\{\mu_t\}_{t\geq 0}$ with $\mu_0 = \mu^0 \in \mathcal{P}^2(\mathbb{R}^d)$, and the initial state of player i satisfies $X_0^i \sim \mu^0 \perp \sigma(W_t, t \geq 0)$. Moreover, μ_t is the limiting empirical distribution of players' states, and by strong law of large numbers $\mu_t = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i} = Law(X_t^i)$ for all $t \in (0, T]$.

In the cost function, $f : [0, \infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ is the running cost. Moreover μ^0 is the initial condition which is assumed to have a density denoted by m^0. In the state dynamics $\sigma > 0$ is a constant diffusion coefficient and the drift term $b : [0, \infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies appropriate conditions. These conditions ensure that there exists a unique solution

$\{X_t^i\}_{t\geq 0}$ for the state dynamics such that for any $t \geq 0$, $\mu_t = Law(X_t^i) \in \mathcal{P}^2(\mathbb{R}^d)$ ([151] and [74]).

We will denote by $m(t,\cdot)$ the density function of μ_t for any $t \geq 0$, and with slight abuse of notation, for any $\mu \in \mathcal{P}(\mathbb{R}^d)$ with density function m, denote $b(t,x,\mu,\alpha) := b(t,x,m,\alpha)$ and $f(t,x,\mu,\alpha) := f(t,x,m,\alpha)$.

Definition 4.1. *A control and mean-field pair $(\{\alpha_t^*\}_{t\geq 0}, \{\mu_t^*\}_{t\geq 0})$, with initial distribution $\mu_0^* = \mu^0$, is called the solution to* (MFG) *if the following conditions hold.*

- *(Optimal control) Under $\{\mu_t^*\}_{t\geq 0}$, $\{\alpha_t^*\}_{t\geq 0}$ solves the following optimal control problem that for $s \in [0,T]$ and $x \in \mathbb{R}$,*

$$u(s,x;\{\mu_t^*\}) = \inf_{\alpha\in\mathcal{A}} E\left[\int_s^T f(t,X_t^i,\mu_t^*,\alpha_t)dt \,\middle|\, X_s^i = x\right]$$

subject to

$$dX_t^i = b(t,X_t^i,\mu_t^*,\alpha_t)dt + \sigma dW_t^i,\ X_0 \sim \mu^0.$$

- *(Consistency) $\{\mu_t^*\}_{t\geq 0}$ is the flow of probability distribution of the optimally controlled process, i.e., $\mu_t^* = Law(X_t^{i,*})$ for $t \geq 0$, where $X^{i,*}$ is given by the following stochastic differential equation,*

$$dX_t^{i,*} = b(t,X_t^{i,*},\mu_t^*,\alpha_t^*)dt + \sigma dW_t^i,\ X_0^{i,*} \sim \mu^0.$$

The solution of this MFG (MFG) can be characterized by the following PDE system,

$$\left.\begin{aligned} \partial_s u(s,x) + \frac{\sigma^2}{2}\Delta_x u(s,x) + H\left(s,x,\nabla_x u(s,x)\right) = 0, \\ u(T,x) \equiv 0; \end{aligned}\right\} \tag{HJB}$$

$$\left.\begin{aligned} &\partial_s m(s,x) + div\left[m(s,x)b(s,x,m(s,x),\alpha^*)\right] = \frac{\sigma^2}{2}\Delta_x m(s,x), \\ &\quad m(t,\cdot) \geq 0,\ \int_{\mathbb{R}^d} m(t,x)dx = 1,\ \forall t \in [0,T]; \quad \int_{\mathbb{R}^d} m(0,x)dx = \int_{\mathbb{R}^d} m^0(dx). \end{aligned}\right\} \tag{FP}$$

Here the Hamiltonian $H(s,x,p)$ in (HJB) is given by

$$H\left(s,x,p\right) = \min_{\alpha\in\mathbb{R}^d}\left\{b(s,x,m(s,x),\alpha)p + f(s,x,m(s,x),\alpha)\right\} s \in (0,T),\ x,p \in \mathbb{R}^d,$$

and α^* in (FP) is the optimal control, with

$$\alpha_t^* = \arg\min_{\alpha\in\mathbb{R}^d}\left\{b(t,x,m(t,x),\alpha)\nabla_x u(t,x) + f(t,x,m(t,x),\alpha)\right\}. \tag{4.1}$$

Note that from (4.1), the optimal control is determined by the value function u.

4.1.2 Basics of OT

Theory of optimal transport (OT), originated from Monge [135], studies the optimization problem of transporting one given initial distribution to another given terminal distribution so that the transport cost functional is minimized. Deeply rooted in linear programming, many theoretical works on the existence and uniqueness of an optimal transport plan focus on the duality gap between the primal optimization problem and its dual form. For instance, the Kantorovich-Rubinstein duality in [171] characterizes sufficient conditions of the existence of optimal transport plans (i.e., when there is no duality gap), and [81] studies the multi-marginal case.

Martingale optimal transport problem is motivated mostly by problems from finance, starting with the problem of super-hedging, see for instance [67], [126], and [143]. [17] establishes a complete duality theory for generic martingale OT problems via a quasi-sure formulation of the dual problem; and [86] studies a continuous-time martingale OT problem and establishes the duality theory via the S-topology and the dynamic programming approach.

To compute for the optimal transport plan, [85] proposes a computational method for martingale optimal transport problem based on linear programming via proper relaxation of the martingale constraint and discretization of the marginal distributions; [70] proposes a deep learning algorithm to solve multi-step, multi-marginal optimal transport problem via its dual form with an appropriate penalty term.

It is worth pointing out that OT theory has been applied to analyze and improve the stability of GANs training, see, for instance, [84], [154], and [55]. There are earlier studies connecting GANs and OT, by different approaches and from different perspectives. [153] defines a novel divergence based on solutions of three associated optimal transport problems. This new divergence is then used to replace the JS divergence for the vanilla GANs. In [124], an interpretation of Wasserstein GANs (WGANs) from the geometric perspective of optimal transport is provided. In our work, we provide sufficient conditions for which the minimax game of general GANs, including WGANs, can be reformulated analytically in the framework of OT. (See Remark 4.2 for more detailed discussions).

Mathematically, the OT is defined as follows (see [171]).

Definition 4.2. *Take a Polish space $\mathcal{X}$ with metric $d : \mathcal{X} \times \mathcal{X} \to [0, \infty)$. Let $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \bigcup \{+\infty\}$ be a lower semi-continuous function such that $c(x, y) \geq a(x) + b(y)$, where a and b are upper semi-continuous functions on $\mathcal{X}$. For any $\mu, \nu \in \mathcal{P}(\mathcal{X})$, the OT problem between μ and ν with cost function c is defined as*

$$W_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} c(x, y)\pi(dx, dy), \tag{OT}$$

$\Pi(\mu, \nu)$ is the set of all possible couplings between μ and ν.

The well-definedness of this OT, i.e., the existence of an optimal cost W_c, is guaranteed by Theorem 4.1 of [171].

The dual problem of this OT goes as follows.

Definition 4.3. *Let $\mu, \nu \in \mathcal{P}(\mathcal{X})$. The dual Kantorovich problem of* (OT) *is*

$$D_c(\mu,\nu) = \sup_{\psi \in L^1(\mu), \phi \in L^1(\nu)} \left\{ \int_{\mathcal{X}} \phi(x)\nu(dx) - \int_{\mathcal{X}} \psi(x)\mu(dx) \middle| \phi(x) - \psi(y) \leq c(x,y), \, \forall (x,y) \in \mathcal{X} \times \mathcal{X} \right\}.$$

It is easy to see that $D_c(\mu,\nu) \leq W_c(\mu,\nu)$. The following Kantorovich-Rubinstein duality provides sufficient conditions under which the equality holds.

Theorem 4.4 (Theorem 5.10(i) in [171])**.** *Take $\mu, \nu \in \mathcal{P}(\mathcal{X})$. Let $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \bigcup \{+\infty\}$ be a lower semi-continuous function such that $c(x,y) \geq a(x) + b(y)$ where $a \in L^1(\mu)$ and $b \in L^1(\nu)$ are upper semi-continuous functions on $\mathcal{X}$. Then,*

$$W_c(\mu,\nu) = D_c(\mu,\nu) = \sup_{\psi \in L^1(\mu)} \int_{\mathcal{X}} \psi^c(x)\nu(dx) - \int_{\mathcal{X}} \psi(x)\mu(dx).$$

Here $\psi : \mathcal{X} \to \mathbb{R} \bigcup \{+\infty\}$ is taken from the set of all c-convex functions: ψ is not constantly $+\infty$ and there exists a function $\zeta : \mathcal{X} \to \mathbb{R} \bigcup \{+\infty\}$ such that

$$\psi(x) = \sup_{y \in \mathcal{X}} \left[\zeta(y) - c(x,y) \right], \quad \forall x \in \mathcal{X}.$$

$\psi^c : \mathcal{X} \to \mathbb{R} \bigcup \{-\infty\}$ is its c-transform

$$\psi^c(y) = \inf_{x \in \mathcal{X}} \left[\psi(x) + c(x,y) \right], \quad \forall y \in \mathcal{X}.$$

If the transport cost c takes the particular form of $c = d^p$ for some $p \geq 1$, then the corresponding optimal cost gives rise to the Wasserstein distance between μ and ν of order p, or simply the Wasserstein-p distance,

$$W_p(\mu,\nu) = \left[\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X} \times \mathcal{X}} l(x,y)^p \pi(dx,dy) \right]^{\frac{1}{p}}.$$

Note that for $p = 1$, the Wasserstein-1 distance is adopted in Wasserstein GANs (WGANs) in [6] to improve the stability of GANs.

4.2 MFGs as GANs.

As stochastic differential games with a continuum of players, MFGs in general are different from 2-player minimax games such as GANs. However, there is a class of MFGs that indeed can be view as minimax games. It becomes a spark to explore the conceptual connection between GANs and MFGs of a broader class. In this section, we first establish the conceptual connection between MFGs and GANs, we then present a class of MFGs for which such connection to GANs is explicit. Based on this connection, we proposal a GANs-based algorithm (MFGANs) for solving MFGs.

4.2.1 A class of MFGs with minimax structure

As pointed out by [56], a class of periodic MFGs on flat torus $\mathbb{T}^d$ and a finite time horizon $[0, T]$ admits an explicit minimax structure. Consider such an MFG that minimizes the following cost,

$$J_m(t, \alpha) = \mathbb{E}\left[\int_t^T L(X_t^\alpha, \alpha(X_t^\alpha)) + f(X_t^\alpha, m(X_t^\alpha))dt\right], \ t \in [0, T] \tag{4.2}$$

where $X^\alpha = (X_t^\alpha)_t$ is a d-dimensional process with dynamics

$$dX_t^\alpha = \alpha(X_t^\alpha)dt + \sqrt{2\epsilon}dW_t.$$

Here α is a control policy, L and f constitute the running cost and $m(t, \cdot)$, for $t \in [0, T]$, denotes the probability density of X_t^α at time t. We then introduce the convex conjugate of the running cost L, namely,

$$H_0(x, p) = \sup_{\alpha \in \mathbb{R}^d} \{\alpha \cdot p - L(x, \alpha)\},$$

and denote $F(x, m) = \int^m f(x, z)dz$. From a PDE perspective, this class of MFGs can be characterized by the following coupled PDE system as illustrated in [56],

$$\begin{cases} -\partial_s u - \epsilon\Delta_x u + H_0(x, \nabla_x u) = f(x, m), \\ \partial_s m - \epsilon\Delta_x m - div\,(m\nabla_p H_0(x, \nabla u)) = 0, \\ m > 0,\ m(0, \cdot) = m^0(\cdot),\ u(T, \cdot) = u^T(\cdot), \end{cases} \tag{4.3}$$

where the first equation is an HJB equation governing the value function and the second is an FP equation governing the evolution of the optimally controlled state process; here m^0 and u^T are the initial functions for $m(t, \cdot)$ and $u(t, \cdot)$, respectively. The system of equations (4.3) is equivalent to the following minimax game

$$\inf_{u \in \mathcal{C}^2([0,T]\times\mathbb{T}^d)} \sup_{m \in \mathcal{C}^2([0,T]\times\mathbb{T}^d)} \Phi(m, u), \tag{4.4}$$

where

$$\begin{aligned} \Phi(m, u) = &\int_0^T \int_{\mathbb{T}^d} [m(-\partial_t u - \epsilon\Delta_x u) + mH_0(x, \nabla_x u) - F(x, m)]\, dxdt \\ &+ \int_{\mathbb{T}^d} \left[m(T, x)u(T, x) - m^0(x)u(0, x) - m(x, T)u^T(x)\right] dx. \end{aligned}$$

From (4.4), one can see that the connection between GANs and MFGs is transparent.

4.2.2 Conceptual connection

Having this transparent connection between GANs with a particular class of MFGs, one may continue asking if such a connection can be extended to a broader class of MFGs. We claim that

MFGs in the form of (MFG) *are GANs.*

To see this connection more precisely, recall that in classical GANs, the generator G is to mimic the sample data to generate new one to minimize the difference between the true distribution $\mathbb{P}_r$ and $\mathbb{P}_\theta$. The discriminator D measures the performance of the generator by some divergence between $\mathbb{P}_\theta$ and $\mathbb{P}_r$. Meanwhile, observe that in MFGs:

- The latent space $\mathcal{Z} = \mathbb{R}^d$ and sample x of latent variable Z are drawn from the probability distribution $\mathbb{P}_z = \mu^0$.

- The generator $G = u$ maps the element x into $\mathbb{R}$ so that it mimics the optimal cost and its gradient dictates the optimal strategy in the equilibrium state of the MFGs. We can then define a loss function

  $$L_G(u, m) = L_{HJB}(u, m) + \beta_G L_{term}(u, m),$$

  where $\beta_G > 0$ denotes the weight on the penalty of the terminal condition,

  $$L_{HJB}(u, m) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \left[\partial_s u(s, x; m) + \frac{\sigma^2}{2}\Delta_x u(s, x; m) + H\left(s, x, \nabla_x u(s, x; m)\right)\right]^2 \mu^0(dx)ds,$$

  and

  $$L_{term} = \int_{\mathbb{R}^d} u(T, x; m)^2 \mu^0(dx).$$

  Here we use $u(\cdot, \cdot; m)$ to denote the value function u, as from the PDE system (HJB)-(FP), the value function u is coupled with the mean information m; later on we will also use $m(\cdot, \cdot; u)$ to denote the mean information m.

- The equilibrium state of the MFGs, just as the true distribution $\mathbb{P}_r$ in GANs, exists but is not explicitly available. The characterization of the equilibrium is through a *consistency condition* between value function and the controlled dynamics. The discriminator $D = m$ helps to measure the distance from the current state process to the equilibrium state process by checking if m is indeed the density function of the state dynamic (4.1.1) under the optimal control given by the generator. Its loss, in place of the divergence function between $\mathbb{P}_\theta$ and $\mathbb{P}_r$, is defined as

  $$L_D(u, m) = L_{FP}(u, m) + \beta_D L_{init}(u, m),$$

where $\beta_D > 0$ denotes the weight on the penalty of the initial condition,

$$L_{FP}(u,m) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\partial_s m(s,x;u) + div\left[m(s,x;u)b(s,x,m(s,x;u),\alpha^*(u,m))\right] - \frac{\sigma^2}{2}\Delta_x m(s,x;u)\Big]^2 \mu^0(dx)ds,$$

and

$$L_{init} = \int_{\mathbb{R}^d} \Big[m(0,x;u) - m^0(x)\Big]^2 \mu^0(dx).$$

Here $\alpha^*(u,m)$ denotes the optimal control solved under current u and m.

- In MFG the generator solves the HJB equation via an NN and the discriminator computes an appropriate differential residue of the FP equation via another NN.

The comparisons of the roles of generator and discriminator between general MFGs and GANs are summarized in Table 4.1.

Table 4.1: A first link between GANs and MFGs

	GANs	**MFGs**
Generator G	NN for approximating the map $G : \mathcal{Z} \mapsto \mathcal{X}$	NN for solving HJB
Characterization of $\mathbb{P}_r$	Sample data	FP equation for consistency
Discriminator D	NN measuring divergence between $\mathbb{P}_\theta$ and $\mathbb{P}_r$	NN for measuring differential residual from the FP equation

In fact there is more than one way to see this connection between MFGs and GANs. Alternatively, one can switch the roles of the generator and discriminator and view the mean-field term as a generator and the value function as a discriminator.

4.2.3 Computing MFGs via GANs

The above discussion points to a new computational approach for MFGs using NNs, assuming that the equilibrium of MFGs can be computed via the coupled HJB-FB system.

That is, one can compute MFGs using two neural networks in an adversarial way:

- u_θ being the NN approximation of the unknown value function u for the HJB equation,

- m_ω being the NN approximation for the unknown mean information function m.

This new computational algorithm for MFGs is summarized in Algorithm 1.

Note that Algorithm 1 can be adapted for broader classes of dynamical systems with variational structures. Such GANs structures have been exploited in [179] and [180] to synthesize complex systems governed by physical laws.

4.3 GANs as MFGs.

Having established MFGs as GANs, we next show that GANs are MFGs, under the Pareto Optimality criterion.

Theorem 4.5. *GANs in [80] are MFGs under the Pareto Optimality criterion, assuming that the latent variables Z and true data X are both i.i.d. sampled, respectively, with* $\mathbb{E}[|\log(D(X))|], \mathbb{E}[|\log(1-D(G(Z)))|] < \infty$ *for all possible D and G.*

Proof. Let $\mathbb{P}_r$ denote the probability distribution from which the real data is sampled on the sample space $\mathcal{X} \subset \mathbb{R}^d$, and let $\mathbb{P}_z$ be the prior distribution of the input on $\mathcal{Z} \subset \mathbb{R}^k$. A generator G maps any $z \in \mathcal{Z}$ to $G(z) \in \mathcal{X}$. A discriminator D, on the other hand, takes any sample $x \in \mathcal{X}$ and returns some probability of x being sampled from $\mathbb{P}_r$. The objective function of this GAN is

$$\min_G \max_D \mathbb{E}_{X\sim\mathbb{P}_r}\left[\log D(X)\right] + \mathbb{E}_{Z\sim\mathbb{P}_z}\left[\log\left(1 - D(G(Z))\right)\right], \tag{4.5}$$

where G and D are selected from appropriate functional spaces.

Consider a group of N indistinguishable players, each holding an initial belief distributed as $\mathbb{P}_z$, i.e., $Z_i \overset{i.i.d}{\sim} \mathbb{P}_z$ for $i = 1, \ldots, N$. Players can access the sample data from a masked model $\mathbb{P}_r$, independent from $\mathbb{P}_z$; each one is asked to find a strategy transforming the initial belief into a mimic version of the sample data so that on average the group can fool the best discriminator.

First, define the set of admissible strategies and the candidate pool for discriminators. Denote the set of admissible strategies as $\mathcal{G}$, which is the collection of mappings from $\mathcal{Z}$ to $\mathcal{X}$ and let the collection of possible discriminators $\mathcal{D}$ be the collection of mappings from $\mathcal{X}$ to $(0, 1]$. Fix any $i \in \{1, \ldots, N\}$, let Z_i be player i's initial belief and suppose $Z_i \in \mathcal{Z}$. Let X_j, $j = 1, \ldots, M$, be the sample data. When player k chooses strategy $G_k \in \mathcal{G}$, $k = 1, \ldots, N$, each player is subject to the same cost

$$J(\mathbf{G}) = \max_{D\in\mathcal{D}} \frac{\sum_{k=1}^N \sum_{j=1}^M \log\left[D(X_j)\left(1 - D(G_k(Z_k))\right)\right]}{N \cdot M},$$

where $\mathbf{G} = (G_1, \ldots, G_N) \in \bigotimes_{k=1}^N \mathcal{G}$ denotes the profile of strategies for all N players.

Recall that a profile of strategies $\mathbf{G}^*$ is called a **Pareto optimal point** (PO) if $J(\mathbf{G}^*) \leq J(\mathbf{G})$, for all $\mathbf{G} \in \bigotimes_{k=1}^N \mathcal{G}$.

Algorithm 1 MFGANs

At $k = 0$, initialize θ and ω. Let N_θ and N_ω be the number of training steps of the inner-loops and K be that of the outer-loop. Let $\beta_i > 0$, $i = 1, 2$.

for $k \in \{0, \ldots, K-1\}$ **do**

Let $m = 0$, $n = 0$.

Sample $\{(s_i, x_i)\}_{i=1}^{B_d}$ on $[0,T] \times \mathbb{R}^d$ according to a predetermined distribution p_{prior}, where B_d denotes the number of training samples for updating loss related to FP residual.

Let $\hat{L}_D(\theta, \omega) = \hat{L}_{FP}(\theta, \omega) + \beta_D \hat{L}_{init}(\omega)$, with

$$\begin{aligned}\hat{L}_{FP} &= \frac{1}{B_d}\left\{\sum_{i=1}^{B_d}\left[\partial_s m_\omega(s_i, x_i) + div\left[m_\omega(s_i, x_i) b(s_i, x_i, m(s_i, x_i), \alpha^*_{\theta,\omega}(s_i, x_i))\right]\right.\right.\\ &\qquad\left.\left. - \frac{\sigma^2}{2}\Delta_x m_\omega(s_i, x_i)\right]^2\right\},\\ \hat{L}_{init} &= \frac{\sum_{i=1}^{B_d}\left[m_\omega(0, x_i) - m^0(x_i)\right]^2}{B_d},\end{aligned}$$

where m^0 is a known density function for the initial distribution of the states and $\beta_D > 0$ is the weight for the penalty on the initial condition of m.

for $m \in \{0, \ldots, N_\omega - 1\}$ **do**

$\omega \leftarrow w - \alpha_d \nabla_\omega \hat{L}_D$ with learning rate α_d.

Increase m.

end for

Sample $\{(s_j, x_j)\}_{j=1}^{B_g}$ on $[0,T] \times \mathbb{R}$ according to a predetermined distribution p_{prior}, where B_g denotes the number of training samples for updating loss related to HJB residual.

Let $\hat{L}_G(\theta, \omega) = \hat{L}_{HJB}(\theta, \omega) + \beta_G \hat{L}_{term}(\theta)$, with

$$\begin{aligned}\hat{L}_{HJB} &= \frac{1}{B_g}\left\{\sum_{j=1}^{B_g}\left[\partial_s u_\theta(s_j, x_j) + \frac{\sigma^2}{2}\Delta_x u_\theta(s_j, x_j) + H_\omega\left(s_j, x_j, \nabla_x u_\theta(s_j, x_j)\right)\right]^2\right\},\\ \hat{L}_{term} &= \frac{\sum_{j=1}^{B_g} u_\theta(T, x_j)^2}{B_g},\end{aligned}$$

where $\beta_G > 0$ is the weight for the penalty on the terminal condition of u.

for $n \in \{0, \ldots, N_\theta - 1\}$ **do**

$\theta \leftarrow \theta - \alpha_g \nabla_\theta \hat{L}_G$ with learning rate α_g.

Increase n

end for

Increase k.

end for

Return θ, ω

Notice that the players are indistinguishable. Then there must be a symmetric PO consisting of the same strategy for all the players, provided that a PO exists. Let $\mathbb{S} \subset \bigotimes_{k=1}^{N} \mathcal{G}$ denote the set of symmetric strategies, i.e.,

$$\min_{\mathbf{G}\in\bigotimes_{k=1}^{N}\mathcal{G}} J(\mathbf{G}) = \min_{\mathbf{G}\in\mathbb{S}} J(\mathbf{G}) = \min_{G\in\mathcal{G}} \max_{D\in\mathcal{D}} \frac{\sum_{k=1}^{N}\sum_{j=1}^{M} \log\left[D(X_j)\left(1-D(G(Z_k))\right)\right]}{N\cdot M}.$$

When the number of players as well as the size of the sample data becomes large, by strong law of large numbers, almost surely we have

$$\frac{\sum_{k=1}^{N}\sum_{j=1}^{M} \log\left[D(X_j)\left(1-D(G(Z_k))\right)\right]}{N\cdot M} \to \mathbb{E}_{X\sim\mathbb{P}_r}\left[\log D(X)\right] + \mathbb{E}_{Z\sim\mathbb{P}_z}\left[\log\left(1-D(G(Z))\right)\right],$$

Now define $m_N = \frac{1}{N}\sum_{k=1}^{N}\delta_{G(Z_k)}$. Then by the strong law of large numbers, $m_N \overset{N\to\infty}{\Longrightarrow} Law(G(Z))$, with $Z \sim \mathbb{P}_z$. Here, $\mathbb{P}_G = Law(G(Z))$ is called the mean field. Therefore, by strong law of large numbers, sending M and N to ∞ the original loss for vanilla GANs is recovered,

$$\min_{G\in\mathcal{G}} \max_{D\in\mathcal{D}} \mathbb{E}_{X\sim\mathbb{P}_r}\left[\log D(X)\right] + \mathbb{E}_{Y\sim\mathbb{P}_\theta}\left[\log\left(1-D(Y)\right)\right].$$

□

4.4 GANs and OT.

As discussed in Section 4.1, through optimization over discriminators, GANs are essentially minimizing proper divergences between true distribution and the generated distribution over some sample space $\mathcal{X}$. The flexibility of choosing appropriate divergence allows us to connect GANs and OT problems, and to identify sufficient conditions for which GANs can be recast in the framework of OT.

Intuitively, this connection between GANs and OT is very natural: GANs as generative models are minimax games with the goal to minimize the "error" of the generated sample data against the true sample data; this error is measured under appropriate divergence functions between the true distribution and the generated distribution. Now if this error is viewed as a cost of transporting/fitting the generated distribution into the true distribution, GANs become an OT.

Indeed, this connection between GANs and OT is explicit in the case of WGANs.

Theorem 4.6. *Suppose that $\mathbb{P}_r \in L^1(\mathcal{X})$ and $G \in L^1(\mathbb{P}_z)$ where*

$$L^1(\mathbb{P}_z) = \left\{ f : \mathcal{Z} \to \mathbb{R} : \int_{\mathcal{Z}} |f(z)|\mathbb{P}_z(dz) < \infty \right\}.$$

WGAN is an OT problem between $Law(G(Z))$ and $\mathbb{P}_r$.

Proof. Recall the objective function of WGAN introduced in [6],

$$\min_{G} \max_{D \text{ s.t. } \|D\|_L \leq 1} \mathbb{E}_{X\sim\mathbb{P}_r}[D(X)] - \mathbb{E}_{Z\sim\mathbb{P}_z}[D(G(Z))].$$

Define a cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$,

$$c(x, y) = l(x, y), \tag{4.6}$$

with l being the metric of $\mathcal{X}$. Take a fixed generator $G \in L^1(\mathbb{P}_Z)$, then the cost function (4.6) is the cost of transporting mass from a distribution $Law(G(Z)) = \mathbb{P}_\theta$ to a different distribution $\mathbb{P}_r$. Then, consider the following OT problem

$$\inf_{\pi\in\Pi_G} \int_{\mathcal{X}\times\mathcal{X}} c(x, y)\pi(dx, dy), \tag{OT-WGAN}$$

where $\Pi_G = \Pi(Law(G(Z)), \mathbb{P}_r)$ is the set of couplings between $Law(G(Z))$ and $\mathbb{P}_r$, where $Z \sim \mathbb{P}_z$, for a fixed G.

By Theorem 4.4, when $c = l$, any c-convex function is Lipschitz with Lipschitz constant 1, and $\psi^c = \psi$. Therefore, the OT problem (OT-WGAN) becomes

$$\inf_{\pi\in\Pi_G} \int_{\mathcal{X}\times\mathcal{X}} c(x, y)\pi(dx, dy) = \sup_{D \text{ s.t. } \|D\|_L\leq 1} \int_{\mathcal{X}} D(x)\mathbb{P}_\theta(dx) - \int_{\mathcal{X}} D(y)\mathbb{P}_r(dy), \tag{4.7}$$

which is exactly the Wasserstein-1 distance between $\mathbb{P}_\theta$ and $\mathbb{P}_r$. The role of the discriminator is to locate the best coupling among Π_G for (OT-WGAN) under a given G, whereas the role of the generator is to refine the set of possible couplings Π_G so that the infimum in (OT-WGAN) becomes 0 eventually. Therefore, the following equivalence holds,

$$\min_{G} \max_{D \text{ s.t. } \|D\|_L \leq 1} \mathbb{E}_{X\sim\mathbb{P}_r}[D(X)] - \mathbb{E}_{Z\sim\mathbb{P}_z}[D(G(Z))] \Longleftrightarrow \min_{G} W_1(Law(G(Z)), \mathbb{P}_r).$$

□

Remark 4.1 (Sufficient condition)**.** *Rechecking the proof, it is clear that this connection between GANs and OT goes beyond the framework of WGANs. Indeed, take any Polish space $\mathcal{X}$ with metric l, then $\mathcal{X} \times \mathcal{X}$ is also a Polish space with metric l'. Denote $\mathcal{P}(\mathcal{X})$ as the set of all probability distributions over the sample space $\mathcal{X}$. Define a generic divergence function*

$$W : \mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \mapsto \mathbb{R}^+,$$

and take a class of GANs with this divergence W. If W can be written as the optimal cost W_c as in (OT)*, and if such an OT problem has a duality representation. Then GAN is an OT problem with the discriminator locating the best coupling among Π_G for* (OT-WGAN) *under a given G, and with the generator refining the set of possible couplings Π_G to minimize* (OT-WGAN).

Remark 4.2. *Note that there are earlier studies connecting GANs and OT, by different approaches and from different perspectives. [153] defines a novel divergence called the minibatch energy distance, based on solutions of three associated optimal transport problems. This new divergence is then used to replace the JS divergence for the vanilla GANs. Note that this minibatch energy distance itself is not an optimal transport cost. In [124], an interpretation of Wasserstein GANs (WGANs) from the perspective of optimal transport is provided: the latent random variable from the latent space is mapped to the sample space via an optimal mass transport so that the resulted distribution can minimize its Wasserstein distance against the true distribution. In our work, we provide sufficient conditions for which the minimax game of GANs, including WGANs, can be reformulated in the framework of OT.*

4.5 Experiments.

We now assess the quality of the proposed Algorithm 1, with a class of ergodic MFGs, for both one-dimension and high-dimension cases. This class of MFGs is chosen because of their explicit solution structures, which facilitate numerical comparison.

4.5.1 A class of ergodic MFGs.

Specifically, take (MFG) and consider the following long-run average cost,

$$\hat{J}_m(\alpha) = \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_t^T L(X_t^\alpha, \alpha(X_t^\alpha)) + f(X_t^\alpha, m(X_t^\alpha))dt\right], \tag{4.8}$$

where the cost of control and running cost are given by

$$L(x,\alpha) = \frac{1}{2}|\alpha|^2 + \tilde{f}(x), \quad f(x,m) = \ln(m), \quad \epsilon = \frac{1}{2},$$

with,

$$\tilde{f}(x) = 2\pi^2\left[-\sum_{i=1}^d \sin(2\pi x_i) + \sum_{i=1}^d |\cos(2\pi x_i)|^2\right] - 2\sum_{i=1}^d \sin(2\pi x_i).$$

Then the PDE system (HJB)–(FP) becomes

$$\begin{cases} -\epsilon\Delta u + H_0(x,\nabla u) = f(x,m) + \bar{H}, \\ -\epsilon\Delta m - div\,(m\nabla_p H_0(x,\nabla u)) = 0, \\ \int_{\mathbb{T}^d} u(x)dx = 0;\ m > 0,\ \int_{\mathbb{T}^d} m(x)dx = 1, \end{cases} \tag{4.9}$$

where the convex conjugate H_0 is given by

$$H_0(x,p) = \sup_\alpha\{\alpha\cdot p - \frac{1}{2}|\alpha|^2\} - \tilde{f}(x).$$

Here, the periodic value function u, the periodic density function m, and the unknown $\bar{H}$ can be explicitly derived. Indeed, assuming the existence of a smooth solution $(m, u, \bar{H})$, m in the second equation in (4.9) can be written as

$$m(x) = \frac{e^{2u(x)}}{\int_{\mathbb{T}^d} e^{2u(x')} dx'}. \tag{4.10}$$

Hence the solution to (4.9) is given by

$$u(x) = \sum_{i=1}^{d} \sin(2\pi x_i)$$

and

$$\bar{H} = \ln \left(\int_{\mathbb{T}^d} e^{2 \sum_{i=1}^{d} \sin(2\pi x_i)} dx \right).$$

The optimal control policy is also explicitly given by

$$\begin{aligned} \alpha^* &= \arg\max_{\alpha} \{\nabla_x u \cdot \alpha - L(x, \alpha)\} \\ &= \nabla_x u = 2\pi \begin{pmatrix} \cos(2\pi x_1) & \dots & \cos(2\pi x_d) \end{pmatrix} \in \mathbb{R}^d. \end{aligned}$$

4.5.2 Experiment setup

We will compute the above MFGs in Section 4.5.1 by exploiting its GANs structure and by using Algorithm 1.

Implementation. Both the value function u and the density function m are computed via NN with parameters θ and ω respectively. Moreover,

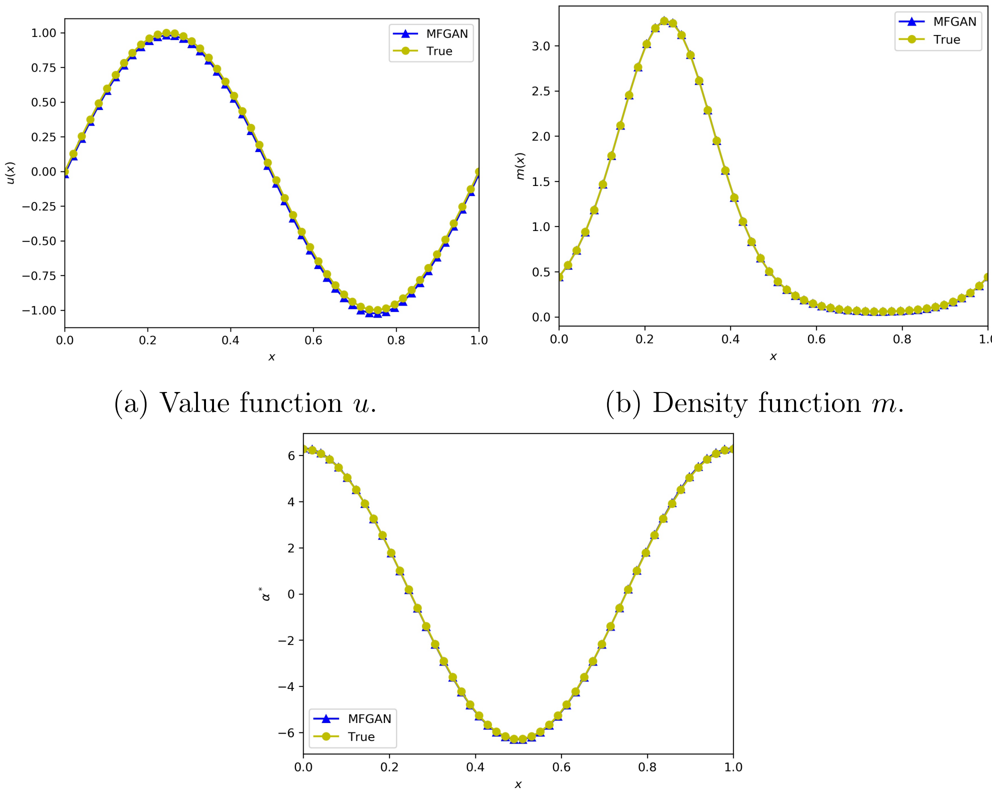
- The NN approximate m_ω is assumed to be a maximum entropy probability distribution, i.e., $m_\omega \propto \exp f_\omega$. This is due to the lack of information about the density function m. (See also [77] for the use of maximum entropy probability distribution).
- The network architecture for implementing both u_θ and f_ω adopts the Deep Galerkin Method (DGM), proposed in [160]. The DGM architecture is known to be useful for solving PDEs numerically. (See for instance [50]).

Adaptation. Since the MFG in Section 4.5.1 is of an ergodic type with a specified periodicity, Algorithm 1 is adapted accordingly. More precisely,

- To accommodate the periodicity given by the domain flat torus $\mathbb{T}^d$, for any data point $x_i = (x_{i,1}, \dots, x_{i,d}) \in \mathbb{R}^d$, we use

$$\begin{aligned} y_i = &(\sin(2\pi x_{i,1}), \dots, \sin(2\pi x_{i,d}), \\ &\cos(2\pi x_{i,1}), \dots, \cos(2\pi x_{i,d})) \end{aligned}$$

  as input. The $x_i's$ and $y_i's$ here are the latent variables in the vanilla GANs.



(a) Value function u.

(b) Density function m.

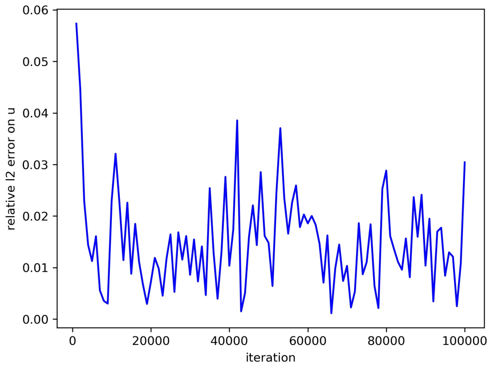
(c) Optimal control α^*.

Figure 4.1: One-dimensional test case.

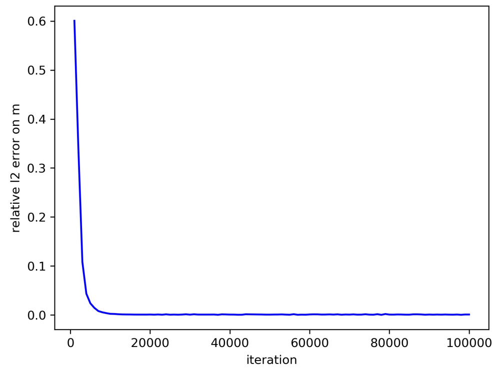
- An additional trainable variable $\bar{H}$ is introduced in the graphical model.
- The loss functions L_{HJB} and L_{FP} are modified according to the first and second equations of (4.9). The generator penalty becomes

$$\hat{L}_{term} = \left[\frac{\sum_{i=1}^{B_g} u_\theta(y^i)}{B_g} \right]^2 .$$

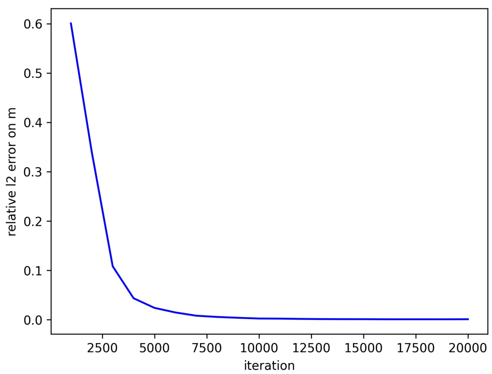
  Due to the structure m_ω, the discriminator penalty on m_ω being a probability density function can be ignored, i.e. $\beta_D = 0$.
- We train the generator first; and in each inner loop we take more SGD steps and with a larger learning rate compared with those in the discriminator. Note that this is opposite to the typical GANs training.



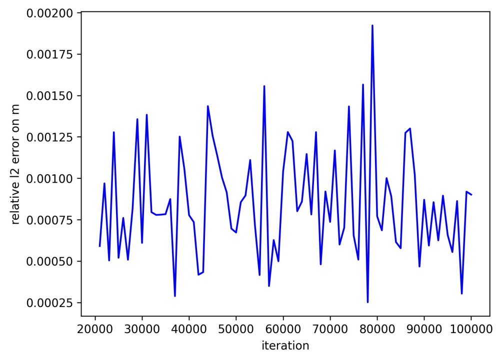
(a) Relative l_2 error of u.



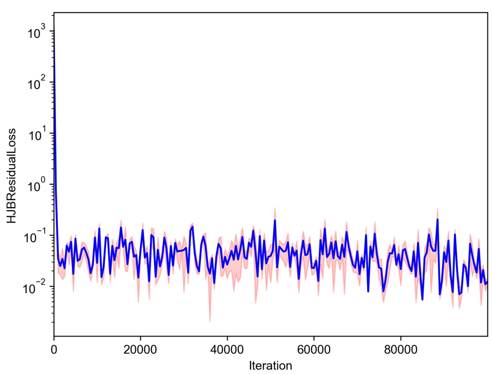
(b) Relative l_2 error of m.



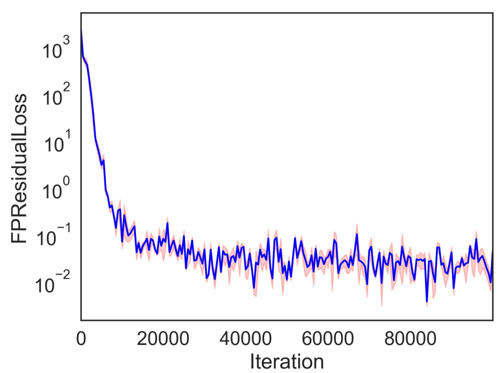
(c) Error of m first 20k iterations.



(d) Error of m after 20k iterations.



(e) HJB residual loss.

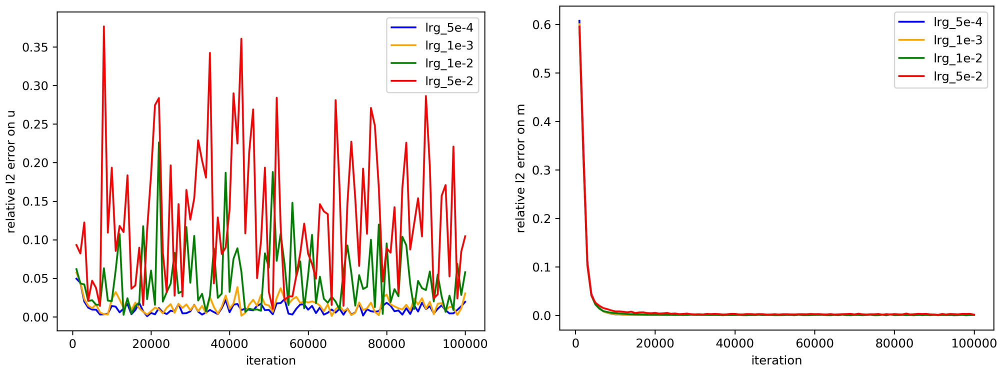


(f) FP residual loss.

Figure 4.2: Losses and errors in the one-dimensional test case.

Performance evaluations. To assess the performance of our algorithm, the following procedure is adopted.

- Given the explicit solution to the MFG (4.9), we compare the learned value function,



(a) Relative l_2 error of u.

(b) Relative l_2 error of m.

Figure 4.3: Impact of generator learning rate on relative l_2 error.

the learned density function and the learned optimal control against their perspective analytical form.

- We adopt the evolution of relative l_2 errors between the learnt and true value and density functions. The relative l_2 error of a function f against another function g, with $f, g : \mathbb{T}^d \to \mathbb{R}$ and g not constant 0, is given by

$$err_{rel-l_2}(f, g) = \sqrt{\frac{\int_{\mathbb{T}^d} [f(x) - g(x)]^2 dx}{\int_{\mathbb{T}^d} g(x)^2 dx}}.$$

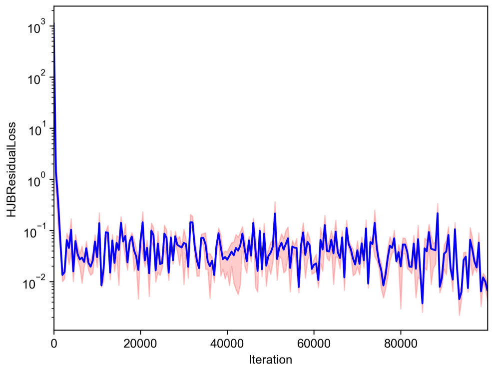
Moreover, to facilitate comparisons for broader classes of MFGs whose analytical solutions may not be available, additional loss functions are adopted. Here we take differential residuals of both the HJB and the FP equations as measurement of the performance.

4.5.3 Result of one-dimensional case.

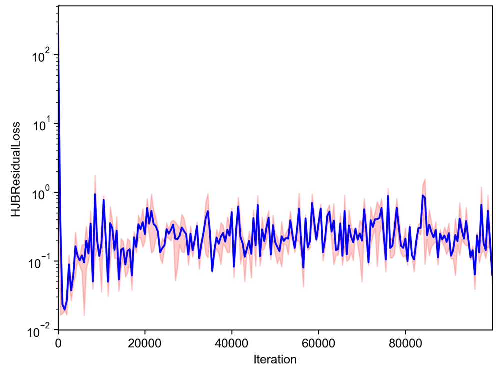
We first conduct numerical experiment with one-dimensional input.

- The DGM network for both u_θ and f_ω contains 1 hidden layer with 4 nodes. The activation function for u_θ is hyperbolic tangent function and that of f_ω is sigmoid function.

- Within each iteration of training, i.e., one complete outer loop, SGD is performed to update parameters of both the generator and the discriminator; the inputs of the SGD steps are mini-batches of size $B_g = B_d = 32$, where B_g and B_d denote the batch sizes for the generator and the discriminator updates, respectively.

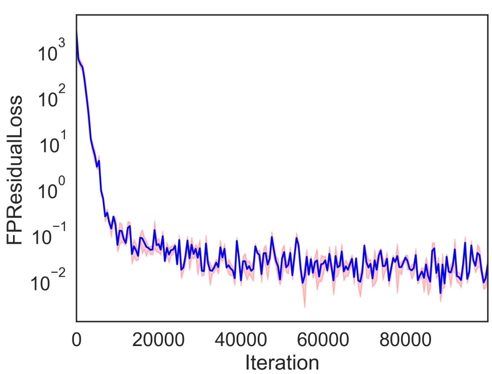


(a) $\alpha_g^0 = 5 \times 10^{-4}$.

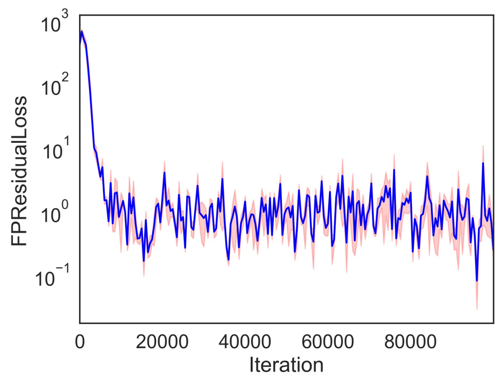


(b) $\alpha_g^0 = 5 \times 10^{-2}$.

Figure 4.4: HJB residual loss under different generator learning rate.



(a) $\alpha_g^0 = 5 \times 10^{-4}$.



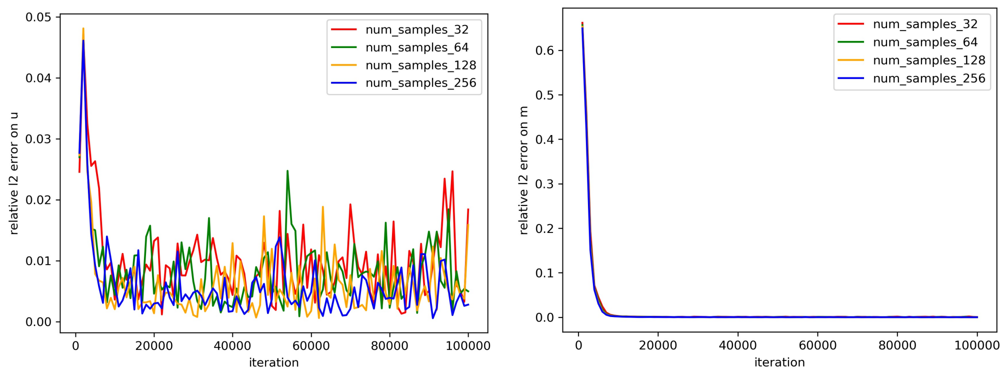
(b) $\alpha_g^0 = 5 \times 10^{-2}$.

Figure 4.5: FP residual loss under different generator learning rate.

- As mentioned in the adaptation, the number of SGD steps for the generator is $N_\theta = 5$ with initial learning rate 1×10^{-3} , whereas the number of SGD steps for the discriminator is $N_\omega = 2$ with initial learning rate 1×10^{-4}. The number of total iterations, i.e., the number of outer loops is $K = 10^5$. Adam optimizer is used for the updates.

- The weight for the generator penalty is $\beta_G = 1$.

The result is summarized in Figures 4.1 and 4.2. Figures 4.1a and 4.1b show the learnt functions of u and m against the true ones, respectively, and 4.1c shows the optimal control. Both show the accuracy of the learnt functions versus the true ones. This strong performance is supported by the plots of loss in Figures 4.2a and 4.2b, depicting the evolution of relative l_2 error as the number of outer iterations grows to K. Within 10^5 iterations, the relative l_2 error of u oscillates around 3×10^{-2}, and the relative l_2 errors of m decreases below 10^{-3}.



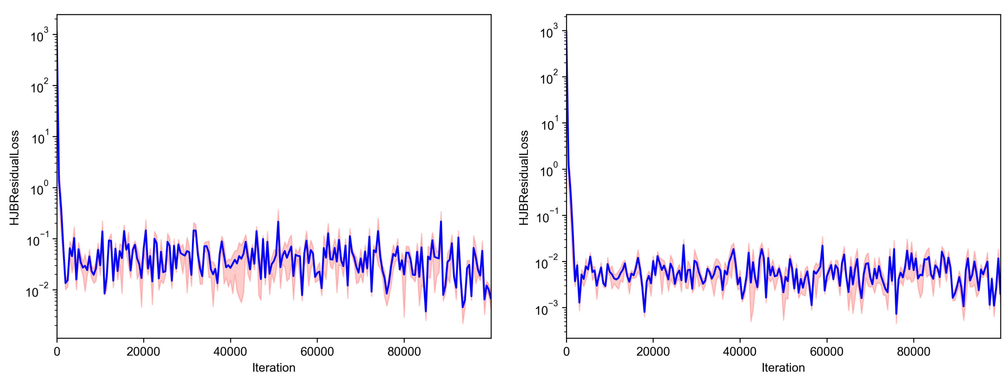
(a) Relative l_2 error of u.

(b) Relative l_2 error of m.

Figure 4.6: Impact of minibatch size on relative l_2 errors.

To facilitate comparisons for broader classes of MFGs whose analytical solutions are not necessarily available, we also take differential residuals of both the HJB and the FP equations to measure the performance. The evolution of the HJB and FP differential residual loss is shown in Figures 4.2e and 4.2f, respectively. In theses figures, the solid line is the average loss among 3 experiments, with standard deviation captured by the shadow around the line. Both differential residuals first rapidly descend to the magnitude of 10^{-2} and then the descent slows down accompanied by oscillation.

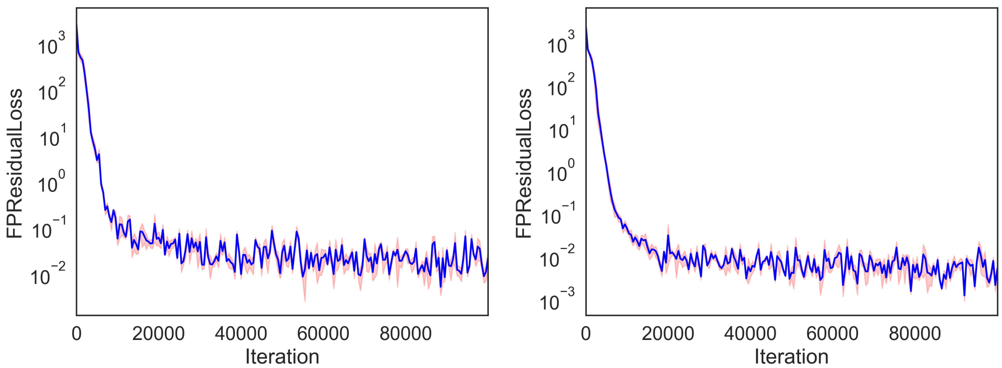
One may notice the difference between the training results of u and m. One reason is that u and m are implemented using different neural networks. The other is that different loss functions are adopted for training u and m.



(a) $B_g = B_d = 32$.

(b) $B_g = B_d = 256$.

Figure 4.7: HJB residual loss under different minibatch size.



(a) $B_g = B_d = 32$. (b) $B_g = B_d = 256$.

Figure 4.8: FP residual loss under different minibatch size.

Ablation study. To understand possible contributing factors for the oscillation in the loss, especially for u, an ablation study on the learning rate of the generator α_g is conducted. In our test, the initial learning rate for the Adam Optimizer α_g^0 takes the values of 5×10^{-4}, 1×10^{-3}, 1×10^{-2} and 5×10^{-2}, respectively.

From Figures 4.2a and 4.2b, the relative l_2 error on u oscillates more than that of m. Similar phenomenon is observed in Figure 4.3. In particular, from Figure 4.3a, a drastic decrease in oscillation can be seen as the generator learning rate α_g decreases.

Turning to the differential residual losses, one can observe from Figures 4.4 and 4.5 that, if decreasing α_g^0 from 5×10^{-2} to 5×10^{-4}, the residual losses for both HJB and FP decrease to a lower level with less oscillation.

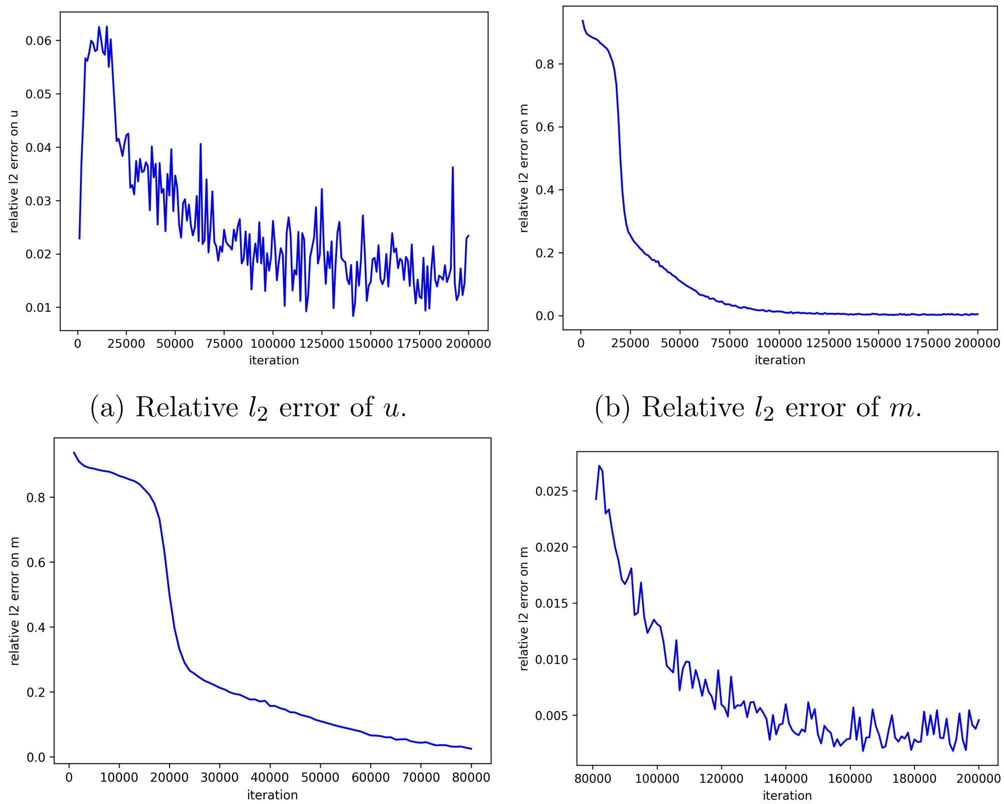
Another parameter of interest is the number of samples in each minibatch, i.e., B_g and B_d in Algorithm 1. Setting $B_g = B_d$, the cases of 32, 64, 128 and 256 are tested. Figure 4.6a shows that the relative l_2 error of u oscillates less as B_g and B_d increases from 32 to 256. Moreover, comparing the case of $B_g = B_d = 32$ and $B_g = B_d = 256$, the residual losses for both HJB and FP decrease to a lower level with less oscillation as minibatch size increases, as shown in Figures 4.7 and 4.8.

4.5.4 Results of multi-dimensional case.

We next test with input of dimension 4 and relative l_2 errors are shown in Figure 4.9.

- Just as in the one-dimensional case, the DGM network for both u_θ and f_ω contains 1 hidden layer with 4 nodes. The activation function for u_θ is hyperbolic tangent function and that of f_ω is sigmoid function.

- Within each iteration of training, i.e., one complete outer loop, SGD is performed to update parameters of both the generator and the discriminator; the inputs of the SGD



(a) Relative l_2 error of u.

(b) Relative l_2 error of m.

(c) Error of m first 80k iterations.

(d) Error of m after 80k iterations.

Figure 4.9: Input of dimension 4.

steps are mini-batches of size $B_g = B_d = 32$.

- The number of SGD steps for the generator is $N_\theta = 5$ with initial learning rate $\alpha_g = 5 \times 10^{-4}$, whereas the number of SGD steps for the discriminator is $N_\omega = 2$ with initial learning rate $\alpha_d = 1 \times 10^{-4}$. The number of outer loops is increased to $K = 2 \times 10^5$. Adam optimizer is used for the updates.

- The weight for the generator penalty is $\beta_2 = 1$.

Within 2×10^5 iterations, the relative l_2 error of u decreases below 2×10^{-2} and that of m decreases to 4×10^{-3}.

Finally, it is worth noting that similar experiment for dimension 4 has been conducted in [50]; see Test Case 4. In comparison, their algorithms need significantly larger number of iterations: 10^6 of iterations vs our 2×10^5 to achieve the same level of accuracy.

4.6 Conclusion and remarks

This chapter focuses on the connection of GANs and MFGs. In particular, the algorithm 1 is shown to be an effective computational method for MFGs. It is worth noticing that this algorithm is based on the HJB-FP approach for solving MFGs. The potential implementation of probabilistic approach such as FBSDEs via GANs, the development of GANs-based algorithms for MFGs with singular and impulse types of controls, as well as the application of Algorithm 1 on real-world problems such as price impact models and principal agent problems, will be exciting follow-up works. Also, as the popularity of big data in financial industry grows, we could formulate the impact of incorporating large amount of new data into prevailing models, methods and algorithms in the industry as an optimal transport problem and use GANs as computational tools to provide proper measurement and even feasible method of amendment.

Chapter 5

Approximation and Convergence of GANs Training: An SDE Approach

As introduced in Section 1.2, generative adversarial networks (GANs) [80] are generative models between two competing neural networks: a generator network G and a discriminator network D. The generator network G attempts to fool the discriminator network by converting random noise into sample data, while the discriminator network D tries to identify whether the input sample is faked or true. Since its introduction to the machine learning community, the popularity of GANs has grown exponentially with a wide range of applications.

Despite the empirical success of GANs, there are well recognized issues in GANs training, such as the vanishing gradient when the discriminator significantly outperforms the generator [5], the mode collapse which is believed to be linked with gradient exploding [152], and the challenge of GANs convergence [12].

In response to these issues, there has been a growing research interest in the theoretical understanding of GANs training. [24] proposed a novel visualization method for the GANs training process through the gradient vector field of loss functions. In a deterministic GANs training framework, [132] demonstrated that regularization improved the convergence performance of GANs. [57] and [68] analyzed a generic zero-sum minimax game including that of GANs, and connected the mixed Nash equilibrium of the game with the invariant measure of Langevin dynamics. In addition, various approaches have been proposed for amelioration, including different choices of network architectures, loss functions, and regularization. See for instance, a comprehensive survey on these techniques [174] and the references therein.

In this chapter, we will establish approximations for the training of GANs under stochastic gradient algorithms (SGAs), with precise error bound analysis. The approximations are in the form of coupled stochastic differential equations (SDEs). It then demonstrates the convergence of GANs training via invariant measures of SDEs under proper conditions. This work builds theoretical foundation for GANs training and provides analytical tools to study its evolution and stability. In particular,

a) the SDE approximations characterize precisely the distinction between GANs with alternating update and GANs with simultaneous update, in terms of the interaction between the generator and the discriminator; the error bound analysis for the SDEs supports the claim that GANs with alternating update converges faster and are more stable than GANs with simultaneous update;

b) the drift terms in the SDEs show the direction of the parameters evolution; the diffusion terms prescribes the ratio between the batch size and the learning rate in order to modulate the fluctuations of SGAs in GANs training;

c) regularity conditions for the coefficients of the SDEs provide constraints on the growth of the loss function with respect to the model parameters, necessary for avoiding the explosive gradient encountered in the training of GANs; they also explain mathematically some well known heuristics in GANs training, and confirm the importance of appropriate choices for network depth and of the introduction of gradient clipping and gradient penalty;

d) the dissipative property of the training dynamics in the form of SDE ensures the existence of the invariant measures, hence the convergence of GANs training; it underpins the practical tactic of adding regularization term to the GANs objective to improve the stability of training;

e) the invariant measures for the SDEs give rise to the dynamics of training loss and the fluctuation-dissipation relations (FDRs) for GANs. These FDRs reveal the trade-off of the loss landscape between the generator and the discriminator and can be used to schedule the learning rate.

Our analysis on the approximation and the convergence of GANs training is inspired by [125] and [127]. The former established the SDE approximation for the parameter evolution in SGAs applied to pure minimization problems (see also [101] on the similar topic); the latter surveyed theoretical analysis of deep learning from two perspectives: propagation of chaos through neural networks and training process of deep learning algorithms.

Throughout this chapter, the following notations will be adopted.

- The transpose of a vector $x \in \mathbb{R}^d$ is denoted by x^T and the transpose of a matrix $A \in \mathbb{R}^{d_1 \times d_2}$ is denoted by A^T.
- The set of k continuously differentiable functions over some domain $\mathcal{X} \subset \mathbb{R}^d$ is denoted by $\mathcal{C}^k(\mathcal{X})$ for $k = 0, 1, 2, \ldots$; in particular when $k = 0$, $\mathcal{C}^0(\mathcal{X}) = \mathcal{C}(\mathcal{X})$ denotes the set of continuous functions.
- Let $p \geq 1$. $L^p_{loc}(\mathbb{R}^d)$ denotes the set of functions f defined on $\mathbb{R}^d$ such that for any compact subset $\mathcal{X}$, $\int_{\mathcal{X}} \|f(x)\|_p^p dx < \infty$.

- Let $J = (J_1, \dots, J_d)$ be a d-tuple multi-index of order $|J| = \sum_{i=1}^d J_i$. For a function $f \in L^1_{loc}(\mathbb{R}^d)$, its J^{th}-weak derivative $D^J f \in L^1_{loc}(\mathbb{R}^d)$ is a function such that for any smooth and compactly supported test function g,

$$\int_{\mathbb{R}^d} D^J f(x) g(x) dx = (-1)^{|J|} \int_{\mathbb{R}^d} f(x) \nabla^J g(x) dx.$$

- The Sobolev space $W^{k,p}_{loc}(\mathbb{R}^d)$ is a set of functions f on $\mathbb{R}^d$ such that for any d-tuple multi-index J with $|J| \le k$, $D^J f \in L^p_{loc}(\mathbb{R}^d)$.

5.1 GANs training

GANs fall into the category of generative models to approximate an unknown probability distribution $\mathbb{P}_r$. GANs are minimax games between two competing neural networks, the generator G and the discriminator D. The neural network for the generator G maps a latent random variable Z with a known distribution $\mathbb{P}_z$ into the sample space to mimic the true distribution $\mathbb{P}_r$. Meanwhile, the other neural network for the discriminator D will assign a score between 0 to 1 to the generated sample. A higher score from the discriminator D indicates that the sample is more likely to be from the true distribution. GANs are trained by optimizing G and D iteratively until D can no longer distinguish between true samples and generated samples.

GANs training is performed on a data set $\mathcal{D} = \{(z_i, x_j)\}_{1 \le i \le N, 1 \le j \le M}$, where $\{z_i\}_{i=1}^N$ are sampled from $\mathbb{P}_z$ and $\{x_j\}_{j=1}^M$ are real image data following the unknown distribution $\mathbb{P}_r$. Let G_θ denote the generator parametrized by the neural network with the set of parameters $\theta \in \mathbb{R}^{d_\theta}$, and let D_ω denote the discriminator parametrized by the other neural network with the set of parameters $\omega \in \mathbb{R}^{d_\omega}$. Then the objective of GANs is to solve the following minimax problem

$$\min_\theta \max_\omega \Phi(\theta, \omega), \tag{5.1}$$

for some cost function Φ, with Φ of the form

$$\Phi(\theta, \omega) = \frac{\sum_{i=1}^N \sum_{j=1}^M J(D_\omega(x_j), D_\omega(G_\theta(z_i)))}{N \cdot M}. \tag{5.2}$$

For instance, Φ in the vanilla GANs model [80] is given by

$$\Phi(\theta, \omega) = \frac{\sum_{i=1}^N \sum_{j=1}^M \log D_\omega(x_j) + \log(1 - D_\omega(G_\theta(z_i)))}{N \cdot M},$$

while Φ in Wasserstein GANs [6] takes the form

$$\Phi(\theta, \omega) = \frac{\sum_{i=1}^N \sum_{j=1}^M D_\omega(x_j) - D_\omega(G_\theta(z_i))}{N \cdot M}.$$

In practice, stochastic gradient algorithm (SGA) is performed in order to solve the minimax problem (5.1), where the full gradients of Φ with respect to θ and ω are estimated over a mini-batch $\mathcal{B}$ of batch size B. One way of sampling $\mathcal{B}$ is to choose B samples out of a total of $N \cdot M$ samples without putting back, another is to take B i.i.d. samples. The analyses for both cases are similar, here we adopt the second sampling scheme. More precisely, let $\mathcal{B} = \{(z_{I_k}, x_{J_k})\}_{k=1}^B$ be i.i.d. samples from $\mathcal{D}$. Let g_θ and g_ω be the full gradients of Φ with respect to θ and ω such that

$$\begin{aligned} g_\theta(\theta,\omega) &= \nabla_\theta \Phi(\theta,\omega) = \frac{\sum_{i=1}^N \sum_{j=1}^M g_\theta^{i,j}(\theta,\omega)}{N \cdot M}, \\ g_\omega(\theta,\omega) &= \nabla_\omega \Phi(\theta,\omega) = \frac{\sum_{i=1}^N \sum_{j=1}^M g_\omega^{i,j}(\theta,\omega)}{N \cdot M}. \end{aligned} \tag{5.3}$$

Here $g_\theta^{i,j}$ and $g_\omega^{i,j}$ denote $\nabla_\theta J(D_\omega(x_j), D_\omega(G_\theta(z_i)))$ and $\nabla_\omega J(D_\omega(x_j), D_\omega(G_\theta(z_i)))$, respectively, with differential operators defined as $\nabla_\theta := \begin{pmatrix} \partial_{\theta_1} & \cdots & \partial_{\theta_{d_\theta}} \end{pmatrix}^T$ and $\nabla_\omega := \begin{pmatrix} \partial_{\omega_1} & \cdots & \partial_{\omega_{d_\omega}} \end{pmatrix}^T$. Then, the estimated gradients for g_θ and g_ω corresponding to the mini-batch $\mathcal{B}$ are

$$g_\theta^{\mathcal{B}}(\theta,\omega) = \frac{\sum_{k=1}^B g_\theta^{I_k,J_k}(\theta,\omega)}{B}, \quad g_\omega^{\mathcal{B}}(\theta,\omega) = \frac{\sum_{k=1}^B g_\omega^{I_k,J_k}(\theta,\omega)}{B}. \tag{5.4}$$

Moreover, let $\eta_t^\theta > 0$ and $\eta_t^\omega > 0$ be the learning rates at iteration $t = 0, 1, 2, \ldots$, for θ and ω respectively, then solving the minimax problem (5.1) with SGA and *alternating parameter update* implies descent of θ along g_θ and ascent of ω along g_ω at each iteration, i.e.,

$$\begin{cases} \omega_{t+1} = \omega_t + \eta_t^\omega g_\omega^{\mathcal{B}}(\theta_t, \omega_t), \\ \theta_{t+1} = \theta_t - \eta_t^\theta g_\theta^{\mathcal{B}}(\theta_t, \omega_{t+1}). \end{cases} \tag{5.5}$$

Furthermore, within each iteration, the minibatch gradient for θ and ω are calculated on different batches. In order to emphasize this difference, we use $\bar{\mathcal{B}}$ to represent the minibatch for θ and $\mathcal{B}$ for that of ω, with $\bar{\mathcal{B}} \overset{i.i.d.}{\sim} \mathcal{B}$. That is,

$$\begin{cases} \omega_{t+1} = \omega_t + \eta_t^\omega g_\omega^{\mathcal{B}}(\theta_t, \omega_t), \\ \theta_{t+1} = \theta_t - \eta_t^\theta g_\theta^{\bar{\mathcal{B}}}(\theta_t, \omega_{t+1}). \end{cases} \tag{ALT}$$

Some practical training of GANs uses *simultaneous parameter update* between the discriminator and the generator, corresponding to a similar yet subtly different form

$$\begin{cases} \omega_{t+1} = \omega_t + \eta_t^\omega g_\omega^{\mathcal{B}}(\theta_t, \omega_t), \\ \theta_{t+1} = \theta_t - \eta_t^\theta g_\theta^{\mathcal{B}}(\theta_t, \omega_t). \end{cases} \tag{SML}$$

For the ease of exposition, we will assume throughout the chapter, an constant learning rates $\eta_t^\theta = \eta_t^\omega = \eta$, with η viewed as the time interval between two consecutive parameter updates.

5.2 Approximation and error bound analysis of GANs training

In this section, we will establish continuous time approximations and error bounds for the GANs training process prescribed by (ALT) and (SML). The approximations are in the form of coupled SDEs.

To get an intuition of how the form of SDEs emerges, let us start by some basic properties embedded in the training process. First, let I and J denote the indices independently and uniformly drawn from $\{1,\dots,N\}$ and $\{1,\dots,M\}$, respectively, then

$$\mathbb{E}[g_\theta^{I,J}(\theta,\omega)] = g_\theta(\theta,\omega), \quad \mathbb{E}[g_\omega^{I,J}(\theta,\omega)] = g_\omega(\theta,\omega).$$

Denote the correspondence covariance matrices as

$$\Sigma_\theta(\theta,\omega) = \frac{\sum_i\sum_j[g_\theta^{i,j}(\theta,\omega)-g_\theta(\theta,\omega)][g_\theta^{i,j}(\theta,\omega)-g_\theta(\theta,\omega)]^T}{N\cdot M},$$
$$\Sigma_\omega(\theta,\omega) = \frac{\sum_i\sum_j[g_\omega^{i,j}(\theta,\omega)-g_\omega(\theta,\omega)][g_\omega^{i,j}(\theta,\omega)-g_\omega(\theta,\omega)]^T}{N\cdot M},$$

then as the batch size B gets sufficiently large, the classical central limit theorem leads to

$$\mathbb{E}_\mathcal{B}[g_\theta^\mathcal{B}(\theta,\omega)] = \mathbb{E}\left[\frac{\sum_{k=1}^B g_\theta^{I_k,J_k}(\theta,\omega)}{B}\right] = g_\theta(\theta,\omega),$$
$$\mathbb{E}_\mathcal{B}[g_\omega^\mathcal{B}(\theta,\omega)] = \mathbb{E}\left[\frac{\sum_{k=1}^B g_\omega^{I_k,J_k}(\theta,\omega)}{B}\right] = g_\omega(\theta,\omega),$$
$$Var_\mathcal{B}(g_\theta^\mathcal{B}(\theta,\omega)) = Var_\mathcal{B}\left(\frac{\sum_{k=1}^B g_\theta^{I_k,J_k}(\theta,\omega)}{B}\right) = \frac{1}{B}\Sigma_\theta(\theta,\omega),$$
$$Var_\mathcal{B}(g_\omega^\mathcal{B}(\theta,\omega)) = Var_\mathcal{B}\left(\frac{\sum_{k=1}^B g_\omega^{I_k,J_k}(\theta,\omega)}{B}\right) = \frac{1}{B}\Sigma_\omega(\theta,\omega),$$

as well as the following approximation of (ALT),

$$\begin{cases}\omega_{t+1} = \omega_t + \eta g_\omega^\mathcal{B}(\theta_t,\omega_t) \approx \omega_t + \eta g_\omega(\theta_t,\omega_t) + \dfrac{\eta}{\sqrt{B}}\Sigma_\omega^{\frac{1}{2}}(\theta_t,\omega_t)Z_t^1, \\ \theta_{t+1} = \theta_t - \eta g_\theta^\mathcal{B}(\theta_t,\omega_{t+1}) \approx \theta_t - \eta g_\theta(\theta_t,\omega_{t+1}) + \dfrac{\eta}{\sqrt{B}}\Sigma_\theta^{\frac{1}{2}}(\theta_t,\omega_{t+1})Z_t^2, \end{cases} \tag{5.6}$$

with independent random variables $Z_t^1 \sim N(0, I_{d_\omega})$ and $Z_t^2 \sim N(0, I_{d_\theta})$, $t = 0,1,2,\dots$.

If ignoring the difference between t and $t+1$, then the approximation could be written in the following form

$$d\begin{pmatrix}\Theta_t \\ \mathcal{W}_t\end{pmatrix} = \begin{pmatrix}-g_\theta(\Theta_t,\mathcal{W}_t) \\ g_\omega(\Theta_t,\mathcal{W}_t)\end{pmatrix}dt + \sqrt{2\beta^{-1}}\begin{pmatrix}\Sigma_\theta(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}} & 0 \\ 0 & \Sigma_\omega(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}}\end{pmatrix}dW_t, \tag{5.7}$$

with $\beta = \frac{2B}{\eta}$ and $\{W_t\}_{t\geq 0}$ be standard $(d_\theta + d_\omega)$-dimensional Brownian motion. This would be the approximation for GANs training of (SML).

If emphasizing the difference between t and $t+1$ thus the interaction between the generator and the discriminator, then the precise approximation for the GANs training process of (ALT) should be

$$d\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix} = \left[\begin{pmatrix}-g_\theta(\Theta_t,\mathcal{W}_t)\\ g_\omega(\Theta_t,\mathcal{W}_t)\end{pmatrix} + \frac{\eta}{2}\begin{pmatrix}\nabla_\theta g_\theta(\Theta_t,\mathcal{W}_t) & -\nabla_\omega g_\theta(\Theta_t,\mathcal{W}_t)\\ -\nabla_\theta g_\omega(\Theta_t,\mathcal{W}_t) & -\nabla_\omega g_\omega(\Theta_t,\mathcal{W}_t)\end{pmatrix}\begin{pmatrix}-g_\theta(\Theta_t,\mathcal{W}_t)\\ g_\omega(\Theta_t,\mathcal{W}_t)\end{pmatrix}\right]dt$$
$$+\sqrt{2\beta^{-1}}\begin{pmatrix}\Sigma_\theta(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}} & 0\\ 0 & \Sigma_\omega(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}}\end{pmatrix}dW_t. \tag{5.8}$$

Equations (5.7) and (5.8) can be written in more compact forms

$$d\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix} = b_0(\Theta_t,\mathcal{W}_t)dt + \sigma(\Theta_t,\mathcal{W}_t)dW_t, \tag{SML-SDE}$$
$$d\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix} = b(\Theta_t,\mathcal{W}_t)dt + \sigma(\Theta_t,\mathcal{W}_t)dW_t. \tag{ALT-SDE}$$

where $b(\theta,\omega) = b_0(\theta,\omega) + \eta b_1(\theta,\omega)$, with

$$b_0(\theta,\omega) = \begin{pmatrix}-g_\theta(\theta,\omega)\\ g_\omega(\theta,\omega)\end{pmatrix}, \tag{5.9}$$
$$b_1(\theta,\omega) = \frac{1}{2}\begin{pmatrix}\nabla_\theta g_\theta(\theta,\omega) & -\nabla_\omega g_\theta(\theta,\omega)\\ -\nabla_\theta g_\omega(\theta,\omega) & -\nabla_\omega g_\omega(\theta,\omega)\end{pmatrix}\begin{pmatrix}-g_\theta(\theta,\omega)\\ g_\omega(\theta,\omega)\end{pmatrix}$$
$$= -\frac{1}{2}\nabla b_0(\theta,\omega)b_0(\theta,\omega) - \begin{pmatrix}\nabla_\omega g_\theta(\theta,\omega)g_\omega(\theta,\omega)\\ 0\end{pmatrix}, \tag{5.10}$$
$$\text{and } \sigma(\theta,\omega) = \sqrt{2\beta^{-1}}\begin{pmatrix}\Sigma_\theta(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}} & 0\\ 0 & \Sigma_\omega(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}}\end{pmatrix}. \tag{5.11}$$

Note the term $-\begin{pmatrix}\nabla_\omega g_\theta(\theta,\omega)g_\omega(\theta,\omega)\\ 0\end{pmatrix}$ for (ALT-SDE), which highlights the interaction between the generator and the discriminator in GANs training process.

We will show that these coupled SDEs are indeed the continuous time approximations of GANs training processes, with precise error bound analysis. Our error bound analysis helps to explain why GANs with alternating update tend to be more stable and converge faster than GANs with simultaneous update.

More precisely, we have the following theorems.

Theorem 5.1. *Fix an arbitrary time horizon $\mathcal{T} > 0$ and take the learning rate $\eta \in (0, 1\wedge\mathcal{T})$ and the number of iterations $N = \left\lfloor\frac{\mathcal{T}}{\eta}\right\rfloor$. Suppose that*

1. *$g_\omega^{i,j}$ is twice continuously differentiable, and $g_\theta^{i,j}$ and $g_\omega^{i,j}$ are Lipschitz, for any $i = 1,\ldots,N$ and $j = 1,\ldots,M$;*

2. *Φ is of $\mathcal{C}^3(\mathbb{R}^{d_\theta+d_\omega})$, $\Phi \in W^{4,1}_{loc}(\mathbb{R}^{d_\theta+d_\omega})$, and for any multi-index $J = (J_1, \ldots, J_{d_\theta+d_\omega})$ with $|J| = \sum_{i=1}^{d_\theta+d\omega} J_i \leq 4$, there exist $k_1, k_2 \in \mathbb{N}$ such that*
$$|D^J\Phi(\theta,\omega)| \leq k_1 \left(1 + \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2^{2k_2}\right)$$
for $\theta \in \mathbb{R}^{d_\theta}$, $\omega \in \mathbb{R}^{d_\omega}$ almost everywhere;
3. *$(\nabla_\theta g_\theta)g_\theta$, $(\nabla_\omega g_\theta)g_\omega$, $(\nabla_\theta g_\omega)g_\theta$ and $(\nabla_\omega g_\omega)g_\omega$ are all Lipschitz.*

Then, given any initialization $\theta_0 = \theta$ and $\omega_0 = \omega$, for any test function $f \in \mathcal{C}^3(\mathbb{R}^{d_\theta+d_\omega})$ such that for any multi-index J with $|J| \leq 3$ there exist $k_1, k_2 \in \mathbb{N}$ satisfying
$$|\nabla^J f(\theta,\omega)| \leq k_1 \left(1 + \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2^{2k_2}\right),$$
we have the following weak approximation,
$$\max_{t=1,\ldots,N} |\mathbb{E}f(\theta_t,\omega_t) - \mathbb{E}f(\Theta_{t\eta}, \mathcal{W}_{t\eta})| \leq C\eta^2 \tag{5.12}$$
for constant $C \geq 0$, where (θ_t,ω_t) and $(\Theta_{t\eta}, \mathcal{W}_{t\eta})$ are given by (ALT) *and* (ALT-SDE)*, respectively.*

Theorem 5.2. *Fix an arbitrary time horizon $\mathcal{T} > 0$, take the learning rate $\eta \in (0, 1 \wedge \mathcal{T})$ and the number of iterations $N = \left\lfloor \frac{\mathcal{T}}{\eta} \right\rfloor$. Suppose*

1. *$\Phi(\theta,\omega)$ is continuously differentiable, $\Phi \in W^{3,1}_{loc}(\mathbb{R}^{d_\theta+d_\omega})$ and for any multi-index $J = (J_1, \ldots, J_{d_\theta+d_\omega})$ with $|J| = \sum_{i=1}^{d_\theta+d\omega} J_i \leq 3$, there exist $k_1, k_2 \in \mathbb{N}$ such that $D^J\Phi$ satisfies*
$$|D^J\Phi(\theta,\omega)| \leq k_1 \left(1 + \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2^{2k_2}\right)$$
for $\theta \in \mathbb{R}^{d_\theta}$, $\omega \in \mathbb{R}^{d_\omega}$ almost everywhere;
2. *$g_\theta^{i,j}$ and $g_\omega^{i,j}$ are Lipschitz for any $i = 1, \ldots, N$ and $j = 1, \ldots, M$.*

Then, given any initialization $\theta_0 = \theta$ and $\omega_0 = \omega$, for any test function $f \in \mathcal{C}^2(\mathbb{R}^{d_\theta+d_\omega})$ such that for any multi-index J with $|J| \leq 2$ there exist $k_1, k_2 \in \mathbb{N}$ satisfying
$$|\nabla^J f(\theta,\omega)| \leq k_1 \left(1 + \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2^{2k_2}\right),$$
we have the following weak approximation,
$$\max_{t=1,\ldots,N} |\mathbb{E}f(\theta_t,\omega_t) - \mathbb{E}f(\Theta_{t\eta}, \mathcal{W}_{t\eta})| \leq C\eta \tag{5.13}$$
for constant $C \geq 0$, where (θ_t,ω_t) and $(\Theta_{t\eta}, \mathcal{W}_{t\eta})$ are given by (SML) *and* (SML-SDE)*, respectively.*

Detailed proofs of Theorems 5.1 and 5.2 will be deferred to the Section 5.5.1.

Implications for GANs. Approximations of GANs training by the SDEs (ALT-SDE) and (SML-SDE) enable analyzing the evolution of GANs parameters. For instance,

i. the difference between GANs with alternating update and GANs with simultaneous update can be seen in two aspects: first is the term $-\begin{pmatrix}\nabla_\omega g_\theta(\theta,\omega)g_\omega(\theta,\omega)\\ 0\end{pmatrix}$ for (ALT-SDE) which highlights the interaction between the generator and the discriminator; the second is the difference in the orders of error bounds between (5.12) and (5.13), which explains why in practice GANs with alternating update converges faster and are more stable than GANs with simultaneous update;

ii. the drift terms in the SDEs show the direction of the parameters evolution; the diffusion terms represent the fluctuations of the learning curves for these parameters; the form of SDEs prescribes the ratio between the batch size and the learning rate in order to modulate the fluctuations of SGAs in GANs training;

iii. the regularity conditions for the drift, the volatility, and the derivatives of loss function Φ, on one hand ensure mathematically the well-posedness of (ALT-SDE), on the other hand provide constraints on the growth of the loss function with respect to the model parameters, necessary for avoiding the explosive gradient encountered in the training of GANs; these regularity conditions explain mathematically some well known heuristics in GANs training, and confirm the importance of appropriate choices for network depth and of the introduction of gradient clipping and gradient penalty.

5.3 Convergence of GANs training via invariant measure of SDE

5.3.1 Convergence of GANs training

In addition to the evolution of parameters in GANs, the convergence of GANs training can be derived through these SDEs (ALT-SDE) and (SML-SDE). This is by analyzing the limiting behavior of SDEs, characterized by their invariant measures. Recall the following definition of invariant measures in [60].

Definition 5.3. *A probability measure $\mu^* \in \mathcal{P}(\mathbb{R}^{d_\theta+d_\omega})$ is called an invariant measure for a stochastic process $\left\{\begin{pmatrix}\Theta_t & \mathcal{W}_t\end{pmatrix}^T\right\}_{t\geq 0}$ if for any measurable bounded function f,*

$$\int \mathbb{E}\left[f(\Theta_t,\mathcal{W}_t)|\Theta_0=\theta,\mathcal{W}_0=\omega\right]\mu^*(d\theta,d\omega)=\int f(\theta,\omega)\mu^*(d\theta,d\omega).$$

Following [170], we have

Theorem 5.4. *Assume the following conditions hold for* (ALT-SDE).

1. *both b and σ are bounded and smooth and have bounded derivatives of any order;*
2. *there exist some positive real numbers r and M_0 such that for any $\begin{pmatrix}\theta & \omega\end{pmatrix}^T \in \mathbb{R}^{d_\theta+d_\omega}$,*

$$\begin{pmatrix}\theta & \omega\end{pmatrix} b(\theta,\omega) \leq -r \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2, \ \text{if} \ \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2 \geq M_0;$$

3. *$\mathcal{A}$ is uniformly elliptic, i.e., there exists $l > 0$ such that for any $\begin{pmatrix}\theta \\ \omega\end{pmatrix}, \begin{pmatrix}\theta' \\ \omega'\end{pmatrix} \in \mathbb{R}^{d_\theta+d_\omega}$,*

$$\begin{pmatrix}\theta' & \omega'\end{pmatrix}^T \sigma(\theta,\omega)\sigma(\theta,\omega)^T \begin{pmatrix}\theta' \\ \omega'\end{pmatrix} \geq l \left\| \begin{pmatrix}\theta' \\ \omega'\end{pmatrix} \right\|_2^2,$$

then (ALT-SDE) *admits a unique invariant measure μ^* with an exponential convergence rate. Similar results hold for the invariant measure of* (SML-SDE) *with b replaced by b_0.*

The proof of the Theorem 5.4 is deferred to the Section 5.5.2.

Implications for GANs. Condition 2 is a dissipative property of the training dynamics (ALT-SDE): the drift should drive the parameters towards a compact region. It ensures the existence of the invariant measure, hence the convergence of GANs training. This condition underpins the practical tactic of adding regularization term to the GANs objective to improve the stability of training.

5.3.2 Dynamics of training loss and FDR

In fact, one can further analyze the dynamics of the training loss based on the SDE approximation; and derive a fluctuation-dissipation relation (FDR) for the GANs training, given the existence of the invariant measure.

To see this, let $\mu = \{\mu_t\}_{t\geq 0}$ denote the flow of probability measures for $\left\{\begin{pmatrix}\Theta_t \\ \mathcal{W}_t\end{pmatrix}\right\}_{t\geq 0}$ given by (ALT-SDE).

Itô's formula to the smooth function Φ (see [151] for more details) gives the following dynamics of training loss,

$$\Phi(\Theta_t, \mathcal{W}_t) = \Phi(\Theta_s, \mathcal{W}_s) + \int_s^t \mathcal{A}\Phi(\Theta_r, \mathcal{W}_r)dr + \int_s^t \sigma(\Theta_r, \mathcal{W}_r)\nabla\Phi(\Theta_r, \mathcal{W}_r)dW_r; \tag{5.14}$$

where

$$\mathcal{A}f(\theta,\omega) = b(\theta,\omega)^T\nabla f(\theta,\omega) + \frac{1}{2}Tr\left(\sigma(\theta,\omega)\sigma(\theta,\omega)^T\nabla^2 f(\theta,\omega)\right), \tag{5.15}$$

is the infinitesimal generator for (ALT-SDE), given any test function $f : \mathbb{R}^{d_\theta+d_\omega} \to \mathbb{R}$. The existence of the unique invariant measure μ^* for (ALT-SDE) suggests that $\begin{pmatrix}\Theta_t \\ \mathcal{W}_t\end{pmatrix}$ in

(ALT-SDE) converges as $t \to \infty$ to some $\begin{pmatrix} \Theta^* \\ \mathcal{W}^* \end{pmatrix} \sim \mu^*$. By the Definition 5.3 of invariant measure, from (5.14) we have

$$\mathbb{E}_{\mu^*}[\mathcal{A}\Phi(\Theta^*, \mathcal{W}^*)] = 0.$$

By (5.15),

$$\begin{aligned}
\mathcal{A}\Phi(\theta,\omega) =& b_0(\theta,\omega)^T \nabla \Phi(\theta,\omega) + \eta b_1(\theta,\omega)^T \nabla\Phi(\theta,\omega) + \frac{1}{2} Tr(\sigma(\theta,\omega)\sigma(\theta,\omega)^T \nabla^2 \Phi(\theta,\omega)) \\
=& - \|\nabla_\theta \Phi(\theta,\omega)\|_2^2 + \|\nabla_\omega \Phi(\theta,\omega)\|_2^2 \\
& - \frac{\eta}{2}\Big[\nabla_\theta\Phi(\theta,\omega)^T \nabla_\theta^2 \Phi(\theta,\omega)\nabla_\theta \Phi(\theta,\omega) + \nabla_\omega \Phi(\theta,\omega)^T \nabla_\omega^2 \Phi(\theta,\omega)\nabla_\omega\Phi(\theta,\omega)\Big] \\
& + \beta^{-1} Tr\Big(\Sigma_\theta(\theta,\omega)\nabla_\theta^2\Phi(\theta,\omega) + \Sigma_\omega(\theta,\omega)\nabla_\omega^2 \Phi(\theta,\omega)\Big).
\end{aligned}$$

In other words, the evolution of loss function (5.14) leads to the following FRD for GANs training.

Theorem 5.5. *Assume the existence of an invariant measure μ^* for* (ALT-SDE)*, then*

$$\begin{aligned}
\mathbb{E}_{\mu^*}\Big[\|\nabla_\theta \Phi(\Theta^*,\mathcal{W}^*)\|_2^2 - \|\nabla_\omega\Phi(\Theta^*,\mathcal{W}^*)\|_2^2\Big] =& \beta^{-1}\mathbb{E}_{\mu^*}\Big[Tr\Big(\Sigma_\theta(\Theta^*,\mathcal{W}^*)\nabla_\theta^2\Phi(\Theta^*,\mathcal{W}^*) \\
+ \Sigma_\omega(\Theta^*,\mathcal{W}^*)\nabla_\omega^2\Phi(\Theta^*,\mathcal{W}^*)\Big)\Big] &- \frac{\eta}{2}\mathbb{E}_{\mu^*}\Big[\nabla_\theta\Phi(\Theta^*,\mathcal{W}^*)^T\nabla_\theta^2\Phi(\Theta^*,\mathcal{W}^*)\nabla_\theta\Phi(\Theta^*,\mathcal{W}^*) \\
+ \nabla_\omega\Phi(\Theta^*,\mathcal{W}^*)^T\nabla_\omega^2\Phi(\Theta^*,\mathcal{W}^*)\nabla_\omega\Phi(\Theta^*,\mathcal{W}^*)\Big]&.
\end{aligned} \tag{FDR1}$$

The corresponding FDR for the simultaneous update case of (SML-SDE) *is*

$$\begin{aligned}
&\mathbb{E}_{\mu^*}\Big[\|\nabla_\theta\Phi(\Theta^*,\mathcal{W}^*)\|_2^2 - \|\nabla_\omega\Phi(\Theta^*,\mathcal{W}^*)\|_2^2\Big] = \\
&\quad \beta^{-1}\mathbb{E}_{\mu^*}\Big[Tr\Big(\Sigma_\theta(\Theta^*,\mathcal{W}^*)\nabla_\theta^2\Phi(\Theta^*,\mathcal{W}^*) + \Sigma_\omega(\Theta^*,\mathcal{W}^*)\nabla_\omega^2\Phi(\Theta^*,\mathcal{W}^*)\Big)\Big].
\end{aligned}$$

Implications for GANs. This FDR relation for the minimax games in GANs connects the microscopic fluctuation caused by the noise of SGA with the macroscopic dissipation phenomena related to the loss function under a stationary status. The quantity $Tr(\Sigma_\theta\nabla_\theta^2\Phi + \Sigma_\omega\nabla_\omega^2\Phi)$ demonstrates the link between noise covariance matrices from SGAs and the loss landscape of Φ. It reveals the trade-off of the loss landscape between the generator and the discriminator. Note that this FDR relation is the counterpart of that for stochastic gradient descent (SGD) algorithm on a pure minimization problem in [178] and [127], which exposes the direct evaluation of the loss landscape such as gradient and Hessian.

Further analysis of the invariant measure can lead to a different type of FDR that will be practically useful for learning rate scheduling.

For example, applying Itô's formula to the squared norm of the parameters $\left\|\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix}\right\|_2^2$, we have the following dynamics

$$d\left\|\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix}\right\|_2^2 = 2\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix}^T d\begin{pmatrix}\Theta_t\\ \mathcal{W}_t\end{pmatrix} + Tr\Big(\sigma(\Theta_t,\mathcal{W}_t)\sigma(\Theta_t,\mathcal{W}_t)^T\Big)dt.$$

Theorem 5.6. *Assume the existence of an invariant measure μ^* for* (SML-SDE)*, then*

$$\mathbb{E}_{\mu^*}\Big[\Theta^{*,T}\nabla_\theta\Phi(\Theta^*,\mathcal{W}^*) - \mathcal{W}^{*,T}\nabla_\omega\Phi(\Theta^*,\mathcal{W}^*)\Big] = \beta^{-1}\mathbb{E}_{\mu^*}\Big[Tr(\Sigma_\theta(\Theta^*,\mathcal{W}^*) + \Sigma_\omega(\Theta^*,\mathcal{W}^*))\Big] \tag{FDR2}$$

Proofs of Theorems 5.5 and 5.6 will be deferred to the Section 5.5.3.

Implications for GANs. Notice that the quantities in (FDR2), including the parameters (θ,ω) and first-order derivatives of the loss function g_θ, g_ω, $g_\theta^{i,j}$ and $g_\omega^{i,j}$, are computationally inexpensive and can be evaluated on the fly. Therefore, instead of using a predetermined scheduling of learning rate such as Adam or RMSprop optimizer, one can customize the scheduling based on (FDR2).

For instance, recall that $g_\theta^{\mathcal{B}}$ and $g_\omega^{\mathcal{B}}$ are respectively unbiased estimators for g_θ and g_ω, and

$$\hat{\Sigma}_\theta(\theta,\omega) = \frac{\sum_{k=1}^{B}[g_\theta^{I_k,J_k}(\theta,\omega) - g_\theta^{\mathcal{B}}(\theta,\omega)][g_\theta^{I_k,J_k}(\theta,\omega) - g_\theta^{\mathcal{B}}(\theta,\omega)]^T}{B-1},$$
$$\hat{\Sigma}_\omega(\theta,\omega) = \frac{\sum_{k=1}^{B}[g_\omega^{I_k,J_k}(\theta,\omega) - g_\omega^{\mathcal{B}}(\theta,\omega)][g_\omega^{I_k,J_k}(\theta,\omega) - g_\omega^{\mathcal{B}}(\theta,\omega)]^T}{B-1}$$

are respectively unbiased estimators of $\Sigma_\theta(\theta,\omega)$ and $\Sigma_\omega(\theta,\omega)$; now to improve GANs training result with the simultaneous update, one can introduce two tunable parameters $\epsilon > 0$ and $\delta > 0$ to have the following scheduling:

$$\text{if } \left|\frac{\Theta^T g_\theta^{\mathcal{B}}(\Theta_t,\mathcal{W}_t) - \mathcal{W}_t^T g_\omega^{\mathcal{B}}(\Theta_t,\mathcal{W}_t)}{\beta^{-1}Tr(\hat{\Sigma}_\theta(\Theta_t,\mathcal{W}_t)+\hat{\Sigma}_\omega(\Theta_t,\mathcal{W}_t))} - 1\right| < \epsilon, \text{ then update } \eta \text{ by } (1-\delta)\eta.$$

5.4 Verifiability of the assumptions.

Theorems 5.1, 5.2, and 5.4 provide assumptions under which the convergence of GANs training could be established via the SDEs approximation. These assumptions are essentially assumptions on the gradients of the objective functions with respect to the parameters. These assumptions are easy to verify for many choices of GANs structures for a wide range of applications. We illustrate this via the example of WGANs for image processing:

1. **Smoothness and boundedness of drift and volatility.** Given that sample data in image processing problems are supported on compact domain, these assumptions are easily satisfied with proper prior distribution and activation function: first, the prior distribution $\mathbb{P}_z$ such as the uniform distribution is naturally compactly supported; next, take $D_\omega = \tanh(\omega \cdot x)$, $G_\theta(z) = \tanh(\theta \cdot z)$, and the objective function

$$\Phi(\theta, \omega) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} D_\omega(x_j) - D_\omega(G_\theta(z_i))}{N \cdot M}.$$

   Then the assumptions of Lipschitz continuity, differentiability and boundedness are guaranteed by boundedness of the data $\{(z_i, z_j)\}_{1 \le i \le N, 1 \le j \le M}$ and property of

$$\psi(y) = \tanh y = \frac{e^y - e^{-y}}{e^y + e^{-y}} = 1 - \frac{2}{e^{2y} + 1} \in (-1, 1).$$

   More precisely, the first and second order derivatives of ψ are

$$\psi'(y) = \frac{4}{(e^y + e^{-y})^2} \in (0, 1], \quad \psi''(y) = -8 \frac{e^y - e^{-y}}{(e^y + e^{-y})^3} = -2\psi(y)\psi'(y) \in (-2, 2).$$

   Any higher order derivatives can be written as functions of $\psi(\cdot)$ and $\psi'(\cdot)$ and therefore bounded.

2. **Dissipative property.** The dissipative property specified by second assumption in Theorem 5.4 essentially prevents the evolution of the parameters from being driven to infinity. In fact the weights clipping technique in WGANs, for instance, is consistent with this assumption.

3. **Elliptic condition**. The elliptic property of volatility term is trivially satisfied given its expression in (5.11).

5.5 Detailed Proofs

5.5.1 Proofs of Theorems 1 and 2

In this section we will provide a detailed proof of Theorem 1; proof of Theorem 2 is a simple analogy. Note that in this work, we establish the approximation of GANs training through SDEs with error bound analysis. We will tailor the methodology from [125] to our analysis of GANs training. Here we highlight the adaptation we make, which mostly concentrates on the preliminary analysis part.

Preliminary analysis

One-step difference. Recall that under the alternating update scheme and constant learning rate η, the GANs training is as follows,

$$\begin{cases} \omega_{t+1} = \omega_t + \eta g_\omega^{\mathcal{B}}(\theta_t, \omega_t), \\ \theta_{t+1} = \theta_t - \eta g_\theta^{\bar{\mathcal{B}}}(\theta_t, \omega_{t+1}), \end{cases} \tag{ALT}$$

where $\mathcal{B}$ and $\bar{\mathcal{B}}$ are i.i.d., emphasizing the fact that the evalutions of gradients are performed on different mini-batches when updating θ and ω alternatively.

Let (θ, ω) denote the initial value for (θ_0, ω_0) and

$$\Delta = \Delta(\theta, \omega) = \begin{pmatrix} \theta_1 - \theta \\ \omega_1 - \omega \end{pmatrix} \tag{5.16}$$

be the one-step difference. Let $\Delta^{i,j}$ denote the tuple consisting of the i-th and j-th component of one-step difference of θ and ω, respectively, with $i = 1, \ldots, d_\theta$ and $j = 1, \ldots, d_\omega$.

Lemma 5.7. *Assume that $g_\theta^{i,j}$ is twice continuously differentiable for any $i = 1, \ldots, N$ and $j = 1, \ldots, M$.*

1. *The first moment is given by*

$$\mathbb{E}[\Delta^{i,j}] = \eta \begin{pmatrix} -g_\theta(\theta, \omega)_i \\ g_\omega(\theta, \omega)_j \end{pmatrix} + \eta^2 \begin{pmatrix} \{-\nabla_\omega [g_\theta(\theta, \omega)_i]\}^T g_\omega(\theta, \omega) \\ 0 \end{pmatrix} + O(\eta^3).$$

2. *The second moment is given by*

$$\begin{aligned}\mathbb{E}[\Delta^{i,j}(\Delta^{k,l})^T] =& \eta^2 \left[\frac{1}{B} \begin{pmatrix} \Sigma_\theta(\theta, \omega)_{i,k} & 0 \\ 0 & \Sigma_\omega(\theta, \omega)_{j,l} \end{pmatrix} + \begin{pmatrix} -g_\theta(\theta, \omega)_i \\ g_\omega(\theta, \omega)_j \end{pmatrix} \begin{pmatrix} -g_\theta(\theta, \omega)_k \\ g_\omega(\theta, \omega)_l \end{pmatrix}^T \right] \\ &+ O(\eta^3),\end{aligned}$$

where $\Sigma_\theta(\theta, \omega)_{i,k}$ and $\Sigma_\omega(\theta, \omega)_{j,l}$ denote the element at position (i, k) and (j, l) of matrices $\Sigma_\theta(\theta, \omega)$ and $\Sigma_\omega(\theta, \omega)$, respectively.

3. *The third moments are all of order $O(\eta^3)$.*

Proof. By a second-order Taylor expansion, we have

$$\Delta(\theta, \omega) = \eta \begin{pmatrix} -g_\theta^{\bar{\mathcal{B}}}(\theta, \omega) \\ g_\omega^{\mathcal{B}}(\theta, \omega) \end{pmatrix} + \eta^2 \begin{pmatrix} -\nabla_\omega g_\theta^{\bar{\mathcal{B}}}(\theta, \omega)) g_\omega^{\mathcal{B}}(\theta, \omega) \\ 0 \end{pmatrix} + O(\eta^3). \tag{5.17}$$

Then,

$$\Delta^{i,j}(\theta,\omega) = \eta \begin{pmatrix} -g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_i \\ g_\omega^{\mathcal{B}}(\theta,\omega)_j \end{pmatrix} + \eta^2 \begin{pmatrix} \left\{-\nabla_\omega[g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_i]\right\}^T g_\omega^{\mathcal{B}}(\theta,\omega) \\ 0 \end{pmatrix} + O(\eta^3), \tag{5.18}$$

$$\Delta^{i,j}(\theta,\omega)[\Delta^{k,l}(\theta,\omega)]^T = \eta^2 \begin{pmatrix} g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_i g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_k & -g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_i g_\omega^{\mathcal{B}}(\theta,\omega)_l \\ -g_\theta^{\bar{\mathcal{B}}}(\theta,\omega)_k g_\omega^{\mathcal{B}}(\theta,\omega)_j & g_\omega^{\mathcal{B}}(\theta,\omega)_j g_\omega^{\mathcal{B}}(\theta,\omega)_l \end{pmatrix} + O(\eta^3), \tag{5.19}$$

and higher order polynomials are of order $O(\eta^3)$. Notice that $\bar{\mathcal{B}} \perp \mathcal{B}$ and recall the definition of Σ_θ and Σ_ω. The conclusion follows. □

Now consider the following SDE,

$$d\begin{pmatrix} \Theta_t \\ \mathcal{W}_t \end{pmatrix} = b(\Theta_t,\mathcal{W}_t)dt + \sigma(\Theta_t,\mathcal{W}_t)dW_t, \tag{ALT-SDE}$$

where $b(\theta,\omega) = b_0(\theta,\omega) + \eta b_1(\theta,\omega)$, with

$$b_0(\theta,\omega) = \begin{pmatrix} -g_\theta(\theta,\omega) \\ g_\omega(\theta,\omega) \end{pmatrix}, \tag{5.20}$$

$$\begin{aligned} b_1(\theta,\omega) &= \frac{1}{2}\begin{pmatrix} \nabla_\theta g_\theta(\theta,\omega) & -\nabla_\omega g_\theta(\theta,\omega) \\ -\nabla_\theta g_\omega(\theta,\omega) & -\nabla_\omega g_\omega(\theta,\omega) \end{pmatrix}\begin{pmatrix} -g_\theta(\theta,\omega) \\ g_\omega(\theta,\omega) \end{pmatrix} \\ &= -\frac{1}{2}\nabla b_0(\theta,\omega)b_0(\theta,\omega) - \begin{pmatrix} \nabla_\omega g_\theta(\theta,\omega)g_\omega(\theta,\omega) \\ 0 \end{pmatrix}, \end{aligned} \tag{5.21}$$

$$\text{and } \sigma(\theta,\omega) = \sqrt{2\beta^{-1}}\begin{pmatrix} \Sigma_\theta(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}} & 0 \\ 0 & \Sigma_\omega(\Theta_t,\mathcal{W}_t)^{\frac{1}{2}} \end{pmatrix}. \tag{5.22}$$

With the same initialization like (5.16), define the corresponding one-step difference for (ALT-SDE),

$$\tilde{\Delta} = \tilde{\Delta}(\theta,\omega) = \begin{pmatrix} \Theta_{1\times\eta} - \theta \\ \mathcal{W}_{1\times\eta} - \omega \end{pmatrix}. \tag{5.23}$$

Let $\tilde{\Delta}_k$ be the k-th component of $\tilde{\Delta}$, $k = 1,\ldots,d_\theta + d_\omega$ and $\tilde{\Delta}^{i,j}$ be the tuple consisting of the i-th and j-th component of one-step difference of Θ and $\mathcal{W}$, respectively, with $i = 1,\ldots,d_\theta$ and $j = 1,\ldots,d_\omega$.

Lemma 5.8. *Suppose b_0, b_1 and σ, given by* (5.20),(5.21) *and* (5.22)*, are from $\mathcal{C}^3(\mathbb{R}^{d_\theta+d_\omega})$ such that for any multi-index J of order $|J| \le 3$, there exist $k_1, k_2 \in \mathbb{N}$ satisfying*

$$\max\{|\nabla^J b_0(\theta,\omega)|, |\nabla^J b_1(\theta,\omega)|, |\nabla^J \sigma(\theta,\omega)|\} \le k_1\left(1 + \left\|\begin{pmatrix} \theta \\ \omega \end{pmatrix}\right\|_2^{2k_2}\right)$$

and they are all Lipschitz. Then

1. *The first moment is given by*

$$\mathbb{E}[\tilde{\Delta}^{i,j}] = \eta \begin{pmatrix} -g_\theta(\theta,\omega)_i \\ g_\omega(\theta,\omega)_j \end{pmatrix} + \eta^2 \begin{pmatrix} \{-\nabla_\omega[g_\theta(\theta,\omega)_i]\}^T g_\omega(\theta,\omega) \\ 0 \end{pmatrix} + O(\eta^3).$$

2. *The second moment is given by*

$$\begin{aligned}\mathbb{E}[\tilde{\Delta}^{i,j}(\tilde{\Delta}^{k,l})^T] =& \eta^2 \left[\frac{1}{B} \begin{pmatrix} \Sigma_\theta(\theta,\omega)_{i,k} & 0 \\ 0 & \Sigma_\omega(\theta,\omega)_{j,l} \end{pmatrix} + \begin{pmatrix} -g_\theta(\theta,\omega)_i \\ g_\omega(\theta,\omega)_j \end{pmatrix} \begin{pmatrix} -g_\theta(\theta,\omega)_k \\ g_\omega(\theta,\omega)_l \end{pmatrix}^T \right] \\ &+ O(\eta^3).\end{aligned}$$

3. *The third moments are all of order $O(\eta^3)$.*

Proof. Let $\psi : \mathbb{R}^{d_\theta+d_\omega} \to \mathbb{R}$ be any smooth test function. Under the dynamic (ALT-SDE), define the following operators

$$\begin{aligned}\mathcal{L}_1\psi(\theta,\omega) &= b_0(\theta,\omega)^T \nabla\psi(\theta,\omega), \\ \mathcal{L}_2\psi(\theta,\omega) &= b_1(\theta,\omega)^T \nabla\psi(\theta,\omega), \\ \mathcal{L}_3\psi(\theta,\omega) &= \frac{1}{2} Tr\Big(\sigma(\theta,\omega)\sigma(\theta,\omega)^T \nabla^2\psi(\theta,\omega)\Big).\end{aligned}$$

Apply Itô's formula to $\psi(\Theta_t, \mathcal{W}_t)$, $\mathcal{L}_i\psi(\Theta_t,\mathcal{W}_t)$ for $i = 1, 2, 3$, and $\mathcal{L}_1^2\psi(\Theta_t,\mathcal{W}_t)$, we have the following,

$$\psi(\Theta_\eta,\mathcal{W}_\eta) = \psi(\theta,\omega) + \int_0^\eta (\mathcal{L}_1 + \eta\mathcal{L}_2 + \mathcal{L}_3)\psi(\Theta_t,\mathcal{W}_t)dt + \int_0^\eta [\nabla\psi(\Theta_t,\mathcal{W}_t)]^T \sigma(\Theta_t,\mathcal{W}_t)dW_t$$

$$= \psi(\theta,\omega) + \eta\Big(\mathcal{L}_1 + \mathcal{L}_3\Big)\psi(\theta,\omega) + \eta^2\Big(\frac{1}{2}\mathcal{L}_1^2 + \mathcal{L}_2\Big)\psi(\theta,\omega) \tag{5.24}$$

$$\left.\begin{aligned} &+ \int_0^\eta\int_0^t\int_0^s \mathcal{L}_1^3\psi(\Theta_u,\mathcal{W}_u)dudsdt + \int_0^\eta\int_0^t \Big(\mathcal{L}_3\mathcal{L}_1 + \mathcal{L}_1\mathcal{L}_3 + \mathcal{L}_3^2\Big)\psi(\Theta_s,\mathcal{W}_s)dsdt \\ &\qquad + \eta\int_0^\eta\int_0^t \Big(\mathcal{L}_2\mathcal{L}_1 + \mathcal{L}_1\mathcal{L}_2 + \mathcal{L}_3\mathcal{L}_2 + \mathcal{L}_2\mathcal{L}_3\Big)\psi(\Theta_s,\mathcal{W}_s)dsdt \\ &\qquad + \eta^2\int_0^\eta\int_0^t \mathcal{L}_2^2\psi(\Theta_s,\mathcal{W}_s)dsdt \end{aligned}\right\} \tag{5.25}$$

$$+ M_\eta, \tag{5.26}$$

where M_η denotes the remaining martingale term with mean zero. Given the regularity conditions of b_0, b_1 and σ, [113, Theorem 9 in Section 2.5] implies that (5.25) is of order $O(\eta^3)$. Therefore,

$$\mathbb{E}\Big[\psi(\Theta_\eta,\mathcal{W}_\eta)\Big|\Theta_0 = \theta, \mathcal{W}_0 = \omega\Big] = \psi(\theta,\omega) + \eta\Big(\mathcal{L}_1 + \mathcal{L}_3\Big)\psi(\theta,\omega) + \eta^2\Big(\frac{1}{2}\mathcal{L}_1^2 + \mathcal{L}_2\Big)\psi(\theta,\omega).$$

Take $\psi(\Theta_\eta,\mathcal{W}_\eta)$ as $\tilde{\Delta}_i$, $\tilde{\Delta}_i\tilde{\Delta}_j$ and $\tilde{\Delta}_i\tilde{\Delta}_j\tilde{\Delta}_k$ for arbitrary indices $i, j, k = 1, \ldots, d_\theta + d_\omega$, then the conclusion follows. □

Estimate of moments. In this section, we will bound the moments of GANs parameters under (ALT).

Lemma 5.9. *Fix an arbitrary time horizon $\mathcal{T} > 0$ and take the learning rate $\eta \in (0, 1 \wedge \mathcal{T})$ and the number of iterations $N = \left\lfloor \frac{\mathcal{T}}{\eta} \right\rfloor$. Suppose that $g_\theta^{i,j}$ and $g_\omega^{i,j}$ are all lipschitz, i.e. there exists $L > 0$ such that*

$$\max_{i,j}\{|g_\theta^{i,j}(\theta,\omega)|, |g_\omega^{i,j}(\theta,\omega)|\} \leq L\left(1 + \left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2\right).$$

Then for any $m \in \mathbb{N}$, $\max_{t=1,\dots,N} \mathbb{E}\left[\left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m\right]$ is uniformly bounded, independent from η.

Proof. Throughout the proof, positive constants C and C' may vary from line to line. The Lipschitz assumption suggests that

$$\max\{|g_\theta^{\mathcal{B}}(\theta,\omega)|, |g_\omega^{\mathcal{B}}(\theta,\omega)|\} \leq L\left(1 + \left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2\right).$$

For any $k = 1, \dots, m$

$$\max\{|g_\theta^{\mathcal{B}}(\theta,\omega)|^k, |g_\omega^{\mathcal{B}}(\theta,\omega)|^k\} \leq L \cdot k\binom{k}{\lfloor \frac{k}{2} \rfloor} \cdot \left(1 + \left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2^k\right),$$

and

$$\left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2^k + \left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2^m \leq 2\left(1 + \left\|\begin{pmatrix}\theta\\ \omega\end{pmatrix}\right\|_2^m\right).$$

For any $t = 0, \dots, N-1$,

$$\begin{aligned}
\left\|\begin{pmatrix}\theta_{t+1}\\ \omega_{t+1}\end{pmatrix}\right\|_2^m &\leq \left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m + \sum_{k=1}^m \binom{m}{k}\left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^{m-k} \eta^k \left\|\begin{pmatrix}-g_\theta^{\mathcal{B}}(\theta_t,\omega_t)\\ g_\omega^{\mathcal{B}}(\theta_t,\omega_t)\end{pmatrix}\right\|_2^k \\
&\leq \left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m + C\eta \sum_{k=1}^m \binom{m}{k}\left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^{m-k}\left(1 + \left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m\right) \\
&\leq \left(1 + C\eta\right)\left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m + C'\eta.
\end{aligned}$$

Denote $a_t^m = \left\|\begin{pmatrix}\theta_t\\ \omega_t\end{pmatrix}\right\|_2^m$. Then, $a_{t+1}^m \leq (1 + C\eta)a_t^m + C'\eta$ that leads to

$$\begin{aligned}
a_t^m &\leq (1 + C\eta)^t\left(a_0^m + \frac{C'}{C}\right) - \frac{C'}{C} \\
&\leq (1 + C\eta)^{\frac{\mathcal{T}}{\eta}}\left(a_0^m + \frac{C'}{C}\right) - \frac{C'}{C} \\
&\leq e^{C\mathcal{T}}\left(a_0^m + \frac{C'}{C}\right) - \frac{C'}{C}.
\end{aligned}$$

The conclusion follows. □

Mollification. Notice that in Theorem 1 (and Theorem 2), the condition about the differentiability of loss function Φ is in the weak sense. For the ease of analysis, we will adopt the following mollfication, given in [74].

Definition 5.10 (Mollifier)**.** *Define the following function* $\nu : \mathbb{R}^{d_\theta+d_\omega} \to \mathbb{R}$,

$$\nu(u) = \begin{cases} C \exp\left\{ -\frac{1}{\|u\|_2^2-1} \right\}, & \|u\|_2 < 1; \\ 0, & \|u\|_2 \geq 1, \end{cases}$$

such that $\int_{\mathbb{R}^{d_\theta+d_\omega}} \nu(u)du = 1$. *For any* $\epsilon > 0$, *define* $\nu^\epsilon(u) = \frac{1}{\epsilon^{d_\theta+d_\omega}} \nu\left(\frac{u}{\epsilon}\right)$.

Note that the mollifier $\nu \in \mathcal{C}^\infty(\mathbb{R}^{d_\theta+d_\omega})$ and for any $\epsilon > 0$, $supp(\nu^\epsilon) = B_\epsilon(0)$ where $B_\epsilon(0)$ denotes the ϵ ball around the origin in the Euclidean space $\mathbb{R}^{d_\theta+d_\omega}$.

Definition 5.11 (Mollification)**.** *Let* $f \in \mathcal{L}^1_{loc}(\mathbb{R}^{d_\theta+d_\omega})$ *be any locally integrable function. For any* $\epsilon > 0$, *define* $f^\epsilon = \nu^\epsilon * f$ *such that*

$$f^\epsilon(u) = \int_{\mathbb{R}^{d_\theta+d_\omega}} \nu^\epsilon(u-v)f(v)dv = \int_{\mathbb{R}^{d_\theta+d_\omega}} \nu^\epsilon(v)f(u-v)dv.$$

By a simple change of variables and integration by part, one could derive that for any multi-index J,

$$\nabla f^\epsilon = \nu^\epsilon * [D^J f].$$

Here we quote some well-known results about this mollification from [74, Theorem 7 of Appendix C.4].

Lemma 5.12.

1. $f^\epsilon \in \mathcal{C}^\infty(\mathbb{R}^{d_\theta+d_\omega})$.
2. $f^\epsilon \longrightarrow f$ *almost everywhere as* $\epsilon \longrightarrow 0$.
3. *If* $f \in \mathcal{C}(\mathbb{R}^{d_\theta+d_\omega})$, *then* $f^\epsilon \longrightarrow f$ *uniformly on compact subsets of* $\mathbb{R}^{d_\theta+d_\omega}$.
4. *If* $f \in \mathcal{L}^p_{loc}(\mathbb{R}^{d_\theta+d_\omega})$ *for some* $1 \leq p < \infty$, *then* $f^\epsilon \longrightarrow f$ *in* $\mathcal{L}^p_{loc}(\mathbb{R}^{d_\theta+d_\omega})$.

To give a convergence rate for the pointwise convergence in Lemma 5.12, we have the following proposition.

Lemma 5.13. *Assume* $f \in W^{1,1}_{loc}(\mathbb{R}^{d_\theta+d_\omega})$ *and there exist* k_1, k_2 *such that* $|Df(u)| \leq k_1(1 + \|u\|_2^{2k_2})$, *then for any* $u \in \mathbb{R}^{d_\theta+d_\omega}$, *there exists* $\rho : \mathbb{R}^+ \to \mathbb{R}$ *that* $\lim_{\epsilon\to 0} \rho(\epsilon) = 0$ *and* $|f^\epsilon(u) - f(u)| \leq \rho(\epsilon)$.

Proof.

$$\begin{aligned}|f^\epsilon(u)-f(u)| &= \left|\int_{B_\epsilon(0)} \nu^\epsilon(v)[f(u-v)-f(u)]\right|dv \\ &= \left|\int_{B_\epsilon(0)} \nu^\epsilon(v)\int_0^1 [Df(u-hv)^T v]dhdv\right|. \\ &\leq \epsilon \int_{B_\epsilon(0)} \nu^\epsilon(v)\int_0^1 |Df(u-hv)|dhdv\end{aligned}$$

Since there exist k_1, k_2 such that $|Df(u)| \leq k_1(1+\|u\|_2^{2k_2})$,

$$\begin{aligned}|f^\epsilon(u)-f(u)| &\leq \epsilon \int_{B_\epsilon(0)} \nu^\epsilon(v)\int_0^1 \left[k_1(1+\|u-hv\|_2^{2k_2})\right]dhdv \\ &\leq \epsilon \int_{B_\epsilon(0)} \nu^\epsilon(v)\int_0^1 \left[k_1(1+\|u\|_2^{2k_2}+h^{2k_2}\|v\|_2^{2k_2})\right]dhdv \\ &\leq \epsilon \int_{B_\epsilon(0)} \nu^\epsilon(v)\left[k_1(1+\|u\|_2^{2k_2})+\frac{k_1}{2k_2+1}\|v\|_2^{2k_2}\right]dv \\ &\leq \epsilon[k_1(1+\|u\|_2^{2k_2})]+\frac{k_1}{2k_2+1}\epsilon^{2k_2+1}.\end{aligned}$$

Let $\rho(\epsilon) = \epsilon[k_1(1+\|u\|_2^{2k_2})]+\frac{k_1}{2k_2+1}\epsilon^{2k_2+1}$. Then $\rho(\epsilon) \longrightarrow 0$ as $\epsilon \longrightarrow 0$. □

It is also straightforward to see that mollification preserves Lipschitz conditions.

Consider the following SDE under componentwise mollification of coefficients,

$$d\begin{pmatrix}\Theta_t^\epsilon \\ \mathcal{W}_t^\epsilon\end{pmatrix} = [b_0^\epsilon(\Theta_t^\epsilon, \mathcal{W}_t^\epsilon)dt + \eta b_1^\epsilon(\Theta_t^\epsilon, \mathcal{W}_t^\epsilon)] + \sigma^\epsilon(\Theta_t^\epsilon, \mathcal{W}_t^\epsilon)dW_t. \tag{SDE-MLF}$$

Lemma 5.14. *Assume b_0, b_1 and σ are all Lipschitz. Then*

$$\mathbb{E}\left[\max_{t=1,\dots,N} \| \begin{pmatrix}\Theta_{t\eta}^\epsilon \\ \mathcal{W}_{t\eta}^\epsilon\end{pmatrix} - \begin{pmatrix}\Theta_{t\eta} \\ \mathcal{W}_{t\eta}\end{pmatrix} \|_2^2\right] \overset{\epsilon\to 0}{\longrightarrow} 0,$$

where $\begin{pmatrix}\Theta_{t\eta}^\epsilon \\ \mathcal{W}_{t\eta}^\epsilon\end{pmatrix}$ *and* $\begin{pmatrix}\Theta_{t\eta} \\ \mathcal{W}_{t\eta}\end{pmatrix}$ *are given by* (SDE-MLF) *and* (ALT-SDE)*, respectively.*

Proof. With Lemma 5.13, the conclusion follows from [113, Theorem 9 in Section 2.5]. □

Remaining proof

Given the conditions of Theorem 1 and the fact that mollification preserves Lipschitz conditions, b_0^ϵ, b_1^ϵ and σ^ϵ inherit regularity conditions from Theorem 1. Therefore, the conclusion from Lemma 5.8 holds. Lemmas 5.7, 5.8, 5.9 and 5.13 verify the condition in [125, Theorem 3]. Therefore, for any test function $f \in \mathcal{C}^3(\mathbb{R}^{d_\theta+d_\omega})$ such that for any multi-index J with $|J| \leq 3$ there exist $k_1, k_2 \in \mathbb{N}$ satisfying

$$|\nabla^J f(\theta,\omega)| \leq k_1 \left(1 + \left\| \begin{pmatrix}\theta \\ \omega\end{pmatrix} \right\|_2^{2k_2}\right),$$

we have the following weak approximation,

$$\max_{t=1,\ldots,N} \left|\mathbb{E} f(\theta_t,\omega_t) - \mathbb{E} f(\Theta^\epsilon_{t\eta}, \mathcal{W}^\epsilon_{t\eta})\right| \leq C[\eta^2 + \rho(\epsilon)] \tag{5.27}$$

for constant $C \geq 0$, where (θ_t, ω_t) and $(\Theta_{t\eta}, \mathcal{W}_{t\eta})$ are given by (ALT) and (SDE-MLF), respectively, and ρ is given as in Lemma 5.13.

Finally, taking ϵ to 0, Lemma 5.14 and the explicit form of ρ lead to the conclusion.

The proof of Theorem 2 can be executed in a similar fashion.

5.5.2 Proof of Theorem 3

In this section, we will prove Theorem 3. One of the key components is to identify a suitable Lyapunov function given the conditions of Theorem 3. The associated Lyapunov condition leads to the existence of an invariant measure for the dynamics of the parameters. We highlight this very technique since it can be used in the analysis of broader classes of dynamical systems, for both stochastic and deterministic cases; see for instance [116].

Consider the following function $V : [0,\infty) \times \mathbb{R}^{d_\theta+d_\omega} \to \mathbb{R}$,

$$V(t,u) = \exp\{\delta t + \epsilon \|u\|_2\}, \quad \forall u \in \mathbb{R}^{d_\theta+d_\omega}, \tag{Lyapunov}$$

where the parameters $\delta, \epsilon > 0$ will be determined later. Note that V is a smooth function, and

$$\lim_{\|u\|_2\to\infty} \inf_{t\geq 0} V(t,u) = +\infty, \tag{5.28}$$

for any fixed $\delta, \epsilon > 0$. Under (ALT-SDE), applying Itô's formula to V gives

$$\begin{aligned} dV(t,\Theta_t,\mathcal{W}_t) = {} & V(t,\Theta_t,\mathcal{W}_t)\left[\epsilon \frac{(\Theta_t \;\; \mathcal{W}_t)\, b(\Theta_t,\mathcal{W}_t)}{\left\|(\Theta_t \;\; \mathcal{W}_t)^T\right\|_2} + \delta + \frac{1}{2} Tr\left(\sigma(\Theta_t,\mathcal{W}_t)\sigma(\Theta_t,\mathcal{W}_t)^T \right.\right. \\ & \times \left.\left.\left[\frac{\epsilon \left\|(\Theta_t \;\; \mathcal{W}_t)^T\right\|_2^2 I + (\epsilon^2 \left\|(\Theta_t \;\; \mathcal{W}_t)^T\right\|_2 - \epsilon)\,(\Theta_t \;\; \mathcal{W}_t)^T (\Theta_t \;\; \mathcal{W}_t)}{\left\|(\Theta_t \;\; \mathcal{W}_t)^T\right\|_2^3}\right]\right)\right] dt \\ & + \epsilon V(t,\Theta_t,\mathcal{W}_t) \frac{(\Theta_t \;\; \mathcal{W}_t)\, \sigma(\Theta_t,\mathcal{W}_t)}{\left\|(\Theta_t \;\; \mathcal{W}_t)^T\right\|_2} dW_t. \end{aligned}$$

Define the Lyapunov operator

$$\mathcal{L}V(t,u) = V(t,u)\left[\epsilon\frac{u^T b(u)}{\|u\|_2} + \delta + \frac{1}{2}Tr\left(\sigma(u)\sigma(u)^T\frac{\epsilon\|u\|_2^2 I + (\epsilon^2\|u\|_2 - \epsilon)uu^T}{\|u\|_2^3}\right)\right].$$

Given the boundedness of σ, i.e. there exists $K > 0$ such that $\|\sigma\|_F \leq K$, and dissipative property given by condition 2, i.e. there exists $l, M_0 > 0$ such that for any $u \in \mathbb{R}^{d_\theta + d_\omega}$ with $\|u\|_2 > M_0$,

$$u^T b(u) \leq -l\|u\|_2,$$

we have that

$$\mathcal{L}V(t,u) \leq V(t,u)\left[\delta - l\epsilon + \frac{1}{2}\left(\epsilon\frac{\|\sigma\|_F^2}{\|u\|_2} + \epsilon^2\|\sigma\|_F^2\right)\right] \leq V(t,u)\left[\delta + \frac{K^2\epsilon^2}{2} - \left(l - \frac{K^2}{2\|u\|_2}\right)\epsilon\right].$$

Now take $M > \max\left\{\frac{K^2}{2l}, M_0\right\}$, $0 < \epsilon < \frac{2l}{K^2} - \frac{1}{M}$ and $\delta = -\frac{1}{2}\left[\frac{K^2\epsilon^2}{2} + \left(\frac{K^2}{2M} - l\right)\epsilon\right] > 0$, then for any $\|u\|_2 > M$,

$$\mathcal{L}V(t,u) \leq -\delta V(t,u).$$

Therefore,

$$\lim_{\|u\|_2\to\infty} \inf_{t\geq 0} \mathcal{L}V(t,u) = -\infty. \tag{5.29}$$

Following [108, Theorem 2.6], (5.28) and (5.29) ensure the existence of a invariant measure μ^* for (ALT-SDE). By the uniform elliptic condition, uniqueness follows from [100, Theorem 2.3]. The exponential convergence rate follows from [170, Main result].

5.5.3 Proofs of Theorems 4 and 5

The proofs of Theorems 4 and 5 are relatively simple, as we have already derived the infinitesimal generator for (ALT-SDE), The conclusions of Theorems 4 and 5 follow from a direct computation.

5.6 Conclusion and remarks

This chapter analyzes the convergence of GANs training via the approximation by SDEs. There are a few problems that remain open. The first one is to establish similar approximations when generator and discriminator are trained under different time-scales. Currently we consider the case where both networks are updated once every iterations. It will be interesting to see the difference in the SDE approximation if they are updated under two frequencies within an iteration. The second one is the explicit characterization of the invariant measure. These will be interesting potential topics to explore. Also, given the interactions between the generator and discriminator during the training process, it would be a delicate issue to choose proper filtration when developing the continuous-time approximation in the form of SDEs. The filtration enlargemenet problem associated with SDE approximation of GANs training will be a challenging and profound problem that is worth probing.

Bibliography

[1] Y. Achdou and I. Capuzzo-Dolcetta. "Mean field games: numerical methods". In: *SIAM Journal on Numerical Analysis* 48.3 (2010), pp. 1136–1162.

[2] R. Aïd, M. Basei, G. Callegaro, L. Campi, and T. Vargiolu. "Nonzero-sum stochastic differential games with impulse controls: a verification theorem with applications". In: *Mathematics of Operations Research* 45.1 (2020), pp. 205–232.

[3] R. Aïd, M. Basei, G. Callegaro, L. Campi, and T. Vargiolu. "Nonzero-sum stochastic differential games with impulse controls: a verification theorem with applications". In: *Mathematics of Operations Research* 45.1 (2020), pp. 205–232.

[4] A. Alfonsi and P. Blanc. "Dynamic optimal execution in a mixed-market-impact Hawkes price model". In: *Finance and Stochastics* 20.1 (2016), pp. 183–218.

[5] M. Arjovsky and L. Bottou. "Towards principled methods for training generative adversarial networks". In: *International Conference on Learning Representations*. Toulon, 2017.

[6] M. Arjovsky, S. Chintala, and L. Bottou. "Wasserstein generative adversarial networks". In: *International Conference on Machine Learning*. 2017, pp. 214–223.

[7] K. J. Arrow et al. "An extension of the basic theorems of classical welfare economics". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California. 1951.

[8] R. Atar and A. Budhiraja. "Singular control with state constraints on unbounded domain". In: *The Annals of Probability* 34.5 (2006), pp. 1864–1909.

[9] P. Azimzadeh. "A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions". In: *Applied Mathematics & Optimization* 79.2 (2019), pp. 483–514.

[10] F. Bagagiolo and D. Bauso. "Mean-field games and dynamic demand management in power grids". In: *Dynamic Games and Applications* 4.2 (2014), pp. 155–176.

[11] M. Bardi. "Explicit solutions of some linear-quadratic mean field games". In: *Networks and Heterogeneous Media* 7.2 (2011), pp. 243–261.

[12] S. A. Barnett. "Convergence problems with generative adversarial networks (GANs)". In: *arXiv preprint arXiv:1806.11382* (2018).

[13] M. Basei. "Optimal price management in retail energy markets: an impulse control problem with asymptotic estimates". In: *Mathematical Methods of Operations Research* 89.3 (2019), pp. 355–383.

[14] M. Basei, H. Cao, and X. Guo. "Nonzero-sum stochastic games and mean-field games with impulse controls". In: *Mathematics of Operations Research* to appear (2020).

[15] D. Bauso, H. Tembine, and T. Baar. "Robust mean field games with application to production of an exhaustible resource". In: *IFAC Proceedings Volumes* 45 (2012), pp. 454–459.

[16] E. Bayraktar, T. Emmerling, and J.-L. Menaldi. "On the impulse control of jump diffusions". In: *SIAM Journal on Control and Optimization* 51.3 (2013), pp. 2612–2637.

[17] M. Beiglböck, M. Nutz, and N. Touzi. "Complete duality for martingale optimal transport on the line". In: *The Annals of Probability* 45.5 (2017), pp. 3038–3074.

[18] V. E. Beneš, L. A. Shepp, and H. S. Witsenhausen. "Some solvable stochastic control problems". In: *Stochastics* 4.1 (1980), pp. 39–83.

[19] A. Bensoussan and J. L. Lions. *Impulse control and quasi-variational inequalities.* Paris: Bordas, 1982.

[20] A. Bensoussan, K. C. J. Sung, S. C. P. Yam, and S. P. Yung. "Linear-quadratic mean feld games". In: *Journal of Optimization Theory and Applications* 169.2 (2016), pp. 496–529.

[21] A. Bensoussan and B. Chevalier-Roignant. "Sequential capacity expansion options". In: *Operations Research* 67.1 (2019), pp. 33–57.

[22] A. Bensoussan, J. Frehse, and P. Yam. *Mean Field Games and Mean Field Type Control Theory.* SpringerBriefs in Mathematics. New York, NY: Springer New York, 2013.

[23] A. Bensoussan and A. Friedman. "Nonzero-sum stochastic differential games with stopping times and free boundary problems". In: *Transactions of the American Mathematical Society* 231.2 (1977).

[24] H. Berard, G. Gidel, A. Almahairi, P. Vincent, and S. Lacoste-Julien. "A closer look at the optimization landscape of generative adversarial networks". In: *International Conference on Learning Representations.* 2020.

[25] G. Bertola, W. J. Runggaldier, and K. Yasuda. "On classical and restricted impulse stochastic control for the exchange rate". In: *Applied Mathematics & Optimization* 74.74 (2016).

[26] C. Bertucci. "Fokker-Planck equations of jumping particles and mean field games of impulse control". In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire* (2020).

[27] T. R. Bielecki and S. R. Pliska. "Risk sensitive asset management with transaction costs". In: *Finance and Stochastics* 4.1 (2000), pp. 1–33.

[28] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[29] D. Bridge and S. Shreve. "Multi-dimensional finite-fuel singular stochastic control". In: *Applied Stochastic Analysis*. 1992, pp. 38–58.

[30] S. Browne and W. Whitt. "Piecewise-linear diffusion processes". In: *Advances in Queueing: Theory, Methods, and Open Probelms*. Ed. by J. H. Dshalalow. CRC Press, 1995. Chap. 18, pp. 463–480.

[31] G. Brunick and S. Shreve. "Mimicking an Itô process by a solution of a stochastic differential equation". In: *The Annals of Applied Probability* 23.4 (2013), pp. 1584–1628.

[32] A. Cadenillas, P. Lakner, and M. Pinedo. "Optimal control of a mean-reverting inventory". In: *Operations Research* 58.6 (2010), pp. 1697–1710.

[33] A. Cadenillas and F. Zapatero. "Classical and impulse stochastic control of the exchange rate using interest rates and reserves". In: *Mathematical Finance* 10.2 (2000), pp. 141–156.

[34] A. Cadenillas, T. Choulli, M. Taksar, and L. Zhang. "Classical and impulse stochastic control for the optimization of the dividend and risk policies of an insurance firm". In: *Mathematical Finance* 16.1 (2006), pp. 181–202.

[35] L. Campi and D. De Santis. "Nonzero-sum stochastic differential games between an impulse controller and a stopper". In: *arXiv:1904.00059* (2019).

[36] L. Campi and D. De Santis. "Nonzero-sum stochastic differential games between an impulse controller and a stopper". In: *Journal of Optimization Theory and Applications* (2020), pp. 1–37.

[37] H. Cao and X. Guo. "Approximation and convergence of GANs training: an SDE approach". In: *arXiv preprint arXiv:2006.02047* (2020).

[38] H. Cao and X. Guo. "MFGs for partially reversible investment". In: *Stochastic Processes and their Applications* to appear (2020).

[39] H. Cao, X. Guo, and M. Laurière. "Connecting GANs, MFGs and OT". In: *arXiv preprint arXiv:2002.04112* (2020).

[40] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games: (AMS-201)*. Vol. 201. Princeton University Press, 2019.

[41] P. Cardaliaguet and C.-A. Lehalle. "Mean field game of controls and an application to trade crowding". In: *Mathematics and Financial Economics* 12.3 (2018), pp. 335–363.

[42] P. Cardaliaguet and A. Porretta. "Long time behavior of the master equation in mean-field game theory". In: *Analysis & PDE* 12.6 (2019), pp. 1397–1453.

[43] P. Cardaliaguet and C. Rainer. "On the (in)efficiency of MFG equilibria". In: *arXiv preprint arXiv:1802.06637* (2018).

[44] E. Carlini and F. J. Silva. "A fully discrete semi-Lagrangian scheme for a first order mean field game problem". In: *SIAM Journal on Numerical Analysis* 52.1 (2014), pp. 45–67.

[45] R. Carmona and F. Delarue. "Probabilistic analysis of mean field games". In: *SIAM Journal on Control and Optimization* 51.4 (2013), pp. 2705–2734.

[46] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games*. Springer, 2018.

[47] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications II: Mean Field Games with Common Noise and Master Equations*. Springer, 2018.

[48] R. Carmona, J.-P. Fouque, and L.-H. Sun. "Mean field games and systemic risk". In: *Communications in Mathematical Sciences* 13.4 (2015), pp. 911–933.

[49] R. Carmona and M. Laurière. *Convergence Analysis of Machine Learning Algorithms for the Numerical Solution of Mean Field Control and Games: I - The Ergodic Case*. Preprint. 2019.

[50] R. Carmona and M. Laurière. *Convergence Analysis of Machine Learning Algorithms for the Numerical Solution of Mean Field Control and Games: II - The Finite Horizon Case*. Preprint. 2019.

[51] P. Casgrain and S. Jaimungal. "Mean-field games with differing beliefs for algorithmic trading". In: *Mathematical Finance* 30.3 (2020), pp. 995–1034.

[52] D. Chang, H. Wang, and Z. Wu. "Maximum principle for non-zero sum differential games of BSDEs involving impulse controls". In: *Chinese Control Conference*. 2013.

[53] D. Chang and Z. Wu. "Stochastic maximum principle for non-zero sum differential games of FBSDEs with impulse controls and its application to finance". In: *Journal of Industrial and Management Optimization* 11.1 (2014), pp. 27–40.

[54] L. Chen, M. Pelger, and J. Zhu. "Deep learning in asset pricing". In: *Available at SSRN 3350138* (2019).

[55] C. Chu, K. Minami, and K. Fukumizu. "Smoothness and Stability in GANs". In: *International Conference on Learning Representations*. 2020.

[56] M. Cirant and L. Nurbekyan. "The variational structure and time-periodic solutions for mean-field games systems". In: *arXiv preprint arXiv:1804.08943* (2018).

[57] G. Conforti, A. Kazeykina, and Z. Ren. "Game on random environment, mean-field Langevin system and neural networks". In: *arXiv preprint arXiv:2004.02457* (2020).

[58] G. M. Constantinides and S. F. Richard. "Existence of optimal simple policies for discounted-cost inventory and cash management in continuous time". In: *Operations Research* 26.4 (1978), pp. 620–636.

[59] A. Cosso. “Stochastic differential games involving impulse controls and double-obstacle quasi-variational inequalities”. In: *SIAM Journal on Control and Optimization* 51.3 (2013), pp. 2102–2131.

[60] G. Da Prato. *An Introduction to Infinite-Dimensional Analysis*. Springer Science & Business Media, 2006.

[61] M. H. Davis and A. R. Norman. “Portfolio selection with transaction costs”. In: *Mathematics of Operations Research* 15.4 (1990), pp. 676–713.

[62] M. H. Davis and M. Zervos. “A problem of singular stochastic control with discretionary stopping”. In: *The Annals of Applied Probability* (1994), pp. 226–240.

[63] G. Debreu. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. 17. Yale University Press, 1959.

[64] E. L. Denton, S. Chintala, A. Szlam, and R. Fergus. “Deep generative image models using a Laplacian pyramid of adversarial networks”. In: *Advances in Neural Information Processing Systems*. 2015, pp. 1486–1494.

[65] J. Dianetti and G. Ferrari. “Nonzero-sum submodular monotone-follower games: existence and approximation of Nash equilibria”. In: *SIAM Journal on Control and Optimization* 58.3 (2020), pp. 1257–1288.

[66] A. K. Dixit, R. K. Dixit, and R. S. Pindyck. *Investment under Uncertainty*. Princeton university press, 1994.

[67] Y. Dolinsky and H. M. Soner. “Martingale optimal transport and robust hedging in continuous time”. In: *Probability Theory and Related Fields* 160.1-2 (2014), pp. 391–427.

[68] C. Domingo-Enrich, S. Jelassi, A. Mensch, G. M. Rotskoff, and J. Bruna. “A mean-field analysis of two-player zero-sum games”. In: *arXiv preprint arXiv:2002.06277* (2020).

[69] J. F. Eastham and K. J. Hastings. “Optimal impulse control of portfolios”. In: *Mathematics of Operations Research* 13.4 (1988), pp. 588–605.

[70] S. Eckstein and M. Kupper. “Computation of optimal transport and related hedging problems via penalization and neural networks”. In: *Applied Mathematics & Optimization* (2018), pp. 1–29.

[71] N. El Karoui and I. Karatzas. “Probabilistic aspects of finite-fuel, reflected follower problems”. In: *Acta Applicandae Mathematica* 11.3 (1988), pp. 223–258.

[72] N. El Karoui and S. Méléard. “Martingale measures and stochastic calculus”. In: *Probability Theory and Related Fields* 84.1 (1990), pp. 83–101.

[73] N. El Karoui, D. Nguyen, and M. Jeanblanc-Picqué. “Compactification methods in the control of degenerate diffusions: existence of an optimal control”. In: *Stochastics* 20.3 (1987), pp. 169–219.

[74] L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[75] K. Fan. "Fixed-point and minimax theorems in locally convex topological linear spaces". In: *Proceedings of the National Academy of Sciences of the United States of America* 38.2 (1952), pp. 121–126.

[76] G. Ferrari and T. Koch. "On a strategic model of pollution control". In: *Annals of Operations Research* 275.2 (2019), pp. 297–319.

[77] C. Finn, S. Levine, and P. Abbeel. "Guided cost learning: Deep inverse optimal control via policy optimization". In: *International Conference on Machine Learning*. 2016, pp. 49–58.

[78] A. Friedman. "Stochastic games and variational inequalities". In: *Archive for Rational Mechanics and Analysis* 51.5 (1973), pp. 321–346.

[79] G. Fu and U. Horst. "Mean field games with singular controls". In: *SIAM Journal on Control and Optimization* 55.6 (2017), pp. 3833–3868.

[80] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. "Generative adversarial nets". In: *Advances in Neural Information Processing Systems*. 2014, pp. 2672–2680.

[81] C. Griessler. "C-cyclical monotonicity as a sufficient criterion for optimality in the multimarginal Monge-Kantorovich problem". In: *Proceedings of the American Mathematical Society* 146.11 (2018), pp. 4735–4740.

[82] O. Guéant, J.-M. Lasry, and P.-L. Lions. "Application of mean field games to growth theory". In: *Paris–Princeton Lectures on Mathematical Finance 2010*. Springer Berlin, 2011.

[83] O. Guéant, J.-M. Lasry, and P.-L. Lions. "Mean field games and applications". In: *Paris-Princeton Lectures on Mathematical Finance 2010*. Springer Berlin Heidelberg, 2011, pp. 205–266.

[84] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville. "Improved training of Wasserstein GANs". In: *Advances in Neural Information Processing Systems*. 2017, pp. 5767–5777.

[85] G. Guo, J. Obłój, et al. "Computational methods for martingale optimal transport problems". In: *The Annals of Applied Probability* 29.6 (2019), pp. 3311–3347.

[86] G. Guo, X. Tan, and N. Touzi. "Tightness and duality of martingale transport on the Skorokhod space". In: *Stochastic Processes and their Applications* 127.3 (2017), pp. 927–956.

[87] X. Guo, J. Hong, T. Lin, and N. Yang. "Relaxed Wasserstein with applications to GANs". In: *arXiv preprint arXiv:1705.07164* (2017).

[88] X. Guo, A. Hu, R. Xu, and J. Zhang. "Learning mean-field games". In: *Advances in Neural Information Processing Systems*. 2019, pp. 4967–4977.

[89] X. Guo, P. Kaminsky, P. Tomecek, and M. Yuen. “Optimal spot market inventory strategies in the presence of cost and price risk”. In: *Mathematical Methods of Operations Research* 73.1 (2011), pp. 109–137.

[90] X. Guo and J. S. Lee. “Stochastic games and mean field games with singular controls”. In: *arXiv preprint arXiv:1703.04437* (2018).

[91] X. Guo and H. Pham. “Optimal partially reversible investment with entry decision and general production function”. In: *Stochastic Processes and their Applications* 115.5 (2005), pp. 705–736.

[92] X. Guo and G. Wu. “Smooth fit principle for impulse control of multi-dimensional diffusion processes”. In: *SIAM Journal on Control and Optimization* 48.2 (2010), pp. 594–617.

[93] X. Guo and R. Xu. “Stochastic games for fuel follower problem: N versus mean field game”. In: *SIAM Journal on Control and Optimization* 57.1 (2019), pp. 659–692.

[94] J. M. Harrison, T. M. Sellke, and A. J. Taylor. “Impulse control of Brownian motion”. In: *Mathematics of Operations Research* 8.3 (1983), pp. 454–466.

[95] J. M. Harrison and M. I. Taksar. “Instantaneous control of Brownian motion”. In: *Mathematics of Operations Research* (1983).

[96] J. M. Harrison and R. J. Williams. “Multidimensional reflected Brownian motions having exponential stationary distributions”. In: *The Annals of Probability* 15.1 (1987), pp. 115–137.

[97] U. Haussmann and J. Lepeltier. “On the existence of optimal controls”. In: *SIAM Journal on Control and Optimization* 28.4 (1990), pp. 851–902.

[98] D. Hernández-Hernández, J. L. Pérez, and K. Yamazaki. “Optimal of refraction strategies for spectrally negative Lévy processes”. In: *SIAM Journal on Control and Optimizations* 54.3 (2016), pp. 1126–1156.

[99] S. Hochreiter and J. Schmidhuber. “Long short-term memory”. In: *Neural Computation* 9.8 (1997), pp. 1735–1780.

[100] J. Hong and X. Wang. “Invariant Measures for Stochastic Differential Equations”. In: *Invariant Measures for Stochastic Nonlinear Schödinger Equations: Numerical Approximations and Symplectic Structures*. Singapore: Springer Singapore, 2019, pp. 31–61.

[101] W. Hu, C. J. Li, L. Li, and J.-G. Liu. “On the diffusion approximation of nonconvex stochastic gradient descent”. In: *Annals of Mathematical Sciences and Applications* 4.1 (2019), pp. 3–32.

[102] M. Huang, R. P. Malhamé, and P. E. Caines. “Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle”. In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.

[103] X. Huang, S. Jaimungal, and M. Nourian. "Mean-field game strategies for optimal execution". In: *Applied Mathematical Finance* 0.0 (2019), pp. 1–33.

[104] A. Jack, T. C. Johnson, and M. Zervos. "A singular control model with application to the goodwill problem". In: *Stochastic Processes and their Applications* 118.11 (2008), pp. 2098–2124.

[105] M. Jeanblanc-Picque. "Impulse control method and exchange rate". In: *Mathematical Finance* 3.2 (1993), pp. 161–177.

[106] M. Jeanblanc-Picque and A. N. Shiryaev. "Optimization of the flow of dividends". In: *Russian Mathematical Surveys* 50.2 (1995), pp. 257–277.

[107] I. Karatzas. "A class of singular stochastic control problems". In: *Advances in Applied Probability* 15.2 (1983), pp. 225–254.

[108] R. Khasminskii. *Stochastic Stability of Differential Equations*. 2nd ed. Vol. 66. Springer Science & Business Media, 2011.

[109] R. Korn. "Some applications of impulse control in mathematical finance". In: *Mathematical Methods of Operations Research* 50.2 (1999), pp. 493–528.

[110] R. Korn. "Portfolio optimisation with strictly positive transaction costs and impulse control". In: *Finance and Stochastics* 2 (1998), pp. 85–114.

[111] R. Korn. "Some applications of impulse control in mathematical finance". In: *Mathematical Methods of Operations Research* 50 (1999), pp. 493–518.

[112] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in Neural Information Processing Systems*. 2012, pp. 1097–1105.

[113] N. V. Krylov. *Controlled Diffusion Processes*. Springer Science & Business Media, 2008.

[114] V. Kulharia, A. Ghosh, A. Mukerjee, V. Namboodiri, and M. Bansal. "Contextual RNN-GANs for abstract reasoning diagram generation". In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 2017, pp. 1382–1388.

[115] H. D. Kwon and H. Zhang. "Game of singular stochastic control and strategic exit". In: *Mathematics of Operations Research* 40.4 (2015), pp. 869–887.

[116] M. Laborde and A. Oberman. "A Lyapunov analysis for accelerated gradient methods: From deterministic to stochastic case". In: *International Conference on Artificial Intelligence and Statistics*. 2020, pp. 602–612.

[117] D. Lacker. "A general characterization of the mean field limit for stochastic differential games". In: *Probability Theory and Related Fields* 165 (2016), pp. 581–648.

[118] D. Lacker. "Mean field games via controlled martingale problems: existence of Markovian equilibria". In: *Stochastic Processes and their Applications* 125 (2015), pp. 2856–2894.

[119] D. Lacker. "On the convergence of closed-loop Nash equilibria to the mean field game limit". In: *Annals of Applied Probability* 30.4 (2020), pp. 1693–1761.

[120] D. Lacker and T. Zariphopoulou. "Mean field and N-agent games for optimal investment under relative performance criteria". In: *Mathematical Finance* 29.4 (2019), pp. 1003–1038.

[121] J.-M. Lasry and P.-L. Lions. "Mean field games". In: *Japanese Journal of Mathematics* 2.1 (2007), pp. 229–260.

[122] C. Lea, R. Vidal, A. Reiter, and G. D. Hager. "Temporal convolutional networks: A unified approach to action segmentation". In: *European Conference on Computer Vision*. Springer. 2016, pp. 47–54.

[123] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. "Photo-realistic single image super-resolution using a generative adversarial network". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 4681–4690.

[124] N. Lei, K. Su, L. Cui, S.-T. Yau, and X. D. Gu. "A geometric view of optimal transportation and generative model". In: *Computer Aided Geometric Design* 68 (2019), pp. 1–21.

[125] Q. Li, C. Tai, and W. E. "Stochastic modified equations and dynamics of stochastic gradient algorithms I: mathematical foundations". In: *Journal of Machine Learning Research* 20.40 (2019), pp. 1–47.

[126] T. Lim. "Multi-martingale optimal transport". In: *arXiv preprint arXiv:1611.01496* (2016).

[127] G.-H. Liu and E. A. Theodorou. "Deep learning theory review: An optimal control and dynamical systems perspective". In: *arXiv preprint arXiv:1908.10920* (2019).

[128] P. Luc, C. Couprie, S. Chintala, and J. Verbeek. "Semantic segmentation using adversarial networks". In: *arXiv preprint arXiv:1611.08408* (2016).

[129] V. Ly Vath, M. Mnif, H. Pham, V. Ly Vath, H. Pham, and M. Mnif. "A model of optimal portfolio selection under liquidity risk and price impact". In: *Finance and Stochastics* 11 (2007), pp. 51–90.

[130] D. C. Mauer and A. J. Triantis. "Interactions of corporate financing and investment decisions: a dynamic framework". In: *The Journal of Finance* 49.4 (1994), pp. 1253–1277.

[131] A. Merhi and M. Zervos. "A Model for reversible investment capacity expansion". In: *SIAM Journal on Control and Optimization* 46.3 (2008), pp. 839–876.

[132] L. Mescheder, A. Geiger, and S. Nowozin. "Which training methods for GANs do actually converge?" In: *International Conference on Machine Learning*. 2018, pp. 3481–3490.

[133] T. Mikolov, S. Kombrink, L. Burget, J. Černockỳ, and S. Khudanpur. "Extensions of recurrent neural network language model". In: *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2011, pp. 5528–5531.

[134] D. Mitchell, H. Feng, and K. Muthuraman. "Impulse control of interest rates". In: *Operations Research* 62.3 (2014), pp. 602–615.

[135] G. Monge. "Mémoire sur la théorie des déblais et des remblais". In: *Histoire de l'Académie Royale des Sciences de Paris* (1781).

[136] A. J. Morton and S. R. Pliska. "Optimal portfolio management with fixed transaction costs". In: *Mathematical Finance* 5.4 (1995), pp. 337–356.

[137] G. Mundaca and B. Øksendal. "Optimal stochastic intervention control with application to the exchange rate". In: *Journal of Mathematical Economics* 29.2 (1998), pp. 225–243.

[138] G. Mundaca and B. Oksendal. "Optimal stochastic intervention control with application to the exchange rate". In: *Economics* 29 (1998), pp. 225–243.

[139] J. F. Nash. "Equilibrium points in n-person games". In: *Proceedings of the National Academy of Sciences* 36.1 (1950), pp. 48–49.

[140] J. F. Nash. "Non-cooperative games". In: *Annals of Mathematics* (1951), pp. 286–295.

[141] R. Nock, Z. Cranko, A. K. Menon, L. Qu, and R. C. Williamson. "f-GANs in an information geometric nutshell". In: *Advances in Neural Information Processing Systems*. 2017, pp. 456–464.

[142] M. Nutz, J. San Martin, X. Tan, et al. "Convergence to the mean field game limit: a case study". In: *The Annals of Applied Probability* 30.1 (2020), pp. 259–286.

[143] M. Nutz, F. Stebegg, and X. Tan. "Multiperiod martingale transport". In: *Stochastic Processes and their Applications* 130.3 (2020), pp. 1568–1615.

[144] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, 2003.

[145] B. Øksendal and A. Sulem. *Applied Stochasic Control of Jump Diffusions*. 2nd. Berlin: Springer, 2007.

[146] B. Øksendal and A. Sulem. "Optimal consumption and portfolio with both fixed and proportional transaction costs". In: *SIAM Journal on Control and Optimizations* 40.6 (2002), pp. 1765–1790.

[147] E. Pardoux and S. Peng. "Backward stochastic differential equations and quasilinear parabolic partial differential equations". In: *Stochastic Partial Differential Equations and their Applications*. Springer, 1992, pp. 200–217.

[148] W. Poundstone. *Prisoner's Dilemma*. Doubleday, 1992.

[149] A. Radford, L. Metz, and S. Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks". In: *arXiv preprint arXiv:1511.06434* (2015).

[150] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee. "Generative Adversarial Text to Image Synthesis". In: *33rd International Conference on Machine Learning*. 2016, pp. 1060–1069.

[151] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales Volume 2: Itô Calculus*. Cambridge University Press, 2000.

[152] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. "Improved techniques for training GANs". In: *Advances in Neural Information Processing Systems*. 2016, pp. 2234–2242.

[153] T. Salimans, H. Zhang, A. Radford, and D. Metaxas. "Improving GANs Using Optimal Transport". In: *International Conference on Learning Representations*. 2018.

[154] M. Sanjabi, J. Ba, M. Razaviyayn, and J. D. Lee. "On the convergence and robustness of training GANs with regularized optimal transport". In: *Advances in Neural Information Processing Systems*. 2018, pp. 7091–7101.

[155] H. Scarf. *The optimality of (s,S) policies for the dynamic inventory problem*. 1960.

[156] L. A. Shepp and A. N. Shiryaev. "A new look at pricing of the "Russian option "". In: *Theory of Probability & Its Applications* 39.1 (1995), pp. 103–119.

[157] S. E. Shreve and H. M. Soner. "Optimal investment and consumption with transaction costs". In: *The Annals of Applied Probability* 4 (1994), pp. 609–692.

[158] S. E. Shreve and H. M. Soner. "Regularity of the value function for a two-dimensional singular stochastic control problem". In: *SIAM Journal on Control and Optimization* 27.4 (1989), pp. 876–907.

[159] M. Sion. "On general minimax theorems." In: *Pacific Journal of Mathematics* 8.1 (1958), pp. 171–176.

[160] J. Sirignano and K. Spiliopoulos. "DGM: A deep learning algorithm for solving partial differential equations". In: *Journal of Computational Physics* 375 (2018), pp. 1339–1364.

[161] A. Srivastava, K. Greenewald, and F. Mirzazadeh. "BreGMN: scaled-Bregman generative modeling networks". In: *arXiv preprint arXiv:1906.00313* (2019).

[162] J. H. Steg. "Irreversible investment in oligopoly". In: *Finance and Stochastics* 16.2 (2012), pp. 207–224.

[163] D. W. Stroock and S. S. Varadhan. *Multidimensional diffusion processes*. Springer, 2007.

[164] A. Sulem. "A solvable one-dimensional model of a diffusion inventory system". In: *Mathematics of Operations Research* 11.1 (1986), pp. 125–133.

[165] D. Svozil, V. Kvasnicka, and J. Pospichal. "Introduction to multi-layer feed-forward neural networks". In: *Chemometrics and Intelligent Laboratory Systems* 39.1 (1997), pp. 43–62.

[166] A. J. Triantis and J. E. Hodder. "Valuing flexibility as a complex option". In: *The Journal of Finance* 45.2 (1990), pp. 549–565.

[167] K. Uchida. "On existence of a Nash equilibrium point in N-person nonzero sum stochastic differential games". In: *SIAM Journal on Control and Optimization* 16.1 (1978).

[168] S. R. Varadhan and R. J. Williams. "Brownian motion in a wedge with oblique reflection". In: *Communications on Pure and Applied Mathematics* 38.4 (1985), pp. 405–443.

[169] P. Varaiya. "N-player stochastic differential games". In: *SIAM Journal on Control and Optimizations* 14.3 (1976).

[170] A. Y. Veretennikov. "Bounds for the mixing rate in the theory of stochastic equations". In: *Theory of Probability & Its Applications* 32.2 (1988), pp. 273–281.

[171] C. Villani. *Optimal Transport: Old and New*. Springer Science & Business Media, 2008.

[172] J. Von Neumann. "On the theory of games of strategy". In: *Contributions to the Theory of Games* 4 (1959), pp. 13–42.

[173] C. Vondrick, H. Pirsiavash, and A. Torralba. "Generating videos with scene dynamics". In: *Advances in Neural Information Processing Systems*. 2016, pp. 613–621.

[174] M. Wiatrak, S. V. Albrecht, and A. Nystrom. "Stabilizing generative adversarial networks: a survey". In: *arXiv preprint arXiv:1910.00927* (2019).

[175] M. Wiese, L. Bai, B. Wood, J. P. Morgan, and H. Buehler. "Deep hedging: learning to simulate equity option markets". In: *arXiv preprint arXiv:1911.01700* (2019).

[176] M. Wiese, R. Knobloch, R. Korn, and P. Kretschmer. "Quant GANs: deep generation of financial time series". In: *Quantitative Finance* (2020), pp. 1–22.

[177] O. Q. Wu and H. Chen. "Optimal control and equilibrium behavior of production-inventory systems". In: *Management Science* 56.8 (2010), pp. 1362–1379.

[178] S. Yaida. "Fluctuation-dissipation relations for stochastic gradient descent". In: *International Conference on Learning Representations*. 2019.

[179] L. Yang, D. Zhang, and G. E. Karniadakis. "Physics-informed generative adversarial networks for stochastic differential equations". In: *SIAM Journal on Scientific Computing* 42.1 (2020), A292–A317.

[180] Y. Yang and P. Perdikaris. "Adversarial uncertainty quantification in physics-informed neural networks". In: *Journal of Computational Physics* 394 (2019), pp. 136–152.

[181] R. Yeh, C. Chen, T. Y. Lim, M. Hasegawa-Johnson, and M. N. Do. "Semantic image inpainting with perceptual and contextual losses". In: *arXiv preprint arXiv:1607.07539* 2.3 (2016).

[182] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang. "Stock market prediction based on generative adversarial network". In: *Procedia Computer Science* 147 (2019), pp. 400–406.

[183] B. Zhu, J. Jiao, and D. Tse. "Deconstructing generative adversarial networks". In: *IEEE Transactions on Information Theory* (2020).

[184] J.-Y. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros. "Generative visual manipulation on the natural image manifold". In: *European Conference on Computer Vision*. Springer. 2016, pp. 597–613.

Appendix A

Preliminary on neural networks

Neural networks (NNs) are building blocks of any GANs model. Though there are numerous variations of networks structures, the following three types are among the most commonly used ones, namely feedforward networks (FFNs) [165], recurrent neural networks (RNNs) [133] and convolutional neural networks (CNNs) [112].

A.1 Feedforward networks and multilayer perceptron

NNs are introduced as function approximators. FFNs are a class of NNs that are presented by directed acyclic graphs. The most common type of FFNs is the multilayer perceptron (MLP). As a graphical model, the basic components of MLP are nodes and edges. Nodes are more commonly referred as neurons. The neurons are arranged in layers from left to right, denoting input and output sides respectively. The number of layers is often referred to as the depth of the NN. The edges only link neurons from two adjacent layers and the direction of the edges are from left to right. For a MLP with $L+1$ layers, $L = 1, 2, \dots,$ the leftmost layer is the input layer, layer 0, while the rightmost, layer L, is the output layer. The layers in between are called hidden layers. Suppose neuron i in layer $l-1$ denoted by $n_{l-1,i}$ is linked with neuron j in layer l denoted by $n_{l,j}$, $l = 1, \dots, L$. The output of $n_{l-1,i}$ is the input of $n_{l,j}$, denoted by $h_{l-1,i}$; if $l = 1$, then the $h_{0,i} = x_i$. The computation happening at $n_{l,j}$ is

$$h_{l,j} = \sigma_l \left(\sum_{i:n_{l-1,i} \to n_{l,j}} w_{l,i,j} h_{l-1,i} + b_{l,j} \right), \tag{A.1}$$

where σ_l is a nonlinear function called an activation function, $w_{l,i,j}$ is the weight and $b_{l,j}$ is the bias. If $l = L$ then $h_{L,j}$ becomes a part of the output of the MLP; otherwise, it is taken as input in the computations for the next layer. The input $x = (x_1, \dots, x_N)$ is mapped through all the layers and output as the image of some nonlinear function.

A.2 Recurrent neural networks and long-short term memory.

RNNs are a class neural networks that takes sequential data as input. If FFNs are interpreted as one-dimensional computation along the depth direction, the computation of RNNs is two-dimensional involving both the depth direction and the temporal direction. At each time step, along the depth direction, similar computation like (A.1) occurs; the difference is that not only information at the current time step is taken into consideration but also the information from the previous time steps. There are different ways to incorporate this past information manifested through different network architectures. One particular way is through the long-short term memory (LSTM) units [99]. At time step t, an LSTM unit produces an output h_t and also keeps track of a quantity called cell state, denoted by c_t. The LSTM unit takes the data at t, x_t, as well as the output and cell state of LSTM unit from the previous time step, h_{t-1} and c_{t-1}, as the input. The computation of the output at current time step t involves three regulators, the input gate i_t, the output o_t and the forget gate f_t,

$$\begin{aligned} f_t &= \sigma_g\left(W_{f,t}x_t + U_{f,t}h_{t-1} + b_{f,t}\right), && \text{(forget gate)} \\ i_t &= \sigma_g\left(W_{i,t}x_t + U_{i,t}h_{t-1} + b_{i,t}\right), && \text{(input gate)} \\ o_t &= \sigma_g\left(W_{o,t}x_t + U_{o,t}h_{t-1} + b_{o,t}\right), && \text{(output gate)} \end{aligned}$$

where σ_g is the activation function chosen for the computation for these regulators, $W.$ and $U.$ are weight matrices for current input x_t and past output h_{t-1}, respectively and $b.$ is the bias. Having computed the regulators, the current cell state is given by

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c\left(W_{c,t}x_t + U_{c,t}h_{t-1} + b_{c,t}\right), \qquad \text{(cell state)}$$

and the output is given by

$$h_t = o_t \circ \sigma_h\left(c_t\right), \qquad \text{(output)}$$

where σ_c and σ_h are activation functions chosen for the computations of cell state and output, respectively, and the operator $\circ$ means element-wise multiplication.

A.3 Convolutional neural networks and temporal convolutional networks.

When dealing with high dimensional data such as images, fully connected MLPs will suffer from high variance due to the large number of weights. CNNs can help ameliorate this issue. The distinct characteristic of CNNs is the introduction of convolutional layer and pooling layers.

Consider an input of dimension $W \times H \times D$, denoted by $X = (x_{i,j,l})$. In a convolutional layer, a kernel of size $w \times h \times D$ is introduced, denoted by $K = (k_{i,j,l})$. The output is a matrix $Y = (y_{i,j}) = X * K$ of size $(W - w + 1) \times (H - h + 1)$ given by

$$y_{i,j} = [X * K]_{i,j} = \sum_{a=1}^{w} \sum_{b=1}^{h} \sum_{l=1}^{D} x_{i+a-1,j+b-1,l} \times k_{a,b,l}. \tag{A.2}$$

Multiple kernels can be introduced in which case their outputs will be stacked together and each individual output will be referred to as a channel. The computation in (A.2) has the filter stride across the input by a single tuple. To save memory, it can stride by multiple tuples. To control the size of the output from the convolution, we can apply zero-padding, that is, to add suitable number of zero terms around the original input.

The pooling layer, a hidden layer, is to downsample the previous layer. In a pooling layer, the size of kernel and stride are to specified. Instead of performing convolution to get each element of the output matrix, a representative element of the corresponding sub-matrix of the output from the previous layer will be chosen. For instance, in max-pooling the maximum will be chosen and in average-pooling the average value will be recorded.

Unlike CNNs, in temporal convolutional networks (TCNs) [122] the convolution happens along the temporal direction. The inputs for TCNs are sequential data. The fundamental building block is causal convolutional layer where a computation called dilated causal convolution takes place. Denote the input as $x = (x_1, \ldots, x_T)$ where $x_t \in \mathbb{R}^{n_I}$ for $t = 1, \ldots, T$. The kernel of the causal convolutional layer is $W = (w_{i,j,l})_{i,j,l} \in \mathbb{R}^{K \times n_I \times n_O}$ where K is the kernel size and let $W_i \in \mathbb{R}^{n_I \times n_O}$ denote the 2-dimensional matrix $(w_{i,j,l})_{j,l}$ for $i = 1, \ldots, K$. The output of this causal convolutional layer with dilation D is a sequence of length $T - D(K-1)$, denoted by $y = (y_{D(K-1)+1}, \ldots, y_T)$, where $y_t \in \mathbb{R}^{n_O}$, for $t \in \{D(K-1)+1, \ldots, T\}$, is given by

$$y_t = \sum_{i=1}^{K} W_i^T x_{t-D(K-i)}.$$

A vanilla TCN consists of multiple such causal convolutional layers combined with activation functions. Note that the depth of the TCN and the kernel size K and dilation D at each layer are properly chosen so that the output sequence has length no less than 1.

Appendix B

More on Nash equilibrium and Pareto optimality

The two major components of this thesis, MFGs and GANs, are centered around games. Among the various concepts of equilibrium, Nash equilibrium (NE) (see Definitions 2.25 and 3.6) and Pareto optimality (see Section 4.3) are most commonly seen in literature. We will use several simple 2-player games to make some more illustrations.

B.1 Difference between NE and PO

The difference between NE and PO can be summarized as follows.

Definition B.1. *For a generic N-player game, consider a set of strategies $\mathbf{s} = (s^1, \dots, s^N)$.*

- *If no players has the incentive to make a unilateral deviation from $\mathbf{s}$, then $\mathbf{s}$ is called an NE for the game.*
- *If there is* ***no other set of strategies*** *$\mathbf{s}' = (s'^1, \dots, s'^N)$ such that, compared with $\mathbf{s}$, $\mathbf{s}'$ can strictly improve the benefit of one player without harming the other players, then $\mathbf{s}$ is called PO point for the game.*

To make a more concrete distinction, let us consider a classical Prisoners' Dilemma problem (see [148] for more details).

Ted and Jack are arrested because of a crime they committed together. They will be facing interrogation next morning. If Ted confesses and Jack remains silent, then Jack will be sentenced to 5 years in prison while Ted can be freed and vice versa. If, one the other hand, both of them confess, then both get a sentence of 3 years in prison; if both of them keep silent, then both get a sentence of 1 year in prison. Suppose they are not allowed to communicate prior to the interrogation. Looking at this game from Jack's perspective, no matter ted confesses or not, according to the payoff described in Table B.1, making the confession is a better strategy for Jack. So is ted. Therefore, by the description of NE in

Table B.1: Prisoner's dilemma

	Jack confesses	**Jack silent**
Ted confesses	(-3, -3)	(0, -5)
Ted silent	(-5, 0)	(-1, -1)

Definition B.1, both of them making the confession is an NE and in this case, the unique NE. However, if given the chance of communication, then Jack and Ted may quickly realize that both of them remaining silent will get them the lightest sentence, thus it is indeed the unique PO point of this game according to Definition B.1. This Prisoners' Dilemma case is a straightforward illustration on the difference between NE and Po.

B.2 (Non-)uniqueness of NE and PO

The sensational works of Nash [139, 140] establish the existence of NE for non-cooperative games. However, uniqueness of NE is not a guarantee; in Chapter 3, we have provided multiple NE for the 2-player impulse game. PO is extensively studied in the field of welfare economics, see for instance the classical works [7, 63]. Again, uniqueness is an exception for PO. The non-uniqueness can be better seen from the following two cases.

B.2.1 Case 1: unique NE and multiple PO

Consider a 2-player game where each player can his/her strategy from a binary set $\{0, 1\}$. The utility functions for the players are given by

$$u_1(s_1, s_2) = (s_1 + s_2) - 1.5s_1 = s_2 - 0.5s_1, \quad u_2(s_1, s_2) = (s_1 + s_2) - 1.5s_2 = s_1 - 0.5s_2.$$

The possible outcomes of utilities are given by Table B.2 as follows. To maximize individual

Table B.2: Utilities of players

	$\mathbf{s_2 = 0}$	$\mathbf{s_2 = 1}$
$\mathbf{s_1 = 0}$	(0, 0)	(1, -0.5)
$\mathbf{s_1 = 1}$	(-0.5, 1)	(0.5, 0.5)

utility, $s_1 = 0$ is a better strategy for player 1 than $s_1 = 1$ no matter which strategy player 2 chooses. So is for player 2. Then according to Definition B.1, $(0, 0)$ is the unique NE in this 2-player game. To examine the all 4 set of strategies according to the description of PO in

Definition B.1, $(0,1)$, $(1,0)$ and $(1,1)$ are all PO points while $(0,0)$ is strictly dominated by $(1,1)$ therefore not a PO point.

B.2.2 Case 2: multiple NE and unique PO

Consider a game of distributing fruits between Alex and Bob. There are 100 apples and 100 bananas. In the end, each of them will have to take away exactly 100 fruits. Their preferences of the fruits (evaluated in real numbers) are given by Table B.3 as follows. Player i's strategy,

Table B.3: Individual preferences

	An apple	**A banana**
Alex	1	-1
Bob	-1	1

with $i = 1, 2$, can be characterized as $s^i = (n_1^i, n_2^i)$, where n_1^i and n_2^i denote the numbers of apples and bananas player i takes, respectively; $i = 1$ stands for Alex and $i = 2$ stands for Bob. A set of strategies (s^1, s^2) is **admissible** if $n_1^i + n_2^i = 100$ and $n_j^1 + n_j^2 = 100$ for all $i = 1, 2$ and $j = 1, 2$. Consider only the set of admissible strategies, denoted by $\mathcal{S}$. Then immediately any $s \in \mathcal{S}$ is an NE for the game, since unilateral deviation is not possible if restricted to $\mathcal{S}$. For PO, on the other hand, $(s^{*,1}, s^{*,2})$ where $s^{*,1} = (100, 0)$ and $s^{*,2} = (0, 100)$ is the unique PO point.