Thèse de Doctorat

Université de Limoges

Université de Limoges
ED 653 – École Doctorale Sciences et Ingénierie (SI)
Faculté des Sciences et Techniques – Institut de Recherche XLIM

Thèse pour obtenir le grade de
Docteur de l'Université de Limoges
Mathématiques appliquées

Présentée et soutenue par
Ahmad Zorkot
Le 5 Juin 2024

APPROXIMATION DE JEUX À CHAMP MOYEN

Thèse dirigée par Francisco J. Silva et Elisabetta Carlini

JURY:

Président:
M. Samir Adly, Professeur – Université de Limoges

Rapporteurs:
M. Fabio Camilli, Professore ordinario – Università di Roma I, La Sapienza
M. Diogo Gomes, Full professor – The King Abdullah University of Science and Technology

Examinateurs:
M. Yves Achdou, Professeur – Université Paris Cité
Mme. Justina Gianatti, Profesora adjunta – Universidad Nacional de Rosario
M. Cristian Mendico, Ricercatore – Università di Roma II, Tor Vergata

Directeurs de thèse:
M. Francisco J. Silva Alvarez, Maître de conférences HDR – Université de Limoges, **directeur de thèse**
Mme. Elisabetta Carlini, Professore associato – Università di Roma I, La Sapienza, **co-directrice de thèse**

To my parents, Assaad and Ghada

Remerciements

First of all, I would like to express my deepest gratitude to Francisco Silva and Elisabetta Carlini for the chance they gave me to be a part in the amazing world of mean field games, and for their generous advice, time, and knowledge, their assistance, availability, patience, and continuous support throughout my thesis. I was very fortunate to have them as supervisors. Working with them has been extremely enriching on both scientific and personal levels. I cannot imagine better PhD advisors than these two amazing persons.

I also express my gratitude to Fabio Camilli and Diogo Gomes for the honor of reviewing my thesis. I also thank Yves Achdou, Samir Adly, Justina Gianatti, and Cristian Mendico for agreeing to be examiners for this thesis. I would like also to thank Laurent Pfeiffer for accepting to be a part in the monitoring committee of my last year of thesis.

I extend my sincere thanks to my colleagues from the MOD team, as well as the staff of the Mathematics and Computer Science Department at the University of Limoges. A big thank you also to Débora and Sophie for your kindness and continuous help. I thank my friends and colleagues from Xlim for the good times we shared, especially my first officemate Hamza.

I would also like to thank my Lebanese friends in Limoges. We formed a wonderful, close-knit family and created many amazing memories together. There are too many names to mention, Youssef, Hassan, Ali S, Mohamad and Jihane, Ali and Maya, Jamil and Wafaa, Ali and Carine, Ahmad and Nour, Jad, Lara, Mohamad K, Khalil, Ali D, Ali G, Mohamad Z, Mariam, Mira Maya and Marina, Mohamad M, thank you a lot.

A special thanks also to my Bordeaux friends (or I should say brothers): Mahdi, Hassan, Kassem and Mahmoud, who I made the best memories with.

The greatest thanks go to my parents (Assaad and Ghada) and my four sisters Fatima, Doha, Batoul and Aya. Their unconditional support and love have helped me realize what I am now. Finally, to my beloved lady Fatima, who keeps always making sure that I'm surrounded by her unconditional love, this thesis rests on your support and encouragement from the beginning to the very end.

I have tried to translate in simple words the infinite gratitude I have and will always have to Fatima and my parents for being my fixed point in my life. Obviously, I failed.

Table of contents

1 Introduction

Sommaire

The main objective of this thesis is to study numerical approximation methods for different classes of mean field games.
The theory of mean field games, abbreviated as MFG, offers the possibility to analyze systems involving a multitude of interacting participants. In this framework, players are considered indistinguishable and rational. Each of them pursues a goal and acts optimally to achieve it while facing the presence of other players. Specifically, interactions among players occur exclusively through couplings that depend on measures induced by other participants, such as the measure induced by their state or their control actions.
The main idea of such a theory is to draw inspiration from statistical physics, borrowing the general principle of a mean field approach to describe equilibriums in a system of many interacting particles, such as the stars in a galaxy or subatomic particles (see e.g [102]).
In classic game theory, the focus is on Nash equilibria with a finite number of players. The applications of Nash equilibria with a large number of players are numerous and varied, including economics, population dynamics, social networks, finance, and the economics of fossil and renewable energies. However, their theoretical analysis is often complex, and their numerical solution is challenging or even impossible due to the large number of players. Mean field game theory simplifies the study of the statistical behavior of players by transitioning to an infinite number of players. Thus, the notion of mean field equilibrium is justified as the limit, as N tends to infinity, of Nash equilibria for games with N players, assuming that players are symmetrical and rational.
Mean field game theory is a relatively recent field that emerged in the years 2005-2006. It is the result of independent work conducted by Jean-Michel Lasry and Pierre-Louis Lions, as well as by Minyi Y. Huang, Peter E. Caines, and Roland E. Malhamé.

1.1 The mean field games theory

1.1.1 N-player games

We introduce here the most straightforward mathematical framework. In the context of a differential game, time and space variables are considered continuous. The number of players is represented by $N > 1$, and the time horizon is denoted as $T > 0$. We denote by $\mathcal{P}(\mathbb{R}^d)$ the space of probability measures on $\mathbb{R}^d$.

Each player, denoted as $i \in \{1, 2, \ldots, N\}$, has a variable state $X^i(t)$ that follows a stochastic dynamics in the time interval $[0, T]$. This process can be described by a stochastic differential equation:

$$dX^i(t) = \alpha^i dt + \sqrt{2\nu} dW^i(t) \quad \text{for } t \in]0, T[\,, \quad X_0^i = x^i \in \mathbb{R}^d,$$

where $\nu \geq 0$, $(W^1, \ldots, W^N)$ is a vector of independent Brownian motions, and α^i represents the control process chosen by player i. In what follows, we consider as admissible controls those in feedback form and such that $X^i(t)$ is well defined.
Each player with index i seeks to minimize an individual cost, given by:

$$J(\alpha^i, (\alpha^j)_{j \neq i}) = \mathbb{E}\left[\int_0^T L(X_t^i, \alpha_t^i) + F(X_t^i, m_t^i)\mathrm{d}t + G(X_T^i, m_T^i)\right]. \tag{1.1.1}$$

Here, $L \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $F \colon \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ represent the instantaneous cost functions, $G \colon \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is the final cost function, and $m_t^i := \sum_{j \neq i} \delta_{X_t^j}$ is the empirical average of other players' strategies at time t.

Definition 1.1.1 *A Nash equilibrium (in closed loop) of the N-player system described above is a N-tuple $(\alpha^1, \ldots, \alpha^N)$ of measurable functions from $[0, T] \times (\mathbb{R}^d)^N$ to $\mathbb{R}^d$, such that:*

$$J\left(\alpha^i, (\alpha^j)_{j \neq i}\right) \leq J\left(\alpha, (\alpha^j)_{j \neq i}\right), \tag{1.1.2}$$

for all $i \in 1, \ldots, N$ and $\alpha : [0, T] \times (\mathbb{R}^d)^N \to \mathbb{R}^d$ measurable.

In this context, Nash equilibrium describes the configuration where each player's chosen strategy α^i is optimal given the strategies chosen by all the other players $(\alpha^j)_{j \neq i}$.
The condition $J(\alpha^i, (\alpha^j)_{j \neq i}) \leq J(\alpha, (\alpha^j)_{j \neq i})$ for every admissible control α highlights that, at this equilibrium, no player has an incentive to unilaterally deviate from her/his strategy.

Define the Hamiltonian of the $N-$players system as the (modified) Legendre transform of L w.r.t α, i.e.

$$H(x, p) = \sup_{\alpha \in \mathbb{R}^d} \Big\{ -\langle p, \alpha \rangle - L(x, \alpha) \Big\} \quad \text{for } x, p \in \mathbb{R}^d.$$

Then, the Nash system with N players is defined by the following equations:

$$\begin{cases} -\partial_t u^{N,i} - \nu \sum_{j=1}^{N} \Delta_{x^j} u^{N,i} + H(x^i, D_{x^i} u^{N,i}) + \sum_{j\neq i} D_{x^i} u^{N,i} H_p(x^j, D_{x^i} u^{N,j}) = F(x^i, m_x^{N,i}) \\ u^{N,i}(T,x) = G(x^i, m_x^{N,i}), \end{cases} \tag{1.1.3}$$

where $u^{N,i} : [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}$ and $m_x^{N,i} = \sum_{j\neq i} \delta_{x^j}$ for $i \in \{1,\dots,N\}$.
When a solution to system (1.1.3) exists and is regular, a Nash equilibrium is given by the control functions $\alpha^i(t,x) = -H_p(x, D_{x^i} u^{N,i}(t,x))$, and, in this case, the stochastic equation satisfied by the state of a player becomes:

$$dX_t^i = -H_p(x, D_{x^i} u^{N,i}(t,x))dt + \sqrt{2\nu} dW_t^i, \quad X_0^i = x^i \in \mathbb{R}^d.$$

When players use these controls, they achieve their optimal cost $u^{N,i}$, also called value function.
System (1.1.3) consists of N coupled Hamilton-Jacobi equations. For a detailed study of the Hamilton-Jacobi equations and the framework in which they are well-posed, we refer to the works of M. G. Crandall and P. L. Lions [70], M. Bardi and I. Capuzzo-Dolcetta [23] and G. Barles [26] on viscosity solution theory. For a specific study of the system (1.1.3), we refer to the works by A. Bensoussan and J. Frehse [29], [30].
The analysis of system (1.1.3) is complicated due to the coupling between the equations. Furthermore, the numerical approximation of its solutions is an impossible task with classical methods if N is large. Finally, similar to statistical physics, the dependency of data on the empirical mean and the symmetric nature of the system suggest the transition to the limit as $N \to \infty$.
This idea was the basis of the mean field game theory, which was simultaneously developed in the seminal works of Jean-Michel Lasry and Pierre-Louis Lions [111]–[113] on the one hand, and Minyi Y. Huang, Peter E. Caines, and Roland E. Malhamé [97], [99] on the other hand.

1.1.2 The mean field games system

Let us introduce the mean field games system that describes the limit game of the differential games described above when the number of players tends to infinity.
The trajectory $X(t)$ of a typical player whose initial condition is $x \in \mathbb{R}^d$ is determined by the following stochastic differential equation (SDE):

$$dX_t = \alpha_t dt + \sqrt{2\nu} dW_t, \quad X_0 = x, \tag{1.1.4}$$

where $(W_t)_{t\in[0,T]}$ is a standard $d-$dimensional Brownian motion and $(\alpha_t)_{t\in[0,T]}$ is the control.
Given a forecast $(m(t))_{t\in[0,T]} \subset \mathcal{P}(\mathbb{R}^d)$ of the time evolution of the distribution of the population, the aim of a typical player is to minimize a cost having the form

$$J(\alpha, m) = \mathbb{E}\left[\int_0^T L(\alpha_t, X_t) + F(X_t, m(t))\, dt + G(X_T, m(T))\right]. \tag{1.1.5}$$

The value function $u\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$ is defined as

$$u(s,x) = \inf_{(\alpha_t)_{t\in[s,T]}} \mathbb{E}\left[\int_s^T L(\alpha_t, X_t) + F(X_t, m(t))dt + G(X_T, m(T))\right], \tag{1.1.6}$$

where $(X_t)_{t\in[0,T]}$ is solution to (1.1.4), $X_s = x$, and the infimum is taken over all adapted controls $(\alpha_t)_{t\in[s,T]}$. The value function satisfies the following dynamic programming equation: for any $s \in [0,T[$, $\varepsilon \in]0, T-s[$, we have

$$u(s,x) = \inf_{(\alpha_t)_{t\in[s,T]}} \mathbb{E}\left[\int_s^{s+\epsilon} L(\alpha_t, X_t) + F(X_t, m(t))dt + u(X_{s+\epsilon}, s+\varepsilon)\right]. \tag{1.1.7}$$

From the dynamic programming equation, we can deduce that the value function satisfies the Hamilton-Jacobi-Bellman (HJB) equation (see e.g [134, Chapter 4]):

$$\begin{cases} -\partial_t u - \nu\Delta u + H(x, D_x u) = F(x, m(t)) \quad \text{in } [0,T]\times\mathbb{R}^d, \\ u(T,x) = G(x, m(T)) \quad \text{in } \mathbb{R}^d. \end{cases} \tag{1.1.8}$$

Under suitable assumptions, the value function u is the unique solution that satisfies the (HJB) equation (1.1.8) in the viscosity sense (e.g see [23]). The optimal control for the problem (1.1.5) can be expressed in terms of the value function as

$$\alpha^*(x,t) = -D_p H(x, D_x u(x,t)).$$

We now discuss the evolution of the population density. Denote by $m_0 \in \mathcal{P}(\mathbb{R}^d)$ the initial distribution of states for our system and suppose that all the agents implement their optimal control. Then the trajectory of a typical player whose initial condition is distributed as m_0 is governed by the following SDE

$$dX_t = -D_p H(X_t, D_x u(t, X_t))dt + \sqrt{2\nu}dW_t, \quad \mathcal{L}(X_0) = m_0. \tag{1.1.9}$$

where for every random variable Y, $\mathcal{L}(Y)$ represents it's law. We denote by $\tilde{m}(t)$ the distribution of the players at time t. Since $\tilde{m}(0) = m_0$, it follows from (1.1.9) that

$\{\tilde{m}(t) \,|\, t \in [0,T]\}$ solves the following Fokker-Planck equation:

$$\begin{cases} \partial_t \tilde{m} - \nu\Delta\tilde{m} - \mathrm{div}\,(\tilde{m} D_p H(x, D_x u) = 0 \\ \tilde{m}(0,\cdot) = m_0 \in \mathcal{P}(\mathbb{R}^d). \end{cases} \tag{1.1.10}$$

In an equilibrium state, the agents' anticipation $m(t)$ of the distribution aligns with the actual distribution $\tilde{m}(t)$, i.e

$$\tilde{m}(t) = m(t) \quad \text{for all } t \in [0,T]. \tag{1.1.11}$$

Hence, the equilibrium in a mean field game is characterized by all pairs (u,m) that simultaneously satisfy the coupled HJB and Fokker-Planck equations. In other words, (u,m) is the solution to the *mean field game system*

$$\begin{cases} -\partial_t u - \nu\Delta u + H(x, D_x u) = F(x, m(t)) & \text{in } (0,T)\times\mathbb{R}^d, \\ \partial_t m - \nu\Delta m - \mathrm{div}(D_p H(x, D_x u) m) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ m(0,x) = m_0(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{1.1.12}$$

System (1.1.12) introduces an additional complexity: it is no longer an ordinary evolution system where time progresses in a single direction. Indeed, we can see that the (HJB) equation is backward in time and the Fokker-Planck equation is forward in time. However, this system possesses a distinctive structure: the Fokker-Planck equation can be viewed as the dual of the linearized Hamilton-Jacobi-Bellman equation. This particular structure facilitates the analysis of solutions to (1.1.12), providing a priori estimates, among other advantages.

Remark 1.1.1 *Two types of coupling are considered in the mean field game literature:*

- *Nonlocal Coupling* $(F = F(x, m(t)))$*: in this case,* $F(x,\cdot)$ *depends on the whole distribution* $m(t)$ *of the population at time* t*.*
- *Local Coupling* $(F = F(m(t,\cdot)))$*: in this case, the coupling depends on the density of the population which explains its local character.*

We avoid delving into the case of first-order models ($\nu = 0$*) with local coupling because it's only understood under specific structural conditions (e.g see [49], [120]).*

The existence of a solution that satisfies the MFG system (1.1.12) is established by employing a fixed point argument in the reasoning process. Indeed, let $u[m]$ (for a given $m \in \mathcal{P}(\mathbb{R}^d)$) be the solution to the HJB equation (1.1.8), then define the map T as:

$$T(m) = \tilde{m},$$

where $\tilde{m}$ is the solution to the Fokker Planck equation (1.1.10) with $u = u[m]$. Thus, the existence of MFG solution is equivalent to find a fixed point m for the map T.
Arguing similarly, one can see the mean field game system (1.1.12) as a fixed point point problem on u. A summary on existence results for solutions to (1.1.12), under structural assumptions over H, F and G, can be found, for instance, in [5], [51] and the references therein.

Differently from the stochastic case ($\nu \neq 0$), the deterministic case ($\nu = 0$) requires a particular attention because one does not expect to obtain the existence of smooth solutions. We will delve into these two notions of solutions in subsequent subsections. However, for existence results of solutions to the system (1.1.12) with $\nu = 0$, one can refer to [46], [49], [54], [79], [91], [120]. Regarding uniqueness, a monotonicity criterion introduced by Lasry and Lions [112] has played an important role in previous works.

Definition 1.1.2 *We say that functions F and G are monotone in the sense of Lasry-Lions, if for every $t \in [0, T]$ and $m_1, m_2 \in \mathcal{P}(\mathbb{R}^d)$,*

$$\int_{\mathbb{R}^d} F(x, m_1) - F(x, m_2)\, d(m_1 - m_2)(x) \geq 0, \tag{1.1.13}$$

$$\int_{\mathbb{R}^d} G(x, m_1) - G(x, m_2)\, d(m_1 - m_2)(x) \geq 0. \tag{1.1.14}$$

The impact of the theory

The impact of mean field games is substantial, especially in the context of numerical approximations for partial differential equations (PDEs). These approximations offer cost-effective computations for equilibria in complex systems. Additionally, mean field games have demonstrated significant applications across diverse fields, including finance [53], [64], [82], [88], [107], autonomous vehicles [95], [132], energy production and management [14], [68], [74], [103], epidemic control [22], [110], [118], macroeconomic models [11], [73], [75], [90], as well as in security and communication [105], [123].
We refer to the monograph [85] for an overview on economic models and mean field games.

It is also insightful to understand how the system (1.1.12) can be regarded as the limit, in a certain sense, of the Nash system (1.1.3) when the number of players tends towards infinity.
The fact that a solution to the mean field game system provides a good approximation of a Nash equilibrium for N players when N is large has been extensively discussed in [59], [96], [98], [106].

Considering the convergence of Nash equilibria towards solutions of the mean field game system as the number of players tends to infinity, a lot of work has been undertaken to delve into this area. In [50], the authors investigate the master equation, which is a non-local, nonlinear partial differential equation posed in the infinite-dimensional space of probability measures. The system (1.1.12) can be derived from the master equation using the method of characteristics in an infinite-dimensional setting along the distribution measures of the states. Consequently, the existence and uniqueness of a regular solution to the master equation are closely linked to those of the system (1.1.12) and its stability concerning the initial condition m_0. It is demonstrated in [50] that, under certain assumptions, a solution to the master equation exists, is unique and regular, and it constitutes the limit of the solutions of the system (1.1.3) (see [50, Theorem 2.4.8]). A similar result is presented in [61, Theorem 6.28], allowing treatment of the linear-quadratic case. However, the previous results rely on the uniqueness of solutions to the system (1.1.12) Specifically, in the non-monotone case, the master equation approach has not been shown to yield the desired convergence. D. Lacker [108] proposes a probabilistic definition of solution to the mean field game system and shows that when the control space is compact, Nash equilibria for N players converge to such weak solutions.
In their article [79], Fischer and Silva prove the convergence of symmetric Nash equilibria for N players to solutions of the mean field game in Lagrangian form for a class of first-order finite horizon mean field games. They then establish a connection between these Lagrangian solutions and those obtained through the usual mean field game system.

A brief

The main concepts concerning mean field games discussed this far are summarized in Figure 1.1, and we refer to the survery [87], the lectures [5] and the monographs [60], [61], [86] for a thorough overview on MFGs.

Except for a few specific cases, such as the linear-quadratic case [24], [92], mean field game systems generally lack of explicit solutions, and hence they need numerical methods for their resolution. In the following subsection, we present a concise overview of the historical development of numerical methods for MFGs.

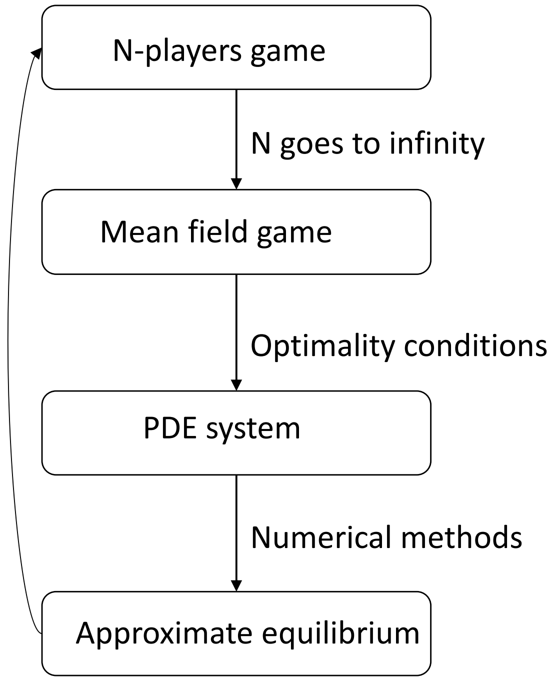


Figure 1.1: The MFG methodology.

1.1.3 Numerical methods

The study of numerical methods for solving mean field game system is an active area of research due to the vast range of models and scenarios that have not been fully covered yet. A discrete analogous of mean field games was proposed by Gomes, Mohr and Souza in [84], where they have studied the existence of a Nash equilibrium via a fixed point approach and investigated the long-term behavior of the game. For a review of numerical methods to solve mean field game problems, the reader is referred to [7], [114]. Here, we provide a concise overview of some contributions.

Let us begin with the case of second order mean field games, i.e., when $\nu > 0$ in system (1.1.12):

- In [4], Achdou and Capuzzo-Dolcetta proposed a finite differences approximation of (1.1.12). This discrete system preserves the main properties of the continuous one. The convergence towards a solution to (1.1.12) has been established by Achdou, Camilli and Capuzzo-Dolcetta in [2].
- In [58], Carlini and Silva, proposed a fully discrete semi-Lagrangian scheme and proved convergence towards a solution to (1.1.12).

- In a recent work, Cacace, Camilli, and Goffi proposed a policy iteration method for separable Hamiltonians in [40]. In [116], an extension of this method to the case of non-separable Hamiltonians is considered.
- Another approach based on numerical optimization methods for a specific case of mean field games, known as *variational (or potential) mean field games*, involve the use of augmented Lagrangian methods [27] and primal-dual methods [37], [38].
- In [36], Bonnans et al. proposed a new finite differences scheme, relying on the theta-method, for solving monotone second order mean field game system and they show a convergence rate for their method.
- In [117], Lavigne and Pfeiffer proposed the generalized conditional gradient algorithm, which is an extension of the Frank-Wolfe algorithm, to solve variational second order mean field game problems.

Regarding the case of deterministic mean field games ($\nu = 0$ in system (1.1.12)) we have:

- A semi-discrete semi-Lagrangian scheme proposed by Camilli and Silva in [42]. Then a fully-discrete scheme has been proposed by Carlini and Silva in [57]. The convergence of this approximation has been shown in the one dimensional case. In addition, an extension of this scheme to deal with fractional and non-local operators was proposed in [66], and an application to solve a price formation MFG model has been accomplished in [20].
- In [80], [94], the authors propose a semi-Lagrangian scheme without interpolation, leveraging on the specific structure of the dynamics and a relaxed definition of the mean field games equilibrium.
- In [15], [89], Gomes et al. proposed gradient flow methods for solving deterministic mean field games in infinite horizon. Their main idea is that the solution to the system of PDEs can be recast as a zero of a monotone operator, and can thus be found by following the related gradient flow.
- In [122], [126], approximations based on the Fourier method are employed.

Let us also mention that there has been a recent surge in interest in machine learning techniques applied to mean field game problems, such as deep learning and reinforcement learning methods [19], [62], [63]. For an overview on learning methods for mean field games, we refer to [115].

As mentioned before, the first order mean field game system requires the definition of a weaker notion of solution. In the next section, we provide a brief review of the main properties of these systems.

1.2 First order MFG systems

For a given $m_0^* \in L^p(\mathbb{R}^d)(p \in]1,\infty])$, a first order MFG system has the form

$$\begin{cases} -\partial_t u + H(x, D_x u) = F(x, m(t)) & \text{in } (0,T)\times\mathbb{R}^d, \\ \partial_t m - \operatorname{div}(D_p H(x, D_x u) m) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ m(0,x) = m_0^*(x) & \text{in } \mathbb{R}^d, \\ u(T,x) = G(x, m(T)) & \text{in } \mathbb{R}^d, \end{cases} \tag{1.2.1}$$

where we recall that F and $G\colon \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$ are the coupling terms and H is the Hamiltonian given by

$$H(x,p) = \sup_{a\in\mathbb{R}^d} \big(\langle a, p\rangle - L(x,a)\big) \quad \text{for all } x,\ p \in \mathbb{R}^d, \tag{1.2.2}$$

where $L\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is of class C^2.
Let us recall the definition of a solution (u,m) of (1.2.1) (see [111], [112]).

Definition 1.2.1 *The pair $(u,m) \in W^{1,\infty}([0,T]\times\mathbb{R}^d,\mathbb{R}) \times L^{\mathsf{p}}([0,T]\times\mathbb{R}^d,\mathbb{R})$ is a solution of (1.2.1) if the first equation is satisfied in the viscosity sense, while the second one is satisfied in the distributional sense.*

We will recall the definition of viscosity solution later. We say m satisfies the continuity equation in (1.2.1), in distribution sense, if for every test function $\varphi \in C_c^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$ we have

$$\int_{\mathbb{R}^d} \varphi(x,0) m_0^*(x)\mathrm{d}x + \int_0^T \int_{\mathbb{R}^d} \left[\partial_t \varphi(t,x) - \langle D_x\varphi(t,x), D_p H(x, D_x u(t,x))\rangle\right] m(t,x)\mathrm{d}x\mathrm{d}t = 0. \tag{1.2.3}$$

The existence of solutions to the system (1.2.1) is ensured under assumptions on couplings F, G, the Hamiltonian H, and the initial measure m_0^*. Before introducing our assumptions, let us recall the definition of semi-concavity, which is an essential aspect in the study of first order mean field game systems.

Definition 1.2.2 *We say that a function $v : \mathbb{R}^d \to \mathbb{R}$ is semi-concave if there exists $C > 0$ such that*

$$v(x+y) + v(x-y) - 2v(x) \leq C|y^2| \quad \textit{for all } x, y \in \mathbb{R}^d. \tag{1.2.4}$$

We assume the following:

(FG) Set $h = F, G$. We suppose that h continuous, bounded, and for every $\zeta \in \mathcal{P}_1(\mathbb{R}^d)$, $h(\cdot, \zeta)$ satisfies the semiconcavity property (1.2.4) with C indepenedent of ζ . Moreover, there exists $C > 0$ such that, for every x, $y \in \mathbb{R}^d$ and $\zeta \in \mathcal{P}_1(\mathbb{R}^d)$,

$$|h(x, \zeta) - h(y, \zeta)| \leq C|x - y|.$$

(L) For every x, $a \in \mathbb{R}^d$, we have

$$\begin{aligned} L(x, a) &\leq C_1(|a|^2 + 1), \\ |D_x L(x, a)| &\leq C_2(1 + |a|^2), \\ C_3|b|^2 &\leq \langle D^2_{aa} L(x, a)b, b\rangle \quad \text{for all } b \in \mathbb{R}^d, \\ \langle D^2_{xx} L(x, a)y, y\rangle &\leq C_4(1 + |a|^2)|y|^2 \quad \text{for all } y \in \mathbb{R}^d, \end{aligned}$$

with $C_i > 0$ for $i = 1, \ldots, 4.$

(IC) The initial condition m_0^* satisfies $\mathrm{supp}(m_0^*) \subset \overline{B}_\infty(0, C^*)$ for some $C^* > 0,$ where $\mathrm{supp}(m_0^*)$ denotes the support of m_0^*.

Remark 1.2.1 *Notice that we do not impose differentiability on F and G in assumption **(FG)**. Under assumption **(L)**, it follows that H belongs to class C^1. Additionally, this assumption permits the consideration of Hamiltonians in the form*

$$H(x, p) = \kappa(x)|p|^2 + \langle b(x), p\rangle. \tag{1.2.5}$$

Here, $\kappa : \mathbb{R}^d \to \mathbb{R}$ is a C^2 function with bounded first and second order derivatives, such that $\underline{\kappa} \leq \kappa(x) \leq \overline{\kappa}$ for all $x \in \mathbb{R}^d$, $\underline{\kappa}$, $\overline{\kappa} \in (0, \infty)$, and $b : \mathbb{R}^d \to \mathbb{R}^d$ is bounded, a C^2 function, and possesses bounded first and second order derivatives.

Theorem 1.2.1 *Assume **(FG), (L)** and **(IC)** and let $p \in (1, \infty]$. Then system* (1.2.1) *admits a solution (u^*, m^*), in the sense of Definition 1.2.1, such that $m^* \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d)) \cap L^p([0, T] \times \mathbb{R}^d)$ and, for $\widetilde{C} > 0$*

$$\|m^*(t, \cdot)\|_{L^p(\mathbb{R}^d)} \leq \widetilde{C}\|m_0^*\|_{L^p(\mathbb{R}^d)} \quad \textit{for all } t \in [0, T]. \tag{1.2.6}$$

If, in addition, the monotonicity conditions (1.1.13)-(1.1.14) *hold, and that, for all $\zeta \in \mathcal{P}_1(\mathbb{R}^d)$, the functions $F(\cdot, \zeta)$ and $G(\cdot, \zeta)$ are differentiable, then (u^*, m^*) is unique.*

A proof of existence, under slightly different assumptions, can be found, for instance, in [55, Section 1.3.4]. The L^p estimate (1.2.6) follows from the results of Chapter 2.

We present next the key properties of solutions to HJB and the continuity equations under our assumptions within the MFG system (1.2.1). In our numerical approach to solve (1.2.1), we will ensure that our scheme preserves these properties.

Let us start with the definition of viscosity solution to the HJB equation. Viscosity solutions were introduced by Crandall and Lions [70] (also see Crandall, Evans, and Lions [69]) for equations of the HJB type. This approach stemmed from a method called *vanishing viscosity*, aiming to define a weak version of solutions suitable for nonlinear partial differential equations. These weak solutions align with classical solutions if they possess sufficient regularity. The HJB equation

$$-\partial_t u + H(x, D_x u(t,x)) = 0 \quad \text{in } [0,T] \times \mathbb{R}^d \tag{1.2.7}$$

does not admit, in general, a smooth solution even for smooth Hamiltonians. Crandall and Lions [70] showed that, as $\varepsilon \downarrow 0$ the limit of smooth solutions $\{u^\varepsilon \,|\, \varepsilon > 0\}$ of the perturbed equation

$$-\partial_t u^\varepsilon + H(x, D_x u^\varepsilon(t,x)) = \varepsilon \Delta u^\varepsilon \quad \text{in } [0,T] \times \mathbb{R}^d$$

should satisfy a set of conditions which were used to define the notion of viscosity solutions. A comparison principle was also shown under rather general assumptions, implying the uniqueness of such a solution.

Definition 1.2.3 *A function $u \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ is a viscosity solution of* (1.2.7) *if:*

- *u is a subsolution, meaning for every $(t,x) \in [0,T] \times \mathbb{R}^d$ and each test function $\phi \in C^1([0,T] \times \mathbb{R}^d)$ such that $u - \phi$ has a strict maximum at a point (t^*, x^*), we have*

$$-\partial_t \phi(t^*, x^*) + H(x^*, D_x \phi(t^*, x^*)) \leq 0.$$

- *u is a supersolution, meaning for every $(t,x) \in [0,T] \times \mathbb{R}^d$ and each test function $\phi \in C^1([0,T] \times \mathbb{R}^d)$, such that $u - \phi$ has a strict minimum at a point (t_*, x_*), we have*

$$-\partial_t \phi(t_*, x_*) + H(x_*, D_x \phi(t_*, x_*)) \geq 0.$$

Given $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, we consider the HJB equation

$$\begin{cases} -\partial_t u(t,x) + H(x, D_x u(t,x)) = F(x,\mu(t)) & \text{for } (t,x) \in]0,T[\times\mathbb{R}^d, \\ u(x,T) = G(x,\mu(T)) & \text{for } x \in \mathbb{R}^d, \end{cases} \tag{HJB}$$

where H is given by (1.2.2).

Under assumptions **(FG)** and **(L)**, it follows from [23], [71] that equation (HJB) admits a unique viscosity solution $u[\mu]$. Moreover, for every $t \in [0,T[$, $x \in \mathbb{R}^d$

$$\begin{aligned} u[\mu](t,x) = \inf_{\alpha \in L^2([t,T];\mathbb{R}^d)} \int_t^T & \big(L(X^{x,t}[\alpha](s), \alpha(s)) + F(X^{x,t}[\alpha](s), \mu(s))\big)\,\mathrm{d}s \\ & + G(X^{x,t}[\alpha](T), \mu(T)) \end{aligned} \tag{1.2.8}$$

where $X^{x,t}[\varphi](s) = x + \int_t^s \alpha(r)dr$.

We present now the main properties of $u[\mu]$.

Proposition 1.2.1 *Let $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. Then, under assumptions* **(FG)** *and* **(L)**, *there exist $C_{\text{Lip}}, C_{\text{sc}} > 0$, independent of μ, such that*

$$\big|u[\mu](t,x) - u[\mu](t,y)\big| \le C_{\text{Lip}}|x-y| \tag{1.2.9}$$

$$u[\mu](t,x+y) - 2u[\mu](t,x) + u[\mu](t,x-y) \le C_{\text{sc}}|y|^2, \tag{1.2.10}$$

for all $t \in [0,T]$, x, $y \in \mathbb{R}^d$.

Under **(FG)** and **(L)**, we show the existence of an optimal control $\alpha^{t,x}$ of the value function u in (1.2.8). Thereafter, using (1.2.9), we show the existence of $C > 0$, independent of (μ,t,x), such that

$$\|\alpha^{t,x}\|_{L^\infty([0,T];\mathbb{R}^d)} \le C^*. \tag{1.2.11}$$

And hence, by (1.2.11), $u[\mu]$ is also characterized by the HJB equation

$$\begin{cases} -\partial_t u(t,x) + H_{\text{b}}(x, D_x u(t,x)) = F(x,\mu(t)) & \text{for } (t,x) \in (0,T)\times\mathbb{R}^d, \\ u(x,T) = G(x,\mu(T)) & \text{for } x \in \mathbb{R}^d, \end{cases} \tag{1.2.12}$$

where

$$H_{\text{b}}(x,p) = \sup_{a \in \overline{B}(0,C^*)} \big\{\langle a,p\rangle - L(x,a)\big\} \quad \text{for all } x,\ p \in \mathbb{R}^d. \tag{1.2.13}$$

This reformulation of the HJB equation will allow us to simplify the analysis of the scheme that we will discuss in section 1.4.

We now consider the continuity equation

$$\begin{cases} \partial_t m - \operatorname{div}(m H_p(x, D_x u[\mu]) = 0 & \text{for } (t,x) \in (0,T) \times \mathbb{R}^d \\ m(0,x) = m_0^*(x) & \text{for } x \in \mathbb{R}^d. \end{cases} \tag{C.E}$$

Proposition 1.2.2 *Assume **(FG), (L)** and **(IC)**. Then,* (C.E) *admits a solution* $m \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d)) \cap L^{\mathsf{p}}([0,T] \times \mathbb{R}^d)$ *in the sense of* (1.2.3)*. Moreover, there exists* $\widetilde{C} > 0$ *such that*

$$\|m(t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \widetilde{C} \|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \textit{for all } t \in [0,T]. \tag{1.2.14}$$

If, in addition, for every $t \in [0,T]$*, the functions* $F(\cdot, \mu(t))$ *and* $G(\cdot, \mu(T))$ *are differentiable, then the solution* m *to* (C.E) *is unique.*

We recall, again, that the L^{p} estimate (1.2.14) is a consequence of results presented in Chapter 2 where we introduce a Lagrange-Galerkin scheme to solve (1.2.1). The main argument to prove the convergence of this scheme is based on showing an estimate of the form (1.2.14) for the discrete version of (C.E).

1.3 Our contribution

The main results of this thesis are divided into three chapters, each corresponding to a different approach aimed to address a specific approximation of mean field game problems.
In Chapter 2, we introduce a Lagrange-Galerkin discretization approach for first-order mean field games and provide a convergence result for the scheme's solutions in general state dimensions. Additionally, in seeking an implementable version of the scheme, we discuss its connections with previous contributions in the literature.
Beyond researchers focusing on the numerical approximation of mean field games, this chapter might also appeal to those working on approximating continuity equations involving irregular vector fields, and it has been published in 2023 on SIAM Journal of numerical analysis [56].

In Chapter 3, our focus remains on the deterministic problem of mean field games, but we consider a more general scenario. Here, the dynamics of a typical agent are nonlinear with respect to the state variable and affine with respect to the control variable.

Additionally, the cost functional exhibits polynomial growth concerning the state variable. First we recall a relaxed definition of the MFG equilibrium based on a Lagrangian point of view to the problem. Then, the MFG problem is approximated using a discrete-time and finite state space MFG introduced by Gomes, Mohr and Souza [84]. To compute solutions, we employ a fictitious play algorithm for solving our discrete MFG problem. This algorithm converges under the assumption of monotonicity on the coupling terms. This chapter builds upon the work of Hadikhanloo and Silva [94] and Gianatti and Silva [80], extending their results to encompass scenarios with polynomial growth in data, and it has been published in 2023 on Journal of dynamics and games [81].

Chapter 4 deals with solving (1.1.12) when $\nu > 0$ and the coupling is local (see Remark 1.1.1) using Newton iterations. The Newton iterations are directly applied to continuous mean field games, yielding a system of two coupled linear partial differential equations to be solved at each iteration. Under some conditions on the date, it is shown in [44] (see also [31] for the case of stationary MFG system) that the Newton iterations exhibit quadratic convergence towards a solution to (1.1.12). We discretize the new system of partial differential equations using two approaches: a finite difference scheme and a semi-Lagrangian scheme. The main purpose of this chapter is to compare these two approaches through numerical tests and it is still a work in progress with E. Carlini and F. J. Silva.

We summarize in the following sections our main contributions.

1.4 A Lagrange-Galerkin scheme for first order mean field games systems

We highlight here the key findings from Chapter 2, aiming to devise a convergent numerical scheme for the first order MFG system in arbitrary dimensions. Our objective is to formulate a scheme that preserves the fundamental properties of solutions to the HJB and continuity equations discussed previously.

To achieve this, we propose a Lagrange-Galerkin scheme for the continuity equation coupled with a semi-Lagrangian scheme for the HJB equation.
The Lagrange-Galerkin method, initially introduced in [100], [128], combines features of both the method of characteristics and the finite element method. It involves discretizing the space-time domain using finite elements to approximate the solution of the PDE while simultaneously tracing characteristic trajectories to capture transport-related information.

The semi-Lagrangian method was initially introduced for Vlasov-Poisson equations by Cheng and Knorr in 1976 [65]. The monograph [76] presents a unified framework of semi-Lagrangian strategies for approximating hyperbolic partial differential equations, with a specific focus on HJB equations.

The relationship between semi-Lagrangian and Lagrange-Galerkin schemes has been analyzed in [77], [78].

In the context of MFG theory, Carlini and Silva present in their paper [57] a fully discrete semi-Lagrangian method for the first order MFG system with a quadratic Hamiltonian, specifically $H(x,p) = |p|^2/2$ for $x, p \in \mathbb{R}^d$. Consequently, the corresponding MFG system is expressed as follows:

$$\begin{cases} -\partial_t u + \frac{1}{2}|Du|^2 = F(x, m(t)) & \text{in } (0,T)\times\mathbb{R}^d, \\ \partial_t m(t,x) - \operatorname{div}(mDu) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ m(0,x) = m_0(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{1.4.1}$$

In the proposed scheme, the HJB equation above is discretized using a semi-Lagrangian approximation, while the continuity equation is approximated by a scheme that is dual to a linearized version of the HJB equation scheme. A convergence result towards a solution to (1.4.1) is established when $d = 1$.

The results concerning the semi-Lagrangian scheme for the HJB equation presented in Chapter 2 follow a similar approach to those outlined in [57] and [76]. However, we suppose that **(FG), (L)** and **(IC)** hold, so no differentiability assumption is imposed on the functions F and G and also we can consider more general Hamiltonians.

Given $\mu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, we first describe the semi-Lagrangian approximation of $u[\mu]$, the viscosity solution to (HJB).
For this, let $N_t, N_s \in \mathbb{N}$ with $N_s \geq N_t$. Define $\Delta t = 1/N_t$ and $\Delta x = 1/N_s$ as the time and space steps, respectively. Moreover, let $\mathcal{I}_{\Delta t} = \{0,\ldots,N_t\}$, $\mathcal{I}^*_{\Delta t} = \mathcal{I}\setminus\{N_t\}$, and define the space grid $\mathcal{G}_{\Delta x} = \{i\Delta x \,|\, i \in \mathbb{Z}^d\}$.

The semi-Lagrangian scheme to approximate $u[\mu]$ constructed from the dynamic programming principle (1.1.7) is given by: find $\{u_k\colon \mathcal{G}_{\Delta x} \to \mathbb{R} \,|\, k \in \mathcal{I}_{\Delta t}\}$ such that

$$\begin{cases} u_{k,i} = \inf_{a\in\mathbb{R}^d} [\Delta t L(x_i, a) + I^1[u_{k+1}](x_i - \Delta t a)] + \Delta t F(x_i, \mu(t_k)), & \text{for all } k \in \mathcal{I}^*_{\Delta t},\ i \in \mathbb{Z}^d, \\ u_{N,i} = G(x_i, \mu(T)) & \text{for all } i \in \mathbb{Z}^d, \end{cases} \tag{1.4.2}$$

where $I^1\colon \mathcal{G}_{\Delta x} \to \mathbb{R}^d$ is the interpolation operator defined by

$$I^1[\phi](x) = \sum_{i\in\mathbb{Z}^d} \beta_i^1(x)\phi_i \quad \text{for all } x \in \mathbb{R}^d,$$

with $\phi_i = \phi(x_i)$ and $\{\beta_i^1 \,|\, i \in \mathbb{Z}^d\}$ is a $\mathbb{Q}^1$ linear basis satisfying $\beta_i^1(x_j) = \delta_{ij}$ for $x_j \in \mathcal{G}_{\Delta x}$. The scheme (1.4.2) is shown to be consistent, monotone, and stable. Given $u_{k,i}$, solution to (1.4.2), and $(\Delta t, \Delta x) \in]0,\infty[^2$, let us set

$$u^{\Delta t,\Delta x}[\mu](t_k, x) = I^1[u_k](x) \quad \text{for all } k \in \mathcal{I}_{\Delta t},\ x \in \mathbb{R}^d.$$

We extend this definition to $[0,T] \times \mathbb{R}^d$, by setting

$$u^{\Delta t,\Delta x}[\mu](t, x) = u^{\Delta t,\Delta x}[\mu](t_k, x) \quad \text{if } t \in [t_k, t_{k+1}[,\ k \in \mathcal{I}^*_{\Delta t}. \tag{1.4.3}$$

Then, we show that $u^{\Delta t,\Delta x}[\mu]$ possesses the key properties of the solution to the HJB equation, namely a Lipschitz property and a discrete semi-concavity property. Indeed, we show the existence of $\widetilde{C}_{\text{Lip}}, \widetilde{C}_{\text{sc}} > 0$ independent of $(\mu, \Delta t, \Delta x)$ such that

$$\left|u^{\Delta t,\Delta x}[\mu](t,x) - u^{\Delta t,\Delta x}[\mu](t,y)\right| \le \widetilde{C}_{\text{Lip}}|x-y| \quad \text{for all } t \in [0,T],\ x,\ y \in \mathbb{R}^d, \tag{1.4.4}$$

and

$$\begin{aligned} u^{\Delta t,\Delta x}[\mu](t, x+x_i) - 2u^{\Delta t,\Delta x}[\mu](t,x) + u^{\Delta t,\Delta x}[\mu](t, x-x_i) \\ \le \widetilde{C}_{\text{sc}}|x_i|^2 \quad \text{for all } t \in [0,T],\ x \in \mathbb{R}^d,\ i \in \mathbb{Z}^d. \end{aligned} \tag{1.4.5}$$

Let $\rho \in C^\infty(\mathbb{R}^d)$ with $\rho \ge 0$ and $\int_{\mathbb{R}^d} \rho(x) = 1$. For $\varepsilon > 0$, let $\rho_\epsilon(x) = \rho(x/\epsilon)\epsilon^d \in \mathbb{R}^d$ and, for $\Delta = (\Delta t, \Delta x, \varepsilon)$ and $t \in [0,T]$, we define

$$u^\Delta[\mu](t,\cdot) = \rho_\varepsilon * u^{\Delta t,\Delta x}[\mu](t,\cdot). \tag{1.4.6}$$

We show also that $u^\Delta[\mu]$ preserve the Lipshitz property (1.2.9) and its Hessian satisfies the following estimate:

$$\left\langle D_x^2 u^\Delta[\mu](t,x)y,y\right\rangle \leq C\left(1+\frac{(\Delta x)^2}{\varepsilon^4}\right)|y|^2 \quad \text{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d. \tag{1.4.7}$$

We consider next the discretization of the continuity equation:

$$\begin{cases} \partial_t m - \operatorname{div}(D_pH(x,D_xu^\Delta[\mu])m) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ m(0,x) = m_0(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{1.4.8}$$

where we recall that by **(IC)**, $m_0 \in L^p(\mathbb{R}^d)$ for some $p\in(1,\infty]$, and there exists $C^* > 0$, such that

$$\operatorname{supp}(m_0) \subset \overline{B}_\infty(0,C^*). \tag{1.4.9}$$

For $s\in[0,T)$, $t\in[s,T]$, and $x\in\mathbb{R}^d$, set $\Phi^\Delta[\mu](s,t,x) = X(t)$, where X is the unique solution of

$$\begin{cases} \dot{X}(r) &= -D_pH(X(r),D_xu^\Delta[\mu](r,X(r))) \quad \text{a.e } r\in(s,T), \\ X(s) &= x. \end{cases}$$

Under **(FG), (L), (IC)**, (1.4.8) admits a unique solution $m^\Delta[\mu]\in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$. Moreover, it follows from [18, Proposition 8.1.8] that, for every Borel function $\varphi\colon \mathbb{R}^d\to\mathbb{R}$ with $\varphi(\Phi^\Delta[\mu](t,\cdot))$ integrable, we have

$$\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}m^\Delta[\mu](x) = \int_{\mathbb{R}^d}\varphi(\Phi^\Delta[\mu](t,x))\mathrm{d}m_0(x). \tag{1.4.10}$$

For $k\in\mathcal{I}^*_{\Delta t}$ and $x\in\mathbb{R}^d$, let $\Phi^\Delta_k[\mu]$ be the explicit Euler approximation of $\Phi^\Delta[\mu](t_k,t_{k+1},x)$, i.e

$$\Phi^\Delta_k[\mu](x) = x - \Delta t D_pH\big(x, D_xv^\Delta[\mu](t_k,x)\big). \tag{1.4.11}$$

The semi-discrete approximation of (1.4.10) (see e.g [42]) is given by

$$\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}m_{k+1}(x) = \int_{\mathbb{R}^d}\varphi(\Phi^\Delta_k[\mu](x))\mathrm{d}m_k(x). \tag{1.4.12}$$

Starting from this semi-discrete approximation, our goal is to find an approximation of $m^\Delta[\mu]$, denoted as $\mathsf{M}^\Delta[\mu]$, using a Galerkin projection, which preserves the main properties of the solution $m^\Delta[\mu]$ to (1.4.8). In other words, given finite element basis

$\{\beta_i \mid i \in \mathbb{Z}^d\}$, we consider approximation $\mathsf{M}^\Delta[\mu]$ having the form

$$\mathsf{M}^\Delta[\mu](t_k, x) = \sum_{j\in\mathbb{Z}^d} m_{k,j}\beta_j(x), \quad \text{for all } k \in \mathcal{I}_{\Delta t},\ x \in \mathbb{R}^d, \tag{1.4.13}$$

where the constants $\{m_{k,j} \mid k \in \mathcal{I}_{\Delta t}, j \in \mathbb{Z}^d\}$ have to be determined. Since we do not expect $m^\Delta[\mu]$ to be regular, we set, for $i \in \mathbb{Z}^d$, $\beta_i := \beta_i^0 = \mathbb{I}_{E_i}$, where

$$E_i = \{x \in \mathbb{R}^d \mid \|x - x_i\|_\infty \leq \Delta x/2\}.$$

In order to determine the constants $m_{k,j}$, we replace m_k and m_{k+1} in (1.4.12) by $\mathsf{M}^\Delta[\mu](t_k, \cdot)$ and $\mathsf{M}^\Delta[\mu](t_{k+1}, \cdot)$, respectively, and, given $i \in \mathbb{Z}^d$, we take $\varphi = \beta_i$ to obtain the following Galerkin-Lagrange scheme:

$$\begin{cases} m_{k+1,i} = \dfrac{1}{(\Delta x)^d} \displaystyle\sum_{j\in\mathbb{Z}^d} m_{k,j} \int_{E_j} \beta_i^0(\Phi_k^\Delta[\mu](x))\mathrm{d}x & \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i \in \mathbb{Z}^d, \\ m_{0,i} = \dfrac{1}{(\Delta x)^d} \displaystyle\int_{E_i} m_0^*(x)\mathrm{d}x & \text{for all } i \in \mathbb{Z}^d. \end{cases} \tag{1.4.14}$$

In order to provide an interpretation of the scheme, notice that the integral term in the first equation of (1.4.14) can be written as

$$\int_{E_j} \beta_i^0(\Phi_k^\Delta[\mu](x))\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{I}_{E_j \cap \Phi_k^\Delta[\mu]^{-1}(E_i)}(x)\mathrm{d}x = \mathcal{L}^d\Big(E_j \cap \Phi_k^\Delta[\mu]^{-1}(E_i)\Big), \tag{1.4.15}$$

where $\mathcal{L}^d$ denotes the Lebesgue measure in $\mathbb{R}^d$. Figure 1.2 illustrates how is computed, in dimension 2, the term (1.4.15) that multiplies $m_{k,j}$ in (1.4.14).

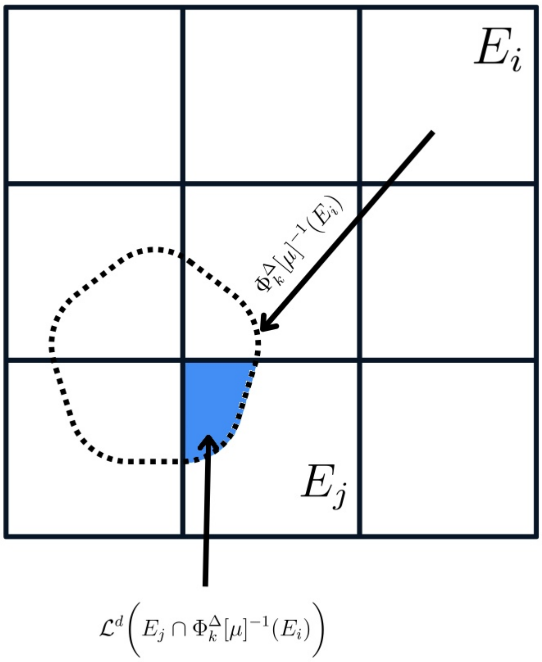


Figure 1.2: The surface of the blue area is given by $\mathcal{L}^2(E_j \cap \Phi^\Delta[\mu]^{-1}(E_i))$.

Remark 1.4.1 *Plugging* (1.4.15) *in the first equation of* (1.4.14) *yields the scheme in* [127, Section 2.2]*. Our primary findings regarding solutions to* (1.4.14)*, outlined in Chapter 2, stand apart from those in* [127], [133]*. As a result, the analysis in Chapter 2 complements the study conducted in* [127], [133] *for the approximation* (1.4.14) *of continuity equations.*

The (LG) scheme (1.4.14) is explicit, yielding a unique solution $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$. Moreover, for each discrete time $k \in \mathcal{I}_{\Delta t}$, the set $\{m_{i,k} \,|\, i \in \mathbb{Z}^d\}$ defines a probability measure over $\mathbb{R}^d$ which, as can be shown from (1.4.9), has a compact support.

The key point to prove the convergence of the discrete mean field game scheme is to show that the LG scheme preserves the L^p uniform estimate (1.2.6). Indeed, given constants $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$ computed with (1.4.14), and extending $\mathsf{M}^\Delta[\mu]$, given by (1.4.13), to $[0,T] \times \mathbb{R}^d$ by

$$\mathsf{M}^\Delta[\mu](t,x) = \left(\frac{t_{k+1}-t}{\Delta t}\right)\mathsf{M}^\Delta[\mu](t_k,x) + \left(\frac{t-t_k}{\Delta t}\right)\mathsf{M}^\Delta[\mu](t_{k+1},x)$$
$$\text{for all } k \in \mathcal{I}^*_{\Delta t},\ t \in [t_k,t_{k+1}),\ x \in \mathbb{R}^d, \quad (1.4.16)$$

then, if $\Delta x = O(\Delta t)$ and $\Delta t = O(\varepsilon^2)$, there exists $\widetilde{C} > 0$, independent of (Δ, μ), such that $\mathsf{M}^\Delta[\mu](t, \cdot) \in L^{\mathsf{p}}(\mathbb{R}^d)$ for all $t \in [0, T]$ and

$$\|\mathsf{M}^\Delta[\mu](t, \cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \widetilde{C}\|m_0\|_{L^{\mathsf{p}}(\mathbb{R}^d)}.$$

The proof relies on the semi-concavity estimate (1.4.7).

Finally, the first order MFG system (1.2.1) is discretized as follows:

$$\text{Find } \mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d)) \text{ such that } \mu = \mathsf{M}^\Delta[\mu](t_k, x_i) \quad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ i \in \mathbb{Z}^d. \tag{MFG$^\Delta$}$$

We establish the existence of solutions to (MFG^Δ) under the condition that $\Delta t/\varepsilon$ is sufficiently small, employing Brouwer's fixed-point theorem. The convergence of (MFG^Δ) is ensured in the general state dimension by the following theorem:

Theorem 1.4.1 *Assume that **(FG), (L)** and **(IC)** hold. Let $\Delta_n = (\Delta t_n, \Delta x_n, \varepsilon_n)$ be such that, as $n \to \infty$, $\Delta x_n/\Delta t_n \to 0$, and $\Delta t_n = O(\epsilon_n^2)$. Consider the corresponding sequence m^n of solutions to equation $(\mathrm{MFG}^{\Delta_n})$ and set $u^n = u^{\Delta_n}[m^n]$. Then there exists (u^*, m^*), solution to system* (1.2.1)*, such that, up to a subsequence, the following hold:*

1. *$(u^n)_{n\in\mathbb{N}}$ converges to u^*, uniformly over compact subsets of $[0,T] \times \mathbb{R}^d$.*

2. *$(m^n)_{n\in\mathbb{N}}$ converges in $C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$ towards m^*. Moreover, the convergence also holds weakly in $L^{\mathsf{p}}([0,T] \times \mathbb{R}^d)$, if $\mathsf{p} < \infty$, and weakly* in $L^\infty([0,T] \times \mathbb{R}^d)$, if $\mathsf{p} = \infty$.*

An implementable version of the scheme

As the LG scheme (1.4.14) involves certain integrals depending on discrete characteristics of the equation, we approximate them using a method known as *area weighting*, first introduced in [124]. The main idea of the area weighting technique is to replace, for each $k \in \mathcal{I}^*_{\Delta t}$, the local nonlinear discrete flow $E_i \ni x \mapsto \Phi_k^\Delta[\mu](x) \in \mathbb{R}^d$ defined in (1.4.11), by the locally affine approximation

$$E_i \ni x \mapsto \overline{\Phi}_k^\Delta[\mu](x) = x - \Delta t D_p H\big(x_i, D_x u^\Delta[\mu](t_k, x_i)\big) \in \mathbb{R}^d.$$

Notice that $\overline{\Phi}_k^\Delta[\mu](x) = x - x_i + \Phi_k^\Delta[\mu](x_i)$ for all $x \in E_i$. Under this approximation, we can compute the integrals in (1.4.14) explicitly, to get

$$\frac{1}{(\Delta x)^d}\int_{E_j} \beta_i^0(\overline{\Phi}_k^\Delta[\mu](y))\mathrm{d}y = \beta_i^1(\Phi_k^\Delta[\mu](x_j)),$$

which yields, surprisingly, to following scheme initially proposed in [57]:

$$\begin{aligned} \overline{m}_{k+1,i} &= \sum_{j\in\mathbb{Z}^d} \overline{m}_{k,j}\beta_i^1(\Phi_k^\Delta[\mu](x_j)) \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i \in \mathbb{Z}^d, \\ \overline{m}_{0,i} &= m_{0,i}^* \quad \text{for all } i \in \mathbb{Z}^d, \end{aligned}$$

where we recall that $\{\beta_i^1 \mid i \in \mathbb{Z}^d\}$ is a $\mathbb{Q}_1$ finite element basis defined on $\mathcal{G}_{\Delta x}$.
A uniform error estimate on the difference between the two schemes in our MFG setting remains an open question.

To solve the discrete problem (MFG^Δ), we use (damped) Picard iterations. Let us point that the convergence of this method has not been shown.

1.5 Approximation of deterministic mean field games under polynomial growth conditions on the data

In all that we have discussed previously, we have considered the simplest dynamic of a typical player, given by $\dot{\gamma}(t) = \alpha(t)$ where α is the control.
In Chapter 3, we tackle a more general case, where, given $m \in C\left([0,T];\mathcal{P}_1(\mathbb{R}^d)\right)$, a typical player positioned at x at time $t=0$, solves an optimal control problem of the form

$$\begin{cases} \inf \displaystyle\int_0^T L\left(s,\alpha(s)\right) + F(\gamma(s),m(s))\,\mathrm{d}s + G(\gamma(T),m(T)) \\ \text{such that} \quad \dot{\gamma}(s) = A(s,\gamma(s)) + B(s)\alpha(s) \quad \text{a.e. } s \in (0,T), \\ \qquad\qquad\quad \gamma(0) = x, \\ \qquad\qquad\quad \gamma \in W^{1,p}([0,T];\mathbb{R}^d),\ \alpha \in L^p([0,T];\mathbb{R}^r). \end{cases} \tag{1.5.1}$$

Unlike Chapter 2, we will not employ the system of partial differential equations to describe the MFG equilibrium. Instead, we will use a relaxed equilibrium notion in the deterministic case, called the *Lagrangian equilibrium* which has been recently studied in [28], [46], [54], [93].

For that, let $\Gamma = C([0,T],\mathbb{R}^d)$, and define

$$\mathcal{P}_{m_0^*}(\Gamma) = \{\xi \in \mathcal{P}_1(\Gamma) \mid e_0\sharp\xi = m_0^*\},$$

where the evolution map $e_t\colon \Gamma \to \mathbb{R}^d$ is given by $e_t(\gamma) = \gamma(t)$ for all $\gamma \in \Gamma$. The Lagrangian MFG equilibrium consists in solving the following problem:

Problem 1 *Find $\xi^* \in \mathcal{P}_{m_0^*}(\Gamma)$ such that $[0,T] \ni t \mapsto e_t\sharp\xi^* \in \mathcal{P}_1(\mathbb{R}^d)$ belongs to $C\left([0,T];\mathcal{P}_1(\mathbb{R}^d)\right)$ and for ξ^*-a.e. $\gamma^* \in \Gamma$ there exists $\alpha^* \in L^p([0,T];\mathbb{R}^r)$ such that (γ^*,α^*) solves* (1.5.1) *with $x = \gamma^*(0)$ and $m(t) = e_t\sharp\xi^*$ for all $t \in [0,T]$.*

The interpretation of a Lagrangian MFG equilibrium is as follows: the measure ξ^* concentrated only over trajectories in $\mathbb{R}^d$, distributed according to m_0^* at the initial time, minimizing a cost that depends on the set of temporal marginals of ξ^* within the interval $[0,T]$.

The main goal of Chapter 3 is to approximate the MFG Lagrangian equilibrium associated to the variational problem (1.5.1). The existence has been demonstrated in [8], [47] under appropriate assumptions on $L, F,$ and G. In [80], Gianatti and Silva proposed an approximation of solution to Problem 1 by analogous problems in discrete time and finite state space (see [84]). In Chapter 3, we consider the same approximation, but we suppose more general asuumptions on the cost function. Specifically, we consider cost functionals that allow polynomial growth with respect to both state and control variables. This type of cost arises in numerous applications, making the study of such problems appealing. A convergence study of the approximated MFG Lagrangian equilibrium is provided.

The discrete setting

In order to keep the introduction short and simple, we will assume the simplest dynamics, i.e., $A = 0$ and $B = I_d$, allowing us to present a clear and understandable idea about our scheme (see [94]). The key point for this is to offer an ad-hoc approximation of the optimal control problem (1.5.1).

First, we show, under appropriate assumptions, the existence of α^* that solves the variational problem (1.5.1) and $C > 0$ such that

$$|\alpha^*(s)| \leq C_b(1+|x|) \quad \text{for all } (s,x) \in [0,T]\times\mathbb{R}^d. \tag{1.5.2}$$

Then, adopting the semi-Lagrangian scheme presented in Chapter 2, and taking into account the estimate (1.5.2), we estimate the value function of a typical player as follows:

$$V_k(x) = \min_{a\in\overline{B}_\infty(0,C_b(1+|x|))} \Big\{\Delta t\big[L(t_k,a)+F(x,m(t_k))\big] + I[V_{k+1}](\Phi(k,x,a))\Big\},$$
$$\text{for all } k\in\mathcal{I}^*_{\Delta t},\ x\in\mathcal{G}_{\Delta x}, \tag{1.5.3}$$
$$V_{N_t}(x) = G(x,m(T)) \quad \text{for all } x\in\mathcal{G}_{\Delta x},$$

where $\overline{B}_\infty(0,R)$ denotes the corresponding closed ball centered at 0 and of radius R, C_b is giving in (1.5.2), I is a $\mathbb{P}_1$ interpolation operator and

$$\Phi(k,x,a) = x+\Delta t a \quad \text{for all } k\in\mathcal{I}^*_{\Delta t},\ x\in\mathcal{G}_{\Delta x},\ a\in\mathbb{R}^d. \tag{1.5.4}$$

Our focus revolves around two primary concepts:

- The estimate (1.5.2) allows us to consider a finite and bounded grid rather.
- Then we can consider controls such that $x+\Delta t a$ is a grid point, i.e controls having the form
$$a = \frac{y-x}{\Delta t}, \quad y\in\mathcal{G}_{\Delta x}, \tag{1.5.5}$$
in order to avoid the interpolation in (1.5.3).

Set $\mathcal{K}_0 = \text{supp}(m_0^*)$. We consider time dependent state grids $\{\mathcal{S}_k \mid k\in\mathcal{I}_{\Delta t}\}$, constructed as follows: let $\alpha(k,x_i,y) = \frac{y-x_i}{\Delta t}$, we define

$$\begin{aligned}\mathcal{S}(x) =& \{y\in\mathcal{G}_{\Delta x} \mid |\alpha(k,x,y)| \le C_b(1+|x|\},\\ \mathcal{S}_0 =& \mathcal{G}_{\Delta x}\cap\mathcal{K}_0,\\ \mathcal{S}_{k+1} =& \bigcup_{x\in\mathcal{S}_k}\mathcal{S}(x) \quad \text{for all } k\in\mathcal{I}^*_{\Delta t}.\end{aligned}$$

Figure 1.3 explains the construction of the time dependent grids $(\mathcal{S}_k)_{k\in\mathcal{I}_{\Delta t}}$.

It follows from Grönwall's inequality that the sequence of grids $(\mathcal{S}_k)_{k\in\mathcal{I}_{\Delta t}}$ is uniformly bounded with respect to the discretization parameters. More precisely, there exists a compact set $\mathcal{K}\subset\mathbb{R}^d$ independent of Δt and Δx such that, if $\Delta x/\Delta t\le 1$, then

$$\mathcal{S}_k\subset K \quad \text{for all } k\in\mathcal{I}_{\Delta t}. \tag{1.5.6}$$

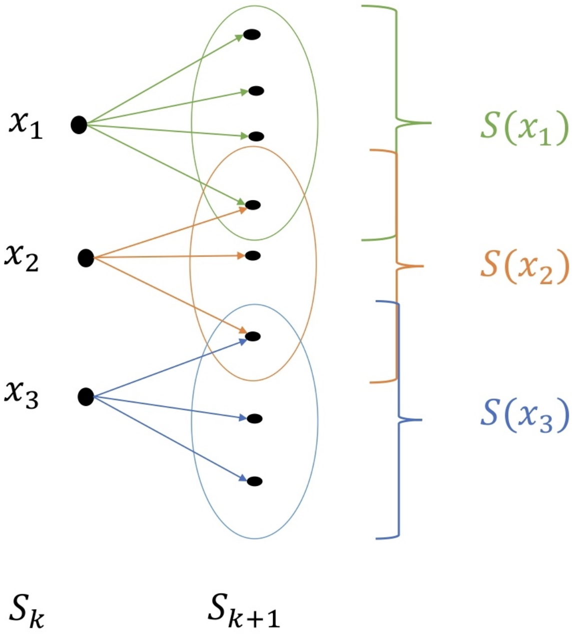


Figure 1.3: The construction of the grid $\mathcal{S}_k$.

We consider the following scheme, which is a variation of (1.5.3)

$$\begin{aligned} \mathcal{V}_k(x) &= \min_{p \in \mathcal{P}(\mathcal{S}_{k+1}(x))} \sum_{y \in \mathcal{S}_{k+1}(x)} p(y) \Big[\Delta t [L(t_k, \alpha(k,x,y)) + F(x, m(t_k))] + \mathcal{V}_{k+1}(y) \Big] \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ x \in \mathcal{S}_k, \\ \mathcal{V}_{N_t}(x) &= G(x, m(T)) \quad \text{for all } x \in \mathcal{S}_{N_t}. \end{aligned} \tag{1.5.7}$$

The finite mean field game approximation

Using the elements outlined in the prior subsection, we introduce an approximation of MFG equilibria associated to the optimal control problem (1.5.1). Let's give the main ideas:

- Define first the space of discrete time marginals $\mathcal{M} = \prod_{k \in \mathcal{I}_{\Delta t}} \mathcal{P}(\mathcal{S}_k)$.

- Let $M \in \mathcal{M}$ and $\varepsilon > 0$. By considering the scheme (1.5.7), we define

$$
\begin{aligned}
V_k^M(x) = \min_{p \in \mathcal{P}(\mathcal{S}_{k+1}(x))} & \left\{ \sum_{y \in \mathcal{S}_{k+1}(x)} p(y) \left[\Delta t L(t_k, \alpha(k,x,y), x) + V_{k+1}^M(y) \right] + \varepsilon p(y) \log(p(y)) \right\} \\
& + \Delta t F(x, M_k) \qquad \text{for all } k \in \mathcal{I}^*,\ x \in \mathcal{S}_k, \\
V_{N_t}^M(x) = G(x, M_{N_t}) & \quad \text{for all } x \in \mathcal{S}_{N_t}.
\end{aligned}
\tag{1.5.8}
$$

  The entropy term in (1.5.8) ensures the existence of a unique solution $p_k^M(x,\cdot)$ for the above optimization problem.

- Given $y \in \mathcal{S}_{k+1}$, we set the transition probabilities

$$
P_k^M(x,y) := \begin{cases} p_k^M(x,y) & \text{if } y \in \mathcal{S}_{k+1}(x), \\ 0 & \text{if } y \in \mathcal{S}_{k+1} \setminus \mathcal{S}_{k+1}(x). \end{cases} \tag{1.5.9}
$$

- We define *the best response map* $\mathbf{br}(M) \in \mathcal{M}$ by

$$
\begin{aligned}
\widehat{M}_{k+1}(y) &= \sum_{x \in \mathcal{S}_k} P_k^M(x,y) \widehat{M}_k(x) \quad \text{for all } k \in \mathcal{I}^*,\ y \in \mathcal{S}_{k+1}, \\
\widehat{M}_0(x) &= m_0^*(E(x)) \quad \text{for all } x \in \mathcal{S}_0.
\end{aligned}
\tag{1.5.10}
$$

- The discretization of the Problem 1 reads as follows:

$$
\text{Find } M \in \mathcal{M} \text{ such that } M = \mathbf{br}(M). \tag{1.5.11}
$$

  The existence of a fixed point of $\mathbf{br}$ follows from Brouwer's fixed point theorem. The uniqueness result under the monotonicity assumptions (1.1.13)-(1.1.14) follow from the arguments in the proof of [80, Proposition 4.2].

Next, we explore the convergence of solutions to (1.5.11) towards a solution to Problem 1 as the discretization parameters Δt, Δx, and ε approach zero.

Let $(N_t^n)_{n\in\mathbb{N}} \subset \mathbb{N}$, $(N_s^n)_{n\in\mathbb{N}} \subset \mathbb{N}$, $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$, and, for every $n \in \mathbb{N}$, set $\Delta t_n = T/N_t^n$, $\Delta x_n = 1/N_s^n$, $\mathcal{I}^n = \{0,\dots,N_t^n\}$, $\mathcal{I}^{n,*} := \mathcal{I}^n \setminus \{N_t^n\}$, $t_k^n = k\Delta t_n$ $(k \in \mathcal{I}^n)$, and $\mathcal{G}^n = \{i\Delta x_n \,|\, i \in \mathbb{Z}^d\}$. We assume that $N_s^n \geq N_t^n$. For $k \in \mathcal{I}^{n,*}$ and $x \in \mathcal{G}^n$, we denote by $\mathcal{S}_{k+1}^n(x)$ the set $\mathcal{S}_{k+1}(x)$ defined in (**??**) associated with Δt_n and Δx_n.

Denote by Γ^n the set of continuous functions $\gamma\colon [0,T] \to \mathbb{R}^d$ such that for each $k \in \mathcal{I}_{\Delta t_n}$, $\gamma(t_k^n) \in \mathcal{S}_k^n$ and, for every $k \in \mathcal{I}_{\Delta t_n}^*$, the restriction of γ to the interval $[t_k^n, t_{k+1}^n]$ is affine.

Finally, let $M^n \in \mathcal{M}$ be a solution to (1.5.11) associated with the previous parameters and, recalling (1.5.9), let us define $\xi^n \in \mathcal{P}(\Gamma)$ as

$$\xi^n = \sum_{\gamma \in \Gamma^n} M_0^n(\gamma(0)) P^n(\gamma) \delta_\gamma \in \mathcal{P}(\Gamma), \quad \text{where} \quad P^n(\gamma) := \prod_{k=0}^{N_t^n - 1} P_k^{M^n}(\gamma(t_k^n), \gamma(t_{k+1}^n)). \tag{1.5.12}$$

We extend M^n to the element in $C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$ defined by

$$[0,T] \ni t \mapsto M^n(t) := e_t \sharp \xi^n \in \mathcal{P}_1(\mathbb{R}^d). \tag{1.5.13}$$

We have the following convergence result.

Theorem 1.5.1 *Assume that, as $n \to \infty$, $N_t^n \to \infty$, $N_s^n \to \infty$, $N_t^n / N_s^n \to 0$, and $\varepsilon_n = o\left(1/(N_t^n \log(N_s^n))\right)$. Then there exists a solution ξ^* to* Problem 1 *such that, up to some subsequence, $\xi^n \to \xi^*$ narrowly in $\mathcal{P}(\Gamma)$ and $M^n \to m^* := e_{(\cdot)} \sharp \xi^*$ in $C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$.*

In addition, if the Lasry-Lions monotonicity condition (1.1.13)-(1.1.14) *holds and for every $m \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$ and m_0^*-a.e. $x \in \mathbb{R}^d$ problem* (1.5.1) *admits a unique solution, then the whole sequence $(\xi^n)_{n \in \mathbb{N}}$ converges narrowly towards the unique solution to Problem 1.*

Given discretization parameters $\Delta t, \Delta x$ and ε, a solution M^* to the corresponding mean field games problem can be computed by using the *Fictitious play method* given by Algorithm 1, which, by [94, Theorem 3.2], satisfies $(\mathsf{M}^N, \overline{\mathsf{M}}^N) \underset{N \to \infty}{\longrightarrow} (M^*, M^*)$ if (1.1.13)-(1.1.14) hold.

Algorithm 1 Fictitious play for deterministic MFG

Input: $\overline{\mathsf{M}}^0 \in \mathcal{M}$ arbitrary and a tolerance parameter δ
Output: Approximation of M solving (1.5.11)
$e \leftarrow \delta + 1$
$n \leftarrow 1$
$\overline{\mathsf{M}}^1 \leftarrow \mathsf{M}^0$
while $e > \delta$ **do**
    $\mathsf{M}^{n+1} \leftarrow \mathbf{br}(\overline{\mathsf{M}}^n)$
    $e \leftarrow \|\mathsf{M}^{n+1} - \overline{\mathsf{M}^n}\|_{L^1}$
    $\overline{\mathsf{M}}^{n+1} \leftarrow \frac{n}{n+1}\overline{\mathsf{M}}^n + \frac{1}{n+1}\mathsf{M}^{n+1}$
end while
return $\overline{\mathsf{M}}^{n+1}$

1.6 Newton iterations for Mean Field Games

We consider the following second order mean field game system with local coupling (see Remark 1.1.1):

$$\begin{cases} -\partial_t u - \nu\Delta u + H(x, Du) = F(m(t,x)) & \text{in } [0,T]\times\mathbb{T}^d, \\ \partial_t m - \nu\Delta m - \operatorname{div}(mH_p(x, Du)) = 0 & \text{in } [0,T]\times\mathbb{T}^d, \\ m(0,x) = m_0(x),\ u(T,x) = G(x) & \text{in } \mathbb{T}^d, \end{cases} \tag{1.6.1}$$

where $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ stands as the d-dimensional torus. We opt to define our MFG system (1.6.1) on $\mathbb{T}^d$ to circumvent technicalities related to boundary conditions.
The objective of Chapter 4 is to develop a new scheme to solve system (1.6.1) through the application of Newton iterations.
Newton iterations were initially introduced in the context of MFGs by Achdou et al. in [1]. The authors introduced a discrete MFG system using the finite differences method and employed Newton iterations to solve the resulting discretized non-linear system. In [40], the authors introduced a *policy iteration* method to solve (1.6.1). They interpreted their approach as a *quasi-Newton* method. In [43], the authors showed that the convergence rate of the policy iteration method is linear. This motivation prompted the same authors to apply in [44], Newton iterations to the continuous MFG system (1.6.1) in the following manner:

- First, define $\mathcal{F}$ as follows:

$$\mathcal{F}(u,m) = 0 \ \Leftrightarrow \ \text{(1.6.1) is satisfied} \tag{1.6.2}$$

- The corresponding Newton iterations for (1.6.2) can be expressed as follows:

$$J\mathcal{F}(u^{n-1}, m^{n-1})((u^n, m^n) - (u^{n-1}, m^{n-1})) = -\mathcal{F}(u^{n-1}, m^{n-1}), \tag{1.6.3}$$

where $J\mathcal{F}$ is the Jacobian of $\mathcal{F}$.

- Subsequently, solve the iterative linear system in the unknown (u^n, m^n):

$$\begin{cases} -\partial_t u^n - \nu\Delta u^n + q^n Du^n = q^n Du^{n-1} - H(x, Du^{n-1}) + F(m^{n-1}) + F'(m^{n-1})(m^n - m^{n-1}) \\ \partial_t m^n - \nu\Delta m^n - \operatorname{div}(m^n q^n) = \operatorname{div}(m^{n-1} H_{pp}(Du^{n-1})(Du^n - Du^{n-1})) \\ m^n(x, 0) = m_0(x), \quad u^n(x, T) = u_T(x), \end{cases} \tag{1.6.4}$$

with $q^n = H_p(x, Du^{n-1})$.

The following existence and convergence result for the Newton iteration method holds (see also [31] for the case of a stationary MFG system).

Proposition 1.6.1 *[44] Under suitable assumptions,on the data, there exists a unique solution (u^n, m^n) to system* (1.6.4)*. Moreover there exists $C > 0$ independent of n such that, if (u, m) solves* (1.6.1)*, then*

$$\|u^n - u\|_{C^{1,0}} + \|m^n - m\|_{C^0} \leq C(\|u^{n-1} - u\|_{C^{1,0}} + \|m^{n-1} - m\|_{C^0})^2. \tag{1.6.5}$$

Remark 1.6.1 *Equation* (1.6.5) *means that if we start with an initial guess (u^0, m^0) close enough to the solution (u, m), then $\|u^n - u\|_{C^{1,0}} + \|m^n - m\|_{C^0} \to 0$ with a quadratic rate of convergence.*

Numerical study

The main goal of Chapter 4 is to study the numerical approximation of (1.6.1) through the discretization of the linear system (1.6.4) at each iteration. On the one hand, with a suitable discretization of the system, one can expect promising results in terms of computing time and accuracy due to the quadratic rate of convergence. On the other hand, the linearity of the system simplifies the discretization process. Two distinct methodologies will be considered. The first approach implements an explicit semi-Lagrangian scheme, readily derivable for linear parabolic equations (see e.g [34]), and called Newton-SL. The scheme is showed to be well-posed and a comparative analysis is conducted against the non-linear semi-Lagrangian scheme for system (4.1.1), proposed in [58] and solved via fixed point iterations, which we will refer to as SL-FP. The second method involves a well posed implicit upwind finite differences scheme called

Newton-FD, which, as established through numerical tests, demonstrates a simpler structure and analogous performance to the Newton scheme studied in [1], [4], which will be referred as the FD-Newton.
The Newton iterations process is explained in Algorithm 2, and the primary objective of the chapter is to undertake a comparative analysis of the four mentioned schemes.

Algorithm 2 Solving the discrete Newton iterations system

Input: Initial guesses U^0, M^0, Q^0 and tolerance τ
Output: Approximate solution to (1.6.4)
$n \leftarrow 0$
repeat
    Compute M^{n+1} and U^{n+1} by solving the discrete system Newton-SL or Newton-FD
    err(M) $\leftarrow \|M^{n+1} - M^n\|_\infty$
    err(U) $\leftarrow \|U^{n+1} - U^n\|_\infty$
    $n \leftarrow n + 1$
    Update Q^n, by the derivative of U^n using central differences
until err(M) $< \tau$ **and** err(U) $< \tau$
return M^{n+1}, U^{n+1}

We begin our comparative analysis with two numerical tests in dimension one.
In the first test, we consider $H(x, p) = |p|^2$ and a reference solution is computed in order to conduct a comparative analysis in terms of the uniform norm of the approximation errors. Additionally, we examine the computational time and the required number of Newton iterations. The analysis shows that Newton-SL has a better performance in terms of computational time and number of iterations, and shows comparable accuracy to SL-FP, FD-Newton, and Newton-FD.
We also note that Newton-FD and FD-Newton exhibit comparable performance in terms of accuracy, computational time, and number of iterations. It is noteworthy that Newton-FD uses directly system (1.6.4), eliminating the necessity to define a numerical Hamiltonian as in the FD-Newton scheme. Consequently, Newton-FD emerges as a considerably simpler variant of FD-Newton.

In the second test, we consider $H(x, p) = |p|^2 + V(x)$ for a given potential V, and we change the diffusion parameter, considering first $\nu = 0.4$ and then $\nu = 0.02$.
In the case $\nu = 0.4$, the results show a similar performance for Newton-SL, Newton-FD and FD-Newton. Taking then $\nu = 0.02$, Newton-SL iterations demonstrate convergence reaching the associated threshold. In contrast, both Newton-FD and FD-Newton iterations encounter breakdowns after only few iterations. This indicates a higher robustness offered by the Newton-SL scheme, in scenarios characterised by small diffusion terms.

Remark 1.6.2 *In [6], the authors solve a finite difference discretization of the MFG system by employing Newton's method combined with a continuation method with respect to the diffusion parameter ν. The latter is particularly useful to deal with the case small diffusion parameters. The problem is solved first for a high value of ν and, subsequently, the authors use this solution as an initial guess to solve, still by using Newton's method, the discrete MFG system with a smaller viscosity. The method proceeds in this manner until reaching the desired (small) viscosity.*

We end Chapter 4 with a two dimensional MFG system that we solve using Newton-SL. Finally, Figure 1.4 summarize the main ideas discussed in Chapter 4.

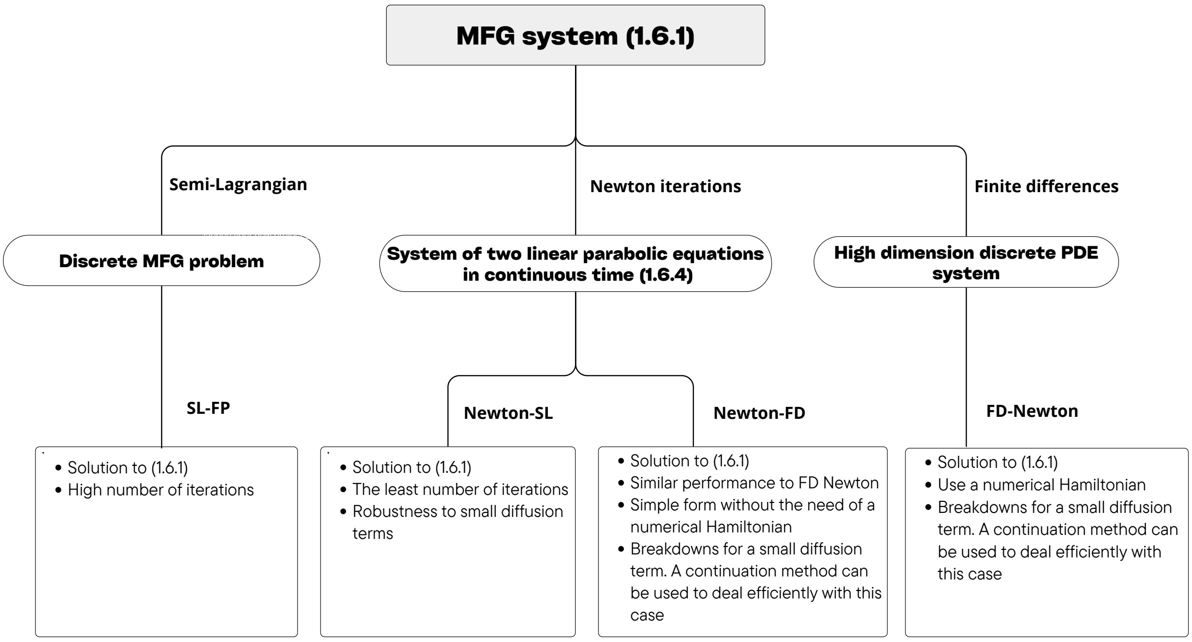


Figure 1.4: Solving the MFG system (1.6.1).

2

A Lagrange-Galerkin scheme for first order mean field games systems

In this chapter, we consider a first order mean field game system with non-local couplings. A Lagrange-Galerkin scheme for the continuity equation, coupled with a semi-Lagrangian scheme for the Hamilton-Jacobi-Bellman equation, is proposed to discretize the mean field games system. The convergence of solutions to the scheme towards a solution to the mean field game system is established in arbitrary space dimensions. The scheme is implemented to approximate two mean field games systems in dimension one and two. This chapter is a joint work with Elisabetta Carlini and Francisco J. Silva [56].

Sommaire

2.1 Introduction

In view of its applications in Economics, Physics, and Social Sciences, the study of optimal control problems and differential games with a large number of agents has attracted the attention of several researchers during the last two decades. An important step in this direction has been achieved with the introduction of the theory of Mean Field Games (MFGs) by J.-M. Lasry-Lions [111]–[113] and, independently, by M. Huang, R.P. Malhamé, and P.E. Caines [99]. The main purpose of this theory is to characterize Nash equilibria for a class of symmetric differential games with a continuum of agents. One of the main applications of MFGs theory is that such equilibria can be used to provide approximate equilibria for the corresponding games with a large, but finite, number of players. In its standard form, MFGs are described by a system of two Partial Differential Equations (PDEs); a Hamilton-Jacobi-Bellman (HJB) equation, describing the optimal cost of a typical player in the game, and a Fokker-Planck (FP) equation, describing the evolution of the initial distribution when all the players act optimally. We refer the reader to the monographs [60], [61], [86], the survey [87], and the lectures [5] for a thorough overview on MFGs.

The numerical approximation of MFGs with nonlocal couplings has been an active research topic in recent years. In the case where the MFGs system includes nondegenerate second order terms, finite-difference schemes have been studied in [2], [4], [9], [15], [92], semi-Lagrangian scheme where investigated in [58], and machine learning methods such as deep learning and reinforcement learning have been analyzed in [19], [62], [63]. In the case where the dynamics of the underlying differential games are deterministic, the resulting MFGs system is of first order and several numerical methods have been proposed to approximate its solutions; see e.g. [42], [57] for semi-Lagrangian discretizations, [80], [94] for the approximation by discrete-time finite state space MFGs (see [83]), and [121], [126] for Fourier analysis techniques. We refer the reader to [7], [114], and the references therein, for an overview on numerical methods to approximate MFGs equilibria including also the case of local couplings and variational methods.

In this paper we focus our attention on the approximation of first order MFGs systems. Namely, we consider the PDE system

$$\begin{aligned} -\partial_t v + H(x, D_x v) &= F(x, m(t)) \quad \text{in }]0,T[\times\mathbb{R}^d, \\ v(T,x) &= G(x, m(T)) \quad \text{in } \mathbb{R}^d, \\ \partial_t m - \operatorname{div}\big(D_p H(x, D_x v)m\big) &= 0 \quad \text{in } \mathbb{R}^d\times]0,T[, \\ m(0) &= m_0^*, \end{aligned} \tag{MFG}$$

where $H\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is convex with respect to its second argument and, denoting by $\mathcal{P}_1(\mathbb{R}^d)$ the set of probability measures over $\mathbb{R}^d$ with finite first order moment, $F\colon \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$, $G\colon \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$, and $m_0^* \in \mathcal{P}_1(\mathbb{R}^d)$. In the article [42], the authors propose a convergent semi-discrete scheme to approximate solutions to (MFG). A fully-discrete version has been proposed in [57]. In the proposed scheme, the HJB equation is discretized by using a semi-Lagrangian approximation of the HJB equation (see e.g. [76]), while the FP equation, or continuity equation, is approximated by a scheme which is dual to a linearized version of the scheme for the HJB equation. The existence of solutions to this approximation is shown and a convergence result to a solution to (MFG) is established when the dimension d of the space variable is equal to one. An extension of this scheme to the case where (MFG) involves non-local and fractional diffusions terms has been studied in [66]. If the resulting system has non-smooth solutions, the convergence of solutions of the scheme is also shown when the space dimension is equal to one.

In order to obtain a convergent scheme for general state dimensions, the key point is to provide a scheme which preserves, under standard conditions on the data (see Section 2.2.2 below), the main properties of solutions to both equations in (MFG). Namely, the boundedness, Lipschitzianity, and semiconcavity of the solution to the HJB equation and a uniform compact support, equicontinuity, and uniform bounds in L^p spaces for solutions to the continuity equation. As shown in [3], [57], [66], the standard semi-Lagrangian scheme for the HJB equation, which is a monotone scheme, enjoys the former properties under suitable assumptions on the discretization parameters. In order to treat the continuity equation, we consider the Lagrange-Galerkin (LG) scheme introduced in [124] and recalled in Section 2.4 below. As we show, it turns out that, for a specific choice of the basis functions, the resulting scheme for the continuity equation, which is explicit and has non-negative coefficients, coincides with the one introduced in [127] and further studied in [133] for Lipschitz velocity fields. The desired properties for the solutions to this scheme are established in Section 2.4.2. In particular, we provide a uniform L^p estimate, not available in the schemes considered in [57], [66] in arbitrary space dimensions, which will play a key role in our main convergence result. Combining the semi-Lagrangian scheme for the HJB equation and the LG scheme for the continuity equation, we obtain a discretization of (MFG) for which the existence of solutions is established and, using stability and compactness arguments, the convergence to a solution to (MFG) is established.

The rest of this article is organized as follows. In Section 2.2 we fix some standard notation, and we state our main assumptions on the data of (MFG). Some important results about solutions to HJB and continuity equations are recalled, as well as existence and uniqueness results for solutions to (MFG). The next two sections deal with the discretization of the HJB and continuity equations in (MFG) separately. Section 2.3 recalls a standard semi-Lagrangian scheme to approximate the solution to the HJB equation in (MFG). Several important properties of this scheme are reviewed and a new semiconcavity estimate for the solution to the scheme is provided in Proposition 2.3.5. This estimate will play a crucial role in Section 2.4, which is devoted to the study of a LG scheme to approximate the continuity equation in (MFG). Notice that, in general, this continuity equation is driven by a non-smooth velocity field. We show that the solutions to the LG scheme inherit the equicontinuity and L^p-stability of the solution to the original equation and we establish in Proposition 2.4.6 a convergence result as the discretization steps tend to zero. In Section 2.5 we couple the schemes studied in the previous sections to obtain a discretization of (MFG). The existence of a solution to the discretized MFG system is provided in Theorem 2.5.1 and the convergence result, valid in arbitrary dimensions, is shown in Theorem 2.5.2. Finally, Section 2.6 is devoted to the numerical implementation of the scheme for the MFGs system. Since the LG scheme for the continuity equation involves some integrals depending on the discrete characteristics of the equation, we approximate them by numerical quadrature and by the so-called area-weighting method introduced in [124]. The performances of these two approximations are compared in a one-dimensional example with an explicit solution, and the area-weighting method is implemented to approximate the solution to a MFGs in a two-dimensional space.

Acknowledgements. F. J. Silva and A. Zorkot where partially supported by l'Agence Nationale de la Recherche (ANR), project ANR-22-CE40-0010. For the purpose of open access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

The three authors were partially supported by KAUST through the subaward agreement ORA-2021-CRG10-4674.6.

2.2 Preliminaries

2.2.1 Notation

Let $d \in \mathbb{N}$. In what follows, $\langle \cdot, \cdot \rangle$ and $|\cdot|$ denote the standard scalar product in $\mathbb{R}^d$ and its induced norm, respectively. We set $|\cdot|_\infty$ for the maximum norm in $\mathbb{R}^d$ and $B_\infty(0,C)$ and $\overline{B}_\infty(0,C)$ for the associated open and closed balls, centered at 0 and of radius $C>0$, respectively. Let $\mathcal{P}(\mathbb{R}^d)$ be the set of probability measures on $\mathbb{R}^d$. For every $\nu \in \mathcal{P}(\mathbb{R}^d)$ we denote by $\operatorname{supp}(\nu)$ its support. Let $\mathcal{P}_1(\mathbb{R}^d) = \{\nu \in \mathcal{P}(\mathbb{R}^d) \,|\, \int_{\mathbb{R}^d} |x| \mathrm{d}\nu(x) < \infty\}$, and, for every ν_1, $\nu_2 \in \mathcal{P}_1(\mathbb{R}^d)$, set

$$d_1(\nu_1,\nu_2) = \inf_{\gamma \in \Pi(\nu_1,\nu_2)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y| \mathrm{d}\gamma(x,y), \tag{2.2.1}$$

where $\Pi(\nu_1,\nu_2)$ denotes the set of probabilities measures on $\mathbb{R}^d \times \mathbb{R}^d$ with first and second marginals given by ν_1 and ν_2, respectively. By the Kantorovich-Rubinstein theorem (see e.g. [18, Section 7.1]) we have

$$d_1(\nu_1,\nu_2) = \sup\left\{ \int_{\mathbb{R}^d} \varphi(x) \mathrm{d}(\nu_1 - \nu_2)(x) \,\Big|\, \varphi \in \mathrm{Lip}_1(\mathbb{R}^d) \right\}, \tag{2.2.2}$$

where $\mathrm{Lip}_1(\mathbb{R}^d)$ denotes the set of all nonexpansive functions on $\mathbb{R}^d$. Given $\nu \in \mathcal{P}(\mathbb{R}^d)$ and a Borel function $\Psi : \mathbb{R}^d \to \mathbb{R}^q$ ($q \in \mathbb{N}$), the *push-forward* measure $\Psi \sharp \nu$, defined on the σ-algebra of Borel sets $\mathbf{B}(\mathbb{R}^q)$, is defined by

$$\Psi \sharp \nu(A) = \nu(\Psi^{-1}(A)) \quad \text{for all } A \in \mathbf{B}(\mathbb{R}^q), \tag{2.2.3}$$

or, equivalently (see e.g. [32, Theorem 3.6.1]), for every $\varphi \colon \mathbb{R}^d \to \mathbb{R}$ such that $\varphi \circ \Psi$ is integrable with respect to ν, one has

$$\int_{\mathbb{R}^q} \varphi(x) \mathrm{d}(\Psi \sharp \nu)(x) = \int_{\mathbb{R}^d} \varphi\big(\Psi(x)\big) \mathrm{d}\nu(x). \tag{2.2.4}$$

2.2.2 Assumptions

Our hypothesis on the data of (MFG) are the following:

(H1) It holds that

$$H(x,p) = \sup_{a \in \mathbb{R}^d} \big(\langle a, p \rangle - L(x,a) \big) \quad \text{for all } x,\, p \in \mathbb{R}^d, \tag{2.2.5}$$

where $L\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is of class C^2, bounded from below, and, for every x, $a \in \mathbb{R}^d$, we have

$$L(x,a) \leq C_{L,1}|a|^2 + C_{L,2}, \tag{2.2.6}$$
$$|D_x L(x,a)| \leq C_{L,3}(1+|a|^2), \tag{2.2.7}$$
$$C_{L,4}|b|^2 \leq \langle D^2_{aa}L(x,a)b,b\rangle \quad \text{for all } b \in \mathbb{R}^d, \tag{2.2.8}$$
$$\langle D^2_{xx}L(x,a)y,y\rangle \leq C_{L,5}(1+|a|^2)|y|^2 \quad \text{for all } y \in \mathbb{R}^d, \tag{2.2.9}$$

for some constants $C_{L,i} > 0$ ($i = 1,\dots,5$).

(H2) The functions F and G are continuous and, for every x, $y \in \mathbb{R}^d$ and $\nu \in \mathcal{P}_1(\mathbb{R}^d)$, we have

$$|F(x,\nu)| \leq C_{F,1}, \tag{2.2.10}$$
$$|G(x,\nu)| \leq C_{G,1}, \tag{2.2.11}$$
$$|F(x,\nu) - F(y,\nu)| \leq C_{F,2}|x-y|, \tag{2.2.12}$$
$$|G(x,\nu) - G(y,\nu)| \leq C_{G,2}|x-y|, \tag{2.2.13}$$
$$F(x+y,\nu) - 2F(x,\nu) + F(x-y,\nu) \leq C_{F,3}|y|^2, \tag{2.2.14}$$
$$G(x+y,\nu) - 2G(x,\nu) + G(x-y,\nu) \leq C_{G,3}|y|^2, \tag{2.2.15}$$

for some constants $C_{F,i} > 0$, $C_{G,i} > 0$ ($i = 1,2,3$).

(H3) The initial condition m_0^* is absolutely continuous with respect to the Lebesgue measure and satisfies:
- (i) There exists $C^* > 0$ such that $\operatorname{supp}(m_0^*) \subset \overline{B}_\infty(0, C^*)$.
- (ii) There exists $\mathsf{p} \in]1,\infty]$ such that the density of m_0^*, still denoted by m_0^*, belongs to $L^{\mathsf{p}}(\mathbb{R}^d)$.

Remark 2.2.1 (i) *Since L is bounded from below, the strong convexity assumption* (2.2.8) *on $L(x,\cdot)$, which is uniform with respect to $x \in \mathbb{R}^d$, and* (2.2.6)*, imply the existence of $C_{L,6} > 0$ and $C_{L,7} > 0$ such that*

$$L(x,a) \geq C_{L,6}|a|^2 - C_{L,7} \quad \textit{for all } x,\ a \in \mathbb{R}^d. \tag{2.2.16}$$

It follows from (2.2.5)*,* (2.2.6)*, and* (2.2.16)*, that there exist $C_{H,i} > 0$ $(i = 1,2,3,4)$ such that*

$$C_{H,1}|p|^2 - C_{H,2} \leq H(x,p) \leq C_{H,3}|p|^2 + C_{H,4} \quad \textit{for all } x,\ p \in \mathbb{R}^d. \tag{2.2.17}$$

Moreover, by (2.2.5)*,* (2.2.16)*, and Danskin's theorem* (*see e.g. [35, Theorem 4.13]*)*, we deduce that* H *is of class* C^1 *and, for every* x*,* $p \in \mathbb{R}^d$*, the following equalities hold*

$$D_a L(x, D_p H(x,p)) = p, \tag{2.2.18}$$
$$D_x H(x,p) = -D_x L(x, D_p H(x,p)). \tag{2.2.19}$$

Since $D_p H(x,p)$ *is the unique maximizer of* $\sup_{a\in\mathbb{R}^d}\left(\langle a,p\rangle - L(x,a)\right)$*,* (2.2.6)*, and* (2.2.16)*, yield the existence of* $C_{H,5} > 0$ *such that*

$$|D_p H(x,p)| \leq C_{H,5}(1+|p|) \quad \textit{for all } x,\, p \in \mathbb{R}^d. \tag{2.2.20}$$

Finally, since L *is of class* C^2*, by* (2.2.8) *and the implicit function theorem applied to* (2.2.18)*, it follows that* $D_p H$ *is of class* C^1 *and hence, by* (2.2.19)*, we obtain that* H *is of class* C^2*.*

A typical example of a function H *satisfying* **(H1)** *is given by* $H(x,p) = \kappa(x)|p|^2 + \langle b(x), p\rangle$*, where* $\kappa\colon \mathbb{R}^d \to \mathbb{R}$ *is of class* C^2*, with bounded first and second order derivatives, there exist* $\underline{\kappa}$*,* $\overline{\kappa}$ $\in]0,\infty[$ *such that* $\underline{\kappa} \leq \kappa(x) \leq \overline{\kappa}$ *for all* $x \in \mathbb{R}^d$*, and* $b\colon \mathbb{R}^d \to \mathbb{R}^d$ *is bounded, of class* C^2*, with bounded first and second order derivatives. In this case we have*

$$L(x,a) = \frac{1}{4\kappa(x)}|a - b(x)|^2 \quad \textit{for all } x,\, a \in \mathbb{R}^d. \tag{2.2.21}$$

(ii) *Assumption* **(H3)**(i) *on the compactness of* $\mathrm{supp}(m_0^*)$ *plays an important role in Section 2.4 dealing with the discretization of the continuity equation in* (MFG)*. Let us point out that in [66] the authors are able to handle initial conditions with unbounded support under a different set of assumptions over* H*. In that framework, we expect that our techniques can be adapted in order to deal with more general initial conditions.*

2.2.3 The first order mean field games system

Given $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, consider the HJB equation

$$\begin{aligned} -\partial_t v(t,x) + H(x, D_x v(t,x)) &= F(x,\mu(t)) \quad \text{for } (t,x) \in]0,T[\times\mathbb{R}^d, \\ v(x,T) &= G(x,\mu(T)) \quad \text{for } x \in \mathbb{R}^d. \end{aligned} \tag{2.2.22}$$

It follows from [23], [71] that (2.2.22) admits a unique viscosity solution $v[\mu]$ and, for every $t \in [0,T[$, $x \in \mathbb{R}^d$, and $\alpha \in L^2\big([t,T];\mathbb{R}^d\big)$, setting $X^{t,x,\alpha}(\cdot) = x - \int_t^{(\cdot)} \alpha(s)\mathrm{d}s$ and

$$J^{t,x}[\mu](\alpha) = \int_t^T \Big(L\big(X^{t,x,\alpha}(s), \alpha(s)\big) + F\big(X^{t,x,\alpha}(s), \mu(s)\big)\Big)\mathrm{d}s + G(X^{t,x,\alpha}(T), \mu(T)), \tag{2.2.23}$$

we have

$$v[\mu](t,x) = \inf\Big\{ J^{t,x}[\mu](\alpha) \,\Big|\, \alpha \in L^2\big([t,T];\mathbb{R}^d\big)\Big\}. \tag{2.2.24}$$

The proof of the following result follows from standard arguments (see e.g. [48]). However, for the sake of completeness, we provide its proof in the appendix of this work.

Proposition 2.2.1 *Assume* **(H1)-(H2)** *and let $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. Then the following hold:*

(i) [Existence of an optimal control] *For every $(t,x) \in [0,T[\times\mathbb{R}^d$, there exists $\alpha^{t,x} \in L^\infty\big([t,T];\mathbb{R}^d\big)$ such that $v[\mu](t,x) = J^{t,x}[\mu](\alpha^{t,x})$. Moreover, there exists $C_{\mathrm{b}} > 0$, independent of (μ,t,x), such that $\|\alpha^{t,x}\|_{L^\infty([0,T];\mathbb{R}^d)} \leq C_{\mathrm{b}}$.*

(ii) [Uniform bound] *We have*

$$|v[\mu](t,x)| \leq C_{\mathrm{v}} \quad \textit{for all } (t,x) \in [0,T]\times\mathbb{R}^d, \tag{2.2.25}$$

where $C_{\mathrm{v}} > 0$ is independent of μ.

(ii) [Lipschitz property] *We have*

$$\big|v[\mu](t,x) - v[\mu](t,y)\big| \leq C_{\mathrm{Lip}}|x-y| \quad \textit{for all } t \in [0,T],\ x,\ y \in \mathbb{R}^d, \tag{2.2.26}$$

where $C_{\mathrm{Lip}} > 0$ is independent of μ.

(vi) [Semi-concavity] *We have*

$$v[\mu](t,x+y) - 2v[\mu](t,x) + v[\mu](t,x-y) \leq C_{\mathrm{sc}}|y|^2 \quad \textit{for all } t \in [0,T],\ x,\ y \in \mathbb{R}^d, \tag{2.2.27}$$

where $C_{\mathrm{sc}} > 0$ is independent of μ.

Remark 2.2.2 *Assertion* (i) *in Proposition 2.2.1 implies that, for every $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, we have*

$$v[\mu](t,x) = \inf\Big\{ J^{t,x}[\mu](\alpha) \,\Big|\, \alpha \in L^\infty([0,T];\mathbb{R}^d),\ \|\alpha\|_{L^\infty([0,T];\mathbb{R}^d)} \leq C_{\mathrm{b}}\Big\} \tag{2.2.28}$$

for all $(t,x) \in [0,T[\times\mathbb{R}^d$. *In particular,* $v[\mu]$ *is also characterized by the HJB equation*

$$\begin{aligned} -\partial_t v(t,x) + H_{\rm b}(x, D_x v(t,x)) = F(x,\mu(t)) \quad & \text{for } (t,x) \in]0,T[\times\mathbb{R}^d, \\ v(x,T) = G(x,\mu(T)) \quad & \text{for } x \in \mathbb{R}^d, \end{aligned} \tag{2.2.29}$$

where

$$H_{\rm b}(x,p) = \sup_{a\in\overline{B}(0,C_{\rm b})} \left\{\langle a,p\rangle - L(x,a)\right\} \quad \text{for all } x,\ p \in \mathbb{R}^d. \tag{2.2.30}$$

Consider the set-valued map $D_x^+ v[\mu] \colon [0,T]\times\mathbb{R}^d \to 2^{\mathbb{R}^d}$ defined by

$$D_x^+ v[\mu](t,x) = \left\{ p\in\mathbb{R}^d \;\middle|\; \limsup_{y\to x} \frac{v[\mu](t,y)-v[\mu](t,x)-\langle p, y-x\rangle}{|y-x|} \leq 0 \right\} \quad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d.$$

It follows from Proposition 2.2.1(iii) and [48, Proposition 3.1.5 and Proposition 3.3.4] that $D_x^+ v[\mu]$ takes nonempty and closed values and its graph is closed. In particular, since Proposition 2.2.1(ii) and [48, Theorem 3.3.6] imply that $D_x^+ v[\mu](t,x) \subset \overline{B}(0,C_{\rm Lip})$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$, by [21, Chapter 1, Corollary 1] we have that $D_x^+ v[\mu]$ is upper-semicontinuous, i.e. for every $M\subset\mathbb{R}^d$ closed, $D_x^+ v[\mu]^{-1}(M)$ is closed. Therefore, $D_x^+ v[\mu]$ is a Borel measurable set-valued map and hence admits a Borel measurable selection (see e.g. [131, Corollary 14.6]). Notice that Proposition 2.2.1(ii), Rademacher's theorem, and [48, Proposition 3.1.5] imply that all the measurable selections of $D_x^+ v[\mu]$ coincide almost everywhere in $[0,T]\times\mathbb{R}^d$ and hence, hereafter, we will denote likewise by $D_x v[\mu]$ any choice among them.

Let $\mathsf{p}\in]1,\infty[$ be as in **(H3)**. We say that $m\in L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$ solves the continuity equation

$$\begin{aligned} \partial_t m - \operatorname{div}\left(D_p H(x, D_x v[\mu])m\right) = 0 \quad & \text{in }]0,T[\times\mathbb{R}^d, \\ m(0) = m_0^* \quad & \text{in } \mathbb{R}^d, \end{aligned} \tag{2.2.31}$$

if, for every $\varphi\in C_0^\infty(\mathbb{R}^d)$ and $t\in[0,T]$, we have

$$\int_{\mathbb{R}^d}\varphi(x)m(t,x)\mathrm{d}x = \int_{\mathbb{R}^d}\varphi(x)m_0^*(x)\mathrm{d}x - \int_0^t\int_{\mathbb{R}^d}\Big\langle D_pH(x,D_xv[\mu](s,x)), D\varphi(x)\Big\rangle m(s,x)\mathrm{d}x\mathrm{d}s. \tag{2.2.32}$$

Proposition 2.2.2 *Assume* **(H1)**-**(H3)** *and let* $\mu\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. *Then* (2.2.32) *admits a solution* $m\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))\cap L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$ *and there exists* $\widetilde{C}>0$ *such* *that*

$$\|m(t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \widetilde{C}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \text{for all } t \in [0,T]. \tag{2.2.33}$$

If, in addition, for every $t \in [0,T]$, the functions $F(\cdot,\mu(t))$ and $G(\cdot,\mu(T))$ are differentiable, then the solution m to (2.2.31) *is unique.*

Proof. The first assertion in the statement follows from Proposition 2.4.6 below, while the second one follows by arguing as in the proof of [55, Lemma 1.10]. The crucial steps in the latter are the use of the superposition principle in [16] for solutions to (2.2.31) and the fact that, under the differentiability assumptions over $F(\cdot,\mu(t))$ and $G(\cdot,\mu(T))$, the optimal control problem $\inf\left\{J^{0,x}[\mu](\alpha) \,\middle|\, \alpha \in L^2\big([0,T];\mathbb{R}^d\big)\right\}$ admits a unique solution for almost every $x \in \mathbb{R}^d$. □

Finally, we say that (v^*,m^*), with $m^* \in C([0,T];\mathcal{P}_1(\mathbb{R}^d)) \cap L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$, solves (MFG) if $v^* = v[m^*]$ and m^* solves (2.2.31) with $\mu = m^*$.

Proposition 2.2.3 *Assume* **(H1)**-**(H3)**. *Then system* (MFG) *admits a solution (v^*,m^*).*

Proof. A proof of this result, under slightly different assumptions, can be found, for instance, in [55, Section 1.3.4]. In the present context, the result follows from Theorem 2.5.2 below. □

A uniqueness result for solutions to (MFG) can be shown under additional assumptions on the coupling terms F and G. A sufficient condition is the so-called Lasry-Lions monotonicity condition which states that, for $h = F,\, G$, it holds

$$\int_{\mathbb{R}^d} \big(h(x,m_1) - h(x,m_2)\big)\mathrm{d}\big(m_1 - m_2\big)(x) \geq 0 \quad \text{for all } m_1,\, m_2 \in \mathcal{P}_1(\mathbb{R}^d). \tag{2.2.34}$$

Proposition 2.2.4 *Assume* **(H1)**-**(H3)***, the monotonicity condition* (2.2.34) *and that, for all $\nu \in \mathcal{P}_1(\mathbb{R}^d)$, the functions $F(\cdot,\nu)$ and $G(\cdot,\nu)$ are differentiable. Then system* (MFG) *admits a unique solution.*

Proof. The existence of a solution to (MFG) follows from Proposition 2.2.3 while, under (2.2.34) and the differentiability assumptions on $F(\cdot,\nu)$ and $G(\cdot,\nu)$, the proof of the uniqueness of the solution follows by arguing as in the proof of [55, Theorem 1.8]. □

2.3 A semi-Lagrangian scheme for the HJB equation

Let $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. In this section we recall a standard semi-Lagrangian scheme to approximate the viscosity solution $v[\mu]$ to (2.2.22). Most of the results for the semi-Lagrangian scheme that will be needed in the remainder of the article follow similarly to those in the monograph [76] and the contributions [57], [58], [66]. The principal differences come from our assumptions on L in **(H1)**, which allow us to consider cost functionals not covered in these references (see e.g. the example in the last paragraph of Remark 2.2.1). Therefore, we confine ourselves to explain the main changes in the proofs of the aforementioned properties. On the other hand, the estimate in Proposition 2.3.5 below seems to be new and will play a key role later in this article. In order to define the scheme, let $N \in \mathbb{N}^*$ be the number of time steps, let $\Delta t = T/N$ be the time step, let $\mathcal{I}_{\Delta t} = \{0,\dots,N\}$, let $\mathcal{I}^*_{\Delta t} = \mathcal{I}_{\Delta t} \setminus \{N\}$, let $t_k = k\Delta t$ for all $k \in \mathcal{I}_{\Delta t}$, and set $\mathcal{G}_{\Delta t} = \{t_k \,|\, k \in \mathcal{I}_{\Delta t}\}$. Given a space step $\Delta x > 0$ and $i = (i_1,\dots,i_d) \in \mathbb{Z}^d$, define $\beta^1_i \colon \mathbb{R}^d \to \mathbb{R}$ as

$$\beta^1_i(z) = \prod_{l=1}^{d} \widehat{\beta}\Big(\frac{z_l}{\Delta x} - i_l\Big) \quad \text{for all } z = (z_1,\dots,z_d) \in \mathbb{R}^d, \tag{2.3.1}$$

where

$$\widehat{\beta}(\xi) = \max\{0, 1 - |\xi|\} \quad \text{for all } \xi \in \mathbb{R}. \tag{2.3.2}$$

Notice that $\beta^1_i \geq 0$, $\sum_{i\in\mathbb{Z}^d} \beta^1_i(x) = 1$ for all $x \in \mathbb{R}^d$ and, setting $x_i = i\Delta x$, we have $\beta^1_i(x_j) = 1$, if $i = j$, and $\beta^1_i(x_j) = 0$, otherwise. Let $\mathcal{G}_{\Delta x} = \{i\Delta x \,|\, i \in \mathbb{Z}^d\}$ be the uniform grid and, given $\phi \colon \mathcal{G}_{\Delta x} \to \mathbb{R}$, define its interpolate as

$$I^1[\phi](x) = \sum_{i\in\mathbb{Z}^d} \beta^1_i(x)\phi_i \quad \text{for all } x \in \mathbb{R}^d,$$

where, for notational simplicity, we have set $\phi_i = \phi(x_i)$. For every $\varphi \colon \mathbb{R}^d \to \mathbb{R}$ denote by $\varphi|_{\mathcal{G}_{\Delta x}}$ its restriction to $\mathcal{G}_{\Delta x}$. If φ is of class C^2 and has bounded second order derivatives, it follows from [129, Remark 3.4.2] that

$$\|\varphi(x) - I[\varphi|_{\mathcal{G}_{\Delta x}}](x)\|_\infty \leq C_\varphi (\Delta x)^2, \tag{2.3.3}$$

where $C_\varphi > 0$ depends only on φ.
We consider the following fully-discrete semi-Lagrangian scheme:

find $\{v_k\colon \mathcal{G}_{\Delta x} \to \mathbb{R} \,|\, k \in \mathcal{I}_{\Delta t}\}$ such that

$$\begin{aligned} v_{k,i} &= \mathcal{S}^{\mathrm{fd}}_{k,i}[\mu](v_{k+1}) \quad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ i \in \mathbb{Z}^d, \\ v_{N,i} &= G(x_i, \mu(T)) \quad \text{for all } i \in \mathbb{Z}^d, \end{aligned} \tag{2.3.4}$$

where, for every $\phi\colon \mathcal{G}_{\Delta x} \to \mathbb{R}$, bounded, $k \in \mathcal{I}_{\Delta t}$, and $i \in \mathbb{Z}^d$,

$$\mathcal{S}^{\mathrm{fd}}_{k,i}[\mu](\phi) = \inf_{a \in \overline{B}(0, C_{\mathrm{b}})} \left[\Delta t L(x_i, a) + I^1[\phi](x_i - \Delta t a)\right] + \Delta t F(x_i, \mu(t_k)). \tag{2.3.5}$$

Notice that, being explicit, scheme (2.3.4) admits a unique solution. By definition, $\mathcal{S}^{\mathrm{fd}}[\mu]$ is *monotone*, i.e. for every ϕ^1, $\phi^2\colon \mathcal{G}_{\Delta x} \to \mathbb{R}$, bounded, with $\phi^1_i \leq \phi^2_i$ for all $i \in \mathbb{Z}^d$, we have that

$$\mathcal{S}^{\mathrm{fd}}_{k,i}[\mu](\phi^1) \leq \mathcal{S}^{\mathrm{fd}}_{k,i}[\mu](\phi^2) \quad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ i \in \mathbb{Z}^d. \tag{2.3.6}$$

Moreover, using **(H1)** and **(H2)**, standard arguments (see e.g. [76, Section 5.2.3]) yield the following *consistency* property for $\mathcal{S}^{\mathrm{fd}}[\mu]$: let $(\mu_n)_{n\in\mathbb{N}} \subset C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, $\big((\Delta t_n, \Delta x_n)\big)_{n\in\mathbb{N}} \subset]0,\infty[^2$, $\big((t_{k_n}, x_{i_n})\big)_{n\in\mathbb{N}} \subset \mathcal{G}_{\Delta t_n} \times \mathcal{G}_{\Delta x_n}$, and $(t,x) \in]0,T[\times\mathbb{R}^d$ such that, as $n \to \infty$, $\mu_n \to \mu$, $(\Delta t_n, \Delta x_n) \to 0$, $(\Delta x_n)^2/\Delta t_n \to 0$, and $(t_{k_n}, x_{i_n}) \to (t,x)$. Then, recalling the definition of H_{b} in (2.2.30), for every $\varphi\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$ of class C^1, with bounded derivatives, we have

$$\lim_{n\to\infty} \frac{1}{\Delta t_n}\Big(\varphi(x_{i_n}, t_{k_n}) - \mathcal{S}^{\mathrm{fd}}_{k_n,i_n}[\mu_n]\big(\varphi(t_{k_n+1}, \cdot)|_{\mathcal{G}_{\Delta x_n}}\big)\Big) = -\partial_t\varphi(x,t) + H_{\mathrm{b}}(x, D_x\varphi(t,x)) - F(x,\mu(t)). \tag{2.3.7}$$

Given $(\Delta t, \Delta x) \in]0,\infty[^2$, let us set

$$v^{\Delta t, \Delta x}[\mu](t_k, x) = I^1[v_k](x) \quad \text{for all } k \in \mathcal{I}_{\Delta t},\ x \in \mathbb{R}^d, \tag{2.3.8}$$

where, for every $k \in \mathcal{I}_{\Delta t}$, $v_k\colon \mathcal{G}_{\Delta x} \to \mathbb{R}$ is computed with (2.3.5). We extend this definition to $[0,T]\times\mathbb{R}^d$, by setting

$$v^{\Delta t, \Delta x}[\mu](t,x) = v^{\Delta t, \Delta x}[\mu](t_k, x) \quad \text{if } t \in [t_k, t_{k+1}[,\ k \in \mathcal{I}^*_{\Delta t}. \tag{2.3.9}$$

The following result provides properties for $v^{\Delta t,\Delta x}[\mu]$ that are analogous to those in Proposition 2.2.1$\mathrm{(ii)}$-$\mathrm{(iv)}$ for $v[\mu]$.

Proposition 2.3.1 *Assume* **(H1)-(H2)**, *let* $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, *and let* $(\Delta t, \Delta x) \in]0,\infty[^2$. *Then the following hold:*

(i) [Stability] *We have*

$$|v^{\Delta x,\Delta t}[\mu](t,x)| \leq \widetilde{C}_{\mathrm{v}} \quad \textit{for all } (t,x) \in [0,T]\times\mathbb{R}^d, \tag{2.3.10}$$

where $\widetilde{C}_{\mathrm{v}} > 0$ *is independent of* $(\mu, \Delta t, \Delta x)$.

(ii) [Lipschitz property] *We have*

$$\left|v^{\Delta t,\Delta x}[\mu](t,x) - v^{\Delta t,\Delta x}[\mu](t,y)\right| \leq \widetilde{C}_{\mathrm{Lip}}|x-y| \quad \textit{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d, \tag{2.3.11}$$

where $\widetilde{C}_{\mathrm{Lip}} > 0$ *is independent of* $(\mu, \Delta t, \Delta x)$.

(iii) [Discrete semi-concavity] *We have*

$$\begin{aligned} v^{\Delta t,\Delta x}[\mu](t,x+x_i) - 2v^{\Delta t,\Delta x}[\mu](t,x) + v^{\Delta t,\Delta x}[\mu](t,x-x_i) \\ \leq \widetilde{C}_{\mathrm{sc}}|x_i|^2 \quad \textit{for all } t\in[0,T],\ x\in\mathbb{R}^d,\ i\in\mathbb{Z}^d, \end{aligned} \tag{2.3.12}$$

where $\widetilde{C}_{\mathrm{sc}} > 0$ *is independent of* $(\mu, \Delta t, \Delta x)$.

Proof. (i): This follows directly from (2.3.4), (2.2.6), (2.2.16), (2.2.10), (2.2.11), and iteration.
(ii): It follows from (2.2.7) that

$$|L(x,a) - L(y,a)| \leq C_{L,3}(1+C_{\mathrm{b}}^2)|x-y| \quad \text{for all } x,\ y\in\mathbb{R}^d,\ a\in\overline{B}(0,C_{\mathrm{b}}). \tag{2.3.13}$$

Using this inequality, (2.2.12), and (2.2.13), the result follows from the same arguments than those in [58, Lemma 3.1(i)] (see also the proof of [66, Lemma 5.3(a)]).
(iii): It follows from (2.2.9) that

$$L(x+y,a) - 2L(x,a) + L(x-y,a) \leq C_{L,5}(1+C_{\mathrm{b}}^2)|y|^2 \quad \text{for all } x,\ y\in\mathbb{R}^d,\ a\in\overline{B}(0,C_{\mathrm{b}}). \tag{2.3.14}$$

In turn, using (2.2.14) and (2.2.15), the result follows by arguing as in the proof of [3, Lemma 4.1] (see also the proof of [66, Lemma 3.2(ii)]). □

Using the monotonicity of $\mathcal{S}^{\mathrm{fd}}[\mu]$, the consistency property in (2.3.7), and the stability result in Proposition 2.3.1(i), the Barles-Souganidis relaxed limit method (see [25]) yields the following convergence result (see [57, Theorem 3.3] for a detailed proof).

Proposition 2.3.2 *Assume* **(H1)-(H2)**, *let* $(\mu_n)_{n\in\mathbb{N}} \subset C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\big((\Delta t_n, \Delta x_n)\big)_{n\in\mathbb{N}} \subset]0,\infty[^2$. *Suppose that, as* $n\to\infty$, $\mu_n\to\mu$, *for some* $\mu\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$,

$(\Delta t_n, \Delta x_n) \to 0$*, and* $(\Delta x_n)^2/\Delta t_n \to 0$*. Then* $(v^{\Delta t_n,\Delta x_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $v[\mu]$ *uniformly over compact subsets of* $[0,T]\times\mathbb{R}^d$*.*

Given $\varepsilon > 0$, consider the mollifier $\mathbb{R}^d \ni x \mapsto \rho_\varepsilon(x) = \rho(x/\varepsilon)/\varepsilon^d \in \mathbb{R}^d$, where $\rho \in C^\infty(\mathbb{R}^d)$ has bounded derivatives of any order and satisfies $\rho(\mathbb{R}^d) \subset [0,\infty[$ and $\int_{\mathbb{R}^d}\rho(x)\mathrm{d}x = 1$. Given $\varphi \in W^{1,\infty}(\mathbb{R}^d)$, a standard computation shows that

$$\sup_{x\in\mathbb{R}^d}\left|(\rho_\varepsilon * \varphi)(x) - \varphi(x)\right| \leq \varepsilon\|D\varphi\|_{L^\infty(\mathbb{R}^d)}, \tag{2.3.15}$$

$$\sup_{x\in\mathbb{R}^d}\left\|D^\ell(\rho_\varepsilon * \varphi)(x)\right\| \leq c_\ell\varepsilon^{1-\ell} \quad \text{for all } \ell\in\mathbb{N}, \tag{2.3.16}$$

where $\left\|D^\ell(\rho_\varepsilon * \varphi)(x)\right\|$ denotes the operator norm of $D^\ell(\rho_\varepsilon * \varphi)(x)$ and $c_\ell > 0$ depends only on ℓ. Let us set $\Delta = (\Delta t, \Delta x, \varepsilon)$ and define

$$v^\Delta[\mu](t,\cdot) = \rho_\varepsilon * v^{\Delta t,\Delta x}[\mu](t,\cdot) \quad \text{for all } t\in[0,T]. \tag{2.3.17}$$

The function $v^\Delta[\mu]$ satisfies similar properties than $v^{\Delta t,\Delta x}[\mu]$, as the following proposition shows.

Proposition 2.3.3 *Assume* **(H1)-(H2)***, let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, and let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in\,]0,\infty[^3$*. Then the following holds:*

(i) [Stability] *We have*

$$|v^\Delta[\mu](t,x)| \leq \widetilde{C}_{\mathrm{v}} \quad \textit{for all } (t,x)\in[0,T]\times\mathbb{R}^d, \tag{2.3.18}$$

with $\widetilde{C}_{\mathrm{v}} > 0$ *being as in Proposition 2.3.1*(i)*.*

(ii) [Lipschitz property] *We have*

$$\left|v^\Delta[\mu](t,x) - v^\Delta[\mu](t,y)\right| \leq \widetilde{C}_{\mathrm{Lip}}|x-y| \quad \textit{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d, \tag{2.3.19}$$

with $\widetilde{C}_{\mathrm{Lip}} > 0$ *being as in Proposition 2.3.1*(ii)*.*

(iii) [Approximate semi-concavity] *We have*

$$\begin{aligned} v^\Delta[\mu](t,x+y) &- 2v^\Delta[\mu](t,x) + v^\Delta[\mu](t,x-y) \\ &\leq \widetilde{C}_{\mathrm{asc}}\left(|y|^2 + (\Delta x)^2 + \frac{(\Delta x)^2}{\varepsilon}\right) \quad \textit{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d, \end{aligned} \tag{2.3.20}$$

where $\widetilde{C}_{\mathrm{asc}} > 0$ *is independent of* (μ,Δ)*.*

Proof. Assertions (i) and (ii) follow directly from (2.3.17) and the corresponding assertions in Proposition 2.3.1. The proof of (iii) follows from Proposition 2.3.1(iii) and arguing as in the proof of [3, Lemma 4.2] (see also the proof of [66, Lemma 5.5(b)]). □
In the following, given $A \subset \mathbb{R}^d$, we denote by $\mathbb{I}_A$ the indicator function of A. The convergence result in the following proposition will play an important role in the next section.

Proposition 2.3.4 *Assume* **(H1)-(H2)**, *let* $(\mu_n)_{n\in\mathbb{N}} \subset C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\big((\Delta t_n, \Delta x_n, \varepsilon_n)\big)_{n\in\mathbb{N}} \subset]0,\infty[^3$. *Set* $\Delta_n = (\Delta t_n, \Delta x_n, \varepsilon_n)$ *and let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. *Suppose that, as* $n \to \infty$, $\mu_n \to \mu$, $\Delta_n \to 0$, *and* $(\Delta x_n)^2/\Delta t_n \to 0$. *Then the following hold:*

(i) $(v^{\Delta_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $v[\mu]$ *uniformly over compact subsets of* $[0,T]\times\mathbb{R}^d$.

(ii) *If, in addition,* $\Delta x_n/\varepsilon_n \to 0$, *then, for every* $K \subset [0,T]\times\mathbb{R}^d$ *compact and* $q \in [1,\infty[$,

$$\mathbb{I}_K D_p H(\cdot, D_x v^{\Delta_n}[\mu_n]) \to \mathbb{I}_K D_p H(\cdot, D_x v[\mu]) \quad \text{in } L^q([0,T]\times\mathbb{R}^d). \tag{2.3.21}$$

Proof. (i): This follows from Proposition 2.3.3(ii), estimate (2.3.15), and Proposition 2.3.2.
(ii): It follows from Proposition 2.3.3(iii) and [3, Remark 6] that

$$\Big\langle D_x v^{\Delta}[\mu](t,y) - D_x v^{\Delta}[\mu](t,x), y-x \Big\rangle \le c\left(|y-x|^2 + \frac{(\Delta x)^2}{\varepsilon^2}\right) \quad \text{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d. \tag{2.3.22}$$

Using this inequality and arguing as in the proof of [57, Theorem 3.5], one deduces that, as $n\to\infty$, $D_x v^{\Delta_n}[\mu_n] \to D_x v[\mu]$ almost everywhere in $[0,T]\times\mathbb{R}^d$. In turn, since H is of class C^2, we get that $D_pH(\cdot, D_x v^{\Delta_n}[\mu_n]) \to D_pH(\cdot, D_x v[\mu])$ almost everywhere in $[0,T]\times\mathbb{R}^d$. Thus, (2.3.21) follows from Proposition 2.3.3(ii) and Lebesgue's dominated convergence theorem. □

We conclude this section with a useful estimate for $D^2 v^{\Delta}[\mu]$.

Proposition 2.3.5 *Assume* **(H1)-(H2)**, *let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^2\times]0,1[$. *Then it holds that*

$$\langle D_x^2 v^{\Delta}[\mu](t,x)y, y\rangle \le \widetilde{C}_{\rm hb}\left(1 + \frac{(\Delta x)^2}{\varepsilon^4}\right)|y|^2 \quad \textit{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d, \tag{2.3.23}$$

where $\widetilde{C}_{\rm hb} > 0$ *is independent of* (μ,Δ).

Proof. Let us fix $t \in [0,T]$ and $x,\, y \in \mathbb{R}^d$. If $y=0$ the result is true and, hence, let us assume that $y \neq 0$ and set $\tau = \Delta x/(\sqrt{\varepsilon}|y|)$. In what follows, $C>0$ denotes a constant, independent of (μ, Δ) which may change from line to line. It follows from Proposition 2.3.3(iii) that

$$v^{\Delta}[\mu](t,x+\tau y) - 2v^{\Delta}[\mu](t,x) + v^{\Delta}[\mu](t,x-\tau y) \leq C\tau^2|y|^2. \tag{2.3.24}$$

On the other hand, a Taylor expansion of order 4 and (2.3.16) imply that

$$\begin{aligned} v^{\Delta}[\mu](t,x+\tau y) \geq v^{\Delta}[\mu](t,x) + \langle D_x v^{\Delta}[\mu](t,x), \tau y\rangle + \frac{1}{2}\langle D_x^2 v^{\Delta}[\mu](t,x)\tau y, \tau y\rangle \\ + \frac{1}{6} D_x^3 v^{\Delta}[\mu](t,x)(\tau y)^3 - \frac{1}{\varepsilon^3} C|\tau y|^4, \\ v^{\Delta}[\mu](t,x-\tau y) \geq v^{\Delta}[\mu](t,x) - \langle D_x v^{\Delta}[\mu](t,x), \tau y\rangle + \frac{1}{2}\langle D_x^2 v^{\Delta}[\mu](t,x)\tau y, \tau y\rangle \\ - \frac{1}{6} D_x^3 v^{\Delta}[\mu](t,x)(\tau y)^3 - \frac{1}{\varepsilon^3} C|\tau y|^4, \end{aligned}$$

Adding both inequalities, using (2.3.24) and the relation $\tau|y| = \Delta x/\sqrt{\varepsilon}$, we get

$$\langle D_x^2 v^{\Delta}[\mu](t,x)\tau y, \tau y\rangle \leq C\left(1 + \frac{(\Delta x)^2}{\varepsilon^4}\right)\tau^2|y|^2. \tag{2.3.25}$$

Dividing by τ^2 yields (2.3.23). $\square$

2.4 A Lagrange-Galerkin type scheme for the continuity equation

Given $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$ and $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$, let $v^{\Delta}[\mu]$ be defined as in (2.3.17). Consider the continuity equation

$$\begin{aligned} \partial_t m - \operatorname{div}\big(D_p H(x, D_x v^{\Delta}[\mu])m\big) &= 0 \qquad \text{in }]0,T[\times\mathbb{R}^d, \\ m(0) &= m_0^* \qquad \text{in } \mathbb{R}^d. \end{aligned} \tag{2.4.1}$$

Since H is of class C^2 and $D_x v^{\Delta}[\mu]$ is bounded and Lipschitz, by [18, Proposition 8.1.8] equation (2.4.1) admits a unique solution $m^{\Delta}[\mu] \in C\left([0,T]; \mathcal{P}_1(\mathbb{R}^d)\right)$, which can be represented as

$$m^{\Delta}[\mu](t) = \Phi^{\Delta}[\mu](0,t,\cdot)\sharp m_0^* \quad \text{for all } t \in [0,T], \tag{2.4.2}$$

where, for all $s \in [0,T)$ and $x \in \mathbb{R}^d$, $\Phi^{\Delta}[\mu](s,\cdot,x)$ denotes the unique solution to

$$\begin{aligned} &\dot{X}(t) = -D_pH(X(t), D_x v^{\Delta}[\mu](t,X(t))) \qquad \text{for a.e. } t \in (s,T),\\ &X(s) = x. \end{aligned} \tag{2.4.3}$$

Relations (2.4.2) and (2.4.3) imply that

$$m^{\Delta}[\mu](t) = \Phi^{\Delta}[\mu](s,t,\cdot)\sharp m^{\Delta}[\mu](s) \quad \text{for all } s,\ t \in [0,T],\ s \leq t. \tag{2.4.4}$$

On the other hand, notice that (2.2.20) and the uniform bound $\|D_x v^{\Delta}[\mu](t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \widetilde{C}_{\mathrm{Lip}}$ for all $t \in [0,T]$, which follows from Proposition 2.3.3(ii), yield the existence of $C_{\mathrm{bf}} > 0$, independent of (μ, Δ), such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left| D_pH(x, D_x v^{\Delta}[\mu](t,x)) \right| \leq C_{\mathrm{bf}}. \tag{2.4.5}$$

Relation (2.4.4) and estimate (2.4.5) have two straightforward consequences. The first one is that, by **(H3)**(i), we have

$$|\Phi^{\Delta}[\mu](0,t,x)| \leq C^* + TC_{\mathrm{bf}} \quad \text{for all } t \in [0,T],\ x \in \mathrm{supp}(m_0^*) \tag{2.4.6}$$

and, hence, by (2.4.2),

$$\mathrm{supp}\left(m^{\Delta}[\mu](t)\right) \subset \overline{B}(0, C^* + TC_{\mathrm{bf}}) \quad \text{for all } t \in [0,T]. \tag{2.4.7}$$

The second one, is the uniform equicontinuity of the familly $\left\{m^{\Delta}[\mu] \,\middle|\, \Delta \in]0,\infty[^3\right\}$ in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. More precisely, using (2.2.2), an easy computation shows that

$$d_1\left(m^{\Delta}[\mu](t), m^{\Delta}[\mu](s)\right) \leq C_{\mathrm{bf}}|t-s| \quad \text{for all } s,\ t \in [0,T]. \tag{2.4.8}$$

The purpose of the following two propositions is to provide some stability estimates for $m^{\Delta}[\mu]$, which, together with (2.4.7) and (2.4.8), motivate the forthcoming analysis for a LG discretization of (2.4.1).

Proposition 2.4.1 *Assume* **(H1)**-**(H3)**, *let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in$

$]0,\infty[^3$. *Then for every* $c>0$ *there exists* $\tilde{c}>0$, *independent of* μ, *such that, if* $\Delta x \leq c\varepsilon^2$, *it holds that*

$$\|m^{\Delta}[\mu](t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \tilde{c}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \textit{for all } t\in[0,T]. \tag{2.4.9}$$

Proof. Let $c>0$ and suppose that $\Delta x \leq c\varepsilon^2$. Proposition 2.3.5 implies that

$$\left\langle D_x^2 v^{\Delta}[\mu](t,x)y,y\right\rangle \leq \widetilde{C}_{\rm hb}\left(1+c^2\right)|y|^2 \quad \text{for all } t\in[0,T],\ x,\ y\in\mathbb{R}^d. \tag{2.4.10}$$

Using this inequality, the proof of (2.4.9) follows from exactly the same arguments than those in the proof of [79, Proposition 4.1]. For the sake of completeness, we sketch the main points of the analysis and refer the reader to [79] for the details. Let $t\in[0,T]$ and suppose first that $\mathsf{p}\in]1,\infty[$. It follows from (2.4.2), **(H3)**, and the change of variable formula, that $m^{\Delta}[\mu](t)$ is absolutely continuous, with density given by

$$m^{\Delta}[\mu](t,x) = \frac{m_0^*\big(\tilde{\Phi}_t^{-1}(x)\big)}{\big|\det\big(D\tilde{\Phi}_t(\tilde{\Phi}_t^{-1}(x))\big)\big|} \quad \text{for a.e. } x\in\mathbb{R}^d,$$

where, for notational convenience, we have set $\tilde{\Phi}_t(\cdot) := \Phi^{\Delta}[\mu](0,t,\cdot)$. Then, by the change of variable formula, we get that

$$\int_{\mathbb{R}^d} |m^{\Delta}[\mu](t,x)|^{\mathsf{p}}\mathrm{d}x = \int_{\mathrm{supp}(m_0^*)} m_0^{*\mathsf{p}}(x)|\mathrm{det}(D_x\tilde{\Phi}_t(x))|^{1-\mathsf{p}}\mathrm{d}x. \tag{2.4.11}$$

On the other hand, thanks to (2.4.6) and [17, Section 2], for a.e. $x\in\mathbb{R}^d$, one has

$$\begin{aligned}|\mathrm{det}(D_x\tilde{\Phi}_t(x))|^{1-\mathsf{p}} &\leq \exp\left((\mathsf{p}-1)\int_0^t \max_{y\in\overline{B}(0,C^*+TC_{\rm bf})}\left[\mathrm{div}\big(D_pH(y,D_xv^{\Delta}[\mu](s,y))\big)\right]_+\mathrm{d}s\right)\\ &\leq \exp\left(\mathsf{p}\int_0^T \max_{y\in\overline{B}(0,C^*+TC_{\rm bf})}\left[\mathrm{div}\big(D_pH(y,D_xv^{\Delta}[\mu](s,y))\big)\right]_+\mathrm{d}s\right).\end{aligned} \tag{2.4.12}$$

Using (2.4.10), it is easy to check that there exists a constant $C_1>0$, independent of t and μ, such that

$$\left[\mathrm{div}\big(D_pH(y,D_xv^{\Delta}[\mu](s,y))\big)\right]_+ \leq C_1 \quad \text{for all } s\in[0,T],\ y\in\overline{B}(0,C^*+TC_{\rm bf}).$$

Thus, it follows from (2.4.11) and (2.4.12) that (2.4.9) holds with $\tilde{c}>0$, independent of t, μ, and p. Finally, the case where $\mathsf{p}=\infty$ follows from the previous one by letting $\mathsf{p}\to\infty$ in (2.4.9). □

Proposition 2.4.2 *Assume* **(H1)**-**(H3)***, let* $(\mu_n)_{n\in\mathbb{N}} \subset C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, and let* $\big((\Delta t_n, \Delta x_n, \varepsilon_n)\big)_{n\in\mathbb{N}} \subset (0,\infty)^3$*. Set* $\Delta_n = (\Delta t_n, \Delta x_n, \varepsilon_n)$ *and let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*. Suppose that, as* $n\to\infty$*,* $\mu_n \to \mu$*,* $\Delta_n \to 0$*,* $(\Delta x_n)^2/\Delta t_n \to 0$*, and* $\Delta x_n = O(\varepsilon_n^2)$*. Then, up to some subsequence, the following hold:*

(i) $(v^{\Delta_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $v[\mu]$*, uniformly over compact subsets of* $[0,T]\times\mathbb{R}^d$*, and, for every* $K\subset[0,T]\times\mathbb{R}^d$ *compact and* $q\in[1,\infty[$*,* $(\mathbb{I}_K D_x v^{\Delta_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $\mathbb{I}_K D_x v[\mu]$ *in* $L^q([0,T]\times\mathbb{R}^d)$*.*

(ii) $\big(m^{\Delta_n}[\mu_n]\big)_{n\in\mathbb{N}}$ *converges in* $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ *towards a solution to* (2.2.31)*. Moreover, the convergence also hold weakly in* $L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$*, if* $\mathsf{p}<\infty$*, and weakly** *in* $L^\infty([0,T]\times\mathbb{R}^d)$*, if* $\mathsf{p}=\infty$*.*

Proof. For every $n\in\mathbb{N}$, let us set $v^n = v^{\Delta_n}[\mu_n]$ and $m^n = m^{\Delta_n}[\mu_n]$.
(i): This follows from Proposition 2.3.4.
(ii): It follows from (2.4.7) that, for every $t\in[0,T]$, $m^n(t)$ belongs to the set $\mathcal{K} = \{\mu\in\mathcal{P}(\mathbb{R}^d) \,|\, \mathrm{supp}(\mu)\subseteq \overline{B}(0, C^* + TC_{\mathrm{bf}})\}$. The set $\mathcal{K}$ is tight and has uniformly integrable moments (see [18, Chapter 5]) and hence, by [18, Proposition 7.1.5], it is a compact subset of $\mathcal{P}_1(\mathbb{R}^d)$. Thus, it follows from (2.4.8) and the Ascoli-Arzelà theorem (see e.g. [125, Theorem 47.1]), that there exists $m^*\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ such that, as $n\to\infty$ and up to some subsequence, $m^n\to m^*$ in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. By Proposition 2.4.1, the convergence also hold weakly in $L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$, if $\mathsf{p}<\infty$, and weakly* in $L^\infty([0,T]\times\mathbb{R}^d)$, if $\mathsf{p}=\infty$. Since m^n solves (2.4.1), for every $t\in[0,T]$, and $\varphi\in C_0^\infty(\mathbb{R}^d)$, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)m^n(t,x)\mathrm{d}x &= \int_{\mathbb{R}^d}\varphi(x)m_0^*(t,x)\mathrm{d}x \\ &\quad + \int_0^t\int_{\mathbb{R}^d}\Big\langle D\varphi(x), D_pH\big(x, D_xv^n(s,x)\big)\Big\rangle m^n(s,x)\mathrm{d}x\mathrm{d}s.\end{aligned} \tag{2.4.13}$$

Thus, by (2.4.2), we can pass to the limit in the previous expression to deduce that m^* solves (2.2.31). □

2.4.1 The Lagrange-Galerkin approximation

The main purpose of the this section is to provide some results in the vein of Propositions 2.4.1 and 2.4.2 for solutions $\mathsf{M}^\Delta[\mu]$ to a LG approximation of (2.4.1) that we proceed to construct.
Let $\mu\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and let $\Delta=(\Delta t,\Delta x,\varepsilon)\in]0,\infty[^3$. For every $k\in\mathcal{I}^*_{\Delta t}$ and $x\in\mathbb{R}^d$,

let $\Phi_k^{\Delta}[\mu](x)$ be the explicit Euler approximation of $\Phi^{\Delta}[\mu](t_k, t_{k+1}, x)$, i.e.

$$\Phi_k^{\Delta}[\mu](x) = x - \Delta t D_p H\big(x, D_x v^{\Delta}[\mu](t_k, x)\big) \quad \text{for all } x \in \mathbb{R}^d. \tag{2.4.14}$$

As in [42], we consider the following semi-discrete approximation of (2.4.4):

$$\begin{aligned} m_{k+1} &= \Phi_k^{\Delta}[\mu]\sharp m_k \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*, \\ m_0 &= m_0^*, \end{aligned} \tag{2.4.15}$$

or, equivalently, for every $k \in \mathcal{I}_{\Delta t}^*$ and every Borel function $\varphi \colon \mathbb{R}^d \to \mathbb{R}$, such that $\varphi\left(\Phi_k^{\Delta}[\mu](\cdot)\right)$ is integrable with respect to m_k,

$$\int_{\mathbb{R}^d} \varphi(x) \mathrm{d}m_{k+1}(x) = \int_{\mathbb{R}^d} \varphi\left(\Phi_k^{\Delta}[\mu](x)\right) \mathrm{d}m_k(x). \tag{2.4.16}$$

Following [124], which mainly deals with a LG approximation of the dual (or transport) equation associated to (2.4.1), let us formaly deduce from (2.4.16) a time-space approximation of (2.4.1). For every $i \in \mathbb{Z}^d$, set

$$E_i = \{x \in \mathbb{R}^d \,|\, |x - x_i|_\infty \leq \Delta x/2\} \tag{2.4.17}$$

and define

$$m_{0,i} = \frac{1}{(\Delta x)^d} \int_{E_i} m_0^*(x) \mathrm{d}x. \tag{2.4.18}$$

Given the regular mesh defined by $\{E_i \,|\, i \in \mathbb{Z}^d\}$, let $\{\beta_i \,|\, i \in \mathbb{Z}^d\}$ be a finite element basis. In the following, we look for an approximation $\mathsf{M}^{\Delta}[\mu]$ of the solution $m^{\Delta}[\mu]$ to (2.4.1) such that

$$\mathsf{M}^{\Delta}[\mu](t_k, x) = \sum_{j \in \mathbb{Z}^d} m_{k,j} \beta_j(x) \quad \text{for all } k \in \mathcal{I}_{\Delta t},\ x \in \mathbb{R}^d, \tag{2.4.19}$$

for some constants $\{m_{k,j} \,|\, k \in \mathcal{I}_{\Delta t},\ j \in \mathbb{Z}^d\}$. In order to determine the latter, we replace m_k and m_{k+1} in (2.4.16) by $\mathsf{M}^{\Delta}(t_k, \cdot)$ and $\mathsf{M}^{\Delta}(t_{k+1}, \cdot)$, respectively, and, given $i \in \mathbb{Z}^d$, we take $\varphi = \beta_i$ to obtain the following equations

$$\sum_{j \in \mathbb{Z}^d} m_{k+1,j} \int_{\mathbb{R}^d} \beta_j(x) \beta_i(x) \mathrm{d}x = \sum_{j \in \mathbb{Z}^d} m_{k,j} \int_{\mathbb{R}^d} \beta_i(\Phi_k^{\Delta}[\mu](x)) \beta_j(x) \mathrm{d}x \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*. \tag{2.4.20}$$

In the context of second-order MFGs with smooth solutions, scheme (2.4.20) is con-

sidered in [41], with a high-order finite element basis and coupled with a high-order semi-Lagrangian scheme for the HJB equation, to provide a high-order scheme for second-order MFGs with smooth solutions. Notice that, in that reference, the authors consider symmetric Lagrangian basis of odd order which preserve the mass but not the positivity of the initial condition $\{m_{0,i} \,|\, i \in \mathbb{Z}^d\}$. In particular, at each time step, the solution to the scheme does not define a probability measure over $\mathbb{R}^d$.
Since we aim to approximate solutions to (2.4.1), which in general are not smooth, from now on we take $\beta_i = \beta_i^0 := \mathbb{I}_{E_i}$ for all $i \in \mathbb{Z}^d$. Under this choice, (2.4.20) and (2.4.18) yield the following LG scheme for (2.4.1):

$$m_{k+1,i} = \frac{1}{(\Delta x)^d} \sum_{j \in \mathbb{Z}^d} m_{k,j} \int_{E_j} \beta_i^0(\Phi_k^\Delta[\mu](x))\mathrm{d}x \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i \in \mathbb{Z}^d, \tag{2.4.21}$$

$$m_{0,i} = \frac{1}{(\Delta x)^d} \int_{E_i} m_0^*(x)\mathrm{d}x \quad \text{for all } i \in \mathbb{Z}^d. \tag{2.4.22}$$

The scheme above is explicit and hence admits a unique solution $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$. Moreover, as shown in Proposition 2.4.3 below, at each time $k \in \mathcal{I}_{\Delta t}$, $\{m_{i,k} \,|\, i \in \mathbb{Z}^d\}$ can be identified with a probability measure over $\mathbb{R}^d$ with compact support. In view of standard compactness results in $\mathcal{P}_1(\mathbb{R}^d)$, this fact will play an important role in our convergence analysis (see the proofs of Proposition 2.4.6 and of Theorem 2.5.2 below). Interestingly, the scheme (2.4.21)-(2.4.22) coincides with the one proposed in [127] (see also [133]) to approximate solutions to continuity equations. Indeed, we have

$$\int_{E_j} \beta_i^0(\Phi_k^\Delta[\mu](x))\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{I}_{E_j \cap \Phi_k^\Delta[\mu]^{-1}(E_i)}(x)\mathrm{d}x = \mathcal{L}^d\Big(E_j \cap \Phi_k^\Delta[\mu]^{-1}(E_i)\Big), \tag{2.4.23}$$

where $\mathcal{L}^d$ denotes the Lebesgue measure in $\mathbb{R}^d$. Plugging this expression in (2.4.21) yields the scheme in [127, Section 2.2]. Notice that our main results for solutions to (2.4.21)-(2.4.22), contained in Propositions 2.4.5 and 2.4.6 below, do not follow from the results in [127], [133]. Therefore, the analysis in this section provides a complementary study to the one in [127], [133] for the approximation (2.4.21)-(2.4.22) of continuity equations.

2.4.2 Properties of LG scheme

We begin with a preliminary result stating that the solution to (2.4.21)-(2.4.22) is supported on a compact set, which is independent of the discretization parameters provided that Δx is of the order of Δt.

Lemma 2.4.1 *Assume* **(H1)**-**(H3)***, let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$*, and let* $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$ *be the solution to* (2.4.21)-(2.4.22)*. Then for every* $c > 0$ *there exists* $\widetilde{C}^* > 0$*, independent of* μ*, such that, if* $\Delta x \leq c\Delta t$*, for every* $k \in \mathcal{I}_{\Delta t}$ *we have* $m_{k,i} = 0$ *if* $x_i \notin \overline{B}_\infty(0, \widetilde{C}^*)$*.*

Proof. Let $c > 0$ and suppose that $\Delta x \leq c\Delta t$. For every $k \in \mathcal{I}^*_{\Delta t}$, set $r_k = \sup\{|x_i|_\infty \,|\, m_{k,i} \neq 0,\ i \in \mathbb{Z}^d\} \in [0,\infty]$. By (2.4.5), (2.4.21), and **(H3)**(i), we have

$$r_{k+1} \leq r_k + \Delta t C_{\text{bf}} + \frac{\Delta x}{2} \leq r_k + \Delta t\left(C_{\text{bf}} + \frac{c}{2}\right) \leq C^* + N_{\Delta t}\Delta t\left(C_{\text{bf}} + \frac{c}{2}\right) = C^* + T\left(C_{\text{bf}} + \frac{c}{2}\right), \tag{2.4.24}$$

for all $k \in \mathcal{I}^*_{\Delta t}$. The result follows by letting $\widetilde{C}^* = C^* + T(C_{\text{bf}} + c/2)$. □

Let $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and let $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$. As a consequence of the previous result, in (2.4.21) it suffices to compute $m_{i,k+1}$ for $i \in \mathcal{I}_{\Delta x}$, where

$$\mathcal{I}_{\Delta x} := \left\{i \in \mathbb{Z}^d \,|\, x_i \in \overline{B}_\infty(0, \widetilde{C}^*)\right\}. \tag{2.4.25}$$

Given the constants $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$, computed with (2.4.21)-(2.4.22), we extend $\mathsf{M}^\Delta[\mu]$, given by (2.4.19), to $[0,T] \times \mathbb{R}^d$ as follows:

$$\mathsf{M}^\Delta[\mu](t,x) = \left(\frac{t_{k+1} - t}{\Delta t}\right)\mathsf{M}^\Delta[\mu](t_k, x) + \left(\frac{t - t_k}{\Delta t}\right)\mathsf{M}^\Delta[\mu](t_{k+1}, x) \quad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ t \in [t_k, t_{k+1}),\ x \in \mathbb{R}^d. \tag{2.4.26}$$

In the following proposition, we state, for later use, some simple properties of the solution to (2.4.21)-(2.4.22).

Proposition 2.4.3 *Assume* **(H1)**-**(H3)***, let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$*, and let* $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\ i \in \mathbb{Z}^d\}$ *be the solution to* (2.4.21)-(2.4.22)*. Then the following hold:*

(i) $\mathsf{M}^\Delta[\mu](t,x) \geq 0$ *for all* $t \in [0,T]$, $x \in \mathbb{R}^d$.

(ii) *Let* $c > 0$*. If* $\Delta x \leq c\Delta t$ *and* $\widetilde{C}^* > 0$ *is given by Lemma 2.4.1, we have*

$$\operatorname{supp}\left(\mathsf{M}^\Delta[\mu](t,\cdot)\right) \subseteq \overline{B}_\infty(0, \widetilde{C}^*) \quad \textit{for all } t \in [0,T]. \tag{2.4.27}$$

(iii) *Let* $a=(a_i)_{i\in\mathbb{Z}^d}\subset\mathbb{R}$ *and set* $\varphi_a(x)=\sum_{i\in\mathbb{Z}^d}a_i\beta_i^0(x)$ *for all* $x\in\mathbb{R}^d$*. Then we have*

$$\int_{\mathbb{R}^d}\varphi_a(x)\mathsf{M}^{\Delta}[\mu](t_{k+1},x)\mathrm{d}x=\int_{\mathbb{R}^d}\varphi_a(\Phi_k^{\Delta}[\mu](x))\mathsf{M}^{\Delta}[\mu](t_k,x)\mathrm{d}x\quad\textit{for all } k\in\mathcal{I}_{\Delta t}^*. \tag{2.4.28}$$

(iv) $\int_{\mathbb{R}^d}\mathsf{M}^{\Delta}[\mu](t,x)\mathrm{d}x=1$ *for all* $t\in[0,T]$.

Proof.
(i): Using that $m_0^*\geq 0$ and $\beta_i^0\geq 0$ for all $i\in\mathbb{Z}^d$, this assertion follows directly from (2.4.21)-(2.4.22).
(ii): This follows from Lemma 2.4.1 and (2.4.26).
(iii): For every $k\in\mathcal{I}_{\Delta t}^*$, by (2.4.21), we have

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi_a\mathsf{M}^{\Delta}[\mu](t_{k+1},x)\mathrm{d}x&=\sum_{i\in\mathbb{Z}^d}\sum_{j\in\mathbb{Z}^d}a_jm_{k+1,i}\int_{\mathbb{R}^d}\beta_i^0(x)\beta_j^0(x)\mathrm{d}x=\sum_{i\in\mathbb{Z}^d}a_im_{k+1,i}(\Delta x)^d\\&=\sum_{i\in\mathbb{Z}^d}a_i\sum_{j\in\mathbb{Z}^d}m_{k,j}\int_{E_j}\beta_i^0(\Phi_k^{\Delta}[\mu](x))\mathrm{d}x=\sum_{j\in\mathbb{Z}^d}m_{k,j}\int_{\mathbb{R}^d}\varphi_a(\Phi_k^{\Delta}[\mu](x))\beta_j^0(x)\mathrm{d}x\\&=\int_{\mathbb{R}^d}\varphi_a(\Phi_k^{\Delta}[\mu](x))\mathsf{M}^{\Delta}[\mu](t_k,x)\mathrm{d}x.\end{aligned}\tag{2.4.29}$$

Notice that the changes of the order of summation above are justified by (2.4.3).
(iv): By (2.4.3), with $\varphi_a(x):=\sum_{i\in\mathbb{Z}^d}\beta_i^0(x)=1$, we obtain the result for $t=t_k$, with $k\in\mathcal{I}_{\Delta t}$. The result for every $t\in[0,T]$ follows from (2.4.26). ☐
In what follows, given $\varphi\colon\mathbb{R}^d\to\mathbb{R}$ we set

$$I^0[\varphi](x)=\sum_{i\in\mathbb{Z}^d}\varphi(x_i)\beta_i^0(x)\quad\text{for all } x\in\mathbb{R}^d. \tag{2.4.30}$$

We will need the following estimate in some of the proofs below.

Lemma 2.4.2 *Assume* **(H1)**-**(H3)***, let* $\mu\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, let* $\Delta=(\Delta t,\Delta x,\varepsilon)\in]0,\infty[^3$*, and, given* $L>0$*, let* $\varphi\colon\mathbb{R}^d\to\mathbb{R}$ *be* L*-Lipschitz. Then, for every* $k\in\mathcal{I}_{\Delta t}^*$*, we have*

$$\left|\int_{\mathbb{R}^d}\varphi(x)\mathsf{M}^{\Delta}[\mu](t_{k+1},x)\mathrm{d}x-\int_{\mathbb{R}^d}\varphi(\Phi_k^{\Delta}[\mu](x))\mathsf{M}^{\Delta}[\mu](t_k,x)\mathrm{d}x\right|\leq L\sqrt{d}\Delta x. \tag{2.4.31}$$

Proof. Since $\sum_{i\in\mathbb{Z}^d}\beta_i^0(x) = 1$ for all $x \in \mathbb{R}^d$, we have

$$\begin{aligned}\left|\varphi(x) - I^0[\varphi](x)\right| &= \left|\sum_{i\in\mathbb{Z}^d}(\varphi(x)-\varphi(x_i))\beta_i^0(x)\right| \\ &\leq \sum_{i\in\mathbb{Z}^d}|\varphi(x)-\varphi(x_i)|\beta_i^0(x) \leq \frac{L\sqrt{d}}{2}\Delta x \quad \text{for all } x\in\mathbb{R}^d. \end{aligned} \tag{2.4.32}$$

It follows that $\|\varphi - I^0[\varphi]\|_\infty \leq (L\sqrt{d}/2)\Delta x$ and, hence,

$$\left|\frac{1}{(\Delta x)^d}\int_{E_i}\varphi(x)\mathrm{d}x - \varphi(x_i)\right| = \left|\frac{1}{(\Delta x)^d}\int_{E_i}(\varphi(x)-\varphi(x_i))\mathrm{d}x\right| \leq \frac{L\sqrt{d}}{2}\Delta x, \tag{2.4.33}$$

from which we deduce that, for every $k \in \mathcal{I}_{\Delta t}^*$ and $j \in \mathbb{Z}^d$,

$$\begin{aligned}&\left|\sum_{i\in\mathbb{Z}^d}\frac{1}{(\Delta x)^d}\int_{E_i}\varphi(x)\mathrm{d}x\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y - \int_{E_j}I^0[\varphi](\Phi_k^\Delta[\mu](x))\mathrm{d}x\right| \\ &= \left|\sum_{i\in\mathbb{Z}^d}\left(\frac{1}{(\Delta x)^d}\int_{E_i}\varphi(x)\mathrm{d}x - \varphi(x_i)\right)\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y\right| \leq \frac{L\sqrt{d}}{2}\Delta x\sum_{i\in\mathbb{Z}^d}\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y \\ &= \frac{L\sqrt{d}}{2}(\Delta x)^{d+1}. \end{aligned} \tag{2.4.34}$$

Therefore, by (2.4.34) and (2.4.32), we obtain

$$\left|\sum_{i\in\mathbb{Z}^d}\frac{1}{(\Delta x)^d}\int_{E_i}\varphi(x)\mathrm{d}x\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y - \int_{E_j}\varphi(\Phi_k^\Delta[\mu](x))\mathrm{d}x\right| \leq L\sqrt{d}(\Delta x)^{d+1}. \tag{2.4.35}$$

Finally, from (2.4.21), (2.4.35), and Proposition 2.4.3(iv), we get

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}\varphi(x)\mathsf{M}^\Delta[\mu](t_{k+1},x)\mathrm{d}x - \int_{\mathbb{R}^d}\varphi(\Phi_k^\Delta[\mu](x))\mathsf{M}^\Delta[\mu](t_k,x)\mathrm{d}x\right| \\ &= \left|\sum_{i\in\mathbb{Z}^d}\int_{E_i}\varphi(x)\mathrm{d}x\frac{1}{(\Delta x)^d}\sum_{j\in\mathbb{Z}^d}m_{k,j}\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y - \sum_{i\in\mathbb{Z}^d}m_{k,i}\int_{E_i}\varphi(\Phi_k^\Delta[\mu](x))\mathrm{d}x\right| \\ &= \left|\sum_{j\in\mathbb{Z}^d}m_{k,j}\left(\sum_{i\in\mathbb{Z}^d}\frac{1}{(\Delta x)^d}\int_{E_i}\varphi(x)\int_{E_j}\beta_i^0(\Phi_k^\Delta[\mu](y))\mathrm{d}y - \int_{E_j}\varphi(\Phi_k^\Delta[\mu](x))\mathrm{d}x\right)\right| \leq L\sqrt{d}(\Delta x), \end{aligned} \tag{2.4.36}$$

which shows (2.4.31). □

In the next result, we study the equicontinuity of the family $\{\mathsf{M}^{\Delta}[\mu] \mid \Delta \in]0,\infty[^3\}$, under the condition that Δx is, at most, of the order of Δt.

Proposition 2.4.4 *Assume* **(H1)**-**(H3)**, *let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$. *Then, for every* $c > 0$, *if* $\Delta x \le c\Delta t$, *we have*

$$d_1(\mathsf{M}^{\Delta}[\mu](t), \mathsf{M}^{\Delta}[\mu](s)) \le (C_{\mathrm{bf}} + c\sqrt{d})|t-s| \quad \textit{for all } t,\ s \in [0,T], \tag{2.4.37}$$

where, for every $t \in [0,T]$, $\mathsf{M}^{\Delta}[\mu](t) \in \mathcal{P}_1(\mathbb{R}^d)$ *denotes the measure* $\mathrm{d}\mathsf{M}^{\Delta}[\mu](t)(x) = \mathsf{M}^{\Delta}[\mu](t,x)\mathrm{d}x$.

Proof. Let $\varphi \in \mathrm{Lip}_1(\mathbb{R}^d)$ and define $\psi_\varphi \colon [0,T] \to \mathbb{R}$ by

$$\psi_\varphi(t) = \int_{\mathbb{R}^d} \varphi(x)\mathsf{M}^{\Delta}[\mu](t,x)\mathrm{d}x \quad \text{for all } t \in [0,T]. \tag{2.4.38}$$

It follows from (2.4.26) that ψ_φ is continuous and affine on every interval $[t_k, t_{k+1}]$ ($k \in \mathcal{I}^*_{\Delta t}$). Thus, $\psi_\varphi \in W^{1,\infty}(]0,T[)$ and

$$\left\|\frac{\mathrm{d}}{\mathrm{d}t}\psi_\varphi\right\|_\infty = \frac{1}{\Delta t}\max_{k\in\mathcal{I}^*_{\Delta}}\left|\int_{\mathbb{R}^d}\varphi(x)\Big(\mathsf{M}^{\Delta}[\mu](t_{k+1},x) - \mathsf{M}^{\Delta}[\mu](t_k,x)\Big)\mathrm{d}x\right|. \tag{2.4.39}$$

In order to estimate the right-hand side of (2.4.39), fix $k \in \mathcal{I}^*_{\Delta t}$ and notice that, by Lemma 2.4.2, (2.4.14), (2.4.5), Proposition 2.4.3(iv), and $\Delta x \le c\Delta t$, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)\Big(\mathsf{M}^{\Delta}[\mu](t_{k+1},x) - \mathsf{M}^{\Delta}[\mu](t_k,x)\Big)\mathrm{d}x &\le \int_{\mathbb{R}^d}\Big(\varphi(\Phi^{\Delta}_k[\mu](x)) - \varphi(x)\Big)\mathsf{M}^{\Delta}[\mu](t_k,x)\mathrm{d}x \\ &\quad + \sqrt{d}\Delta x \\ &\le \int_{\mathbb{R}^d}\Big|\Phi^{\Delta}_k[\mu](x) - x\Big|\mathsf{M}^{\Delta}[\mu](t_k,x)\mathrm{d}x + \sqrt{d}\Delta x \\ &\le (C_{\mathrm{bf}} + c\sqrt{d})\Delta t. \end{aligned} \tag{2.4.40}$$

Changing φ by $-\varphi$ in the previous computation, (2.4.40) implies that

$$\left|\int_{\mathbb{R}^d}\varphi(x)\Big(\mathsf{M}^{\Delta}[\mu](t_{k+1},x) - \mathsf{M}^{\Delta}[\mu](t_k,x)\Big)\mathrm{d}x\right| \le (C_{\mathrm{bf}} + c\sqrt{d})\Delta t. \tag{2.4.41}$$

and hence, by (2.4.39), $\left\|\frac{\mathrm{d}}{\mathrm{d}t}\psi_\varphi\right\|_\infty \leq (C_{\mathrm{bf}} + c\sqrt{d})$. Thus, we deduce that

$$\int_{\mathbb{R}^d} \varphi(x)\Big(\mathsf{M}^\Delta[\mu](t,x) - \mathsf{M}^\Delta[\mu](s,x)\Big)\mathrm{d}x \leq (C_{\mathrm{bf}} + c\sqrt{d})|t-s| \quad \text{for all } t,\, s \in [0,T] \tag{2.4.42}$$

and (2.4.37) follows from (2.2.2). ☐

The following result state a stability property for $\mathsf{M}^\Delta[\mu]$ which is analogous to the one in Proposition 2.4.1 for $m^\Delta[\mu]$.

Proposition 2.4.5 *Assume* **(H1)**-**(H3)***, let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, and let* $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0,\infty[^3$*. Then for every* $c_1,\, c_2 > 0$*, there exists* $\widetilde{C} > 0$*, independent of* μ*, such that, if* Δ *is small enough,* $\Delta x \leq c_1 \Delta t$*, and* $\Delta t \leq c_2 \varepsilon^2$*, we have* $\mathsf{M}^\Delta[\mu](t,\cdot) \in L^{\mathsf{p}}(\mathbb{R}^d)$ *for all* $t \in [0,T]$ *and*

$$\|\mathsf{M}^\Delta[\mu](t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \widetilde{C}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \text{for all } t \in [0,T]. \tag{2.4.43}$$

Proof. Let $c_1,\, c_2 > 0$, suppose that $\Delta x \leq c_1 \Delta t,\ \Delta t \leq c_2 \varepsilon^2$, fix $k \in \mathcal{I}^*_{\Delta t}$, and let $\widetilde{C}^* > 0$ be as in Lemma 2.4.1. Then, by (2.4.5), there exists $R > 0$, independent of (μ, Δ, k), such that $\big(\Phi_k^\Delta[\mu]\big)^{-1}\big(\overline{B}_\infty(0,\widetilde{C}^*)\big) \subset B_\infty(0,R)$. The regularity of H and estimate (2.3.16), with $\ell = 2$, yield the existence of $C_R > 0$, independent of (μ, Δ, k), such that, for every $x \in B_\infty(0,R)$, the norm of the matrix

$$\begin{aligned} &D_x\big(D_p H(x, D_x v^\Delta[\mu](t_k,x))\big) \\ &\quad = D_p^2 H\Big(x, D_x v^\Delta[\mu](t_k,x)\Big) D_x^2 v^\Delta[\mu](t_k,x) + D_{xp}^2 H\Big(x, D_x v^\Delta[\mu](t_k,x)\Big), \end{aligned} \tag{2.4.44}$$

induced by the 2-norm in $\mathbb{R}^d$, is bounded by C_R/ε. In particular, $D_p H(\cdot, D_x v^\Delta[\mu](t_k,\cdot))$ is (C_R/ε)-Lipschitz on $B_\infty(0,R)$. Thus, expression (2.4.14) and the inequality $\Delta t/\varepsilon \leq c_2\varepsilon$ imply that, if ε is small enough, there exists $C_1 > 0$, independent of (μ, Δ, k), such that

$$\big|\Phi_k^\Delta[\mu](x) - \Phi_k^\Delta[\mu](y)\big| \geq C_1|x-y| \quad \text{for all } x,\, y \in B_\infty(0,R), \tag{2.4.45}$$

which implies that $\Phi_k^\Delta[\mu]$ is injective on $B_\infty(0,R)$, and, denoting by I_d the $d \times d$ identity matrix,

$$D_x\Phi_k^\Delta[\mu](x) = I_d - \Delta t D_x[D_p H(x, D_x v^\Delta[\mu](t_k,x))] \tag{2.4.46}$$

is invertible for $x \in B_\infty(0,R)$. In particular, $\Phi_k^\Delta[\mu]$ is a diffeomorphism of $B_\infty(0,R)$ onto $\Phi_k^\Delta[\mu](B_\infty(0,R))$. Let us suppose first that $\mathsf{p} \in]1,\infty[$. By the change of variable formula, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\Big(\mathsf{M}^{\Delta}[\mu]\big(t_{k+1},\Phi_k^{\Delta}[\mu](x)\big)\Big)^{\mathsf{p}}\mathrm{d}x &= \int_{B_\infty(0,R)}\Big(\mathsf{M}^{\Delta}[\mu]\big(t_{k+1},\Phi_k^{\Delta}[\mu](x)\big)\Big)^{\mathsf{p}}\mathrm{d}x\\ &= \int_{\Phi_k^{\Delta}[\mu](B_\infty(0,R))}\big(\mathsf{M}^{\Delta}[\mu](t_{k+1},y)\big)^{\mathsf{p}}\Big|\det\Big(D_x\Phi_k^{\Delta}[\mu]\Big(\Phi_k^{\Delta}[\mu]^{-1}(y)\Big)\Big)\Big|^{-1}\mathrm{d}y. \quad (2.4.47)\end{aligned}$$

Using again that the norm of $D_x\big(D_pH(\cdot,D_xv^{\Delta}[\mu](t_k,\cdot))\big)$ is bounded by C_R/ε on $B_\infty(0,R)$, relation (2.4.46) and a Taylor expansion for the determinant yield the existence of $C_2>0$, independent of (μ,Δ,k), such that

$$\Big|\det\big(D_x\Phi_k^{\Delta}[\mu](x)\big)-\Big(1-\Delta t\,\mathrm{Tr}\,\big(D_x\big(D_pH(x,D_xv^{\Delta}[\mu](t_k,x))\big)\big)\Big)\Big|\leq C_2(\Delta t/\varepsilon)^2 \quad \text{for all } x\in B_\infty(0,R),$$

where, given $B\in\mathbb{R}^{d\times d}$, $\mathrm{Tr}(B)$ denotes its trace. In turn, we get the existence of $C_3>0$, independent of (μ,Δ,k), such that

$$\Big|\big|\det\big(D_x\Phi_k^{\Delta}[\mu](x)\big)\big|^{-1}-\Big(1+\Delta t\,\mathrm{Tr}\,\big(D_x\big(D_pH(x,D_xv^{\Delta}[\mu](t_k,x))\big)\big)\Big)\Big|\leq C_3(\Delta t/\varepsilon)^2, \quad (2.4.48)$$

for all $x\in B_\infty(0,R)$. Since $\Delta x\leq c_1c_2\varepsilon^2$, Proposition 2.3.5 implies that $D_x^2v^{\Delta}[\mu](t_k,x)-\widetilde{C}_{\mathrm{hb}}(1+(c_1c_2)^2)I_d$ is negative semidefinite. Using that $H(\cdot,\cdot)$ is of class C^2 and convex with respect to its second argument, it follows from (2.4.44) and [48, Lemma 1.6.4] that there exists $C_4>0$, independent of (μ,Δ,k), such that

$$\mathrm{Tr}(D_x[D_pH(x,D_xv^{\Delta}[\mu](t_k,x))])\leq C_4 \quad \text{for all } x\in B_\infty(0,R), \quad (2.4.49)$$

which, together with (2.4.48), yields

$$\Big|\det\big(D_x\Phi_k^{\Delta}[\mu](x)\big)\Big|^{-1}\leq 1+C_5\Delta t \quad \text{for all } x\in B_\infty(0,R), \quad (2.4.50)$$

where $C_5=C_4+C_3c_2$. Therefore, by (2.4.47), we get

$$\int_{\mathbb{R}^d}\Big(\mathsf{M}^{\Delta}[\mu]\big(t_{k+1},\Phi_k^{\Delta}[\mu](x)\big)\Big)^{\mathsf{p}}\mathrm{d}x\leq(1+C_5\Delta t)\int_{\mathbb{R}^d}\big(\mathsf{M}^{\Delta}[\mu](t_{k+1},x)\big)^{\mathsf{p}}\mathrm{d}x. \quad (2.4.51)$$

Setting $\mathsf{p}^* = \mathsf{p}/(\mathsf{p}-1)$, it follows from (2.4.28) and Hölder's inequality that

$$\begin{aligned} \|\mathsf{M}^\Delta[\mu](t_{k+1},\cdot)\|^{\mathsf{p}}_{L^{\mathsf{p}}(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \big(\mathsf{M}^\Delta[\mu](t_{k+1},x)\big)^{\mathsf{p}-1}\mathsf{M}^\Delta[\mu](t_{k+1},x)\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \big(\mathsf{M}^\Delta[\mu](t_{k+1},\Phi^\Delta_k[\mu](x))\big)^{\mathsf{p}-1}\mathsf{M}^\Delta[\mu](t_k,x)\mathrm{d}x \\ &\leq \left(\int_{\mathbb{R}^d} \big(\mathsf{M}^\Delta[\mu](t_{k+1},\Phi^\Delta_k[\mu](x))\big)^{\mathsf{p}}\mathrm{d}x\right)^{\frac{1}{\mathsf{p}^*}} \|\mathsf{M}^\Delta[\mu](t_k,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \\ &\leq (1+C_5\Delta t)^{\frac{1}{\mathsf{p}^*}}\|\mathsf{M}^\Delta[\mu](t_{k+1},\cdot)\|^{\frac{\mathsf{p}}{\mathsf{p}^*}}_{L^{\mathsf{p}}(\mathbb{R}^d)}\|\mathsf{M}^\Delta[\mu](t_k,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \\ &\leq (1+C_5\Delta t)\|\mathsf{M}^\Delta[\mu](t_{k+1},\cdot)\|^{\mathsf{p}-1}_{L^{\mathsf{p}}(\mathbb{R}^d)}\|\mathsf{M}^\Delta[\mu](t_k,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)}. \end{aligned} \tag{2.4.52}$$

In turn, we deduce that

$$\|\mathsf{M}^\Delta[\mu](t_{k+1},\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq (1+C_5\Delta t)\|\mathsf{M}^\Delta[\mu](t_k,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)}. \tag{2.4.53}$$

By (2.4.22) and Jensen's inequality, we have $\|\mathsf{M}^\Delta[\mu](0,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)}$, and hence

$$\begin{aligned} \|\mathsf{M}^\Delta[\mu](t_{k+1},\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} &\leq (1+C_5\Delta t)^N\|\mathsf{M}^\Delta[\mu](0,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \\ &\leq e^{C_5T}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)}, \end{aligned} \tag{2.4.54}$$

which, by (2.4.26), shows (2.4.43), with $\widetilde{C} = e^{C_5T}$. If $\mathsf{p} = \infty$, then (2.4.43) holds for every $\mathsf{p}' \in]1,\infty[$. Noticing that $\widetilde{C}$ is independent of p' and that, by Proposition 2.4.3(ii), for every $t \in [0,T]$, the support of $\mathsf{M}^\Delta[\mu](t,\cdot)$ is contained in $\overline{B}_\infty(0,\widetilde{C}^*)$, (2.4.43) for $\mathsf{p}=\infty$ follows by letting $\mathsf{p}' \to \infty$. □

The next result provides the analogous for $\mathsf{M}^\Delta[\mu]$ of Proposition 2.4.2 for $m^\Delta[\mu]$.

Proposition 2.4.6 *Assume* **(H1)**-**(H3)**, *let* $(\mu_n)_{n\in\mathbb{N}} \subset C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, *and let* $\big((\Delta t_n, \Delta x_n, \varepsilon_n)\big)_{n\in\mathbb{N}} \subset (0,\infty)^3$. *Set* $\Delta_n = (\Delta t_n, \Delta x_n, \varepsilon_n)$ *and let* $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. *Suppose that, as* $n\to\infty$, $\mu_n \to \mu$, $\Delta_n \to 0$, $\Delta x_n = o(\Delta t_n)$, *and* $\Delta t_n = O(\varepsilon_n^2)$. *Then, up to some subsequence, the following hold:*

(i) $(v^{\Delta_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $v[\mu]$, *uniformly over compact subsets of* $[0,T]\times\mathbb{R}^d$, *and, for every* $K \subset [0,T]\times\mathbb{R}^d$ *compact and* $q \in [1,\infty[$, $(\mathbb{I}_K D_x v^{\Delta_n}[\mu_n])_{n\in\mathbb{N}}$ *converges to* $\mathbb{I}_K D_x v[\mu]$ *in* $L^q([0,T]\times\mathbb{R}^d)$.

(ii) $\big(\mathsf{M}^{\Delta_n}[\mu_n]\big)_{n\in\mathbb{N}}$ *converges in* $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ *towards a solution* $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d)) \cap L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$ *to* (2.2.31). *Moreover, the convergence also hold weakly in* $L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$, *if* $\mathsf{p} < \infty$, *and weakly** *in* $L^\infty([0,T]\times\mathbb{R}^d)$, *if* $\mathsf{p}=\infty$. *In addi-*

tion, there exists $\widetilde{C} > 0$ *such that*

$$\|m(t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \leq \widetilde{C}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \textit{for all } t \in [0,T]. \tag{2.4.55}$$

Proof. Let us set $v^n := v^{\Delta_n,\varepsilon_n}[\mu_n]$, $\mathsf{M}^n = \mathsf{M}^{\Delta_n,\varepsilon_n}[\mu_n]$, and $\Phi_k^n = \Phi^{\Delta_n,\varepsilon_n}[\mu_n]$ for all $k \in \mathcal{I}^*_{\Delta t_n}$.
(i): This corresponds to Proposition 2.4.2(i)
(ii): Arguing as in the proof of Proposition 2.4.2(ii), it follows from Proposition 2.4.3(ii), [18, Proposition 7.1.5], Proposition 2.4.4, and the Arzelá-Ascoli theorem, that there exists $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ such that, as $n \to \infty$ and up to some subsequence, $\mathsf{M}^n \to m$ in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. Moreover, by Proposition 2.4.5, the convergence holds weakly, if $\mathsf{p} < \infty$, and weakly* in $L^\infty([0,T]\times\mathbb{R}^d)$, if $\mathsf{p} = \infty$, and m satisfies (2.4.55). It remains to show that m solves (2.2.31). Let $t \in]0,T]$ and let $k(n) \in \mathcal{I}^*_{\Delta t_n}$ be such that $t \in]t_{k(n)}, t_{k(n)+1}]$. For every $\varphi \in C_0^\infty(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \varphi(x)\mathsf{M}^n\left(t_{k(n)},x\right)\mathrm{d}x = \int_{\mathbb{R}^d}\varphi(x)\mathsf{M}^n(0,x)\mathrm{d}x + \sum_{k=0}^{k(n)-1}\int_{\mathbb{R}^d}\varphi(x)\big(\mathsf{M}^n(t_{k+1},x) - \mathsf{M}^n(t_k,x)\big)\mathrm{d}x. \tag{2.4.56}$$

Let $k \in \mathcal{I}^*_{\Delta t}$. Since (2.4.14) and (2.4.5) yield $|\Phi_k^n(x) - x| = O(\Delta t_n)$ for all $x \in \operatorname{supp}(\varphi)$, by Lemma 2.4.2, a Taylor expansion, and Proposition 2.4.3(2.4.3), we have

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)\big(\mathsf{M}^n(t_{k+1},x) - \mathsf{M}^n(t_k,x)\big)\mathrm{d}x &= \int_{\mathbb{R}^d}\big(\varphi(\Phi_k^n(x)) - \varphi(x)\big)\mathsf{M}^n(t_k,x)\mathrm{d}x + O(\Delta x_n)\\ &= -\Delta t_n\int_{\mathbb{R}^d}\big\langle D\varphi(x), D_pH(x,D_xv^n(t_k,x))\big\rangle\mathsf{M}^n(t_k,x)\mathrm{d}x\\ &\quad + O(\Delta x_n) + O((\Delta t_n)^2),\end{aligned} \tag{2.4.57}$$

which, combined with (2.4.56), yields

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)\mathsf{M}^n(t_{k(n)},x)\mathrm{d}x = \int_{\mathbb{R}^d}&\varphi(x)\mathsf{M}^n(0,x)\mathrm{d}x\\ &-\Delta t_n\sum_{k=0}^{k(n)-1}\int_{\mathbb{R}^d}\big\langle D\varphi(x), D_pH(x,D_xv^n(t_k,x))\big\rangle\mathsf{M}^n(t_k,x)\mathrm{d}x\\ &+O\left(\frac{\Delta x_n}{\Delta t_n} + \Delta t_n\right).\end{aligned} \tag{2.4.58}$$

Since φ has a compact support, it follows from (2.3.16), with $\ell = 2$, that there exists $C_\varphi > 0$ such that $\big\langle D\varphi(\cdot), D_pH(\cdot, D_xv^n(t_k,\cdot))\big\rangle$ is $(C_\varphi/\varepsilon_n)$-Lipschitz. Thus, by Proposition 2.4.4, for every $k \in \mathcal{I}_{\Delta t}^{*}$, we have

$$\left| \int_{\mathbb{R}^d} \big\langle D\varphi(x), D_p H(x, D_x v^n(t_k, x)) \big\rangle \big(\mathsf{M}^n(s,x) - \mathsf{M}^n(t_k,x)\big) \mathrm{d}x \right| = O\left(\frac{\Delta t_n}{\varepsilon_n}\right)$$

for all $s \in [t_k, t_{k+1}]$.

Recalling that $D_x v^n(s,x) = D_x v^n(t_k,x)$ for all $s \in [t_k, t_{k+1}[$ and $x \in \mathbb{R}^d$, we obtain

$$\begin{aligned} \Delta t_n \int_{\mathbb{R}^d} \big\langle D\varphi(x), D_p H(x, D_x v^n(t_k, x)) \big\rangle \mathsf{M}^n(t_k, x) \mathrm{d}x \\ = \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}^d} \big\langle D\varphi(x), D_p H(x, D_x v^n(s, x)) \big\rangle \mathsf{M}^n(s, x) \mathrm{d}x \mathrm{d}s + O\left(\frac{(\Delta t_n)^2}{\varepsilon_n}\right) \end{aligned} \tag{2.4.59}$$

and hence, in view of (2.4.58), we deduce that, for n large enough,

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) \mathsf{M}^n\big(t_{k(n)}, x\big) \mathrm{d}x \\ = \int_{\mathbb{R}^d} \varphi(x) \mathsf{M}^n(0, x) \mathrm{d}x - \int_0^T \int_{\mathbb{R}^d} \mathbb{I}_{[0, t_{k(n)}]} \big\langle D\varphi(x), D_p H(x, D_x v^n(s, x)) \big\rangle \mathsf{M}^n(s, x) \mathrm{d}x \mathrm{d}s \\ + O\left(\frac{\Delta x_n}{\Delta t_n} + \frac{\Delta t_n}{\varepsilon_n}\right). \end{aligned} \tag{2.4.60}$$

Finally, by (2.4.6),

$$\mathbb{I}_{[0, t_{k(n)}]}(\cdot) \big\langle D\varphi(\cdot), D_p H(x, D_x v^n(\cdot, \cdot)) \big\rangle \underset{n\to\infty}{\longrightarrow} \mathbb{I}_{[0,t]}(\cdot) \big\langle D\varphi(\cdot), D_p H(x, D_x v[m](\cdot, \cdot)) \big\rangle, \tag{2.4.61}$$

in $L^q([0,T] \times \mathbb{R}^d)$, for every $q \in [1, \infty[$, and, hence, we can pass to the limit in (2.4.60) to obtain that m satisfies (2.2.31). □

2.5 A Lagrange-Galerkin scheme for the the mean field games system

In this section, we combine the schemes discussed in Sections 2.3 and 2.4 to obtain a scheme for system (MFG) and we provide a convergence result.
Let $\Delta = (\Delta t, \Delta x, \varepsilon) \in]0, \infty[^3$, let $\widetilde{C}^* > 0$ be as in Lemma 2.4.1, and define

$$\mathfrak{D}^{\Delta t, \Delta x} = \left\{ \mu = (\mu_{k,i}) \,\middle|\, \mu_{k,i} \geq 0, \ \sum_{j \in \mathbb{Z}^d} \mu_{k,j} (\Delta x)^d = 1 \text{ for all } k \in \mathcal{I}_{\Delta t}, \ i \in \mathcal{I}_{\Delta x} \right\}, \tag{2.5.1}$$

where $\mathcal{I}_{\Delta x}$ is defined in (2.4.25). Notice that $\mathfrak{D}^{\Delta t,\Delta x}$ is a convex and compact subset of $\mathbb{R}^{(N_{\Delta t}+1)\times(2N_{\Delta x}+1)^d}$. Given $\mu \in \mathfrak{D}^{\Delta t,\Delta x}$ define $\tilde{\mu} \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ as

$$\mathrm{d}\tilde{\mu}(t)(x) = \left(\frac{t-t_k}{\Delta t}\right)\sum_{i\in\mathcal{I}_{\Delta x}} \mu_{k+1,i}\beta_i^0(x)\mathrm{d}x + \left(\frac{t_{k+1}-t}{\Delta t}\right)\sum_{i\in\mathcal{I}_{\Delta x}} \mu_{k,i}\beta_i^0(x)\mathrm{d}x$$
$$\text{for all } k\in\mathcal{I}^*_{\Delta t},\ t\in[t_k,t_{k+1}[. \quad (2.5.2)$$

The discretization of (MFG) that we propose is the following: find $\mu \in \mathfrak{D}^{\Delta t,\Delta x}$ such that

$$\mu_{k,i} = \mathsf{M}^{\Delta}[\tilde{\mu}](t_k,x_i) \quad \text{for all } k\in\mathcal{I}^*_{\Delta t},\ i\in\mathcal{I}_{\Delta x}, \qquad (\mathrm{MFG}^{\Delta})$$

where we recall that $\mathsf{M}^{\Delta}[\tilde{\mu}]$ is defined in (2.4.26).

Theorem 2.5.1 *Assume that* **(H1)**-**(H3)** *hold. Then, if $\Delta t/\varepsilon$ is small enough, system* (MFG^{Δ}) *admits at least one solution.*

Proof. Consider the application $T\colon \mathfrak{D}^{\Delta t,\Delta x} \to \mathbb{R}^{(N_{\Delta t}+1)\times(2N_{\Delta x}+1)^d}$ defined by

$$(T(\mu))_{k,i} = \mathsf{M}^{\Delta}[\tilde{\mu}](t_k,x_i) \quad \text{for all } k\in\mathcal{I}^*_{\Delta t},\ i\in\mathcal{I}_{\Delta x}.$$

It follows from Proposition 2.4.3(i),(ii),(iv) that $T(\mathfrak{D}^{\Delta t,\Delta x}) \subseteq \mathfrak{D}^{\Delta t,\Delta x}$. Moreover, if $(\mu_n)_{n\in\mathbb{N}} \subset \mathfrak{D}^{\Delta t,\Delta x}$ converges to μ then the continuity of L, F, and G, imply that, as $n\to\infty$, $v_{k,i}^{\Delta t,\Delta x}[\mu_n] \to v_{k,i}^{\Delta t,\Delta x}[\mu]$ for all $k\in\mathcal{I}_{\Delta t}$ and $i\in\mathbb{Z}^d$. Thus, $(v^{\Delta t,\Delta x}[\mu_n])_{n\in\mathbb{N}}$, defined in (2.3.9), converges to $v^{\Delta t,\Delta x}[\mu]$ pointwisely and hence, by Lebesgue's dominated convergence, the sequence $(v^{\Delta}[\mu_n])_{n\in\mathbb{N}}$, defined in (2.3.17), satisfies that $v^{\Delta}[\mu_n] \to v^{\Delta}[\mu]$ and $D_x v^{\Delta}[\mu_n] \to D_x v^{\Delta}[\mu]$ pointwisely. Consequently, given $k\in\mathcal{I}^*_{\Delta t}$, it follows from (2.4.14) that $\Phi_k^{\Delta}[\mu_n] \to \Phi_k^{\Delta}[\mu]$ pointwisely. In particular, $\beta_i^0(\Phi_k^{\Delta}[\mu_n](x)) \to \beta_i^0(\Phi_k^{\Delta}[\mu](x))$ for all $x\in\mathbb{R}^d \setminus \big(\Phi_k^{\Delta}[\mu]^{-1}(\partial E_i)\big)$. If $R>0$ is as in the proof of Proposition 2.4.5 and $\Delta t/\varepsilon$ is small enough, we have that $\Phi_k^{\Delta}[\mu]$ is a diffeomorphism of $B_\infty(0,R)$ onto $\Phi_k^{\Delta}[\mu](B_\infty(0,R))$. Therefore, since $\mathcal{L}^d(\partial E_i)=0$, we have $\mathcal{L}^d\big(\Phi_k^{\Delta}[\mu]^{-1}(\partial E_i)\big)=0$ and hence $\beta_i^0(\Phi_k^{\Delta}[\mu_n](x)) \to \beta_i^0(\Phi_k^{\Delta}[\mu](x))$ for almost every $x\in\mathbb{R}^d$. Therefore, by Lebesgue's dominated convergence,

$$\int_{E_j} \beta_i^0\big(\Phi_k^{\Delta}[\mu_n](x)\big)\mathrm{d}x \underset{n\to\infty}{\longrightarrow} \int_{E_j} \beta_i^0\big(\Phi_k^{\Delta}[\mu](x)\big)\mathrm{d}x \quad \text{for all } k\in\mathcal{I}^*_{\Delta t},\ i,\ j\in\mathbb{Z}^d.$$

Altogether, it follows from (2.4.21) that, as $n\to\infty$, $T(\mu_n)\to T(\mu)$, i.e. T is continuous. Finally, the existence of a solution to (MFG^{Δ}), i.e. of a fixed point of T, follows from

Brouwer's fixed-point theorem. □

In the next result we provide our main result, which shows the convergence, up to some subsequence, of solutions to (MFG^{Δ}) towards a solution to (MFG).

Theorem 2.5.2 *Assume* **(H1)-(H3)***, let* $\big((\Delta t_n, \Delta x_n, \varepsilon_n)\big)_{n\in\mathbb{N}} \subset]0,\infty[^3$*, and set* $\Delta_n = (\Delta t_n, \Delta x_n, \varepsilon_n)$*. Suppose that, as* $n \to \infty$*,* $\Delta_n \to 0$*,* $\Delta x_n = o(\Delta t_n)$*, and* $\Delta t_n = O(\varepsilon_n^2)$*. For every* n*, large enough, let* $m^n \in \mathcal{S}^{\Delta_n}$ *be a solution to* $(\mathrm{MFG}^{\Delta_n})$*, define* $\tilde{m}^n$ *by* (2.5.2)*, and set* $v^n = v^{\Delta_n}[\tilde{m}^n]$*. Then there exists a solution* (v^*, m^*) *to* (MFG) *such that, up to some subsequence, the following hold:*

(i) $(v^n)_{n\in\mathbb{N}}$ *converges to* v^**, uniformly over compact subsets of* $[0,T]\times\mathbb{R}^d$*.*

(ii) $(\tilde{m}^n)_{n\in\mathbb{N}}$ *converges in* $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ *towards* m^**. Moreover, the convergence also hold weakly in* $L^{\mathsf{p}}([0,T]\times\mathbb{R}^d)$*, if* $\mathsf{p} < \infty$*, and weakly* in* $L^{\infty}([0,T]\times\mathbb{R}^d)$*, if* $\mathsf{p} = \infty$*. In addition, there exists* $\widetilde{C} > 0$ *such that*

$$\|m^*(t,\cdot)\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \le \widetilde{C}\|m_0^*\|_{L^{\mathsf{p}}(\mathbb{R}^d)} \quad \textit{for all } t \in [0,T]. \tag{2.5.3}$$

Proof. For all $n \in \mathbb{N}$, large enough, we have $\tilde{m}^n = \mathsf{M}^{\Delta_n}[\tilde{m}^n]$. Arguing as in the proof of Proposition 2.4.6(ii), we obtain the existence of $m^* \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and a subsequence, still labelled by n, such that $(\tilde{m}^n)_{n\in\mathbb{N}}$ converges to m^* in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. It follows from Proposition 2.4.6(ii) that m^* solves (2.2.31), with $\mu = m^*$, i.e. $(v[m^*], m^*)$ solves (MFG). Therefore, assertions(i)-(ii) follow from the corresponding assertions in Proposition 2.4.6. □

Remark 2.5.1 (i) *Theorem 2.5.2 shows, in particular, that system* (MFG) *admits at least one solution* (v^*, m^*)*. If the solution to* (MFG) *is unique, then the entire sequence* (v^n, m^n) *converges to* (v^*, m^*) (*see Theorem 2.2.4 for a sufficient condition ensuring uniqueness*)*.*

(ii) *The condition on* $\Delta x_n = o(\Delta t_n)$ *in Theorem 2.5.2 is stronger than the condition* $(\Delta x_n)^2 = o(\Delta t_n)$ *needed for convergence, when the space dimension is equal to one, in the scheme studied in [57]* (*see also [66]*)*. This can be explained by the estimate* (2.4.31) *in Lemma 2.4.2, which seems difficult to improve, even if* φ *is smooth, and it is in compliance with Assumption 3.1 in [133], which plays an important role in the LG approximation of continuity equations with Lipschitz vector fields.*

(iii) *Let us point out that our method of proof, based on compactness arguments, does not provide rates of convergence for the solutions to the scheme. The establishment of convergence rates for the approximation of solutions to first order*

MFGs *remains as an interesting challenge.*

2.6 Numerical results

In this section, given $\Delta = (\Delta t, \Delta x, \varepsilon)$, we use ($\mathrm{MFG}^\Delta$) to approximate the solutions to two first order MFGs systems. In order to obtain an implementable version, we need to approximate the integrals in the LG scheme (2.4.21)-(2.4.22). We consider two methods. In the first one, the integrals are approximated by numerical quadrature, while, in the second one, we use the so-called *area weighting* technique, introduced in [124] and recalled in Section 2.6.1 below.

In the first example, the state dimension is equal to one and the data of the MFGs system does not satisfy some of the assumptions in Section 2.2.2. On the other hand, the PDE system admits an explicit solution, which allows to compare the quadrature and area weighting methods to solve (MFG^Δ). For comparable accuracies, the area weighting method is less expensive than the quadrature method and, hence, we use the former in order to treat the second example, where the state dimension is equal to two and no explicit solution is known. Let us point out that the data of the second example fulfills all the assumptions in Section 2.2.2.

We solve (MFG^Δ) heuristically by fixed point iterations that are stopped as soon as the uniform norm of the difference between two consecutive iterates is smaller than a given threshold τ, which in the simulations is set to 10^{-3}. In particular, we use the classical Picard iterations in the first test, as in [57], and Picard iterations with damping parameter 0.5 in the second test, as in [66].

All proposed tests are implemented in MATLAB R2023a and run in a computer with MacOS 10.15.7, 2.8 GHz Intel Core i7 Dual Core 16GB RAM.

2.6.1 Area-weighted LG approximation

Let $\mu \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and consider the continuity equation (2.4.1). The main idea of the area-weighting technique is to replace, for each $k \in \mathcal{I}^*_{\Delta x}$, the local nonlinear discrete flow $E_i \ni x \mapsto \Phi^\Delta_k[\mu](x) \in \mathbb{R}^d$, defined by (2.4.14), by the local affine approximation

$$E_i \ni x \mapsto \overline{\Phi}^\Delta_k[\mu](x) = x - \Delta t D_p H\big(x_i, D_x v^\Delta[\mu](t_k, x_i)\big) \in \mathbb{R}^d. \tag{2.6.1}$$

Notice that $\overline{\Phi}^\Delta_k[\mu](x) = x - x_i + \Phi^\Delta_k[\mu](x_i)$ for all $x \in E_i$. Under this approximation, we can compute the integrals in (2.4.21)-(2.4.22) explicitly. Indeed, for all $i = (i_1, \dots, i_d) \in \mathbb{Z}^d$

and $l = 1, \dots, d$, let us set $\mathrm{T}_{i_l} = [(x_i)_l - \Delta x/2, (x_i)_l + \Delta x/2]$, and observe that

$$\beta_i^0(y) = \prod_{l=1}^{d} \mathbb{I}_{\mathrm{T}_{i_l}}(y_l) \quad \text{for all } y = (y_1, \dots, y_d) \in \mathbb{R}^d. \tag{2.6.2}$$

It follows from (2.6.1) and (2.6.2) that, for every $i,\, j \in \mathbb{Z}^d$, we have

$$\begin{aligned}
&\int_{E_j} \beta_i^0(\overline{\Phi}_k^{\Delta}[\mu](y))\mathrm{d}y = \int_{E_j} \beta_i^0(y - x_j + \Phi_k^{\Delta}[\mu](x_j))\mathrm{d}y \\
&= \prod_{l=1}^{d} \int_{(x_j)_l - \Delta x/2}^{(x_j)_l + \Delta x/2} \mathbb{I}_{\mathrm{T}_{i_l}}\left(y_l - (x_j)_l + \left(\Phi_k^{\Delta}[\mu](x_j)\right)_l\right) \mathrm{d}y_l = \prod_{l=1}^{d} \int_{\left(\Phi_k^{\Delta}[\mu](x_j)\right)_l - \Delta x/2}^{\left(\Phi_k^{\Delta}[\mu](x_j)\right)_l + \Delta x/2} \mathbb{I}_{\mathrm{T}_{i_l}}(y_l)\mathrm{d}y_l \qquad (2.6.3) \\
&= \prod_{l=1}^{d} \mathcal{L}^1\Big([(x_i)_l - \Delta x/2, (x_i)_l + \Delta x/2] \cap \left[(\Phi_k^{\Delta}[\mu](x_j))_l - \Delta x/2, (\Phi_k^{\Delta}[\mu](x_j))_l + \Delta x/2\right] \Big).
\end{aligned}$$

On the other hand, for every $l = 1, \dots, d$, it follows from (2.3.2) that

$$\begin{aligned}
&\mathcal{L}^1\Big([(x_i)_l - \Delta x/2, x_i)_l + \Delta x/2] \cap \left[(\Phi_k^{\Delta}[\mu](x_j))_l - \Delta x/2, (\Phi_k^{\Delta}[\mu](x_j))_l + \Delta x/2\right] \Big) \\
&= \begin{cases} \Delta x + \left(\Phi_k^{\Delta}[\mu](x_j)\right)_l - (x_i)_l & \text{if } \left(\Phi_k^{\Delta}[\mu](x_j)\right)_l \in [(x_i)_l - \Delta x, (x_i)_l], \\ \Delta x + (x_i)_l - \left(\Phi_k^{\Delta}[\mu](x_j)\right)_l & \text{if } \left(\Phi_k^{\Delta}[\mu](x_j)\right)_l \in](x_i)_l, (x_i)_l + \Delta x], \\ 0 & \text{otherwise,} \end{cases} \\
&= \Delta x \widehat{\beta}\left(\left(\Phi_k^{\Delta}[\mu](x_j)\right)_l / \Delta x - i_l\right),
\end{aligned}$$

which, combined with (2.3.1) and (2.6.3), yields

$$\frac{1}{(\Delta x)^d} \int_{E_j} \beta_i^0(\overline{\Phi}_k^{\Delta}[\mu](y))\mathrm{d}y = \beta_i^1(\Phi_k^{\Delta}[\mu](x_j)). \tag{2.6.4}$$

Thus, replacing $\Phi_k^{\Delta}[\mu]$ by $\overline{\Phi}_k^{\Delta}[\mu]$ in (2.4.21) and, for every $i \in \mathbb{Z}^d$, denoting by $m_{0,i}^*$ any approximation of $\int_{E_i} m_0^*(x)\mathrm{d}x/(\Delta x)^d$, we obtain the following area-weigthed LG version of (2.4.21)-(2.4.22):

$$\overline{m}_{k+1,i} = \sum_{j \in \mathbb{Z}^d} \overline{m}_{k,j} \beta_i^1(\Phi_k^{\Delta}[\mu](x_j)) \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i \in \mathbb{Z}^d, \tag{2.6.5}$$

$$\overline{m}_{0,i} = m_{0,i}^* \quad \text{for all } i \in \mathbb{Z}^d. \tag{2.6.6}$$

Remark 2.6.1 (i) *Notice that* (2.6.5)- (2.6.6) *corresponds to the scheme proposed in [57] for the continuity equation* (2.4.1)*. Therefore, the latter can be seen as an area-weighted version of the LG scheme of Section 2.4.*

(ii) *Let us point out that we have not been able to provide a satisfactory uniform error estimate for the difference between the solution* $\{m_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\, i \in \mathbb{Z}^d\}$ *to* (2.4.21)-(2.4.22) *and the solution* $\{\overline{m}_{k,i} \,|\, k \in \mathcal{I}_{\Delta t},\, i \in \mathbb{Z}^d\}$ *to* (2.4.21)-(2.4.22)*. The main issue is that, as suggested by the analysis in [124, Section 3], this estimate should depend of the Lipschitz constant of the map* $x \mapsto -D_p H\big(x, D_x v^{\Delta}[\mu](t,x)\big)$*, which, in our case, is not uniform with respect to* ε*.*

2.6.2 Non-local MFG with analytical solution.

We consider system (MFG) with a quadratic Hamiltonian $H(x,p) = \frac{p^2}{2}$, coupling terms

$$F(x,\nu) = \frac{1}{2}\Big(x - \int_{\mathbb{R}^d} y \,\mathrm{d}\nu(y)\Big)^2, \quad G(x,\nu) = 0, \tag{2.6.7}$$

and initial data m_0^* given by the distribution of a d-dimensional Gaussian random variable with mean $\mu^* \in \mathbb{R}^d$ and covariance matrix $\Sigma_0 \in \mathbb{R}^{d\times d}$ assumed, for simplicity, to be diagonal. Notice that the coupling term F in (2.6.7) and the initial distribution m_0^* do not satisfy assumptions **(H2)** and **(H3)**, respectively. On the other hand, the MFG system admits in this case an explicit solution, which allows to compare the performance of quadrature and area-weighting methods to approximate the continuity equation. Indeed, setting

$$\Pi(t) = \left(\frac{e^{2T-t} - e^t}{e^{2T-t} + e^t}\right) I_d, \quad s(t) = -\Pi(t)\mu^*, \quad c(t) = \frac{1}{2}\langle \Pi(t)\mu^*, \mu^*\rangle \quad \text{for all } t \in [0,T],$$

and arguing as in [41, Section 5.2], one finds that (MFG) admits a unique solution (v^*, m^*) given by

$$v^*(t,x) = \frac{1}{2}\langle \Pi(t)x, x\rangle + \langle s(t), x\rangle + c(t) \quad \text{for all } (t,x) \in [0,T]\times\mathbb{R}^d$$

and, for every $t \in [0,T]$, $m^*(t)$ is the joint distribution of d independent Gaussian random variables $\{X_l(t) \,|\, l = 1,\dots,d\}$ with means $\mu_\ell(t)$ and variances $\sigma_\ell^2(t)$ ($l = 1,\dots,d$), given by

$$\mu_\ell(t) = \mu_l^* \quad \text{and} \quad \sigma_\ell^2(t) = \left(\frac{e^{2T-t} + e^t}{e^{2T}+1}\right)^2 (\Sigma_0)_{l,l}.$$

In the numerical test, we take $T = 0.25$, $d = 1$, $\mu^* = 0.1$, and $\Sigma_0 = 0.105$. Since the exact solution $m^*(t)$ does not have a compact support, we approximate the system on the bounded domain $\mathcal{O} =]-2, 2[$ and we impose Dirichlet boundary conditions at $x = -2$ and $x = 2$, which are equal to the values of the exact solution at these points. In order to implement the latter, we proceed as in [76, Section 5.1.5].
We test our scheme for different values of Δx, the time step is chosen as $\Delta t = (\Delta x)^{2/3}/2$, and the mollifier in (2.3.17) is defined with $\mathbb{R} \ni x \mapsto \rho(x) = e^{-x^2/2}/\sqrt{2\pi} \in \mathbb{R}$ and $\varepsilon = \sqrt{\Delta t}$. We denote by $(v^\Delta, \mathsf{M}^\Delta)$ and $(\overline{v}^\Delta, \overline{\mathsf{M}}^\Delta)$ the approximations of solutions to (MFG^Δ) obtained by estimating the integrals in (2.4.21) by numerical quadrature and by the area-weighting method, respectively. For the numerical quadrature of the integrals in the computation of $(v^\Delta, \mathsf{M}^\Delta)$, we divide each interval E_i into $\lfloor 4/\Delta x \rfloor$ subintervals and we use the midpoint rule on each one of them. The initial condition m_0^* being smooth, we use the midpoint rule to approximate the integrals in (2.4.22).
Setting $\mathcal{G}_{\Delta x}(\overline{\mathcal{O}}) := \mathcal{G}_{\Delta x} \cap \overline{\mathcal{O}}$, Tables 2.1 and 2.2 below show the uniform and L^2 relative discrete errors

$$E_\infty(h^\Delta) = \frac{\max_{x_i \in \mathcal{G}_{\Delta x}(\Omega)} |h^\Delta(x_i) - h(x_i)|}{\max_{x_i \in \mathcal{G}_{\Delta x}(\Omega)} |h(x_i)|}, \quad E_2(h^\Delta) = \left(\frac{\sum_{x_i \in \mathcal{G}_{\Delta x}(\Omega)} |h^\Delta(x_i) - h(x_i)|^2}{\sum_{x_i \in \mathcal{G}_{\Delta x}(\Omega)} |h(x_i)|^2} \right)^{\frac{1}{2}}, \tag{2.6.8}$$

for $(h, h^\Delta) = (m^*(T, \cdot), \mathsf{M}^\Delta(T, \cdot))$, $(m^*(T, \cdot), \overline{\mathsf{M}}^\Delta(T, \cdot))$, $(v^*(0, \cdot), v^\Delta(0, \cdot))$, and $(v^*(0, \cdot), \overline{v}^\Delta(0, \cdot))$.
Table 2.1 shows smaller errors for the approximation of m^* computed with numerical quadrature, specially in the uniform norm. Table 2.2 shows that the higher precision obtained by computing an approximation of m^* by numerical quadrature does not significantly affect the approximation of the value function v^*. In Table 2.3, we display the CPU time and number of iterations for both implementations. In particular, it shows that the improvement of precision of the numerical quadrature method is achieved at the expense of a high computational cost compared with the area-weighted approximation. Table 2.3 also shows that, in most of the cases and for both implementations, doubling the space mesh refinement more than triples the calculation time, which indicates that other methods should be used for high dimensional problems.

Δx	$E_\infty(\mathsf{M}^\Delta(T,\cdot))$	$E_\infty(\overline{\mathsf{M}}^\Delta(T,\cdot))$	$E_2(\mathsf{M}^\Delta(T,\cdot))$	$E_2(\overline{\mathsf{M}}^\Delta(T,\cdot))$
4.80 $\cdot 10^{-2}$	8.41 $\cdot 10^{-3}$	3.69 $\cdot 10^{-2}$	6.30 $\cdot 10^{-3}$	1.09 $\cdot 10^{-2}$
2.40 $\cdot 10^{-2}$	6.91 $\cdot 10^{-3}$	3.25 $\cdot 10^{-2}$	4.39 $\cdot 10^{-3}$	1.05 $\cdot 10^{-2}$
1.20 $\cdot 10^{-2}$	3.94 $\cdot 10^{-3}$	2.62$\cdot 10^{-2}$	2.77 $\cdot 10^{-3}$	6.77 $\cdot 10^{-3}$
6.00 $\cdot 10^{-3}$	1.83 $\cdot 10^{-3}$	2.44 $\cdot 10^{-2}$	6.89 $\cdot 10^{-4}$	2.67 $\cdot 10^{-3}$

Table 2.1: Errors for the approximation of $m^*(T,\cdot)$.

Δx	$E_\infty(v^\Delta(0,\cdot))$	$E_\infty(\overline{v}^\Delta(0,\cdot))$	$E_2(v^\Delta(0,\cdot))$	$E_2(\overline{v}^\Delta(0,\cdot))$
4.80 $\cdot 10^{-2}$	7.02 $\cdot 10^{-3}$	7.11 $\cdot 10^{-3}$	6.20 $\cdot 10^{-3}$	6.31 $\cdot 10^{-3}$
2.40 $\cdot 10^{-2}$	5.74 $\cdot 10^{-3}$	5.82 $\cdot 10^{-3}$	4.90 $\cdot 10^{-3}$	5.12 $\cdot 10^{-3}$
1.20 $\cdot 10^{-2}$	4.34 $\cdot 10^{-3}$	4.37 $\cdot 10^{-3}$	3.70 $\cdot 10^{-3}$	3.75 $\cdot 10^{-3}$
6.00 $\cdot 10^{-3}$	3.30 $\cdot 10^{-3}$	3.36 $\cdot 10^{-3}$	2.95 $\cdot 10^{-3}$	3.01 $\cdot 10^{-3}$

Table 2.2: Errors for the approximation of $v^*(0,\cdot)$.

2.6.3 A two-dimensional example

In this test, we consider system (MFG) with $d=2$, a quadratic Hamiltonian $H(x,p)=|p|^2/2$, and coupling terms having the form

$$F(x,\nu)=\gamma\min\{|x-\bar{x}|^2,R\}+(r_\sigma*\nu)(x)\quad\text{and}\quad G(x,\nu)=0, \tag{2.6.9}$$

where $\gamma>0$, $\bar{x}\in\mathbb{R}^2$, $R>0$, and, for $\sigma>0$, $r_\sigma(x)=e^{-|x|^2/2\sigma^2}/(2\pi\sigma^2)$ for all $x\in\mathbb{R}^2$. Given $\ell>0$, $x_0^*\in]0,\ell[^2$, and $\sigma_0>0$, we consider the initial density

$$m_0^*(x)=\frac{\chi(x)}{\int_{[0,\ell]^2}\chi(y)\mathrm{d}y}\quad\text{with}\quad \chi(x)=e^{-|x-x_0^*|^2/2\sigma_0^2}\mathbb{I}_{[0,\ell]^2}(x)\quad\text{for all } x\in\mathbb{R}^d. \tag{2.6.10}$$

Notice that the data above satisfy **(H1), (H2)**, and **(H3)**, with $p=\infty$. In our tests below, we choose $T=1$, $\ell=2$, $x_0^*=(0.75,0.75)$, $\sigma_0=0.07$ in the initial condition, $\bar{x}=(1.75,1.75)$, $R=5$, $\sigma=0.25$, and two values $\gamma=0.5$ and $\gamma=3$ in the running cost F. Since in this two-dimensional example the computational cost to solve (MFG^Δ) is important, in view of the discussion in Section 2.6.2 we implement the area-weighting method of Section 2.6.1 to approximate the integrals in (2.4.21). The integrals in (2.4.22), to approximate the initial condition m_0^*, are computed by using the midpoint rule. We set $\Delta x=0.025$, $\Delta t=\Delta x^{2/3}$, and the mollifier in (2.3.17) is defined with $\mathbb{R}^2\ni x\mapsto\rho(x)=e^{-|x|^2/2}/2\pi\in\mathbb{R}$ and $\varepsilon=\sqrt{\Delta t}/2$. Figure 1 shows the approximation $\overline{\mathsf{m}}^\Delta$ of the exact distribution m^* in the x_1-x_2 plane obtained after solving (MFG^Δ) for $\gamma=0.5$ and $\gamma=3$. On the left, we display the evolution of the initial distribution, concentrated around x_0^*, by overlaying

Δx	CPU_{quad}	CPU_{aw}	It_{quad}	It_{aw}
4.80 $\cdot 10^{-2}$	15.21s	0.93s	5	5
2.40 $\cdot 10^{-2}$	55.35s	2.53s	5	5
1.20 $\cdot 10^{-2}$	200.98s	8.42s	6	5
6.00 $\cdot 10^{-3}$	710.321s	27.71s	6	5

Table 2.3: CPU times (in seconds) and number of iterations to compute $(v^{\Delta}, \mathsf{M}^{\Delta})$ and $(\overline{v}^{\Delta}, \overline{\mathsf{M}}^{\Delta})$. The CPU time and the number of iterations are respectively denoted by CPU_{quad} and It_{quad}, for the numerical quadrature, and by CPU_{aw} and It_{aw}, for the area-weighted approximation.

the distributions $\overline{\mathsf{m}}^{\Delta}(t_k,\cdot)$ for $k \in \mathcal{I}_{\Delta t}$. On the right, we display only the final distribution $\overline{\mathsf{m}}^{\Delta}(T,\cdot)$. The simulation shows the effect of the positive constant γ, which weights the importance of reaching the target point $\bar{x}$. If $\gamma = 0.5$, the aversion to crowed regions, modeled by the second term in the definition of F, has a more relevant impact on the distribution of the players than the term penalizing the distance to $\bar{x}$, while, if $\gamma = 3$, the latter term has a preponderant role in the evolution of the distribution of the agents.

A1. Appendix

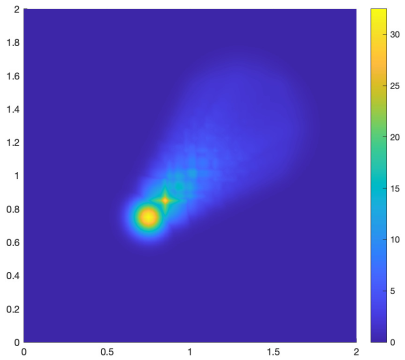
Proof. [Proof of Proposition 3.3.1] Let us fix $(t,x) \in [0,T[\times\mathbb{R}^d$. The existence of $\alpha^{t,x} \in L^2\big([t,T];\mathbb{R}^d\big)$, such that $v[\mu](t,x) = J^{t,x}[\mu](\alpha^{t,x})$, follows from (2.2.16), the continuity assumption on F and G in **(H2)**, and the direct method in the calculus of variations. Setting $\alpha_0(s) = 0$ for all $s \in [t,T]$, the inequalities $J^{t,x}[\mu](\alpha^{t,x}) \leq J^{t,x}[\mu](\alpha_0(s))$, (2.2.6), (2.2.10), (2.2.11), and (2.2.16), imply that

$$\int_t^T |\alpha^{t,x}(s)|^2 \mathrm{d}s \leq \tilde{C} := \frac{T(C_{L,2} + 2C_{F,1} + C_{L,7}) + 2C_{G,1}}{C_{L,6}}. \tag{2.6.11}$$

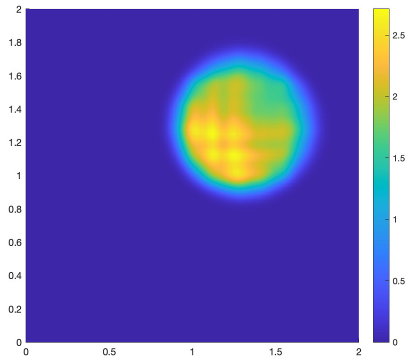
In particular, setting $\mathcal{A}^t = \Big\{\alpha \in L^2\big([t,T];\mathbb{R}^d\big),\ \int_t^T |\alpha(s)|^2 \mathrm{d}s \leq \tilde{C}\Big\}$, we have

$$v[\mu](t,x) = \inf\Big\{J^{t,x}[\mu](\alpha) \,\Big|\, \alpha \in \mathcal{A}^t\Big\}. \tag{2.6.12}$$

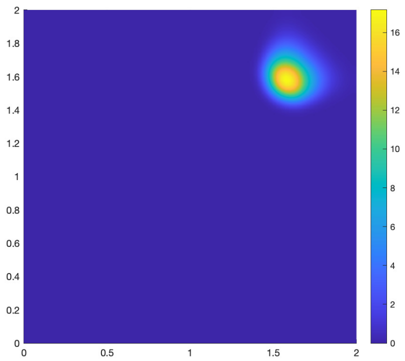
Thus, assertion (2.2.1) follows from (2.2.6), (2.2.16), (2.6.11), (2.2.10), and (2.2.11). Moreover, it follows from conditions (2.2.7), (2.2.12), (2.2.13), and expression (2.6.12)



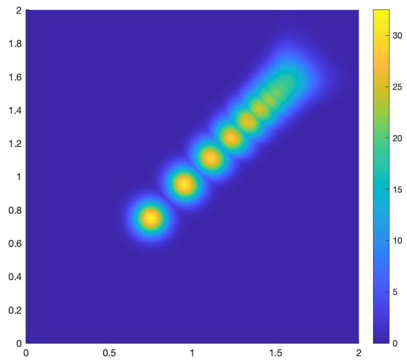
(a) Time evolution $(\overline{\mathrm{m}}^{\Delta}(t_k,\cdot))_{k\in\mathcal{I}_{\Delta t}}$ for $\gamma=0.5$.



(b) Final distribution $\overline{\mathrm{m}}^{\Delta}(T,\cdot)$ for $\gamma=0.5$.



(c) Time evolution $(\overline{\mathrm{m}}^{\Delta}(t_k,\cdot))_{k\in\mathcal{I}_{\Delta t}}$ for $\gamma=3$.



(d) Final distribution $\overline{\mathrm{m}}^{\Delta}(T,\cdot)$ for $\gamma=3$.

Figure 2.1: Approximation of m^* in both cases $\gamma=0.5$ and $\gamma=3$.

that, for every $y\in\mathbb{R}^d$, we have

$$\begin{aligned}
|v[\mu](t,x)-v[\mu](t,y)| &\leq \sup_{\alpha\in\mathcal{A}^t}\left|J^{t,x}[\mu](\alpha)-J^{t,y}[\mu](\alpha)\right| \\
&\leq \sup_{\alpha\in\mathcal{A}^t}\left\{\int_t^T \Big(C_{L,3}(1+|\alpha(s)|^2)+C_{F,2}\Big)|X^{t,x,\alpha}(s)-X^{t,y,\alpha}(s)|\mathrm{d}s\right. \\
&\qquad\qquad \left. + C_{G,2}|X^{t,x,\alpha}(T)-X^{t,y,\alpha}(T)|\right\} \\
&\leq \Big(T(C_{L,3}+C_{F,2})+C_{L,3}\tilde{C}+C_{G,2}\Big)|x-y|,
\end{aligned}$$

which shows (ii). Let us set $\overline{X} = X^{t,x,\alpha^{t,x}}$ and let $s \in [t,T[$. Since $v[\mu]$ satisfies the dynamic programming inequality

$$v[\mu](s,\overline{X}(s)) \leq \int_s^{s+h} \Big(L(X^{s,\overline{X}(s),\alpha}(r),\alpha(r)) + F(X^{s,\overline{X}(s),\alpha}(r),\mu(r)) \Big) \mathrm{d}r \\ + v[\mu]\big(s+h, X^{s,\overline{X}(s),\alpha}(s+h)\big),$$

for all $h \in [0,T-s[$ and $\alpha \in L^2([t,T];\mathbb{R}^d)$, by taking $\alpha = \alpha_0$, the equality

$$v[\mu](s,\overline{X}(s)) = J^{s,\overline{X}(s)}[\mu](\alpha^{t,x}|_{[s,T]}), \tag{2.6.13}$$

the estimates (2.2.16), (2.2.6), (2.2.10), the equality $X^{s,\overline{X}(s),\alpha_0}(s+h) = \overline{X}(s)$, and (2.2.26), imply that

$$\begin{aligned} C_{L,6} \int_s^{s+h} |\alpha^{t,x}(r)|^2 \mathrm{d}r &\leq h(C_{F,1} + C_{L,2}) + h(C_{L,7} + C_{F,1}) \\ &\quad + v[\mu](s+h,\overline{X}(s)) - v[\mu](s+h,\overline{X}(s+h)) \\ &\leq h(2C_{F,1} + C_{L,2} + C_{L,7}) + C_{\mathrm{Lip}} \int_s^{s+h} |\alpha^{t,x}(r)| \mathrm{d}r. \end{aligned}$$

By Young's inequality, we get the existence of $C > 0$, independent of (μ,t,x), such that

$$\int_s^{s+h} |\alpha^{t,x}(r)|^2 \mathrm{d}r \leq Ch$$

and, hence, by the Lebesgue differentiation theorem (see e.g. [32]), we have $\alpha^{t,x} \in L^\infty([0,T];\mathbb{R}^d)$ and $\|\alpha^{t,x}\|_{L^\infty([0,T];\mathbb{R}^d)} \leq \sqrt{C}$, which shows (i).
Finally, in order to show (2.2.1), notice that, for every $y \in \mathbb{R}^d$, (2.2.9) implies that

$$L(x+y,a) - 2L(x,a) + L(x-y,a) \leq C_{L,5}(1+|a|^2)|y|^2 \quad \text{for all } a \in \mathbb{R}^d. \tag{2.6.14}$$

Estimates (2.6.14), (2.2.14), (2.2.15), and (2.6.11), imply

$$\begin{aligned} v(t,x+y)+v(t,x-y) &\leq \int_t^T \Big(L(X^{t,x+y,\alpha^{t,x}}(s),\alpha^{t,x}(s)) + L(X^{t,x-y,\alpha^{t,x}}(s),\alpha^{t,x}(s)) \\ &\quad + F(X^{t,x+y,\alpha^{t,x}}(s),\mu(s)) + F(X^{t,x-y,\alpha^{t,x}}(s),\mu(s))\Big)\mathrm{d}s \\ &\quad + G(X^{t,x+y,\alpha^{t,x}}(T),\mu(T)) + G(X^{t,x-y,\alpha^{t,x}}(T),\mu(T)) \\ &\leq 2\int_t^T \Big(L(X^{t,x,\alpha^{t,x}}(s),\alpha^{t,x}(s)) + F(X^{t,x,\alpha^{t,x}}(s),\mu(s))\Big)\mathrm{d}s \\ &\quad + 2G(X^{t,x,\alpha^{t,x}}(T),\mu(T)) \\ &\quad + \int_t^T \Big(C_{L,5}(1+|\alpha^{t,x}(s)|^2) + C_{F,3}\Big)|y|^2\mathrm{d}s + C_{G,3}|y|^2 \\ &\leq 2v[\mu](t,x) + \Big(T(C_{L,5}+C_{F,3}) + C_{L,5}\tilde{C} + C_{G,3}\Big)|y|^2, \end{aligned} \tag{2.6.15}$$

from which (2.2.27) follows. □

3

Approximation of deterministic mean field games under polynomial growth conditions on the data

We consider a deterministic mean field games problem in which a typical agent solves an optimal control problem where the dynamics is affine with respect to the control and the cost functional has a growth which is polynomial with respect to the state variable. In this framework, we construct a mean field game problem in discrete time and finite state space that approximates equilibria of the original game. Two numerical examples, solved with the fictitious play method, are presented.
This chapter is a joint work with Justina Gianatti and Francisco J. Silva [81].

Sommaire

3.1 Introduction

The theory of Mean Field Game (MFG for short) problems has been introduced in [111]–[113] and, independently, in [99] in order to describe the asymptotic behaviour of Nash equilibria of non-cooperative symmetric differential games with a large number of indistinguishable players, which, individually, have a minor influence on the game. The reader is referred to [5], [60], [61], [86], [87], and the references therein, for an overview on MFG theory including their numerical approximation and applications in crowd motion, economics, and finance. Equilibria in MFGs are usually described in terms of a system of two equations, called MFG *system*; a Hamilton-Jacobi-Bellman equation, describing the optimal cost of a *typical player*, and a Fokker-Planck equation, describing the evolution of the players.

This work deals with the numerical approximation of deterministic mean field game problems, i.e. when the underlying differential game is deterministic. In this setting, a relaxed, also called *Lagrangian*, formulation of equilibria involving a fixed point problem on the space of probability measures over the trajectories of the players, has been introduced in [28], [45], [52]. We assume that the controlled dynamics of a typical player in the MFG has the form

$$\dot{\gamma}(t) = A(t,\gamma(t)) + B(t,\gamma(t))\alpha(t) \quad \text{for a.e. } t \in (0,T). \tag{3.1.1}$$

Here, $T > 0$ denotes the time horizon, γ and α, which take values in $\mathbb{R}^d$ and in $\mathbb{R}^r$, respectively, denote the state and the control of a typical player, and $A\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $B\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times r}$ are given functions. When the typical player controls its velocity, i.e. $\dot{\gamma}(t) = \alpha(t)$ and MFG equilibria are characterized in terms of the MFG system, the reference [42] proposes a semi-discrete scheme which is shown to converge towards a solution to the MFG system. An implementable version of this approximation, including also a discretization of the space variable, has been introduced in [57] and it is shown to converge when the space dimension is equal to one. In the same framework, in [56] the authors propose a Lagrange-Galerkin scheme for the continuity equation appearing in the MFG system and they show the convergence of a fully-discrete approximation in general state dimensions. By adopting the relaxed formulation,

in [94] an approximation of the MFG problem written in terms of a discrete time and finite state MFG [84], hereafter called *finite* MFG, is shown to converge in general dimensions. Moreover, under a monotonicity assumption on the interaction cost terms (see [112, Section 2.3]), an adaptation of the fictitious play method (see [39], [130]) is shown to converge and hence allows to rigorously approximate a solution to the finite MFG. Finally, in the work [80] the authors approximate MFGs by finite MFGs when the dynamics of the typical player takes the general form (3.1.1). The convergence is established in general dimensions, the key point being a careful discretization of the underlying optimal control problem. In particular, the results in [80] cover the approximation of MFGs where the typical player controls its acceleration (see [8], [47]).
In all the references above, the assumptions on the dependence of the cost functional with respect to the state do not allow a polynomial growth. In particular, the cost cannot depend quadratically on the state, which is a typical case considered in the applications. Our aim in this work, which is complementary with [80], is to cover this case. The main difficulty coming from a polynomial growth of the cost is that the value function of a typical player is not globally Lipschitz with respect to the state. In particular, the optimal feedback law is only locally bounded and hence a careful analysis is needed in order to construct a scheme for the value function with good stability properties. Under our assumptions, which include the independence of $B(t,\gamma(t))$ on $\gamma(t)$, the optimal feedback controls have a linear growth with respect to the state. This property still allows us to construct an approximation of the MFG problem where the time marginals of the distributions of the states of the agents are supported on a compact set which is independent of the discretization steps. Next, the analysis in [80] applies and yields the convergence of the scheme as the discretization parameters vanish. As in [80], [94], we adopt in this work the relaxed formulation of the MFG equilibrium, the main point being that, in the convergence study of our approximations, compactness properties for the solutions to the scheme are easier to establish.
The remainder of this article is organized as follows. Section 3.2 fixes the notations and the assumptions in this work. It also recalls the notion of Lagrangian MFG equilibrium and provides an existence and uniqueness result. Section 3.3 is central in this work as it builds the scheme used to approximate the value function of a typical player. We explain the relation between this scheme and a standard semi-Lagrangian scheme (see [76]) and we provide a convergence result. In Section 3.4, we describe the finite MFG that approximates the continuous one and we present existence, uniqueness, and convergence results. Finally, in Section 3.5 we consider two numerical examples where the cost functional depends polynomially on the state variable and the interactions terms are monotone, which allows us to approximate the solutions to the finite MFG problems by using the fictitious play method.

3.2 Preliminaries

Let $n \in \mathbb{N}$. In what follows, $|\cdot|$ will denote the infinity norm in $\mathbb{R}^n$, and, given $R > 0$, $B_\infty(0,R)$ (respectively $\overline{B}_\infty(0,R)$) will denote the corresponding open (respectively closed) ball centered at 0 and of radius R. We denote by $\mathcal{P}(\mathbb{R}^n)$ the set of probability measures over $\mathbb{R}^n$ and, for $\mu \in \mathcal{P}(\mathbb{R}^n)$, we set $\operatorname{supp}(\mu)$ for its support. We define $\mathcal{P}_1(\mathbb{R}^n)$ as the subset of $\mathcal{P}(\mathbb{R}^n)$ consisting on probability measures with finite first order moment, i.e.

$$\mathcal{P}_1(\mathbb{R}^n) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^n) \,\Big|\, \int_{\mathbb{R}^n} |x| \mathrm{d}\mu(x) < \infty \right\}, \tag{3.2.1}$$

which is endowed with the Wasserstein distance

$$d_1(\mu_1,\mu_2) = \inf_{\mu \in \Pi(\mu_1,\mu_2)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y| \mathrm{d}\mu(x,y) \quad \text{for all } \mu_1,\ \mu_2 \in \mathcal{P}_1(\mathbb{R}^n), \tag{3.2.2}$$

where $\Pi(\mu_1,\mu_2)$ denotes the subset of $\mathcal{P}_1(\mathbb{R}^n \times \mathbb{R}^n)$ of probability measures with first and second marginals given by μ_1 and μ_2, respectively. Given $\nu \in \mathcal{P}(\mathbb{R}^n)$ and a Borel function $\Psi \colon \mathbb{R}^n \to \mathbb{R}^q$ ($q \in \mathbb{N}$), the *push-forward* measure $\Psi \sharp \nu$, defined on the σ-algebra of Borel sets $\mathbf{B}(\mathbb{R}^q)$, is defined by

$$\Psi \sharp \nu(A) = \nu(\Psi^{-1}(A)) \quad \text{for all } A \in \mathbf{B}(\mathbb{R}^q). \tag{3.2.3}$$

Let $T > 0$ and $d,\ r \in \mathbb{N}$. The mean field game problem that we will deal with in this article is defined in terms of $\ell \colon [0,T] \times \mathbb{R}^r \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$, $g \colon \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$, $A \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $B \colon [0,T] \to \mathbb{R}^{d \times r}$, and $m_0^* \in \mathcal{P}_1(\mathbb{R}^d)$. We will consider the following assumptions:

(H1) The functions ℓ and g are continuous and, for every $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$, the function $\ell(t,\cdot,x,\mu)$ is convex. Moreover, there exist $p \in (1,\infty)$, $C_{\ell,1}$, $C_{\ell,2}$, $C_{\ell,3}$, $C_{\ell,4} \in (0,\infty)$ and $C_{g,1}$, $C_{g,2}$, $C_{g,3} \in (0,\infty)$ such that, for every $(t,a,x,\mu) \in [0,T] \times \mathbb{R}^r \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$,

$$C_{\ell,1}|a|^p - C_{\ell,2} \leq \ell(t,a,x,\mu) \leq C_{\ell,3}(1+|a|^p+|x|^p), \tag{3.2.4}$$

$$-C_{g,1} \leq g(x,\mu) \leq C_{g,2}(1+|x|^p), \tag{3.2.5}$$

$$|\ell(t,a,x,\mu) - \ell(t,a,y,\mu)| \leq C_{\ell,4}\big(1+|x|^{p-1}+|y|^{p-1}+|a|^{p-1}\big)|x-y| \quad \text{for all } y \in \mathbb{R}^d, \tag{3.2.6}$$

$$|g(x,\mu) - g(y,\mu)| \leq C_{g,3}(1+|x|^{p-1}+|y|^{p-1})|x-y| \quad \text{for all } y \in \mathbb{R}^d. \tag{3.2.7}$$

(H2) The following hold:

(i) The functions A and B are continuous.

(ii) There exists $L_A \in (0,\infty)$ such that, for every $(t,x) \in [0,T] \times \mathbb{R}^d$,

$$|A(t,x) - A(t,y)| \leq L_A|x-y| \quad \text{for all } y \in \mathbb{R}^d.$$

(iii) We have $r \leq d$ and there exists $\{i_1,\dots,i_r\} \subset \{1,\dots,d\}$ such that, for all $t \in [0,T]$, the rows $i_1,\dots,i_r$ of $B(t)$ are linearly independent.

(H3) There exists $C^* \in (0,\infty)$ such that $\operatorname{supp}(m_0^*) \subset \overline{B}_\infty(0,C^*)$.

(H4) The following hold:

(i) The function ℓ satisfies **(H1)** and can be written as

$$\ell(t,a,x,\mu) = \ell_0(t,a,x) + f(t,x,\mu) \quad \text{for all } t \in [0,T],\ a \in \mathbb{R}^r,\ x \in \mathbb{R}^d,\ \mu \in \mathcal{P}_1(\mathbb{R}^d),$$

where $\ell_0\colon [0,T] \times \mathbb{R}^r \times \mathbb{R}^d \to \mathbb{R}$ and $f\colon [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$ is a continuous and bounded function. Moreover, there exists $L_f > 0$ such that, for all $(t,\mu) \in [0,T] \times \mathcal{P}_1(\mathbb{R}^d)$,

$$|f(t,x,\mu) - f(t,y,\mu)| \leq L_f|x-y| \quad \text{for all } x,y \in \mathbb{R}^d.$$

(ii) The functions $f(t,\cdot,\cdot)$, for all $t \in [0,T]$, and g are *monotone*, i.e. for $\Psi = f(t,\cdot,\cdot)$, g it holds that

$$\int_{\mathbb{R}^d} \big(\Phi(x,\mu_1) - \Phi(x,\mu_2)\big) \mathrm{d}(\mu_1 - \mu_2)(x) \geq 0 \quad \text{for all } \mu_1,\, \mu_2 \in \mathcal{P}_1(\mathbb{R}^d). \tag{3.2.8}$$

Assumptions **(H1)**, **(H2)**, and **(H3)** ensure that both the MFG problem defined in Problem 2 below, and its approximation, introduced in Section 3.4, admit at least one solution. Assumption **(H4)** plays an important role in the uniqueness of the equilibrium for both the MFG and its approximation and also in the proof of the convergence of a numerical method to solve the finite MFG problem (see [80]).
Note that **(H2)** implies that

$$|A(t,x)| \leq C_A(1+|x|) \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d, \tag{3.2.9}$$

where $C_A = \max\{\max_{t\in[0,T]} |A(t,0)|, L_A\}$. In what follows, setting $|B(t)|$ for the matrix norm of $B(t)$ induced by the infinity-norm in $\mathbb{R}^r$, we set $C_B = \sup_{t\in[0,T]} |B(t)|$, which is finite by **(H2)**(i).
Let us describe the MFG problem considered in this article. Given $x \in \mathbb{R}^d$ and $m \in C\left([0,T];\mathcal{P}_1(\mathbb{R}^d)\right)$, a typical player positioned at x at time $t=0$ solves an optimal

control problem of the form

$$\begin{cases} \inf \displaystyle\int_0^T \ell\left(s, \alpha(s), \gamma(s), m(s)\right) \mathrm{d}s + g(\gamma(T), m(T)) \\ \text{s.t.} \quad \dot{\gamma}(s) = A(s, \gamma(s)) + B(s)\alpha(s) \quad \text{for a.e. } s \in (0,T), \\ \qquad \gamma(0) = x, \\ \qquad \gamma \in W^{1,p}([0,T];\mathbb{R}^d),\ \alpha \in L^p([0,T];\mathbb{R}^r). \end{cases} \tag{$OC_{x,m}$}$$

Note that assumption **(H1)** states that the cost functional in $(OC_{x,m})$ has a polynomial growth with respect to the state and control variables. In particular, our conditions on the cost functional are more general than those in [80]. On the other hand, regarding the dynamics in $(OC_{x,m})$, in [80] the matrix B can also depend on the state variable. Let us endow $\Gamma := C\left([0,T];\mathbb{R}^d\right)$ with the supremum norm $\|\cdot\|_\infty$ and, for all $t \in [0,T]$, define $e_t \colon \Gamma \to \mathbb{R}^d$ by $e_t(\gamma) = \gamma(t)$ for all $\gamma \in \Gamma$. Let us also set

$$\mathcal{P}_{m_0^*}(\Gamma) = \{\xi \in \mathcal{P}_1\left(\Gamma\right) \mid e_0\sharp\xi = m_0^*\}.$$

The notion of equilibria that we consider is the *Lagrangian* MFG *equilibria*, defined as a solution to the following problem:

Problem 2 *Find $\xi^* \in \mathcal{P}_{m_0^*}(\Gamma)$ such that $[0,T] \ni t \mapsto e_t\sharp\xi^* \in \mathcal{P}_1(\mathbb{R}^d)$ belongs to $C\left([0,T];\mathcal{P}_1(\mathbb{R}^d)\right)$ and for ξ^*-a.e. $\gamma^* \in \Gamma$ there exists $\alpha^* \in L^p([0,T];\mathbb{R}^r)$ such that (γ^*, α^*) solves $(OC_{x,m})$ with $x = \gamma^*(0)$ and $m(t) = e_t\sharp\xi^*$ for all $t \in [0,T]$.*

We have the following result.

Theorem 3.2.1 *Assume that* **(H1)**, **(H2)**, *and* **(H3)** *hold. Then Problem 2 admits at least one solution. Moreover, if* **(H4)** *holds and for every $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and m_0^*-a.e. $x \in \mathbb{R}^d$ problem $(OC_{x,m})$ admits a unique solution, then the MFG equilibrium is unique.*

Proof. The existence of at least one solution follows from Theorem 3.4.2 below while the uniqueness result can be shown by arguing exactly as in the proof of [80, Theorem 2.2]. □

3.3 The value function of a typical player and its approximation

In this section we fix $p \in (1, \infty)$. For every $(t,x) \in [0,T] \times \mathbb{R}^d$ and $\alpha \in L^p([t,T];\mathbb{R}^r)$, note that **(H2)** implies that

$$\dot{\gamma}(s) = A(s, \gamma(s)) + B(s)\alpha(s) \quad \text{for a.e. } s \in (t,T), \quad \gamma(t) = x \tag{3.3.1}$$

admits a unique solution $\gamma^{t,x,\alpha} \in W^{1,p}([0,T];\mathbb{R}^d)$. Given $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, set

$$J^{t,x}[m](\alpha) = \int_t^T \ell(s, \alpha(s), \gamma^{t,x,\alpha}(s), m(s))\mathrm{d}s + g(\gamma(T), m(T)) \quad \text{for all } \alpha \in L^p([t,T];\mathbb{R}^r).$$

The value function $v[m] \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is defined as

$$v[m](t,x) = \inf\left\{ J^{t,x}[m](\alpha) \,|\, \alpha \in L^p([t,T];\mathbb{R}^r)\right\} \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d. \tag{3.3.2}$$

Proposition 3.3.1 *Assume that* **(H1)** *and* **(H2)** *hold, let* $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, and let* $(t,x) \in [0,T]\times\mathbb{R}^d$*. Then there exists* $\alpha^* \in L^p\big([t,T];\mathbb{R}^d\big)$ *such that* $v[m](t,x) = J^{t,x}[m](\alpha^*)$*. Moreover, the following hold:*

(i) *There exists* $C_{\mathrm{Lip}} > 0$*, independent of* (t,x,m)*, such that*

$$\big|v[m](t,x) - v[m](t,y)\big| \leq C_{\mathrm{Lip}}(1 + |x|^{p-1} + |y|^{p-1})|x-y| \quad \textit{for all } y \in \mathbb{R}^d. \tag{3.3.3}$$

(ii) *There exists* $C_{\mathrm{b}} > 0$*, independent of* (t,x,m)*, such that*

$$|\alpha^*(s)| \leq C_{\mathrm{b}}\big(1 + |\gamma^{t,x,\alpha^*}(s)|\big) \quad \textit{for a.e. } s \in [t,T]. \tag{3.3.4}$$

Proof. Let $(\alpha_n)_{n\in\mathbb{N}} \subset L^p([t,T];\mathbb{R}^r)$ be a minimizing sequence for the right-hand side of (3.3.2). Estimate (3.2.4) and (3.2.5) imply that $(\alpha_n)_{n\in\mathbb{N}}$ is bounded in $L^p([t,T];\mathbb{R}^r)$ and hence, up to some subsequence, it converges weakly to some $\alpha^* \in L^p([t,T];\mathbb{R}^r)$. It follows from (3.3.1) and **(H2)**(i)-(ii) that γ^{t,x,α_n} converges uniformly in $[t,T]$ to γ^{t,x,α^*} and hence, by [72, Theorem 3.23], we deduce that $v[m](t,x) = J^{t,x}[m](\alpha^*)$. Denoting by α^0 the null control and $\gamma^0 = \gamma^{t,x,\alpha_0}$, estimate (3.2.9) and Grönwall's lemma imply that $\sup_{s\in[t,T]} |\gamma^0(s)| \leq e^{C_A T}(|x| + C_A T)$. In turn, it follows from (3.2.4), (3.2.5), (3.2.9), and Grönwall's inequality that

$$\begin{aligned}
C_{\ell,1}\int_t^T|\alpha^*(s)|^p\mathrm{d}s &\le TC_{\ell,2}+J^{t,x}[m](\alpha^0)+C_{g,1}\\
&\le \Big(C_{\ell,2}+C_{\ell,3}(1+e^{pC_AT}(|x|+C_AT)^p)\Big)T+C_{g,1}\\
&\quad +C_{g,2}\Big(1+e^{pC_AT}(|x|+C_AT)^p\Big)\\
&\le TC_{\ell,2}+C_{g,1}+(TC_{\ell,3}+C_{g,2})\left(1+2^{p-1}e^{pC_AT}C_A^pT^p\right)\\
&\quad +|x|^p2^{p-1}e^{pC_AT}\left(TC_{\ell,3}+C_{g,2}\right).
\end{aligned}$$

Defining

$$\begin{aligned}
\tilde{C}=\frac{1}{C_{\ell,1}^{\frac{1}{p}}}\max\Big\{&\left(TC_{\ell,2}+C_{g,1}+(TC_{\ell,3}+C_{g,2})\left(1+2^{p-1}e^{pC_aT}C_A^pT^p\right)\right)^{\frac{1}{p}},\\
&2^{\frac{p-1}{p}}e^{C_AT}\left(TC_{\ell,3}+C_{g,2}\right)^{\frac{1}{p}}\Big\},
\end{aligned}$$

which is independent of (t,x,m), we obtain

$$\|\alpha^*\|_{L^p}\le\tilde{C}(1+|x|). \tag{3.3.5}$$

In particular, it holds that

$v[m](t,x)=\inf\left\{J^{t,x}[m](\alpha)\,|\,\alpha\in L^p([t,T];\mathbb{R}^r),\ \|\alpha\|_{L^p}\le\tilde{C}(1+|x|)\right\}$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$.

Let us now show assertions $\mathrm{(i)}$-$\mathrm{(ii)}$.
$\mathrm{(i)}$: Let $y\in\mathbb{R}^d$ and set $\gamma^*=\gamma^{t,x,\alpha^*}$ and $\tilde{\gamma}=\gamma^{t,y,\alpha^*}$. By standard arguments, it follows from **(H2)**$\mathrm{(i)}$, (3.2.9), Grönwall's lemma, Hölder's inequality, and (3.3.5) that

$$\begin{aligned}
\sup_{s\in[t,T]}|\gamma^*(s)| &\le (|x|+C_B\|\alpha^*\|_{L^1}+C_AT)\,e^{C_AT}\\
&\le \left(|x|+C_BT^{\frac{p-1}{p}}\|\alpha^*\|_{L^p}+C_AT\right)e^{C_AT}\\
&\le \hat{C}_1\left(1+|x|\right),
\end{aligned}\tag{3.3.6}$$

where

$$\hat{C}_1=e^{C_AT}\max\left\{C_BT^{\frac{p-1}{p}}\tilde{C}+C_AT,1+C_BT^{\frac{p-1}{p}}\tilde{C}\right\}. \tag{3.3.7}$$

Analogously we obtain

$$\sup_{s\in[t,T]}|\tilde{\gamma}(s)|\le\hat{C}_1\left(1+|x|+|y|\right). \tag{3.3.8}$$

From **(H2)**(ii) and Grönwall's lemma we have,

$$\sup_{s\in[t,T]}|\tilde{\gamma}^*(s)-\gamma^*(s)|\leq e^{L_AT}|x-y|. \tag{3.3.9}$$

In turn, by (3.2.5) and (3.2.6) we have

$$\begin{aligned} v[m](t,y)-v[m](t,x) &\leq \int_t^T \Big(\ell(s,\alpha^*(s),\tilde{\gamma}(s),m(s))-\ell(s,\alpha^*(s),\gamma^*(s),m(s))\Big)\mathrm{d}s \\ &\quad +g(\tilde{\gamma}(T),m(T))-g(\gamma^*(T),m(T)) \\ &\leq C_{\ell,4}\int_t^T \left(1+|\tilde{\gamma}(s)|^{p-1}+|\gamma^*(s)|^{p-1}+|\alpha^*(s)|^{p-1}\right)|\tilde{\gamma}(s)-\gamma^*(s)|\mathrm{d}s \\ &\quad +C_{g,3}\left(1+|\tilde{\gamma}(T)|^{p-1}+|\gamma^*(T)|^{p-1}\right)|\tilde{\gamma}(T)-\gamma^*(T)| \\ &\leq \Big(C_{\ell,4}T+C_{g,3}+C_{\ell,4}\int_t^T|\alpha^*(s)|^{p-1}\mathrm{d}s \\ &\quad +2\hat{C}_1^{p-1}3^{p-1}(1+|x|^{p-1}+|y|^{p-1})(C_{\ell,4}T+C_{g,3})\Big)\,e^{L_AT}|x-y|. \end{aligned} \tag{3.3.10}$$

By Hölder's inequality and (3.3.5), we have

$$\int_t^T|\alpha^*(s)|^{p-1}\mathrm{d}s\leq T^{\frac{1}{p}}\left(\int_t^T|\alpha^*(s)|^p\mathrm{d}s\right)^{\frac{p-1}{p}}\leq T^{\frac{1}{p}}(2\tilde{C})^{p-1}\left(1+|x|^{p-1}\right). \tag{3.3.11}$$

Combining (3.3.6)-(3.3.11) we deduce the existence of $C_{\text{Lip}}>0$ independent of (t,x,y,m), such that

$$v[m](t,y)-v[m](t,x)\leq C_{\text{Lip}}(1+|x|^{p-1}+|y|^{p-1})|x-y|,$$

where

$$C_{\text{Lip}}=e^{L_AT}\left(C_{\ell,4}T+C_{g,3}+C_{\ell,4}T^{\frac{1}{p}}(2\tilde{C})^{p-1}+2\hat{C}_1^{p-1}3^{p-1}(C_{\ell,4}T+C_{g,3})\right).$$

The inequality for $v[m](t,x)-v[m](t,y)$ follows by exchanging the roles of x and y in the previous computation.

(ii): Let $s\in[t,T)$, $h\in[0,T-s)$, and set $y^*=\gamma^*(s)$. Since $v[m]$ satisfies the dynamic programming inequality (see e.g. [23])

$$v[m](s,\gamma^*(s))\leq\int_s^{s+h}\ell(r,\gamma^{s,y^*,\alpha}(r),\alpha(r),m(r))\mathrm{d}r+v[m]\big(s+h,\gamma^{s,y^*,\alpha}(s+h)\big), \tag{3.3.12}$$

for all $\alpha\in L^p([t,T];\mathbb{R}^d)$, with equality for $\alpha=\alpha^*$, by taking $\alpha=\alpha_0$ (the null control) and

$\alpha = \alpha^*$, the equality $v[m](s, y^*) = J^{s,y^*}[m](\alpha^*|_{[s,T]})$, and estimate (3.2.4) yield

$$C_{\ell,1} \int_s^{s+h} |\alpha^*(r)|^p \mathrm{d}r \leq C_{\ell,2} h + C_{\ell,3} \int_s^{s+h} \big(1 + |\gamma^{s,y^*,\alpha_0}(r)|^p\big) \mathrm{d}r \\ + v[m](s+h, \gamma^{s,y^*,\alpha_0}(s+h)) - v[m](s+h, \gamma^*(s+h)). \quad (3.3.13)$$

By (3.2.9) and Grönwall's lemma we have

$$\sup_{r \in [s,T]} |\gamma^{s,y^*,\alpha_0}(r)| \leq e^{C_A T}(|\gamma^*(s)| + C_A T).$$

On the other hand, supposing that $h < 1$, it follows from **(H2)**(ii) and Grönwall's Lemma that

$$|\gamma^*(s+h) - \gamma^{s,y^*,\alpha_0}(s+h)| \leq e^{L_A} C_B \int_s^{s+h} |\alpha^*(r)| \mathrm{d}r.$$

Combining the above two inequalities with (3.3.13) and using (i), we get that

$$C_{\ell,1} \int_s^{s+h} |\alpha^*(r)|^p \mathrm{d}r \leq C_{\ell,2} h + C_{\ell,3} h \big(1 + (e^{C_A T}(|\gamma^*(s)| + C_A T))^p\big) + C_{\mathrm{Lip}} e^{L_A} C_B \int_s^{s+h} |\alpha^*(r)| \mathrm{d}r.$$

By Young's inequality, for $\epsilon > 0$,

$$\int_s^{s+h} |\alpha^*(r)| \mathrm{d}r \leq \frac{\epsilon}{p} \int_s^{s+h} |\alpha^*(r)|^p \mathrm{d}r + \frac{h(p-1)}{\epsilon^{\frac{1}{p-1}} p}.$$

Taking $\epsilon = \frac{p C_{\ell,1}}{2 C_{\mathrm{Lip}} e^{L_A} C_B}$ and defining

$$\hat{C}_2 = \frac{2}{C_{\ell,1}} \max \left\{ C_{\ell,2} + C_{\ell,3}(1 + e^{C_A T} 2^{p-1} (C_A T)^p) + \frac{2^{\frac{1}{p-1}}(p-1)(C_{\mathrm{Lip}} e^{L_A} C_B)^{\frac{p}{p-1}}}{p^{\frac{p}{p-1}} C_{\ell,1}^{\frac{1}{p-1}}}, \right. \\ \left. C_{\ell,3} e^{p C_A T} 2^{p-1} \right\}$$

we obtain

$$\int_s^{s+h} |\alpha^*(r)|^p \mathrm{d}r \leq \hat{C}_2 h (1 + |y^*|^p). \quad (3.3.14)$$

Therefore, setting $C_{\mathrm{b}} = \hat{C}_2^{\frac{1}{p}}$, estimate (3.3.4) follows from the Lebesgue differentiation theorem (see e.g. [32]). □

Remark 3.3.1 *Proposition 3.3.1*(ii) *implies that, for any* $(t,x) \in [0,T] \times \mathbb{R}^d$*,* $v[m](t,x)$ *can be rewritten as*

$$v[m](t,x) = \inf\left\{J^{t,x}[m](\alpha) \,|\, \alpha \in L^\infty([t,T];\mathbb{R}^r),\ |\alpha(s)| \leq C_{\rm b}(1+|\gamma^{t,x,\alpha}(s)|) \text{ for a.e. } s \in [t,T]\right\}. \tag{3.3.15}$$

We consider now the approximation of the value function $v[m]$ given by (3.3.2). Let $N_t \in \mathbb{N}$, $N_s \in \mathbb{N}$, with $N_s \geq N_t$, and set $\Delta t = T/N_t$, $\Delta x = 1/N_s$, $\mathcal{I} = \{0,\dots,N_t\}$, $\mathcal{I}^* = \mathcal{I} \setminus \{N_t\}$, and $\mathcal{G} = \{i\Delta x \,|\, i \in \mathbb{Z}^d\}$. Given a regular mesh $\mathscr{T}$ with vertices in $\mathcal{G}$, let $(\psi_x)_{x\in\mathcal{G}}$ be a $\mathbb{Q}_1$ basis, i.e. for every $x \in \mathcal{G}$, $y = (y_1,\dots,y_d)$, and $i = 1,\dots,d$, the function $\mathbb{R} \ni z_i \mapsto \psi_x(y_1,\dots,z_i,\dots,y_d) \in \mathbb{R}$ is nonnegative and affine on each element of $\mathscr{T}$, $\psi_x(x) = 1$, $\psi_x(y) = 0$ for all $y \in \mathcal{G}$ with $y \neq x$, and $\sum_{x\in\mathcal{G}} \psi_x(z) = 1$ for all $z \in \mathbb{R}^d$. Given $\varphi\colon \mathcal{G} \to \mathbb{R}$, define its interpolant $I[\varphi]\colon \mathbb{R}^d \to \mathbb{R}$ by

$$I[\varphi](x) = \sum_{y\in\mathcal{G}} \psi_y(x)\varphi(y) \quad \text{for all } x \in \mathbb{R}^d.$$

In view of Remark 3.3.1, a standard semi-Lagrangian scheme (see e.g. [76]) to approximate $v[m]$ is given by

$$\begin{aligned} V_k(x) &= \min_{a\in\overline{B}_\infty(0,C_{\rm b}(1+|x|))} \left\{\Delta t \ell(t_k,a,x,m(t_k)) + I[V_{k+1}](\Phi(k,x,a))\right\} \\ &\hspace{8cm} \text{for all } k \in \mathcal{I}^*,\ x \in \mathcal{G}, \\ V_{N_t}(x) &= g(x,m(T)) \quad \text{for all } x \in \mathcal{G}, \end{aligned} \tag{3.3.16}$$

where

$$\Phi(k,x,a) = x + \Delta t(A(t_k,x) + B(t_k)a) \quad \text{for all } k \in \mathcal{I}^*,\ x \in \mathcal{G},\ a \in \mathbb{R}^r. \tag{3.3.17}$$

We now introduce a variation of the previous scheme which exploits the particular structure of the dynamics in (3.3.1). First, notice that, by **(H2)**(iii), without loss of generality we can write

$$A(t,x) = \begin{pmatrix} A_1(t,x) \\ A_2(t,x) \end{pmatrix} \quad \text{and} \quad B(t) = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix} \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d, \tag{3.3.18}$$

where $A_1\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^r$, $A_2\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d-r}$, $B_1\colon [0,T] \to \mathbb{R}^{r\times r}$ is such that $B_1(t)$ is invertible for all $t \in [0,T]$, and $B_2\colon [0,T] \to \mathbb{R}^{(d-r)\times r}$. We partition the coordinates of $x \in \mathbb{R}^d$ accordingly by writing $x = (\mathrm{x}_1,\mathrm{x}_2)$, where $\mathrm{x}_1 \in \mathbb{R}^r$ and $\mathrm{x}_2 \in \mathbb{R}^{d-r}$ denote the first

r and the last $d-r$ components of x, respectively. We also write $\mathcal{G}=\mathcal{G}_r\times\mathcal{G}_{d-r}$, where

$$\mathcal{G}_r=\{i\Delta x \mid i\in\mathbb{Z}^r\} \quad \text{and} \quad \mathcal{G}_{d-r}=\left\{i\Delta x \mid i\in\mathbb{Z}^{d-r}\right\},$$

and we suppose that the basis $(\psi_x)_{x\in\mathcal{G}}$ can be decomposed as the tensorial product of two $\mathbb{Q}_1$ basis $(\eta_{\mathrm{x}_1})_{\mathrm{x}_1\in\mathcal{G}_r}$ and $(\beta_{\mathrm{x}_2})_{\mathrm{x}_2\in\mathcal{G}_{d-r}}$ defined on regular meshes with vertices in $\mathcal{G}_r$ and $\mathcal{G}_{d-r}$, respectively. More precisely, we suppose that for every $x=(\mathrm{x}_1,\mathrm{x}_2)\in\mathcal{G}$ we have

$$\psi_x(y)=\eta_{\mathrm{x}_1}(\mathrm{y}_1)\beta_{\mathrm{x}_2}(\mathrm{y}_2) \quad \text{for all } y=(\mathrm{y}_1,\mathrm{y}_2)\in\mathbb{R}^d. \tag{3.3.19}$$

In what follows, we assume the existence of $C_I>0$, independent of Δx, such that

$$\operatorname{supp}(\beta_{\mathrm{x}_2})\subseteq\{\mathrm{y}_2\in\mathbb{R}^{d-r}\mid |\mathrm{y}_2-\mathrm{x}_2|\leq C_I\Delta x\} \quad \text{for all } \mathrm{x}_2\in\mathcal{G}_{d-r}. \tag{3.3.20}$$

Let $k\in\mathcal{I}^*$ and $x\in\mathbb{R}^d$. In the modified version of (3.3.16), we will only consider controls $a\in\overline{B}_\infty(0,C_{\mathrm{b}}(1+|x|))$ such that, writing $\Phi(k,x,a)=(\Phi_1(k,x,a),\Phi_2(k,x,a))$, we have $\Phi_1(k,x,a)\in\mathcal{G}_r$. Notice that, for every $\mathrm{y}_1\in\mathcal{G}_r$, it holds that

$$\mathrm{y}_1=\Phi_1(k,x,a) \;\Leftrightarrow\; a=B_1(t_k)^{-1}\left[\frac{\mathrm{y}_1-\mathrm{x}_1}{\Delta t}-A_1(t_k,x)\right]. \tag{3.3.21}$$

Thus, setting

$$\begin{aligned}
\alpha(k,x,\mathrm{y}_1) &:= B_1(t_k)^{-1}\left[\tfrac{\mathrm{y}_1-\mathrm{x}_1}{\Delta t}-A_1(t_k,x)\right]\in\mathbb{R}^r,\\
\mathbf{y_2}(k,x,\mathrm{y}_1) &:= \mathrm{x}_2+\Delta t\left[A_2(t_k,x)+B_2(t_k)\alpha(k,x,\mathrm{y}_1)\right]\in\mathbb{R}^{d-r},
\end{aligned} \tag{3.3.22}$$

it is natural to define the sets

$$\begin{aligned}
\mathcal{S}^1_{k+1}(x) &= \left\{\mathrm{y}_1\in\mathcal{G}_r \;\middle|\; |\alpha(k,x,\mathrm{y}_1)|\leq C_{\mathrm{b}}(1+|x|)\right\},\\
\mathcal{S}^2_{k+1}(x,\mathrm{y}_1) &= \left\{\mathrm{y}_2\in\mathcal{G}_{d-r} \;\middle|\; \mathbf{y_2}(k,x,\mathrm{y}_1)\in\operatorname{supp}\beta_{\mathrm{y}_2}\right\} \quad \text{for } \mathrm{y}_1\in\mathcal{S}^1_{k+1}(x),\\
\mathcal{S}_{k+1}(x) &= \left\{(\mathrm{y}_1,\mathrm{y}_2)\in\mathcal{G} \;\middle|\; \mathrm{y}_1\in\mathcal{S}^1_{k+1}(x),\ \mathrm{y}_2\in\mathcal{S}^2_{k+1}(x,\mathrm{y}_1)\right\}.
\end{aligned} \tag{3.3.23}$$

Arguing as in [80, Lemma 3.2], one checks that, if $\Delta x/\Delta t$ is small enough, then $\mathcal{S}_{k+1}(x)\neq\emptyset$. Starting from an initial grid $\mathcal{S}_0=\mathcal{G}\cap\overline{B}_\infty(0,C^*)$, we can then construct the family of time-depending grids

$$\mathcal{S}_{k+1}=\bigcup_{x\in\mathcal{S}_k}\mathcal{S}_{k+1}(x) \quad \text{for all } k\in\mathcal{I}^* \tag{3.3.24}$$

with the property that $(\mathrm{y}_1,\mathrm{y}_2)\in\mathcal{S}_{k+1}$ if and only if there exists $x\in\mathcal{S}_k$ and $a\in\mathbb{R}^r$ such that $|a|\leq C_{\mathrm{b}}(1+|x|)$, $\mathrm{y}_1=\Phi_1(k,x,a)$, and $\Phi_2(k,x,a)\in\operatorname{supp}\beta_{\mathrm{y}_2}$.

On the other hand, notice that, for every $\varphi\colon \mathcal{G} \to \mathbb{R}$, $\mathrm{y}_1 \in \mathcal{G}_r$, and $\mathrm{y}_2 \in \mathbb{R}^{d-r}$, we have

$$\begin{aligned} I[\varphi](\mathrm{y}_1,\mathrm{y}_2) &= \sum_{\mathrm{z}_1\in\mathcal{G}_r, \mathrm{z}_2\in\mathcal{G}_{d-r}} \psi_z(\mathrm{y}_1,\mathrm{y}_2)\varphi(\mathrm{z}_1,\mathrm{z}_2) \\ &= \sum_{\mathrm{z}_1\in\mathcal{G}_r} \eta_{z_1}(\mathrm{y}_1) \sum_{\mathrm{z}_2\in\mathcal{G}_{d-r}} \beta_{\mathrm{z}_2}(\mathrm{y}_2)\varphi(\mathrm{z}_1,\mathrm{z}_2) = \sum_{\mathrm{z}_2\in\mathcal{G}_{d-r}} \beta_{\mathrm{z}_2}(\mathrm{y}_2)\varphi(\mathrm{y}_1,\mathrm{z}_2). \end{aligned} \tag{3.3.25}$$

Altogether, (3.3.16)-(3.3.25) suggest to consider the following variation of (3.3.16):

$$\begin{aligned} \mathcal{V}_k(x) &= \min_{\mathrm{y}_1\in\mathcal{S}^1_{k+1}(x)} \Big\{ \Delta t \ell(t_k, \alpha(k,x,\mathrm{y}_1), x, m(t_k)) \\ &\qquad + I_{\mathcal{S}^2_{k+1}(x,\mathrm{y}_1)}[\mathcal{V}_{k+1}(\mathrm{y}_1,\cdot)](\mathbf{y_2}(k,x,\mathrm{y}_1)) \Big\} \quad \text{for all } k\in\mathcal{I}^*,\ x\in\mathcal{S}_k, \\ \mathcal{V}_{N_t}(x) &= g(x,m(T)) \quad \text{for all } x\in\mathcal{S}_{N_t}, \end{aligned} \tag{3.3.26}$$

where, for $F \subseteq \mathcal{G}_{d-r}$ and $\varphi\colon F \to \mathbb{R}$, we have set

$$I_F[\varphi](\mathrm{y}_2) = \sum_{\mathrm{x}_2\in F} \beta_{\mathrm{x}_2}(\mathrm{y}_2)\varphi(\mathrm{x}_2) \quad \text{for all } \mathrm{y}_2\in\mathbb{R}^{d-r}.$$

Notice that (3.3.26) can be rewritten as

$$\begin{aligned} \mathcal{V}_k(x) &= \min_{p\in\mathcal{P}(\mathcal{S}^1_{k+1}(x))} \Big\{ \sum_{\mathrm{y}_1\in\mathcal{S}^1_{k+1}(x)} p(\mathrm{y}_1)\Big[\Delta t\ell(t_k,\alpha(k,x,\mathrm{y}_1),x,m(t_k)) \\ &\qquad + I_{\mathcal{S}^2_{k+1}(x,\mathrm{y}_1)}[\mathcal{V}_{k+1}(\mathrm{y}_1,\cdot)]\big(\mathbf{y_2}(k,x,\mathrm{y}_1)\big)\Big]\Big\} \quad \text{for all } k\in\mathcal{I}^*,\ x\in\mathcal{S}_k, \\ \mathcal{V}_{N_t}(x) &= g(x,m(T)) \quad \text{for all } x\in\mathcal{S}_{N_t}. \end{aligned} \tag{3.3.27}$$

Remark 3.3.2 (i) *Note that when $r = d$ we have $\mathcal{S}^1_k(x) = \mathcal{S}_k(x)$ for all $x \in \mathbb{R}^d$ and $k \in \mathcal{I}\setminus\{0\}$. In particular, no interpolation is needed in* (3.3.27)*, and the scheme reduces to*

$$\begin{aligned} \mathcal{V}_k(x) &= \min_{p\in\mathcal{P}(\mathcal{S}_{k+1}(x))} \sum_{y\in\mathcal{S}_{k+1}(x)} p(y)\Big[\Delta t\ell(t_k,\alpha(k,x,y),x,m(t_k)) + \mathcal{V}_{k+1}(y)\Big] \\ &\qquad\qquad \textit{for all } k\in\mathcal{I}^*,\ x\in\mathcal{S}_k, \\ \mathcal{V}_{N_t}(x) &= g(x,m(T)) \quad \textit{for all } x\in\mathcal{S}_{N_t}. \end{aligned} \tag{3.3.28}$$

The scheme (3.3.28) *is a slight extension of the one proposed in [94], which deals with the case where* (3.1.1) *has the simple form $\dot\gamma(t) = \alpha(t)$ for a.e. $t \in [0,T]$.*

(ii) *Notice that, for every $k \in \mathcal{I}^*$ and $x \in \mathcal{S}_k$, the set $\{\alpha(k,x,\mathrm{y}_1)\,|\,\mathrm{y}_1 \in \mathcal{S}^1_{k+1}(x)\}$ is a specific discretization of the set $\overline{B}_\infty(0, C_{\mathrm{b}}(1+|x|))$ appearing in* (3.3.16)*. Other*

discretizations of this set are also possible and the analysis of the corresponding schemes for the value function, and also the schemes to solve Problem 2 presented in Section 3.4 below, are analogous to the one considered in this work. An important feature of our scheme is that it takes advantage of the specific form of the dynamics to construct a discretization of the control set $\overline{B}_\infty(0, C_{\rm b}(1+|x|))$ depending on the same grid as the one used to discretize the state space.
Other discretizations of the control space can be useful to deal with control problems and mean field games with dynamics which are nonlinear with respect to the control variable. We intend to consider this extension of our analysis in a future work.
The following result, which shows that the family of time dependent grids $(\mathcal{S}_k)_{k\in\mathcal{I}}$ remains uniformly bounded with respect to the discretization parameters, will play a key role in what follows.

Lemma 3.3.1 *There exists a nonempty compact set $K \subset \mathbb{R}^d$, independent of Δt and Δx as long as $\Delta x/\Delta t \leq 1$, such that*

$$\mathcal{S}_k \subset K \quad \textit{for all } k \in \mathcal{I}^*.$$

Proof. Consider the following family of compact sets: set $K_0 = \overline{B}_\infty(0, C^*)$ and

$$K_{k+1} := K_k + \left(\Delta t\left[\big(C_A + C_B C_{\rm b}\big)\big(1+\sup_{x\in K_k}|x|\big)\right] + C_I \Delta x\right)\overline{B}_\infty(0,1) \quad \text{for all } k \in \mathcal{I}^*, \tag{3.3.29}$$

where we recall that C_A is given in (3.2.9), $C_B = \sup_{t\in[0,T]}|B(t)|$, and C_I satisfies (3.3.20). It follows from (3.3.23) and (3.3.24) that $\mathcal{S}_k \subset K_k$ for all $k \in \mathcal{I}$ and hence it suffices to show that the family $(K_k)_{k\in\mathcal{I}}$ is uniformly bounded. Let $k \in \mathcal{I}$ and set $c_k = \sup_{x\in K_k}|x|$. Equation (3.3.29) yields

$$\begin{aligned} c_{k+1} &\leq c_k + \left(\Delta t\left[\big(C_A + C_B C_{\rm b}\big)\big(1+c_k\big) + C_I \tfrac{\Delta x}{\Delta t}\right]\right) \\ &\leq \big(1+\Delta t(C_A + C_B C_{\rm b})\big)c_k + \Delta t(C_A + C_B C_{\rm b} + C_I), \end{aligned}$$

which, by the discrete Grönwall's lemma, implies that the set $\{c_k \,|\, k\in\mathcal{I}\}$ is uniformly bounded. The result follows. □

Proposition 3.3.2 *Assume that* **(H1)** *and* **(H2)** *hold. Consider three sequences $(N_t^n, N_s^n) \subset \mathbb{N}^2$ and $(m_n)_{n\in\mathbb{N}} \subset C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ such that, as $n\to\infty$, $N_t^n \to \infty$, $N_s^n \to \infty$, $N_t^n/N_s^n \to 0$, and $m_n \to m^*$ for some $m^* \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. Set $\mathcal{I}^n = \{0,\ldots,N_t^n\}$*

and, associated with the parameters (N_t^n, N_s^n) *and* m^n*, define* $\mathcal{S}_k^n$ *as in* (3.3.23) *and denote by* $\mathcal{V}^n$ *the solution to* (3.3.26)*. Then it holds that*

$$\sup \left\{ |\mathcal{V}_k^n(x) - v[m^*](t_k^n, x)| \;\middle|\; k \in \mathcal{I}^n,\ x \in \mathcal{S}_k^n \right\} \underset{n\to\infty}{\longrightarrow} 0, \tag{3.3.30}$$

where $v[m^*]$ *is defined in* (3.3.2)*.*

Proof. [Sketch of the proof] Let $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and consider, as an intermediate step, the following semi-discrete scheme to approximate $v[m]$:

$$\begin{aligned} v_d[m](k,x) &= \min_{\alpha\in\mathbb{R}^r} \{\Delta t \ell(t_k, \alpha, x, m(t_k)) + v_d[m]\,(k+1, x+\Delta t[A(t_k,x)+B(t_k)\alpha])\} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } k \in \mathcal{I}^*,\ x\in\mathbb{R}^d, \\ v_d(N_t, x) &= g(x, m(T)) \quad \text{for all } x \in \mathbb{R}^d. \end{aligned} \tag{3.3.31}$$

Setting v_d^n for the solution to the previous scheme associated with m^n, using the framework developed in [25], and arguing as in [80, Proposition 3.1] one shows that for every compact set $K \subset \mathbb{R}^d$ it holds that

$$\sup_{(k,x)\in\mathcal{I}^n\times K} |v_d^n(k,x) - v[m^*](t_k^n, x)| \underset{n\to\infty}{\longrightarrow} 0. \tag{3.3.32}$$

On the other hand, by adapting to the discrete case the proof of Proposition 3.3.1, one checks that the minimization on the right-hand-side of (3.3.31) can be restricted the set $\{a \in \mathbb{R}^d \,|\, |a| \leq C_{\rm b}(1+|x|)\}$. Using this fact, Lemma 3.3.1, and arguing as in the proof of [80, Lemma 5.3 (ii)] we obtain that

$$\max \{ |v_d^n(k,x) - \mathcal{V}_k^n(x)| \mid k \in \mathcal{I}^n,\ x \in \mathcal{S}_k^n\} \to 0 \quad \text{as } n\to\infty, \tag{3.3.33}$$

and hence (3.3.30) follows from (3.3.32) and (3.3.33). □

3.4 The finite mean field game approximation

In this section, given $N_t \in \mathbb{N}$, $N_s \in \mathbb{N}$, with $N_s \geq N_t$, we approximate Problem 2 by a fixed point problem of a map $\mathbf{br}$, called *best response mapping* defined on the space $\mathcal{M} = \prod_{k\in\mathcal{I}} \mathcal{P}(\mathcal{S}_k)$ of discrete time marginals. The resulting approximation will take the form of a discrete time and finite state MFG (see [84]). In order to construct the map $\mathbf{br}$, let us first introduce some useful definitions. For every $k \in \mathcal{I}$, we identify $p \in \mathcal{P}(\mathcal{S}_k)$ with the probability measure $\sum_{x\in\mathcal{S}_k} p(\{x\})\delta_x \in \mathcal{P}_1(\mathbb{R}^d)$ and, given a finite set F, the (nonpositive) entropy function $\mathcal{E}_F\colon \mathcal{P}(F) \to \mathbb{R}$ is defined by

$$\mathcal{E}_F(p) = \sum_{x\in F} p(x)\log(p(x)) \quad \text{for all } p \in \mathcal{P}(F),$$

with the convention that $p(x) = p(\{x\})$ and $0\log(0) = 0$. Given $M \in \mathcal{M}$ and $\varepsilon > 0$, let us consider the following variation of (3.3.27):

$$\begin{aligned} V_k^M(x) = \min_{p\in\mathcal{P}(\mathcal{S}^1_{k+1}(x))} \Bigg\{ \sum_{\mathrm{y}_1\in\mathcal{S}^1_{k+1}(x)} p(\mathrm{y}_1)\Big[\Delta t \ell(t_k, \alpha(k,x,\mathrm{y}_1), x, M_k) \\ + I_{\mathcal{S}^2_{k+1}(x,\mathrm{y}_1)}[V^M_{k+1}(\mathrm{y}_1,\cdot)](\mathbf{y_2}(k,x,\mathrm{y}_1))\Big] + \varepsilon\mathcal{E}_{\mathcal{S}^1_{k+1}(x)}(p)\Bigg\} \quad (3.4.1) \\ \text{for all } k\in\mathcal{I}^*,\ x\in\mathcal{S}_k, \\ V^M_{N_t}(x) = g(x, M_{N_t}) \quad \text{for all } x\in\mathcal{S}_{N_t}. \end{aligned}$$

Notice that the incorporation of the entropy term in the scheme above implies that, for every $k \in \mathcal{I}^*$ and $x \in \mathcal{S}_k$, the optimization problem defining $V_k^M(x)$ admits a unique solution $p_k^M(x,\cdot)$ which satisfies $p_k^M(x,\mathrm{y}_1) > 0$ for all $\mathrm{y}_1 \in \mathcal{S}^1_{k+1}(x)$. Given $y \in \mathcal{S}_{k+1}$, we also set

$$P_k^M(x,y) := \begin{cases} p_k^M(x,\mathrm{y}_1)\beta_{\mathrm{y}_2}(\mathbf{y_2}(k,x,\mathrm{y}_1)) & \text{if } y\in\mathcal{S}_{k+1}(x), \\ 0 & \text{if } y \in \mathcal{S}_{k+1}\setminus\mathcal{S}_{k+1}(x). \end{cases} \quad (3.4.2)$$

Letting $E(x) = \left\{y\in\mathbb{R}^d \mid |x-y| \leq \Delta x/2\right\}$ for all $x \in \mathcal{G}$, we define $\mathbf{br}(M)$ as the solution to

$$\begin{aligned} \widehat{M}_{k+1}(y) &= \sum_{x\in\mathcal{S}_k} P_k^M(x,y)\widehat{M}_k(x) \quad \text{for all } k\in\mathcal{I}^*,\ y\in\mathcal{S}_{k+1}, \\ \widehat{M}_0(x) &= m_0^*(E(x)) \quad \text{for all } x\in\mathcal{S}_0. \end{aligned} \quad (3.4.3)$$

The discretization of Problem 2 that we consider in this work reads as follows.

Problem 3 *Find $M\in\mathcal{M}$ such that $M = \mathbf{br}(M)$.*

We have the following result.
Theorem 3.4.1 *Assume that* **(H1)**, **(H2)**, *and* **(H3)** *hold. Then Problem 3 admits at least one solution. In addition, if* **(H4)** *holds then the solution is unique.*

Proof. Since, for every $M\in\mathcal{M}$, $k\in\mathcal{I}^*$, and $x\in\mathcal{S}_k$ we have that $p_k^M(x,\cdot)$ is unique, it is easy to check that $\mathbf{br}$ is continuous. In turn, the existence of a fixed point of $\mathbf{br}$ follows from Brouwer's fixed point theorem. The uniqueness result follows from the arguments in the proof of [80, Proposition 4.2], the key point being that, if $\widehat{M} = \mathbf{br}(M)$, then $\widehat{M}_k(x) > 0$ for all $k \in \mathcal{I}$ and $x \in \mathcal{S}_k$. □

Now, let us discuss the convergence of solutions to Problem 3 towards a solution to Problem 2 as the discretization parameters Δt, Δx, and ε tend to zero. Let $(N_t^n)_{n\in\mathbb{N}} \subset \mathbb{N}$, $(N_s^n)_{n\in\mathbb{N}} \subset \mathbb{N}$, $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$, and, for every $n \in \mathbb{N}$, set $\Delta t_n = T/N_t^n$, $\Delta x_n = 1/N_s^n$, $\mathcal{I}^n = \{0,\dots,N_t^n\}$, $\mathcal{I}^{n,*} := \mathcal{I}^n \setminus \{N_t^n\}$, $t_k^n = k\Delta t_n$ $(k \in \mathcal{I}^n)$, and $\mathcal{G}^n = \{i\Delta x_n \,|\, i \in \mathbb{Z}^d\}$. We assume that $N_s^n \geq N_t^n$. For $k \in \mathcal{I}^{n,*}$ and $x \in \mathcal{G}^n$, we denote by $\mathcal{S}_{k+1}^{1,n}(x)$, $\mathcal{S}_{k+1}^{2,n}(x, \mathrm{y}_1)$ ($\mathrm{y}_1 \in \mathcal{S}_{k+1}^{1,n}(x)$), and $\mathcal{S}_{k+1}^n(x)$ the sets defined in (3.3.23) associated with Δt_n and Δx_n. For $k \in \mathcal{I}^n$, the set $\mathcal{S}_k^n$ is defined as in (3.3.24). Denote by Γ^n the set of continuous functions $\gamma\colon [0,T] \to \mathbb{R}^d$ such that for each $k \in \mathcal{I}^n$, $\gamma(t_k^n) \in \mathcal{S}_k^n$ and, for every $k \in \mathcal{I}^{n,*}$, the restriction of γ to the interval $[t_k^n, t_{k+1}^n]$ is affine. Finally, let $M^n \in \mathcal{M}$ be a solution to Problem **??** associated with the previous parameters and, recalling (3.4.2), let us define $\xi^n \in \mathcal{P}(\Gamma)$ as

$$\xi^n = \sum_{\gamma\in\Gamma^n} M_0^n(\gamma(0))P^n(\gamma)\delta_\gamma \in \mathcal{P}(\Gamma), \quad \text{where} \quad P^n(\gamma) := \prod_{k=0}^{N_t^n-1} P_k^{M^n}(\gamma(t_k^n), \gamma(t_{k+1}^n)). \tag{3.4.4}$$

We extend M^n to the element in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ defined by

$$[0,T] \ni t \mapsto M^n(t) := e_t\sharp\xi^n \in \mathcal{P}_1(\mathbb{R}^d). \tag{3.4.5}$$

Lemma 3.4.1 *Assume that* **(H1)**, **(H2)**, *and* **(H3)** *are in force. Then the following hold:*

(i) *The family ξ^n has at least one accumulation point in $\mathcal{P}(\Gamma)$.*

(ii) *The family M^n has at least one accumulation point in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$.*

Proof. (i): Since $\mathrm{supp}(\xi^n) \subset \Gamma^n$, it follows from Lemma 3.3.1 that there exists $C_\infty > 0$ such that

$$\|\gamma\|_\infty \leq C_\infty \quad \text{for all } \gamma \in \mathrm{supp}(\xi^n). \tag{3.4.6}$$

Moreover, if $\gamma \in \mathrm{supp}(\xi^n)$, then γ is absolutely continuous with

$$\dot{\gamma}(t) = \frac{\gamma(t_{k+1}^n) - \gamma(t_k^n)}{\Delta t_n} \quad \text{for all } k \in \mathcal{I}^{n,*},\ t \in]t_k^n, t_{k+1}^n[. \tag{3.4.7}$$

Writing $\gamma(t_k^n) = (\gamma_1(t_k^n), \gamma_2(t_k^n)) \in \mathbb{R}^r \times \mathbb{R}^{d-r}$, the definition of $\mathcal{S}_{k+1}^{1,n}(\gamma(t_k^n))$ and (3.4.6) yield

$$\gamma_1(t_{k+1}^n) = \gamma_1(t_k^n) + \Delta t_n \left[A_1(t_k^n, \gamma(t_k^n)) + B_1(t_k^n)\alpha(k, \gamma(t_k^n), \gamma_1(t_{k+1}^n))\right],$$

with $\left|\alpha(k,\gamma(t_k^n),\gamma_1(t_{k+1}^n))\right| \leq C_{\rm b}(1+C_\infty)$. Thus, using (3.2.9) we deduce that

$$\left|\frac{\gamma_1(t_{k+1}^n)-\gamma_1(t_k^n)}{\Delta t_n}\right| \leq (C_A + C_B C_{\rm b})\,(1+C_\infty)\,,$$

and, by (3.3.22) and (3.3.23), we obtain

$$\left|\frac{\gamma_2(t_{k+1}^n)-\gamma_2(t_k^n)}{\Delta t_n}\right| \leq C_I \frac{\Delta x_n}{\Delta t_n} + (C_A + C_B C_{\rm b})\,(1+C_\infty)\,. \tag{3.4.8}$$

Since $\Delta x_n \leq \Delta t_n$, we deduce from (3.4.7) that there exists $D_\infty > 0$ such that

$$\|\dot\gamma\|_\infty \leq D_\infty \quad \text{for all } \gamma \in \operatorname{supp}(\xi^n)$$

and hence $\operatorname{supp}(\xi^n) \subset \{\gamma \in W^{1,\infty}([0,T];\mathbb{R}^d) \,|\, \|\gamma\|_\infty \leq C_\infty,\ \|\dot\gamma\|_\infty \leq D_\infty\}$, which is a compact subset of $(\Gamma, \|\cdot\|_\infty)$. Thus, the result follows from Prokhorov's theorem (see e.g. [18, Theorem 5.1.3]).
(ii): By (3.4.6), for every $t \in [0,T]$ and $n \in \mathbb{N}$, we have

$$\operatorname{supp}(M^n(t)) \subset \overline{\mathrm{B}}_\infty(0, C_\infty),$$

and, by (3.2.2) and (3.4.8),

$$d_1(M^n(s), M^n(t)) \leq D_\infty |s-t| \quad \text{for all } s,\, t \in [0,T],\, n \in \mathbb{N}.$$

Since $\{\mu \in \mathcal{P}_1(\mathbb{R}^d) \mid \operatorname{supp}(\mu) \subset \overline{\mathrm{B}}_\infty(0,C_\infty)\}$ is compact in $\mathcal{P}_1(\mathbb{R}^d)$ (see e.g. [18, Proposition 7.1.5]), the result follows from the Arzelà-Ascoli theorem. □
Using the previous compactness result and arguing as in the proof of [80, Theorem 5.1], one obtains the following convergence result.
Theorem 3.4.2 *Assume that* **(H1)**, **(H2)**, *and* **(H3)** *hold and that, as $n \to \infty$, $N_t^n \to \infty$, $N_s^n \to \infty$, $N_t^n/N_s^n \to 0$, and $\varepsilon_n = o(1/(N_t^n \log(N_s^n)))$. Then there exists a solution ξ^* to Problem 2 such that, up to some subsequence, $\xi^n \to \xi^*$ narrowly in $\mathcal{P}(\Gamma)$ and $M^n \to m^* := e_{(\cdot)}\sharp\xi^*$ in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$.*
In addition, if **(H4)** *holds and for every $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and m_0^*-a.e. $x \in \mathbb{R}^d$ problem $(OC_{x,m})$ admits a unique solution, then the whole sequence $(\xi^n)_{n\in\mathbb{N}}$ converges narrowly towards the unique solution to Problem 2.*

3.5 Numerical results

In this section we implement our numerical method in two examples. For computational simplicity, we consider here one-dimensional problems, i.e. $d = 1$, and dynamics (3.3.1) having the form $\dot{\gamma} = \alpha$. We refer the reader to [80, Example 2] for the implementation of the scheme in a two-dimensional example, where a typical agent controls its acceleration and the cost functional satisfies **(H1)**. We focus our attention on cost functionals satisfying **(H1)** and **(H4)**, with $\ell_0(t, a, x)$ having polynomial growth on (a, x), and f and g being given by

$$\begin{aligned} f(t,x,\mu) &= \theta_1(\rho_\sigma \star \mu)(x) \quad \text{for all } (t,x,\mu) \in [0,T] \times \mathbb{R} \times \mathcal{P}_1(\mathbb{R}), \\ g(x,\mu) &= g_0(x) + \theta_2(\rho_\sigma \star \mu)(x) \quad \text{for all } (x,\mu) \in \mathbb{R} \times \mathcal{P}_1(\mathbb{R}), \end{aligned} \tag{3.5.1}$$

where θ_1, $\theta_2 \in [0, \infty)$, $\sigma \in (0, \infty)$, $g_0 \colon \mathbb{R} \to \mathbb{R}$ satisfies **(H1)**, and

$$\rho_\sigma(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \quad \text{for all } x \in \mathbb{R}. \tag{3.5.2}$$

Notice that the convolution terms in f and g, which model the aversion of a typical player to crowded areas, satisfy the monotonicity condition in **(H4)**(ii).
Let $(\Delta t, \Delta x) \in (0, \infty)^2$ and $\varepsilon > 0$. Under the assumptions above, the finite MFG Problem **??** associated with these parameters admits a unique solution $M^* \in \mathcal{M}$. In order to approximate M^*, we consider the *fictitious play* sequence

$$\overline{\mathsf{M}}^0 \in \mathcal{M} \ \text{ arbitrary}, \quad (\forall\, n \geq 1) \quad \mathsf{M}^{n+1} = \mathbf{br}(\overline{\mathsf{M}}^n), \quad \overline{\mathsf{M}}^{n+1} = \frac{n}{n+1}\overline{\mathsf{M}}^n + \frac{1}{n+1}\mathsf{M}^{n+1},$$

which, by [94, Theorem 3.2], satisfies $(\mathsf{M}^n, \overline{\mathsf{M}}^n) \underset{n\to\infty}{\longrightarrow} (M^*, M^*)$. Notice that to apply that result, the addition of an entropy term in (3.4.1) (or more generally, a strictly convex term) plays an important role to ensure that the optimization problem defining $V_k^M(x)$ in (3.4.1) admits a unique solution.
In the tests below, setting

$$|\mathbf{br}(\overline{\mathsf{M}}^n) - \overline{\mathsf{M}}^n|_{L^1} := \frac{1}{N_t + 1} \sum_{k=0}^{N_t} \sum_{x \in \mathcal{S}_k} |\mathbf{br}(\overline{\mathsf{M}}^n)_k(x) - \overline{\mathsf{M}}_k^n(x)|$$

and given a tolerance parameter $\delta > 0$, we implement the following fictitious play algorithm:

Algorithm 3 Fictitious play for deterministic MFG

1: **Data:** $\overline{\mathsf{M}}^0 \in \mathcal{M},\ \delta > 0$
2: $e \leftarrow \delta + 1$
3: $n \leftarrow 1$
4: $\overline{\mathsf{M}}^1 \leftarrow \mathsf{M}^0$
5: **while** $e > \delta$ **do**
6: $\quad \mathsf{M}^{n+1} \leftarrow \mathbf{br}(\overline{\mathsf{M}}^n)$
7: $\quad e \leftarrow \|\mathsf{M}^{n+1} - \overline{\mathsf{M}^n}\|_{L^1}$
8: $\quad \overline{\mathsf{M}}^{n+1} \leftarrow \frac{n}{n+1}\overline{\mathsf{M}}^n + \frac{1}{n+1}\mathsf{M}^{n+1}$
9: **end while**
10: **return** $\overline{\mathsf{M}}^{n+1}$

In both examples below, we consider a time horizon $T = 1$, $\sigma = 0.07$, and

$$\Delta t = 1/30, \quad \Delta x = 1/150, \quad \text{and} \quad \varepsilon = 0.002.$$

Given an initial distribution $m_0^* \in \mathcal{P}_1(\mathbb{R})$, we initialize the fictitious play algorithm by defining $\mathsf{M}^0 \in \mathcal{M}$ with constant time marginals given by $\mathsf{M}_k^0 = M_0$, for $k = 1, \dots, 30$, where M_0 is obtained by discretizing the initial distribution m_0^* according to (3.4.3). As it was mentioned above, the algorithm converges for an arbitrary initial condition $\mathsf{M}^0 \in \mathcal{M}$. However, since the term M^0/n is involved in the computation of $(\overline{\mathsf{M}}^n, \mathsf{M}^{n+1})$ the convergence of the algorithm could be slow. In our simulations we have observed an important acceleration of the method by updating the initial condition after some tolerance is achieved. More precisely, given a tolerance parameter $\delta > 0$, we use the resulting approximated equilibrium $\overline{\mathsf{M}}^{n-1}$ as the initial distribution for a subsequent application of the algorithm with a smaller tolerance parameter. In our tests, we update the initial condition twice, taking the tolerance parameters $\delta_1 = 0.1$ and $\delta_2 = 0.01$. We stop the algorithm when the tolerance $\delta_3 = 0.001$ is reached.

3.5.1 Example 1

We consider an absolutely continuous initial distribution $m_0^* \in \mathcal{P}_1(\mathbb{R})$ given by

$$\mathrm{d}m_0^*(x) = \mathbb{I}_{[-1,1]}(x)\frac{e^{-x^2/0.04}}{\int_{-1}^{1} e^{-y^2/0.04}\mathrm{d}y}\mathrm{d}x \quad \text{for all } x \in \mathbb{R},$$

where $\mathbb{I}_{[-1,1]}(x) = 1$ if $x \in [-1,1]$ and $\mathbb{I}_{[-1,1]}(x) = 0$, otherwise. Given ζ_1, $\zeta_2 \in \mathbb{R}$, we define

$$\ell(t,a,x,\mu) = \frac{|a|^4}{4} + \zeta_1|x-0.4|^2|x+0.7|^2 + \theta_1 f(x,\mu) \quad \text{and} \quad g(x,\mu) = \zeta_2|x-0.4|^2|x+0.7|^2 + \theta_2 f(x,\mu).$$

Notice that the functions ℓ and g satisfy **(H1)** for $p = 4$. We run our algorithm for different values of $(\zeta_1, \zeta_2, \theta_1, \theta_2)$. In Figure 3.1 we show the returned distributions for the smallest tolerance parameter δ_3. In Table 3.1, we provide the number of iterations needed for attaining the tolerances δ_1, δ_2, and δ_3.

$(\zeta_1,\zeta_2,\theta_1,\theta_2)$	$\delta_1 = 0.1$	$\delta_2 = 0.01$	$\delta_3 = 0.001$
$(1,1,1,1)$	14	10	9
$(1,1,1,5)$	15	10	17
$(5,1,1,1)$	18	11	9
$(1,0,1,0)$	8	7	6

Table 3.1: Number of iterations to obtain the desired accuracies.

In this example, the initial distribution is concentrated around $x = 0$. The cost functional penalizes the distance to the points -0.7 and 0.4, as well as large values of the speed, and incites a typical player to avoid crowded regions. Since 0.4 is closer to the origin, we notice that most of the agents tend to concentrate around that point. This effect is most pronounced in Figure 3.2 (C), where the impact of this term is most significant. Furthermore, as the congestion term becomes more important, we see how the agents tend to separate from each other. We can observe the role played by this term in the final cost by comparing Figure 3.2 (A) and (B). To see the effect of the final cost, compare figures (A) and (D), where the latter corresponds to $g \equiv 0$.

3.5.2 Example 2

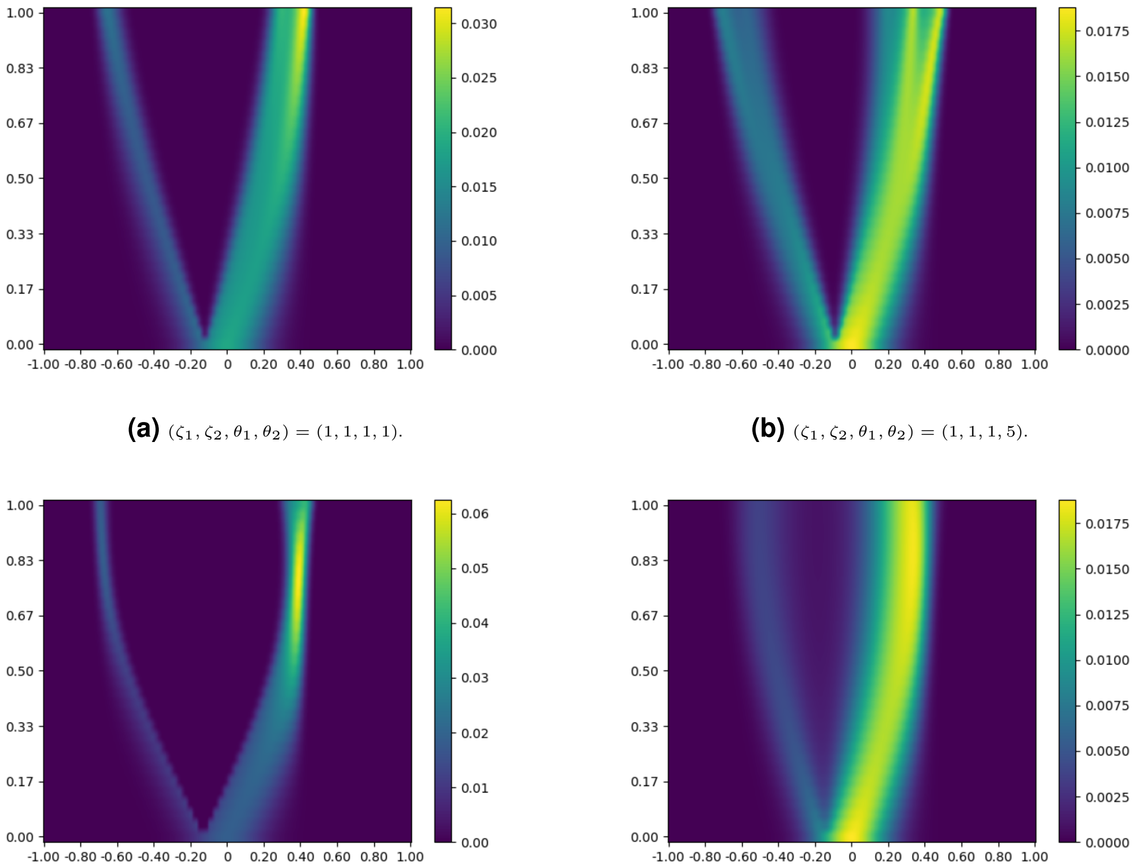
The initial distribution $m_0^* \in \mathcal{P}_1(\mathbb{R})$ is given by

$$\mathrm{d}m_0^*(x) = \mathbb{I}_{[-1,1]}(x)\frac{e^{-(x-0.2)^2/0.01} + e^{-(x+0.2)^2/0.01}}{\int_{-1}^{1}(e^{-(y-0.2)^2/0.01} + e^{-(y+0.2)^2/0.01})\mathrm{d}y}\mathrm{d}x \quad \text{for all } x \in \mathbb{R}.$$

Given ζ_1, $\zeta_2 \in \mathbb{R}$, we define

$$\ell(t,a,x,\mu) = \frac{|a|^4}{4} + \zeta_1|x-0.6|^2|x+0.2|^2 + \theta_1 f(x,\mu) \quad \text{and} \quad g(x,\mu) = \zeta_2|x-0.6|^2|x+0.2|^2 + \theta_2 f(x,\mu).$$

We consider the same parameters as those in Example 1 and we display in Figure 3.2 the distributions obtained for the smallest tolerance $\delta_3 = 0.001$. In Table 3.2, we show



(a) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 1)$. **(b)** $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 5)$.

(c) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (5, 1, 1, 1)$. **(d)** $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 0, 1, 0)$.

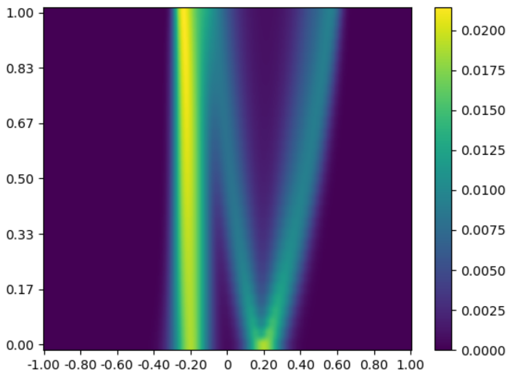
Figure 3.1: Approximate equilibria for the tests in Example 1

the number of iterations required to reach the tolerances δ_1, δ_2, and δ_3.

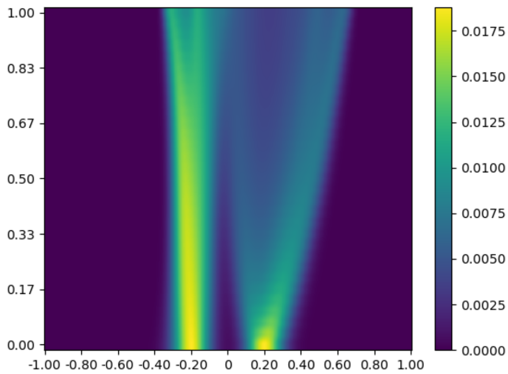
$(\zeta_1, \zeta_2, \theta_1, \theta_2)$	$\delta_1 = 0.1$	$\delta_2 = 0.01$	$\delta_3 = 0.001$
$(1, 1, 1, 1)$	7	6	6
$(1, 1, 1, 5)$	10	12	9
$(5, 1, 1, 1)$	12	10	9
$(1, 0, 1, 0)$	4	6	6

Table 3.2: Number of iterations to obtain the desired accuracies.

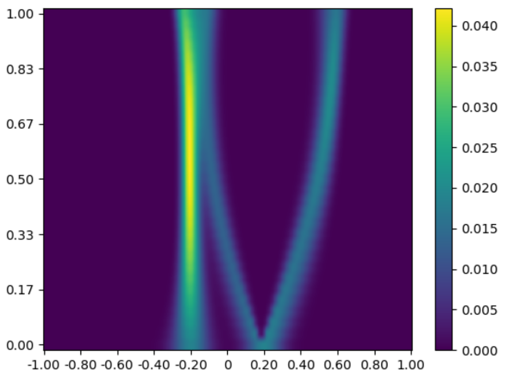
In this example, we start with a distribution symmetrically concentrated around the points -0.2 and 0.2. As in Example 1, the cost functional penalizes large values of the speed and encourages avoiding crowded areas. In addition, it penalizes the distance to the points -0.2 and 0.6. Although these two points are symmetric with respect to 0.2, we see that, in order to avoid crowded regions, most of the agents that are concentrated around 0.2 at $t = 0$ tend to go towards 0.6 instead of -0.2, see for instance Figure 3.2



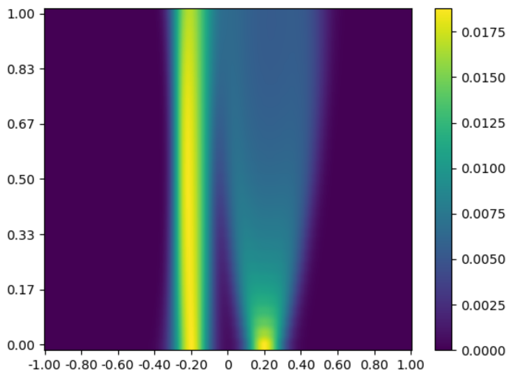
(a) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 1)$.



(b) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 5)$.



(c) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (5, 1, 1, 1)$.



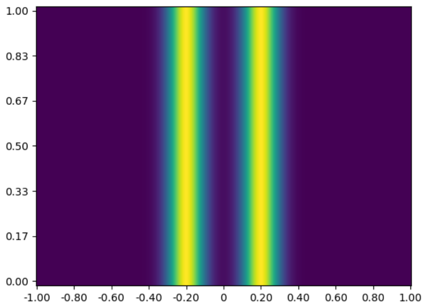
(d) $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 0, 1, 0)$.

Figure 3.2: Approximate equilibria for the tests in Example 2

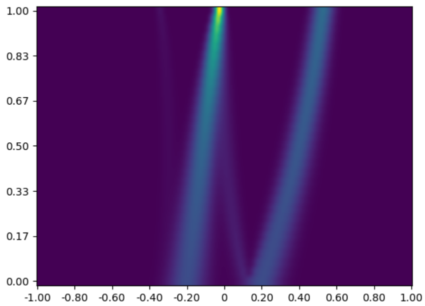
(C). Once again, considering the same cases as in Example 1, we can appreciate the impact of the congestion terms in the final distributions of the agents.
For a better understanding of the fictitious play method, we end this section by displaying in Figure 3.3 the first iterations of the algorithm when $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 5)$. The final distribution in this case is shown in the top right corner of Figure 3.2.

Acknowledgments

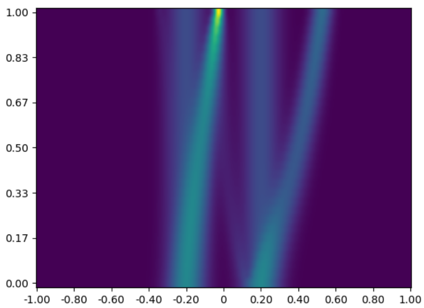
F. J. Silva and A. Zorkot where partially supported by l'Agence Nationale de la Recherche (ANR), project ANR-22-CE40-0010, and by KAUST through the subaward agreements OSR-2017-CRG6-3452.04 and ORA-2021-CRG10-4674.6. For the purpose of open access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.



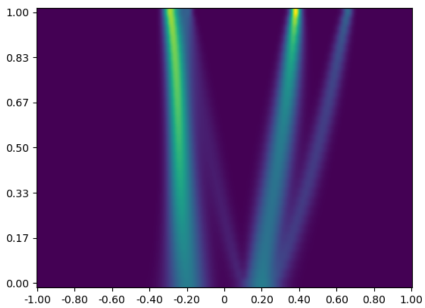
(a) $\overline{\mathsf{M}}^1 = \mathsf{M}^0$.



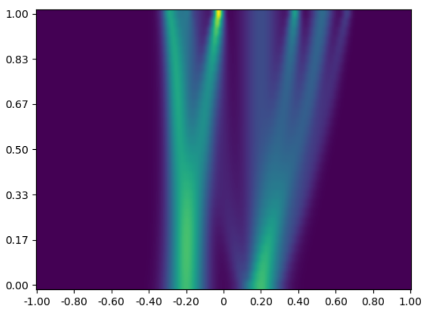
(b) $\mathbf{br}(\overline{\mathsf{M}}^1)$.



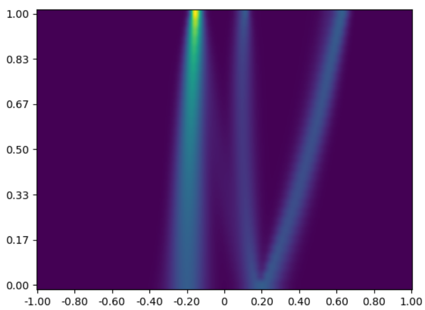
(c) $\overline{\mathsf{M}}^2$.



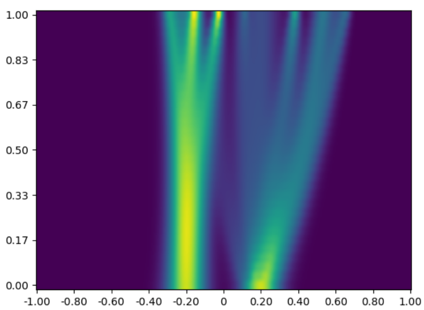
(d) $\mathbf{br}(\overline{\mathsf{M}}^2)$.



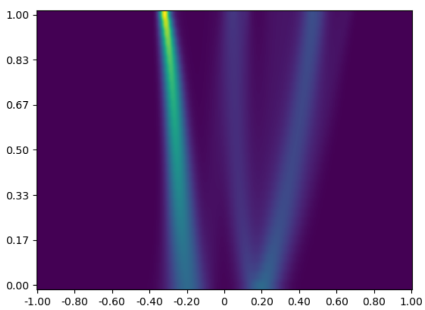
(e) $\overline{\mathsf{M}}^3$.



(f) $\mathbf{br}(\overline{\mathsf{M}}^3)$.



(g) $\overline{\mathsf{M}}^4$.



(h) $\mathbf{br}(\overline{\mathsf{M}}^4)$.

Figure 3.3: First iterations of the algorithm for $(\zeta_1, \zeta_2, \theta_1, \theta_2) = (1, 1, 1, 5)$.

4

Newton iterations for second order mean field game systems

In the preceding chapters, we have explored two numerical methodologies designed to solve first order mean field game problems. This chapter focuses on solving second order MFG systems. We present two numerical methods to solve a system derived after applying Newton iterations to the continuous MFG system: a semi-Lagrangian scheme and a finite difference scheme. We conduct a comparative analysis between these two approaches and other schemes found in the literature through some numerical examples.
This chapter is a work in progress with Elisabetta Carlini and Francisco J. Silva.

Sommaire

4.1 Introduction

We consider the following second order MFG system with local coupling

$$\begin{cases} -\partial_t u - \nu\Delta u + H(x, Du) = F(m(t,x)) & \text{in } [0,T]\times\mathbb{T}^d \\ \partial_t m - \nu\Delta m - \operatorname{div}(mH_p(x,Du)) = 0 & \text{in } [0,T]\times\mathbb{T}^d \\ m(0,x) = m_0(x),\ u(T,x) = G(x) & \text{in } \mathbb{T}^d\ . \end{cases} \tag{4.1.1}$$

Here, $T > 0$ is the finite time horizon, $\nu > 0$ describes the intensity of the noise each agent is submitted to, $F\colon \mathbb{R}^+ \to \mathbb{R}$ is the coupling depending locally on the density, H is a Hamiltonian, convex with respect to its second component, $m_0\colon \mathbb{T}^d \to \mathbb{R}$ is a probability density, $G\colon \mathbb{T}^d \to \mathbb{R}$ is a given function, and $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ denotes the d-dimensional torus, simplifying considerations about boundary conditions.
As discussed in Chapter 1, several numerical methods have emerged to discretize (4.1.1), and various strategies to solve the resulting non linear discrete system have been discussed in the literature. For instance, the authors in [4] have introduced a finite difference method to approximate system (4.1.1). This method retains several favorable properties of the MFG system, and a convergence result has been established in [1] and [9] . However, the resulting scheme yields a high dimensional nonlinear discrete system to be solved. To tackle this difficulty, the authors in [1], [13] employed the Newton method for its numerical solution.
In this chapter, we consider Newton's method in infinite dimension applied to the continuous MFG system (4.1.1), obtaining, at each iteration, a forward-backward linear parabolic system that we solve with suitable numerical methods.
Infinite dimensional Newton iterates for (4.1.1) (see also [31] for the case of a stationary MFG system) have been introduced in [44], where the authors consider a more general case including a non-separable Hamiltonians $H(x, m, p)$. Their key achievement is the proof of a quadratic convergence rate for the continuous Newton iterations towards solutions to (4.1.1). By slightly adjusting the assumptions in the Hamiltonian, their results also cover the case we will consider consisting on a separable Hamiltonian and local coupling.

To apply Newton iterations to system (4.1.1), we define first the map

$$\mathcal{F} : (u,m) \to \begin{pmatrix} -\partial_t u - \Delta u + H(Du) - F(m) \\ \partial_t m - \Delta m - \operatorname{div}(mH_p(Du)) \\ u(T,\cdot) - G(\cdot) \\ m(0,\cdot) - m_0(\cdot) \end{pmatrix},$$

and note that (4.1.1) is equivalent to

$$\mathcal{F}(u,m) = 0.$$

Thus, assuming that the data is sufficiently smooth, the corresponding Newton's iterations read

$$J\mathcal{F}(u^{n-1}, m^{n-1})(u^n - u^{n-1}, m^n - m^{n-1}) = -\mathcal{F}(u^{n-1}, m^{n-1}), \tag{4.1.2}$$

where $J\mathcal{F}$ is the Jacobian of $\mathcal{F}$ is given by

$$J\mathcal{F}(u,m)(v,\rho) = \begin{pmatrix} -\partial_t v - \Delta v + H_p(Du)Dv & -F'(m)\rho \\ -\operatorname{div}(mH_{pp}(Du)Dv) & \partial_t \rho - \Delta\rho - \operatorname{div}(H_p(Du)\rho) \\ v(T,\cdot) & 0 \\ 0 & \rho(0,\cdot) \end{pmatrix} \tag{4.1.3}$$

for all Hölder continuous functions (v,ρ) in Q.

Hence, setting $q^n = H_p(\cdot, Du^{n-1})$ and using (4.1.3), after simplification, identity (4.1.2) reads

$$\begin{cases} -\partial_t u^n - \nu\Delta u^n + q^n Du^n = q^n Du^{n-1} - H(x, Du^{n-1}) + F(m^{n-1}) \\ \qquad\qquad + F'(m^{n-1})(m^n - m^{n-1}) & \text{in } [0,T]\times\mathbb{T}^d, \\ \partial_t m^n - \nu\Delta m^n - \operatorname{div}(m^n q^n) = \operatorname{div}(m^{n-1}H_{pp}(x, Du^{n-1})(Du^n - Du^{n-1})) & \text{in } [0,T]\times\mathbb{T}^d, \\ m^n(x,0) = m_0(x), \quad u^n(x,T) = G(x) & \text{in } \mathbb{T}^d. \end{cases} \tag{4.1.4}$$

The contribution of this chapter is to introduce and compare numerical methods to solve the linearized system (4.1.4). For this purpose, two methodologies are considered.The first approach entails a semi-Lagrangian scheme, readily derivable for linear parabolic equations. A comparative analysis is conducted against the non-linear semi-Lagrangian scheme for system (4.1.1) proposed in [58] and solved via (damped) fix point iterations. The second method involves an upwind finite difference scheme, which, as observed in several numerical tests, has a simpler structure and analogous performance to the Newton scheme proposed of [4].

This chapter is structured as follows. In Section 4.2 we introduce some notations, assumptions, and recall some preliminary results from [44]. In Section 4.3, we discretize system (4.1.4) using a semi-Lagrangian scheme and establish the well-posedness of the resulting discrete system. In Section 4.4, we demonstrate a similar well-posedness result for the upwind finite difference scheme. Finally, in Section 4.5, we provide some numerical simulations for both schemes and compare the results with those obtained by using the schemes in [58] and [4].

4.2 Preliminaries and assumptions

In this section, we briefly recall the framework of [44]. First we set $Q = [0,T] \times \mathbb{T}^d$. We denote by $\mathcal{P}(\mathbb{T}^d)$ the set of Borel probability measures on $\mathbb{T}^d$. For $\alpha \in [0,1]$, the space $C^{0,1}(Q)$ is the set of continuous functions in Q with continuous derivative in the space variable, endowed with the norm

$$\|u\|_{C^{0,1}(Q)} = \|u\|_{C^0(Q)} + \|Du\|_{C^0(Q)}.$$

Let us recall the definition of parabolic Hölder spaces on the torus (see [109] for a more comprehensive discussion). For $\alpha \in]0,1[$, we denote

$$[u]_{C^{\frac{\alpha}{2},\alpha}(Q)} = \sup_{(x_1,t_1),(x_2,t_2)\in Q} \frac{|u(t_1,x_1) - u(t_2,x_2)|}{(|t_1-t_2| + d(x_1,x_2)^2)^{\frac{\alpha}{2}}},$$

where $d(x,y)$ stands for the geodesic distance from x to y in $\mathbb{T}^d$. The parabolic Hölder space $C^{\frac{\alpha}{2},\alpha}(Q)$ is the space of functions $u\colon Q \to \mathbb{R}$ such that $[u]_{C^{\frac{\alpha}{2},\alpha}(Q)} < \infty$, and it is endowed with the norm

$$\|u\|_{C^{\frac{\alpha}{2},\alpha}(Q)} = \|u\|_{C^0(Q)} + [u]_{C^{\frac{\alpha}{2},\alpha}(Q)}.$$

Finally, the spaces $C^{\frac{1+\alpha}{2},1+\alpha}(Q)$ and $C^{1+\frac{\alpha}{2},2+\alpha}(Q)$ denote the spaces of Hölder continuous functions in Q endowed with the norms:

$$\|u\|_{C^{\frac{1+\alpha}{2},1+\alpha}(Q)} = \|u\|_{C^0(Q)} + \sum_{i=1}^{d} \|\partial_{x_i} u\|_{C^{\frac{\alpha}{2},\alpha}(Q)} + \sup_{(t_1,x_1),(t_2,x_2)\in Q} \frac{|u(t_1,x_1)-u(t_2,x_2)|}{|t_1-t_2|^{\frac{1+\alpha}{2}}},$$

$$\|u\|_{C^{1+\frac{\alpha}{2},2+\alpha}(Q)} = \|u\|_{C^0(Q)} + \sum_{i=1}^{d} \left\|\frac{\partial u}{\partial x_i}\right\|_{C^{\frac{1+\alpha}{2},1+\alpha}(Q)} + \left\|\frac{\partial u}{\partial t}\right\|_{C^{\alpha/2,\alpha}(Q)}.$$

The following assumptions hold throughout this chapter:

(H1) $H\colon \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous, twice differentiable in p and there exist positive constants $c_0, \bar{c}_0, C_0$ and $\overline{C}_0$ such that for every $(x,p) \in \mathbb{T}^d \times \mathbb{R}^d$,

$$c_0 I_d \le H_{pp}(x,p) \le C_0 I_d, \quad |H_{px}(x,p)| \le \bar{c}_0(|p|+1), \quad |H_{xx}(x,p)| \le \overline{C}_0(|p|^2+1).$$

(H2) $F\colon \mathbb{R}^+ \to \mathbb{R}$ is bounded and of class C^2. Moreover, there exist positive constants c_1, C_1 and $\overline{C}_1$ such that

$$c_1 m \le F'(m) \le C_1 m, \quad |F''(m)| \le \overline{C}_1 \quad \text{for all } m > 0.$$

(H3) There exists $\alpha \in]0,1[$ and $\eta > 0$, such that

$$m_0 \in C^{2+\alpha}(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d) \quad \text{and } m_0 \ge \eta > 0 \quad \text{for all } x \in \mathbb{T}^d$$
$$G \in C^{2+\alpha}(\mathbb{T}^d).$$

We will consider classical solutions to the MFG system (4.1.1). Recall that a classical solution to (4.1.1) is a couple (u,m) such that u and m belong to $C^{1,2}(Q)$ and (4.1.1) is satisfied for all $(t,x) \in Q$.

Proposition 4.2.1 *Under assumptions **(H1)-(H3)**, the MFG system* (4.1.1) *has a unique classical solution.*

We refer to [44][Proposition 2.4] for the proof of the above proposition. Regarding the Newton iterations system (4.1.4).

Proposition 4.2.2 *Assume that **(H1)-(H3)** hold, then there exists a unique solution* $(u^n, m^n) \in C^{2+\alpha,1+\frac{\alpha}{2}}(Q) \times C^{2+\alpha,1+\frac{\alpha}{2}}(Q)$ *to* (4.1.4)*. Moreover, let* (u,m) *be the unique solution to* (4.1.1)*, there exists a constant* $\eta > 0$ *such that if* $\|u^0-u\|_{C^{0,1}} + \|m^0-m\|_{C^0} \le \eta$*, then* $\|u^n-u\|_{C^{0,1}} + \|m^n-m\|_{C^0} \to 0$ *with a quadratic rate of convergence.*

The existence result follows from [44][Proposition 3.1] and the convergence result follows from [44][Theorem 4.6].

In what follows, to simplify the discussion, we consider the following reference case

$$H(x,p) = \frac{|p|^2}{2} - V(x) \quad \text{for } (x,p) \in \mathbb{T}^d \times \mathbb{R}^d, \tag{4.2.1}$$

where V is a given bounded potential. Additionally, for the sake of simplicity, we consider the case when $d = 2$. Let us note that the two proposed schemes work also for more general Hamiltonians satisfying **(H1)**.

4.3 A semi-Lagrangian scheme

In this section, we discretize the iterative system (4.1.4) by means of a semi-Lagrangian scheme in the 2 dimensional state-space and we prove the well-posedness of the discrete system.
We refer to [58] for the early work on approximating the second-order MFG systems using a semi-Lagrangian scheme, and to [34] for a semi-Lagrangian scheme applied to parabolic equations.

4.3.1 Notations and definitions

Given two positive integers N_t and N_h, we define $\Delta t = \frac{T}{N_t}$ as the time step, $h = \frac{1}{N_h}$ as the space step, and the sets $\mathcal{I}_{\Delta t} := \{0, \dots, N_t\}$, $\mathcal{I}^*_{\Delta t} := \mathcal{I}_{\Delta t} \setminus \{N_t\}$, and $\mathcal{I}_h := \{0, \dots, N_h - 1\}$. We define the discrete time grid $\mathcal{G}_{\Delta t} := \{t_k = k\Delta t \mid k \in \mathcal{I}_{\Delta t}\}$ and discrete space grid $\mathcal{G}_h := \{x_{i,j} = (ih, jh) \mid i, j \in \mathcal{I}_h\}$. We denote by $B(\mathcal{G}_h)$ and $B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$ the two sets of functions defined in $\mathcal{G}_h$ and $\mathcal{G}_{\Delta t} \times \mathcal{G}_h$ respectively.

The objective is to approximate $u^n(t_k, x_{i,j})$ and $m^n(t_k, x_{i,j})$, solution to (4.1.4), respectively by $u^{n,k}_{[i,j]}$ and $m^{n,k}_{[i,j]}$ for all $k \in \mathcal{I}_{\Delta t}$ and $i, j \in \mathcal{I}_h$, by solving the discrete version of (4.1.4) resulting from a semi-Lagrangian approximation, where the index operator $[\cdot,\cdot] = \{(\cdot + N_h, \cdot + N_h)\text{mod } N_h\}$ accounts for the periodic boundary condition.
For notational simplicity, we denote $(u^{n,k}_{i,j}, m^{n,k}_{i,j}) = (u^{n,k}_{[i,j]}, m^{n,k}_{[i,j]})$.

Given a grid function $v\colon \mathcal{G}_h \to \mathbb{R}$, we introduce the first order central differences operators

$$\begin{aligned}(D_1 v)_{i,j} &= \frac{v_{i+1,j} - v_{i-1,j}}{2h} \quad \text{for } i,j \in \mathcal{I}_h,\\ (D_2 v)_{i,j} &= \frac{v_{i,j+1} - v_{i,j-1}}{2h} \quad \text{for } i,j \in \mathcal{I}_h,\end{aligned} \tag{4.3.1}$$

and define the operator D_h as

$$(D_h v)_{i,j} = ((D_1 v)_{i,j}, (D_2 v)_{i,j}) \quad \text{for } i,j \in \mathcal{I}_h. \tag{4.3.2}$$

Let $\beta_i^{(1)}\colon \mathbb{T} \to \mathbb{R}$ be the $\mathbb{P}_1$ piecewise linear basis function associated with the ith nodal point given by:

$$\beta_i^{(1)}(x) = \begin{cases} \frac{x - x_{i-1}}{h} & \text{if } x \in [x_{i-1}, x_i],\\ \frac{x_{i+1} - x}{h} & \text{if } x \in [x_i, x_{i+1}],\\ 0 & \text{otherwise.}\end{cases} \tag{4.3.3}$$

Notice that

$$\beta_i^1 \geq 0 \quad \text{for all } i \in \mathcal{I}_h, \quad \sum_{i \in \mathcal{I}_h} \beta_i^1(x) = 1 \quad \text{for all } x \in \mathbb{T}. \tag{4.3.4}$$

Next, for every $i,j \in \mathcal{I}_h$, we define $\beta_{i,j}\colon \mathbb{T}^2 \to \mathbb{R}$

$$\beta_{i,j}(x) = \beta_i^{(1)}(x_1)\beta_j^{(1)}(x_2) \quad \text{for all } x = (x_1, x_2) \in \mathbb{T}^2.$$

Given $\phi\colon \mathcal{G}_h \to \mathbb{R}$, we define its picewise linear interpolant as

$$I[\phi](x) = \sum_{i,j \in \mathcal{I}_h} \beta_{i,j}(x)\phi_{i,j} \quad \text{for all } x \in \mathbb{T}^2,$$

where $\phi_{i,j} = \phi(x_{i,j})$. For every $\varphi\colon \mathbb{T}^2 \to \mathbb{R}$ denote by $\varphi|_{\mathcal{G}_h}$ its restriction to $\mathcal{G}_h$. If φ is of class C^2 and has bounded second order derivatives, it follows from [129, Remark 3.4.2] that

$$\|\varphi(\cdot) - I[\varphi|_{\mathcal{G}_h}](\cdot)\|_\infty \leq C_\varphi h^2, \tag{4.3.5}$$

where $C_\varphi > 0$ depends only on φ.

4.3.2 Semi-Lagrangian scheme for the backward equation

With the notations of the previous subsection, we are ready to apply a semi-Lagrangian scheme to approximate the iterative system (4.1.4). We start with the backward equation with Hamiltonian given by (4.2.1), which can be written as follows:

$$\begin{cases} -\partial_t u^n - \frac{\sigma^2}{2}\Delta u^n + q^n D u^n - L^n(t,x) = 0 & \text{in } [0,T]\times\mathbb{T}^2, \\ u^n(x,T) = G(x) & \text{in } \mathbb{T}^2, \end{cases} \tag{4.3.6}$$

where $\sigma = \sqrt{2\nu}$, $q^n(t,x) = Du^{n-1}(t,x)$ and

$$L^n(t,x) = \frac{|q^n(t,x)|^2}{2} + F(m^{n-1}(t,x)) + F'(m^{n-1}(t,x))(m^n(t,x) - m^{n-1}(t,x)) - V(x).$$

By the Feynman-Kac formula (see e.g [134]), under assumptions **(H1)-(H2)**, the solution u^n to (4.3.6), admits the following representation: for $(t,x) \in [0,T]\times\mathbb{T}^2$

$$u^n(t,x) = \mathbb{E}\Big[\int_t^T L^n(s, X^{t,x}(s))\mathrm{d}s + G(X^{t,x}(T))\Big], \tag{4.3.7}$$

where $X^{t,x}$ denotes characteristics solving

$$\begin{cases} dX(s) = q^n(s, X(s)) + \sigma dW(s) & \text{for } s \in (t,T), \\ \quad X(t) = x. \end{cases} \tag{4.3.8}$$

We explain now how to construct a SL approximation using the technique shown in [76]. Notice that (4.3.7) imply that for every $k \in \mathcal{I}^*_{\Delta t}$, we have

$$u^n(t_k, x) = \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} L^n(s, X^{t_k,x}(s))\mathrm{d}s + u^n(t_{k+1}, X^{t_k,x}(\Delta t))\Big] \tag{4.3.9}$$

Denote by $Q^{n,k}$ the approximation of the drift term q^n at time t_k, defined as

$$Q^{n,k}_{i,j} := (Q^{n,k}_1, Q^{n,k}_2)_{i,j} = (D_h u^{n-1,k})_{i,j} \quad \text{for all } k \in \mathcal{I}_{\Delta t},\ i,j \in \mathcal{I}_h. \tag{4.3.10}$$

Then, we approximate the expectation in (4.3.9) (see e.g [104]) as

$$\mathbb{E}\left[u^n(t_k, X^{t_k,x_{i,j}}(\Delta t))\right] = \frac{1}{4}\sum_{\ell=1}^{4} u^n(t_k, y^\ell_{i,j}(Q^{n,k})) + O((\Delta t)^2), \tag{4.3.11}$$

where

$$y^\ell_{i,j}(Q^{n,k}) = (x_{i,j} + \Delta t Q^{n,k}_{i,j} + \sqrt{2\Delta t}\sigma e^\ell)_p \quad \text{for } \ell = 1,\dots,4, \tag{4.3.12}$$

with e^ℓ representing four vectors of $\mathbb{R}^2$ with one component equal to ± 1 and the other null, and

$$(z)_p = (z_1 - \lfloor z_1 \rfloor, z_2 - \lfloor z_2 \rfloor) \quad \text{for all } z = (z_1, z_2) \in \mathbb{R}^2$$

denotes the periodic projection on $\mathbb{T}^2$ of $z \in \mathbb{R}^2$.

Finally, combining (4.3.7), (4.3.9), (4.3.11), and the rectangular formula to approximate $\int_{t_k}^{t_{k+1}} L^n(s, X^{t_k,s}(s))\mathrm{d}s$, we define define the semi-Lagrangian scheme for (4.3.6) in the following way:
Given $Q^n \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)^2$ and $m^n, m^{n-1} \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$, find $u^n \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$ such that

$$\begin{cases} u_{i,j}^{n,k} = S_{k,(i,j)}^n(u^{n,k+1}) & \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i,j \in \mathcal{I}_h, \\ u^{n,N_t} = G(x_{i,j}), \end{cases} \tag{4.3.13}$$

where, for every $f \in B(\mathcal{G}_h)$, $k \in \mathcal{I}_{\Delta t}^*$, and $i,j \in \mathcal{I}_h$,

$$S_{k,(i,j)}^n(f) := \frac{1}{4}\sum_{\ell=1}^{4} I[f](y_{i,j}^\ell(Q^{n,k})) + \Delta t L^n(t_k, x_{i,j}).$$

Proposition 4.3.1 *Assume **(H1)-(H3)** and let u^n be a smooth solution, with bounded derivatives, to* (4.3.6)*. Then, for every $k \in \mathcal{I}_{\Delta t}^*$ and $i,j \in \mathcal{I}_h$, the consistency error of scheme* (4.3.13)*, defined as*

$$T_{\Delta t,h}(t_k, x_{i,j}) = \frac{1}{\Delta t}\left(u^n(t_{k+1}, x_{i,j}) - S_{k,(i,j)}^n(u^n(t_{k+1})\right),$$

satisfies

$$T_{\Delta t,h}(t,x) = O\Big((\Delta t)^2 + \frac{h^2}{\Delta t}\Big) \quad (t,x) \in (0,T) \times \mathbb{T}^2. \tag{4.3.14}$$

Proof. Let $\phi \in C_0^2(\mathbb{T}^2)$, and let us denote by C a positive real number which can depend only on ϕ. From (4.3.12) and Taylor expansion (see e.g [34]) we have

$$\Big|\frac{1}{4}\sum_{\ell=1}^{4}\phi(y_{i,j}^\ell(Q^{n,k})) - \Big(\phi(x_{i,j}) + \Delta t\frac{\sigma^2}{2}\Delta\phi(x_{i,j}) + \Delta t Q_{i,j}^{n,k} D\phi(x_{i,j})\Big)\Big| \leq C(\Delta t)^2 \quad \text{for all } i,j \in \mathcal{I}_h. \tag{4.3.15}$$

Thus, using (4.3.5), (4.3.15) in (4.3.13), and assumptions **(H1)-(H3)** yields (4.3.14). □

For $k \in \mathcal{I}_{\Delta t}$, we denote $U^{n,k}$ and U^{N_t} the vectors of $\mathbb{R}^{(N_h)^2}$ such that

$$U^{n,k}_{i+jN_h} = u^{n,k}_{i,j}, \quad U^{N_t}_{i+jN_h} = G(x_{i,j}) \quad \text{for all } i,j \in \mathcal{I}_h, \tag{4.3.16}$$

and, in the same way $M^{n,k}$ and M^0 are the vectors of $\mathbb{R}^{(N_h)^2}$ such that

$$M^{n,k}_{i+jN_h} = m^{n,k}_{i,j}, \quad M^{0}_{i+jN_h} = m_0(x_{i,j}) \quad \text{for all } i,j \in \mathcal{I}_h. \tag{4.3.17}$$

With the above notation, scheme (4.3.13) can be written in the following way

$$\begin{cases} U^{n,k} = \mathcal{A}(Q^{n,k})U^{n,k+1} + \Delta t \mathcal{W}^k M^{n,k} + \Delta t \mathcal{B}^k & \text{for all } k \in \mathcal{I}^*_{\Delta t}, \\ U^{n,N_t} = U^{N_t}, \end{cases} \tag{4.3.18}$$

where

$$(\mathcal{A}(Q^{n,k}))_{iN_h+j,p+N_hq} = \frac{1}{2d}\sum_{\ell=1}^{4} \beta_{p,q}(y^{\ell}_{i,j}(Q^{n,k})), \quad \text{for all } i,j,p,q \in \mathcal{I}_h \tag{4.3.19}$$

$$(\mathcal{W}^k M^{n,k})_{i+N_hj} = F'((M^{n-1,k})_{i+N_hj})(M^{n,k})_{i+N_hj} \quad \text{for all } i,j \in \mathcal{I}_h, \tag{4.3.20}$$

$$\begin{aligned}(\mathcal{B}^k)_{i+N_hj} = &\frac{|(Q^{n,k})_{i,j}|^2}{2} - V(x_{i,j}) \\ &+ F((M^{n-1,k})_{i+N_hj}) - F'((M^{n-1,k})_{i+N_hj})(M^{n-1,k})_{i+N_hj} \quad \text{for all } i,j \in \mathcal{I}_h.\end{aligned} \tag{4.3.21}$$

4.3.3 A semi-Lagrangian scheme for the forward equation

Let us now consider the second equation in system (4.1.4), expressed for the Hamiltonian (4.2.1) as

$$\begin{cases} \partial_t m^n - \frac{\sigma^2}{2}\Delta m^n - \operatorname{div}(m^n q^n) = \operatorname{div}(m^{n-1}(t,x)(Du^n(t,x) - Du^{n-1}(t,x))) & \text{in } [0,T]\times\mathbb{T}^2, \\ m^n(0,x) = m_0(x) & \text{in } \mathbb{T}^2. \end{cases} \tag{4.3.22}$$

Given $U \in B(\mathcal{G}_h)$ and $Q = (Q_1, Q_2) \in B(\mathcal{G}_h)^2$, we define the discrete divergence operator

$$(\operatorname{div}_h(UQ))_{i,j} = \frac{1}{2h}\Big(U_{i+1,j}(Q_1)_{i+1,j} - U_{i-1,j}(Q_1)_{i-1,j} + U_{i,j+1}(Q_2)_{i,j+1} - U_{i,j-1}(Q_2)_{i,j-1}\Big).$$

Let us introduce a semi-Lagrangian scheme to approximate (4.3.22).
Given $Q^n, D_hu^{n-1}, D_hu^n \in B(\mathcal{G}_{\Delta t}\times\mathcal{G}_h)^2$, and $m^{n-1} \in B(\mathcal{G}_{\Delta t}\times\mathcal{G}_h)$, find

$m^n \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$ such that

$$\begin{cases} m_{i,j}^{n,k+1} = (S_{k,(i,j)}^n)^*(m^{n,k}) \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i,j \in \mathcal{I}_h, \\ \quad m_{i,j}^{n,0} = m_0(x_{i,j}), \end{cases} \tag{4.3.23}$$

where, for a given function $f \in B(\mathcal{G}_h)$, $k \in \mathcal{I}_{\Delta t}^*$ and $i, j \in \mathcal{I}_h$

$$(S_{k,(i,j)}^n)^*(f) = \frac{1}{4}\sum_{\ell=1}^{4} I^*[f](y_{i,j}^{\ell}(Q^{n,k})) + \Delta t(\mathrm{div}_h(m^{n-1,k+1}(D_h u^{n-1,k+1} - D_h u^{n,k+1})))_{i,j},$$

where, for every $i, j \in \mathcal{I}_h$, $I^*[f](y_{i,j}^{\ell}(Q^{n,k}))$ is the adjoint operator of $f \to I[f](y_{i,j}^{\ell}(Q^{n,k}))$.

As in (4.3.18), scheme (4.3.23) can be written in matrix form as

$$\begin{cases} M^{n,k+1} := \mathcal{A}(Q^{n,k})^* M^{n,k} + \Delta t \mathcal{Z}^{k+1} U^{n,k+1} + \Delta t \mathcal{C}^{k+1} \quad k \in \mathcal{I}_{\Delta t}^*, \\ M^{n,0} = M^0, \end{cases} \tag{4.3.24}$$

where $\mathcal{A}(Q)^*$ denotes the transpose of $\mathcal{A}(Q)$ given by (4.3.19), for every $k \in \mathcal{I}_{\Delta t}^*$, $\mathcal{Z}^k$ is $(N_h)^2 \times (N_h)^2$ matrix such that

$$\begin{aligned} (\mathcal{Z}^k U^k)_{i+N_h j} = {} & (D_h M^{n-1,k})_{i+N_h j} \cdot (D_h U^k)_{i+N_h j} \\ & + (M^{n-1,k})_{i+N_h j} \,\mathrm{div}_h (D_h U^k)_{i+N_h j} \qquad \text{for } i, j \in \mathcal{I}_h, \end{aligned} \tag{4.3.25}$$

and $\mathcal{C}^k$ is the vector in $\mathbb{R}^{(N_h)^2}$ such that

$$\begin{aligned} (\mathcal{C}^k)_{i+N_h j} = & - (D_h (M^{n-1,k})_{i+N_h j} (D_h U^{n-1,k})_{i+N_h j} \\ & - (M^{n-1,k})_{i+N_h j} \mathrm{div}_h \big(D_h U^{n-1,k}\big)_{i+N_h j} \qquad \text{for } i, j \in \mathcal{I}_h. \end{aligned} \tag{4.3.26}$$

4.3.4 The fully discrete Newton system

The final discrete Newton iteration system is given now as the follow. Given $(U^{n-1,k}, M^{n-1,k})$, set $Q^{n,k} = D_h U^{n-1,k}$, and find $(U^{n,k}, M^{n,k})$ satisfying

$$\begin{cases} U^{n,k} = -\mathcal{A}(Q^{n,k}) U^{n,k+1} + \Delta t \mathcal{W}^k M^{n,k} + \Delta t \mathcal{B}^k & k \in \mathcal{I}_{\Delta t}^*, \\ M^{n,k+1} = -\mathcal{A}(Q^{n,k})^* M^{n,k} + \Delta t \mathcal{Z}^{k+1} U^{n,k+1} + \Delta t \mathcal{C}^{k+1} & k \in \mathcal{I}_{\Delta t}^*, \\ U^{n,N_t} = U^{N_t}, \\ M^{n,0} = M^0, \end{cases} \tag{4.3.27}$$

where $\mathcal{A}(Q^{n,k})$, $\mathcal{W}^k M^{n,k}$, $\mathcal{B}^k$, $\mathcal{Z}^{k+1}$, and $\mathcal{C}^{k+1}$ are given by (4.3.19), (4.3.20), (4.3.21), (4.3.25), and (4.3.26) respectively.
To establish the well-posedness of the system (4.3.27), let us represent it in matrix form. For this purpose, we denote by $\bar{U}$ and $\bar{M}$ the vectors of $\mathbb{R}^{(N_t+1)N_h^2}$ such that

$$(\bar{U}^n)_{kN_h^2+iN_h+j} = (U^{n,k})_{i+N_hj}, \quad (\bar{M}^n)_{kN_h^2+iN_h+j} = (M^{n,k})_{i+N_hj} \quad \text{for } k \in \mathcal{I}_{\Delta t},\ i,j \in \mathcal{I}_h. \tag{4.3.28}$$

Next, we define the matrices $\mathbb{A}$ and $\mathbb{W}$ follows

$$\mathbb{A} = \begin{bmatrix} I_{(N_h)^2} & -\mathcal{A}^0 & 0 & \cdots & 0 \\ 0 & I_{(N_h)^2} & -\mathcal{A}^1 & \cdots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots 0 & \ddots & -\mathcal{A}^{N_t-1} \\ 0 & \cdots & \cdots & \cdots 0 & I_{(N_h)^2} \end{bmatrix}, \quad \mathbb{W} = \Delta t \begin{bmatrix} \mathcal{W}^0 & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{W}^{N_t-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

where I_d is the identity matrix of size $d \times d$. Under Assumption **(H2)**, $F' > 0$ if $m > 0$, the diagonal entries of $\mathbb{W}$ are positive if M^n is positive for any $n \in \mathbb{N}$. We also define the matrix $\mathbb{Z}$

$$\mathbb{Z} = \Delta t \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & \mathcal{Z}^1 & 0 & \cdots & 0 \\ 0 & 0 & \mathcal{Z}^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \mathcal{Z}^{N_t} \end{bmatrix},$$

where for every $k \in \mathcal{I}_{\Delta t}^*$, $\mathcal{Z}^k$ is defined by (4.3.25).
Let $\bar{\mathcal{B}} = \Delta t\Big[\mathcal{B}^0, \dots, \mathcal{B}^{N_t-1}, \frac{1}{\Delta t}U^{N_t}\Big]^*$ and $\bar{\mathcal{C}} = \Delta t\Big[\frac{1}{\Delta t}M^0, \mathcal{C}^0, \dots, \mathcal{C}^{N_t-1}\Big]^*$ with $\mathcal{B}^k$ and $\mathcal{C}^k$ given in (4.3.21) and (4.3.26), respectively.

Finally, system (4.3.27) can be written as

$$\begin{bmatrix} \mathbb{A} & -\mathbb{W} \\ -\mathbb{Z} & -\mathbb{A}^* \end{bmatrix} \begin{bmatrix} \bar{U}^n \\ \bar{M}^n \end{bmatrix} = \begin{bmatrix} \bar{\mathcal{B}} \\ \bar{\mathcal{C}} \end{bmatrix}. \tag{4.3.29}$$

Remark 4.3.1 *Note that the blocks of $\mathbb{W}$ would be dense matrices if $F(m(t,x))$ is replaced by a nonlocal operator $f[m(t,\cdot)](x)\colon \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}^d$. In this case, we use the notation $\frac{\delta f}{\delta m}\colon \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \to \mathbb{R}$ for the flat derivative of f (see e.g [50]), and assumption **(H2)** can be replaced by*

(H2') *The measure derivatives* $\frac{\delta f}{\delta m}$ *is Lipschitz continuous and*

$$\int_{\mathbb{T}^d} (f[m](x) - f[m'](x))\, \mathrm{d}(m - m')(x) \geq 0 \quad \textit{for all } m, m' \in \mathcal{P}(\mathbb{T}^d).$$

A typical example is a nonlocal coupling with smoothing effect

$$f(x, m) = \int_{\mathbb{T}^d} \Phi(z, (\rho * m)(z))\rho(x - z)\mathrm{d}z,$$

where $*$ *denotes the usual convolution product in* $\mathbb{T}^d$*, and* $\Phi : \mathbb{T}^2 \to \mathbb{R}$ *is a smooth map which is nondecreasing with respect to the second variable, and* ρ *is a smooth, even function with compact support. In this case, writing* $\Phi = \Phi(x, \theta)$*, we have*

$$\frac{\delta f}{\delta m}(x, m, y) = \int_{\mathbb{T}^d} \frac{\partial \Phi}{\partial \theta}(z, (\rho * m)(z))\rho(x - z)\rho(z - y)\mathrm{d}z.$$

Proposition 4.3.2 *Assume that* $\bar{M}^n > 0$ *for any* $n \in \mathbb{N}$*. Then, there exists a unique solution* $(\bar{U}^n, \bar{M}^n)$ *to the system* (4.3.29)*.*

Proof. Suppose that $\bar{\mathcal{B}} = 0$ and $\bar{\mathcal{C}} = 0$, then (4.3.29) reads as

$$\begin{cases} \mathbb{A}\, \bar{U}^n - \mathbb{W}\bar{M}^n = 0, \\ -\mathbb{Z}\, \bar{U}^n - \mathbb{A}^* \bar{M}^n = 0 \end{cases} \tag{4.3.30}$$

Multiplying the first equation by $\bar{M}^*$ and the second one by $\bar{U}^*$ one gets

$$\begin{cases} (\bar{M}^n)^* \mathbb{A}\, \bar{U}^n - (\bar{M}^n)^* \mathbb{W}\bar{M}^n = 0 \\ -(\bar{U}^n)^* \mathbb{Z}\, \bar{U}^n - (\bar{U}^n)^* \mathbb{A}^* \bar{M} = 0. \end{cases}$$

Adding both equations, we obtain

$$(\bar{M}^n)^* \mathbb{W}\bar{M}^n + (\bar{U}^n)^* \mathbb{Z}\, \bar{U}^n = 0. \tag{4.3.31}$$

Recall that, by assumption **(H2)**, the block $\mathcal{W}^k$ is positive definite for all $k \in \mathcal{I}_{\Delta t}$. Moreover, $\mathcal{Z}^k$ is positive definite for all $k \in \mathcal{I}_{\Delta t}$, since it is the sum of a positive definite matrix and a skew matrix and then, by [101, Remark 1], it is positive definite.
Hence, it follows from (4.3.31) that

$$(\bar{U}^n)_{kN_h^2 + iN_h + j} = 0 \quad \text{for all } k \in \mathcal{I}^*_{\Delta t},\ i, j \in \mathcal{I}_h, \tag{4.3.32}$$

and

$$(\bar{M}^n)_{kN_h^2+iN_h+j} = 0 \quad \text{for all } k \in \mathcal{I}_{\Delta t} \setminus \{0\},\ i,j \in \mathcal{I}_h. \tag{4.3.33}$$

Finally, replacing (4.3.32) and (4.3.33) in (4.3.30), we get $\bar{U}^n = 0$ and $\bar{M}^n = 0$. □

4.4 A finite differences scheme

In this sections, we consider a variation of the finite difference scheme in [4] which does not need a numerical Hamiltonian. We use an upwind scheme to discretize the first equation in (4.1.4), and then, taking the adjoint of the resulting scheme, we construct an approximation of the second equation in (4.1.4).
For this purpose, let us define the discrete time derivative of $v \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$

$$D_t v_{i,j}^k = \frac{v_{i,j}^{k+1} - v_{i,j}^k}{\Delta t} \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*,\ i,j \in \mathcal{I}_h, \tag{4.4.1}$$

and the five point discrete Laplace operator

$$(\Delta_h v^k)_{i,j} = \frac{1}{h^2}(-4v_{i,j}^k + v_{i+1,j}^k + v_{i-1,j}^k + v_{i,j+1}^k + v_{i,j-1}^k) \quad \text{for all } i,j \in \mathcal{I}_h. \tag{4.4.2}$$

Given $u^{n-1} \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$ and $m^{n-1} \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)$, we define $Q^n \in B(\mathcal{G}_{\Delta t} \times \mathcal{G}_h)^2$ as in (4.3.10) and we discretize the first equation in (4.1.4) using an implicit upwind scheme, which reads

$$\begin{cases} -D_t u_{i,j}^{n,k} - \mu_{i,j}^k (\Delta_h u^{n,k})_{i,j} + Q_{i,j}^{n,k} (D_h u^{n,k})_{i,j} = & \frac{1}{2}|(Q^{n,k})_{i,j}|^2 + V(x_{i,j}) \\ & +\mathcal{F}_{i,j}(m^{n,k+1}, m^{n-1,k+1}), \\ u_{i,j}^{n,T} = G(x_{i,j}), \end{cases} \tag{4.4.3}$$

for all $k \in \mathcal{I}_{\Delta t}^*$ and $i,j \in \mathcal{I}_h$, where $\mu_{i,j}^k = \nu + \frac{h}{2}(|(Q_1^{n,k})_{i,j}| + |(Q_2^{n,k})_{i,j}|)$ and, for every $\mu, \tilde{\mu} \in B(\mathcal{G}_h)$

$$\mathcal{F}_{i,j}(\mu, \tilde{\mu}) = F'(\tilde{\mu}_{i,j})(\mu_{i,j} - \tilde{\mu}_{i,j}) + F(\tilde{\mu}_{i,j}).$$

Then, we consider the following approximation of the forward equation in (4.1.4)

$$\begin{cases} D_t m_{i,j}^{n,k} - \Delta_h(\mu^k m^{n,k+1})_{i,j} - (\mathrm{div}_h(Q^{n,k} m^{n,k+1}))_{i,j} \\ \qquad\qquad\qquad\qquad = \mathrm{div}_h(m^{n-1,k+1}(D_h u^{n,k} - D_h u^{n-1,k}))_{i,j}, \\ m_{i,j}^{n,0} = m_0(x_{i,j}). \end{cases} \tag{4.4.4}$$

Remark 4.4.1 *Notice that the operator* $\mathcal{B}(\mathcal{G}_h) \ni m \to (-\Delta_h(\mu m)_{i,j} - (\mathrm{div}_h(Qm))_{i,j})_{i,j\in\mathcal{I}_h} \in \mathcal{B}(\mathcal{G}_h)$ *is the adjoint of the operator* $\mathcal{B}(\mathcal{G}_h) \ni u \to (-\mu_{i,j}(\Delta_h u)_{i,j} + Q_{i,j}(D_h u)_{i,j})_{i,j\in\mathcal{I}_h} \in \mathcal{B}(\mathcal{G}_h)$.

Combining (4.4.3) and (4.4.4), we consider the following fully discrete upwind scheme for the Newton iterations system (4.1.4)

$$\begin{cases} \mathcal{D}^k U^{n,k} = U^{n,k+1} + \Delta t \mathcal{W}^{k+1} M^{n,k+1} + \Delta t \tilde{\mathcal{B}}^{k+1} & \text{for } k \in \mathcal{I}_{\Delta t}^*, \\ (\mathcal{D}^k)^* M^{n,k+1} = M^{n,k} + \Delta t \mathcal{Z}^k U^{n,k} + \Delta t \tilde{\mathcal{C}}^{k+1} & \text{for } k \in \mathcal{I}_{\Delta t}^*, \\ U^{n,N_t} = U^T, \\ M^{n,0} = M^0, \end{cases} \tag{4.4.5}$$

where $\mathcal{D}^k$ corresponds to the operator

$$(X_{i,j}) \mapsto \left(X_{i,j} - \Delta t \mu_{i,j}^k (\Delta_h X)_{i,j} + \Delta t (D_h X)_{i,j} (Q^{n,k})_{i,j}\right),$$

and $\tilde{\mathcal{B}}^{k+1}$, $\tilde{\mathcal{C}}^{k+1} \in \mathbb{R}^{(N_h)^2}$ are given by

$$\begin{aligned} (\tilde{\mathcal{C}}^{k+1})_{i+N_h j} &= -D_h(M^{n-1,k+1})_{i+N_h j}(D_h U^{n-1,k})_{i+N_h j} \\ &\quad - (M^{n-1,k+1})_{i+N_h j} \mathrm{div}_h \big(D_h U^{n-1,k}\big)_{i+N_h j} \qquad \text{for } i,j \in \mathcal{I}_h, \\ (\tilde{\mathcal{B}}^{k+1})_{i+N_h j} &= \frac{|(Q^{n,k})_{i,j}|^2}{2} - V(x_{i,j}) \\ &\quad + F((M^{n-1,k+1})_{i+N_h j}) - F'((M^{n-1,k+1})_{i+N_h j})(M^{n-1,k+1})_{i+N_h j} \quad \text{for } i,j \in \mathcal{I}_h. \end{aligned}$$

Finally, define $\tilde{U}$ and $\tilde{M} \in \mathbb{R}^{(N_t)N_h^2}$ by

$$\begin{aligned} (\tilde{U}^n)_{kN_h^2 + iN_h + j} &= (U^{n,k})_{i+N_h j} \quad \text{for } k \in \mathcal{I}_{\Delta t}^*,\ i,j \in \mathcal{I}_h, \\ (\tilde{M}^n)_{kN_h^2 + iN_h + j} &= (M^{n,k})_{i+N_h j} \quad \text{for } k \in \mathcal{I}_{\Delta t} \setminus \{0\},\ i,j \in \mathcal{I}_h. \end{aligned}$$

In order to write system (4.4.5) in a matrix way, as in (4.3.29), we define first the matrices

$$\mathbb{F} = \begin{bmatrix} \mathcal{D}^0 & -I_{(N_h)^2} & 0 & \cdots & 0 \\ 0 & \mathcal{D}^1 & -I_{(N_h)^2} & \cdots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots 0 & \ddots & -I_{(N_h)^2} \\ 0 & \cdots & \cdots & \cdots 0 & \mathcal{D}^{N_t-1} \end{bmatrix}, \quad \tilde{\mathbb{W}} = \Delta t \begin{bmatrix} \mathcal{W}^1 & 0 & 0 & \cdots & 0 \\ 0 & \mathcal{W}^2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathcal{W}^{N_t} \end{bmatrix},$$

and

$$\tilde{\mathbb{Z}} = \Delta t \begin{bmatrix} \mathcal{Z}^0 & 0 & 0 & \cdots & 0 \\ 0 & \mathcal{Z}^1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathcal{Z}^{N_t-1} \end{bmatrix}.$$

Hence, (4.4.5) is equivalent to

$$\begin{bmatrix} \mathbb{F} & -\tilde{\mathbb{W}} \\ -\tilde{\mathbb{Z}} & -\mathbb{F}^* \end{bmatrix} \begin{bmatrix} \tilde{U}^n \\ \tilde{M}^n \end{bmatrix} = \begin{bmatrix} \tilde{\mathcal{B}} \\ \tilde{\mathcal{C}} \end{bmatrix}, \tag{4.4.6}$$

where $\tilde{\mathcal{B}} = \Delta t\Big[\mathcal{B}^1, \dots, \mathcal{B}^{N_t} + \frac{1}{\Delta t} U^{N_t}\Big]^*$ and $\tilde{\mathcal{C}} = \Delta t\Big[\frac{1}{\Delta t} M^0 + \mathcal{C}^1, \dots, \mathcal{C}^{N_t}\Big]^*$.

Arguing as in the proof of Proposition 4.3.1, one can show the following well-posedness result.

Proposition 4.4.1 *Suppose that $\tilde{M}^n > 0$ for any $n \in \mathbb{N}$. Then, there exists a unique solution $(\tilde{U}^n, \tilde{M}^n)$ to the system* (4.4.6)*.*

Remark 4.4.2 *The finite difference scheme given by* (4.4.5) *exhibit a different methodology the one in [4]. Indeed, the approach of [4] involves discretizing the MFG system* (4.1.1) *through finite differences, employing a monotone numerical Hamiltonian. Subsequently, Newton iterations are employed to solve the resulting discretized system.*
In contrast, our methodology takes an alternative strategy. We begin the process by considering Newton iterates for the continuous system (4.1.1) *directly. Then we use a simple finite differences scheme to discretize the linear system to be solved at each iteration of Newton's method.*

4.5 Numerical tests

In this section, we assess the performance of our two schemes by conducting tests in both one and two dimensions. The Newton iteration process is stopped once the quantities

$$\|m^{n+1} - m^n\|_\infty \quad \text{and} \quad \|u^{n+1} - u^n\|_\infty \tag{4.5.1}$$

are below a given threshold τ. Systems (4.3.29) and (4.4.6) are solved by Gauss-Seidel iterations and stopped when the uniform norms of the difference between two consecutive solutions are below a threshold δ which is set to 10^{-4}.
Additionally, we conduct a comparative analysis between our semi-Lagrangian scheme (4.3.27) denoted Newton-SL, and the semi-Lagrangian scheme proposed in [58] solved by a (damped) fix point method and denoted SL-FP. Furthermore, we compare the upwind finite differences scheme (4.4.5) denoted Newton-FD, with the finite difference schemes introduced in [4] solved via Newton iterations, denoted FD-Newton.
Algorithm 4 is designed to solve the Newton iteration system (4.1.4) either by Newton-SL or Newton-FD.

Algorithm 4 Newton iterations for mean field games

Input: Initial guesses u^0, m^0, Q^0 and tolerance τ
Output: Solution to the Newton iterations system (4.1.4)
$n \leftarrow 0$
repeat
    Compute m^{n+1} and u^{n+1} by Newton-SL or Newton-FD
    err(m) $\leftarrow \|m^{n+1} - m^n\|_\infty$
    err(u) $\leftarrow \|u^{n+1} - u^n\|_\infty$
    Update Q^n using (4.3.10)
    $n \leftarrow n + 1$
until err(m) $< \tau$ **and** err(u) $< \tau$
return m^{n+1}, u^{n+1}

Remark 4.5.1 *We remind that the semi-Lagrangian schemes that we consider are explicit and do not require a CFL parabolic condition. On the other hand, the finite difference schemes considered here are implicit, and no restriction on the time steps are needed. For the semi-Lagrangian schemes, however, time restriction of the form* $\Delta t = O(h^{3/2})$ *is needed because of accuracy reasons (see [77], [78] for a deeper analysis).*

4.5.1 Dimension 1

Test 1: MFG system with reference solution.

In this test (see [33]), we consider a MFG system in the time-space domain $[0, 0.05]\times]0, 1[$ with periodic boundary conditions, $\nu = 0.1$ and $H(x, p) = \frac{|p|^2}{2}$. The initial condition is given by

$$m_0(x) = \begin{cases} 4\sin^2(2\pi(x - 1/4)) & \text{if } x \in [1/4, 3/4], \\ \\ 0 & \text{otherwise}, \end{cases}$$

and

$$F(m(x)) = 3m_0(x) - 4\min(4, m), \quad G(x) = 0, \quad \text{for all } x \in]0, 1[.$$

In order to compute the errors, we compare the approximated solution with a reference one, computed with SL-FP scheme with $h = 6.67 \cdot 10^{-4}$ and $\Delta t = h^{3/2}/3$. We measure the accuracy of the scheme by computing the errors in the discrete uniform norm. The threshold τ for the Newton stopping iteration criteria is set to 10^{-4}. We set $\Delta t = h^{3/2}/2$ for the two SL schemes, while we set $\Delta t = h/4$ for the finite differences schemes.
Table 4.1 highlights the distinction between Newton-SL and the SL-FP iterations algorithm presented in [58], with the same stopping criteria than Algorithm 3 and 4 (4.5.1), showing CPU time, and the number of Newton iterations. Notably, Newton-SL demonstrates superior performance in terms of CPU time and number of iterations. Additionally, the approximation errors for both the distribution and the value function appear comparable between the two methods.
Table 4.2 shows the results for the same test computed with Newton-FD and FD-Newton. Very similar performance in terms of uniform error, CPU time and iteration count can be observed.
Comparing Tables 4.1 and 4.2, it can be seen that Newton-SL requires the minimum computing time and shows comparable accuracy with respect to the other methods.

Newton-SL with $\Delta t = h^{3/2}/2$				
h	$E_\infty(u)$	$E_\infty(m)$	Time	Iterations
2.50 $\cdot 10^{-2}$	5.51 $\cdot 10^{-2}$	1.64 $\cdot 10^{-1}$	0.61s	6
1.25 $\cdot 10^{-2}$	2.40 $\cdot 10^{-2}$	1.16 $\cdot 10^{-1}$	2.77s	7
6.25 $\cdot 10^{-3}$	1.83 $\cdot 10^{-2}$	6.61 $\cdot 10^{-2}$	13.92s	7
3.125 $\cdot 10^{-3}$	4.50 $\cdot 10^{-3}$	1.41 $\cdot 10^{-2}$	80.60s	7
SL-FP with $\Delta t = h^{3/2}/2$				
h	$E_\infty(u)$	$E_\infty(m)$	Time	Iterations
2.50 $\cdot 10^{-2}$	5.75 $\cdot 10^{-2}$	1.62 $\cdot 10^{-1}$	8.09s	10
1.25 $\cdot 10^{-2}$	2.84 $\cdot 10^{-2}$	1.11 $\cdot 10^{-1}$	40.79s	10
6.25 $\cdot 10^{-3}$	2.15 $\cdot 10^{-2}$	5.84 $\cdot 10^{-2}$	259.72s	12
3.125 $\cdot 10^{-3}$	9.50 $\cdot 10^{-3}$	6.51 $\cdot 10^{-3}$	2793.71s	12

Table 4.1: Errors for the approximation of solution (u, m) using Newton-SL and SL-FP.

FD-Newton with $\Delta t = h/4$				
h	$E_\infty(u)$	$E_\infty(m)$	Time	Iterations
2.50 $\cdot 10^{-2}$	1.23 $\cdot 10^{-1}$	3.11 $\cdot 10^{-2}$	2.23s	7
1.25 $\cdot 10^{-2}$	6.21 $\cdot 10^{-2}$	1.63 $\cdot 10^{-2}$	18.32s	8
6.25 $\cdot 10^{-3}$	3.14 $\cdot 10^{-2}$	8.75 $\cdot 10^{-3}$	92.91s	8
3.125 $\cdot 10^{-3}$	1.77 $\cdot 10^{-2}$	9.54 $\cdot 10^{-3}$	597.21s	8
Newton-FD with $\Delta t = h/4$				
h	$E_\infty(u)$	$E_\infty(m)$	Time	Iterations
2.50 $\cdot 10^{-2}$	1.532 $\cdot 10^{-1}$	3.42 $\cdot 10^{-2}$	1.48s	7
1.25 $\cdot 10^{-2}$	6.71 $\cdot 10^{-2}$	1.83 $\cdot 10^{-2}$	12.27s	7
6.25 $\cdot 10^{-3}$	3.37 $\cdot 10^{-2}$	9.51 $\cdot 10^{-3}$	68.10s	7
3.125 $\cdot 10^{-3}$	1.91 $\cdot 10^{-2}$	7.38 $\cdot 10^{-3}$	436.01s	7

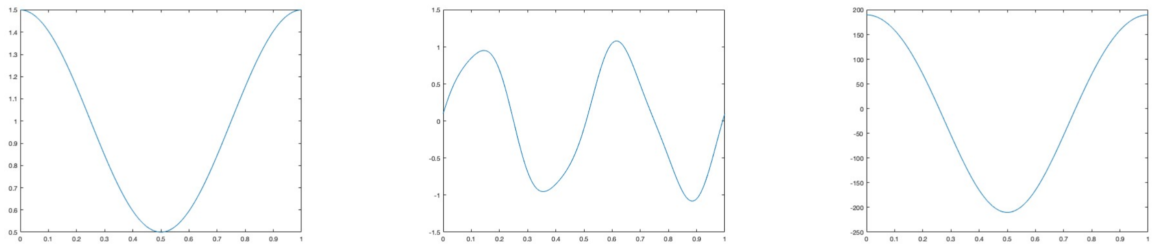
Table 4.2: Errors for the approximation of solution (u, m) using FD-Newton and Newton-FD.

Test 2

In this test, we consider a MFG system, numerically solved in [119], in the time-space domain $[0, 0.01]\times]0, 1[$ with periodic boundary conditions and diffusion coefficient $\nu = 0.4$ and $\nu = 0.02$. We consider the following data, showed in Figure 4.1

$$\begin{aligned} m_0(x) &= 1 + \tfrac{1}{2}\cos(2\pi x), \\ G(x) &= \sin(4\pi x) + 0.1\cos(10\pi x), \\ H(x,p) &= |p|^2 + V(x), \quad V(x) = 200\cos(2\pi x) - 10\cos(4\pi x), \\ F(m) &= m^2. \end{aligned}$$

The threshold τ for the Newton stopping iteration criteria is set to 10^{-4}. We also set $N_h = 160$.

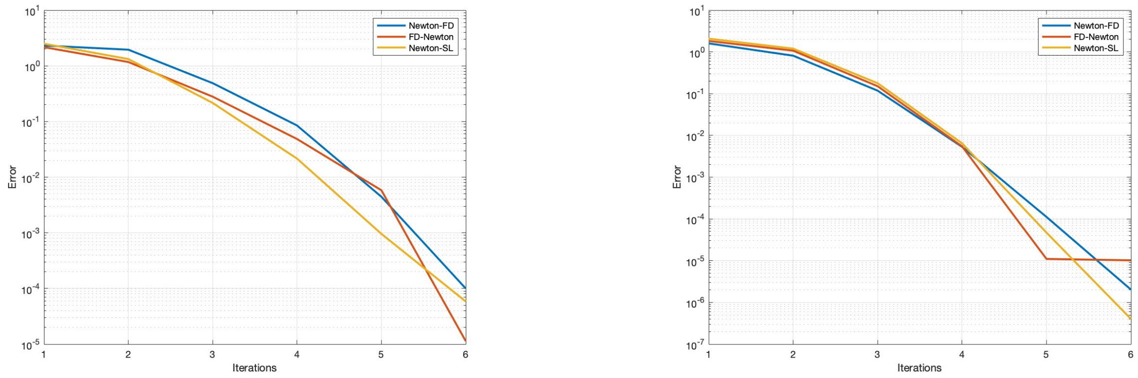


(a) Initial distribution m_0 **(b)** Terminal cost G **(c)** Potential V

Figure 4.1: The initial data.

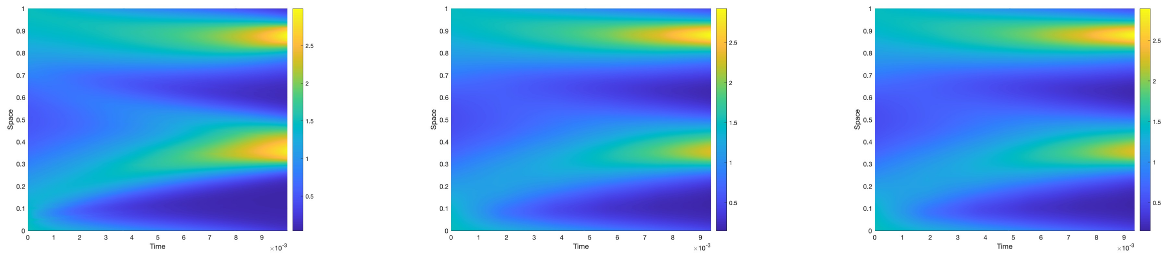
Diffusion coefficient $\nu = 0.4$

We first consider $\nu = 0.4$ and we solve the MFG system using Newton-SL, Newton-FD and FD-Newton. We set $\Delta t = h/4$ for the two finite differences schemes, and $\Delta t = h^{3/2}/2$ for Newton-SL.



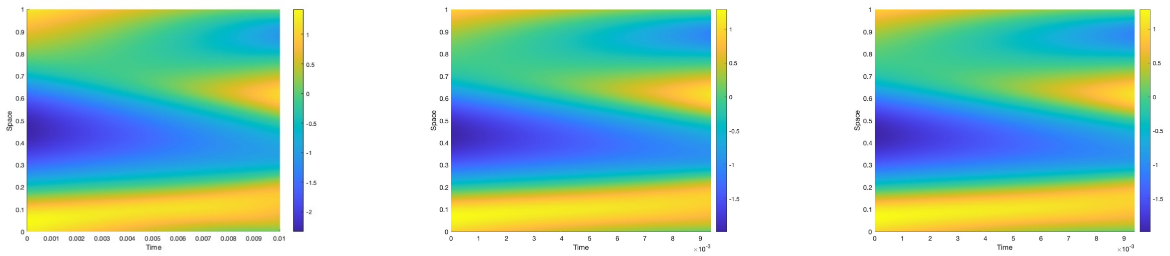
(a) $\|m^{n+1} - m^n\|_\infty$ **(b)** $\|u^{n+1} - u^n\|_\infty$

Figure 4.2: Newton errors for the three schemes.



(a) Newton-SL **(b)** Newton-FD **(c)** FD-Newton

Figure 4.3: The distribution approximated with the three Newton schemes.



(a) Newton-SL **(b)** Newton-FD **(c)** FD-Newton

Figure 4.4: The value function approximated with the three Newton schemes.

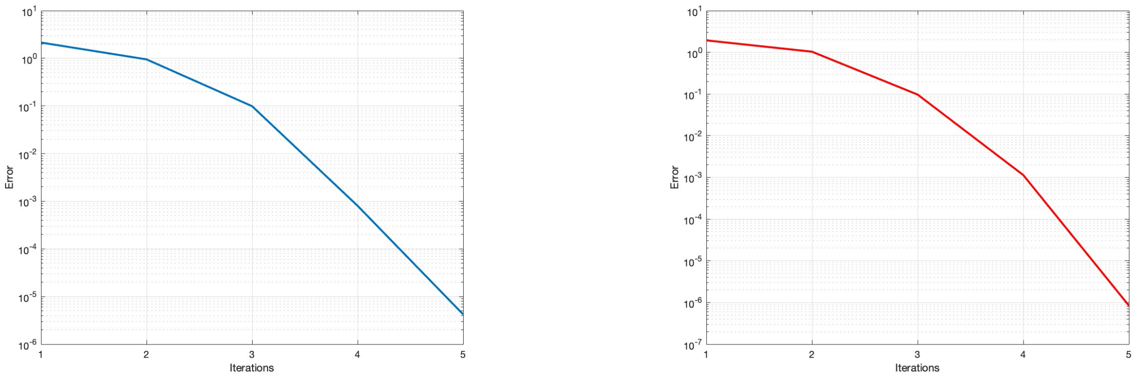
Figure 4.2 shows on the y-axis the Newton iterations errors for the schemes Newton-SL, Newton-FD and FD-Newton in a logarithmic scale with respect to number of Newton iterations. Figure 4.3 shows the plots in the space-time grid of the approximated distribution, computed by the three schemes. We can observe that the distribution concentrates, at final time, near the points where the terminal cost G, showed in Figure 4.1b, reaches its minima.

Finally, Figure 4.4 shows the plots in the space-time grid, of the approximated value function, computed by the three schemes.

Diffusion coefficient $\nu = 0.02$

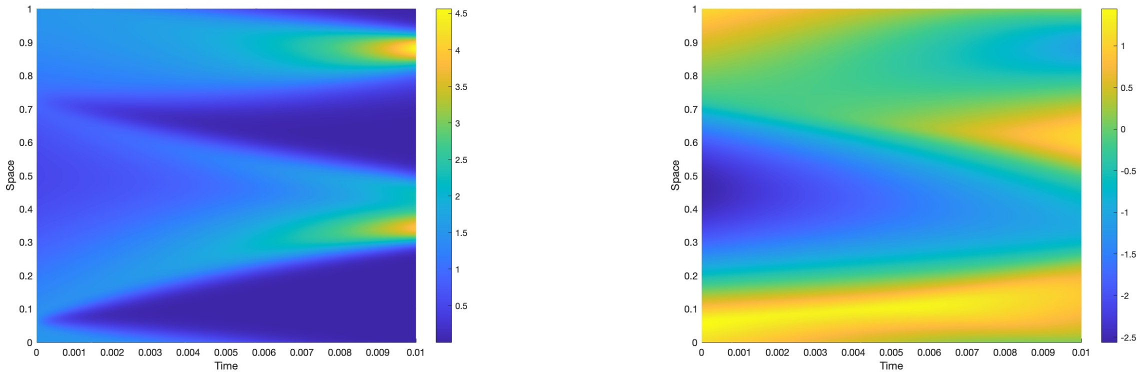
Let us consider now a smaller diffusion term, specifically $\nu = 0.02$. Newton-SL iterations, with $\Delta t = h^2$, demonstrate convergence after 5 iterations showed in Figure 4.5, reaching the associated threshold. In contrast, both Newton-FD and FD-Newton iterations encounter breakdowns after only a few iterations. This indicates a higher robustness offered by the Newton-SL scheme, in scenarios characterised by small diffusion terms. Figure 4.6 shows the approximated distribution and the approximated value function in the time-space domain.

Remark 4.5.2 *In [6], the authors solve a finite difference discretization of the MFG system by employing Newton's method combined with a continuation method with respect to the diffusion parameter ν (see also [10], [12]). The latter is particularly useful to deal with the case small diffusion parameters. The problem is solved first for a high value of ν and, subsequently, the authors use this solution as an initial guess to solve, still by using Newton's method, the discrete MFG system with a smaller viscosity. The method proceeds in this manner until reaching the desired (small) viscosity.*
As a completion on this work, a comparison between the method of continuation applied to our scheme and other methods will be made.



(a) $\|m^{n+1} - m^n\|_\infty$ **(b)** $\|u^{n+1} - u^n\|_\infty$

Figure 4.5: Newton-SL iterations error for $\nu = 0.02$.



(a) Evaluation of m **(b)** Evaluation of u

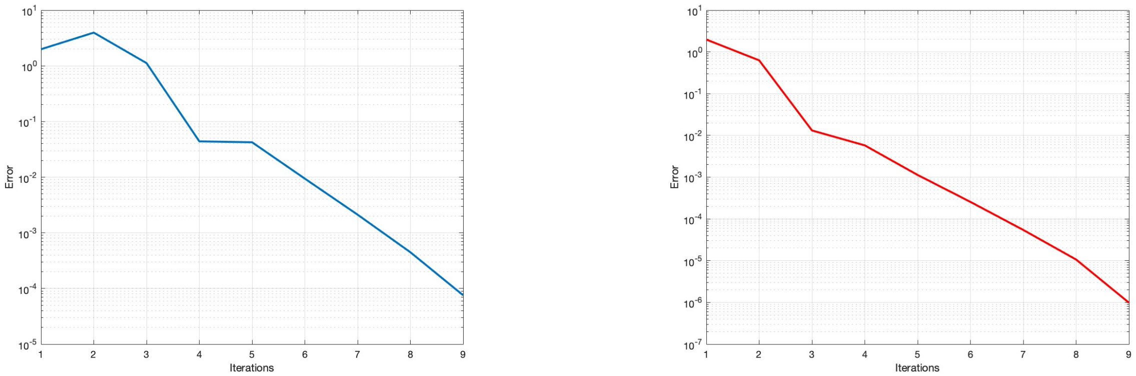
Figure 4.6: Approximated m and u.

4.5.2 Dimension 2

We consider now a MFG system in dimension 2, previously numerically studied in [119]. We set $Q = [0,1]\times]0,1[^2$, $\nu = 1$, and the following data

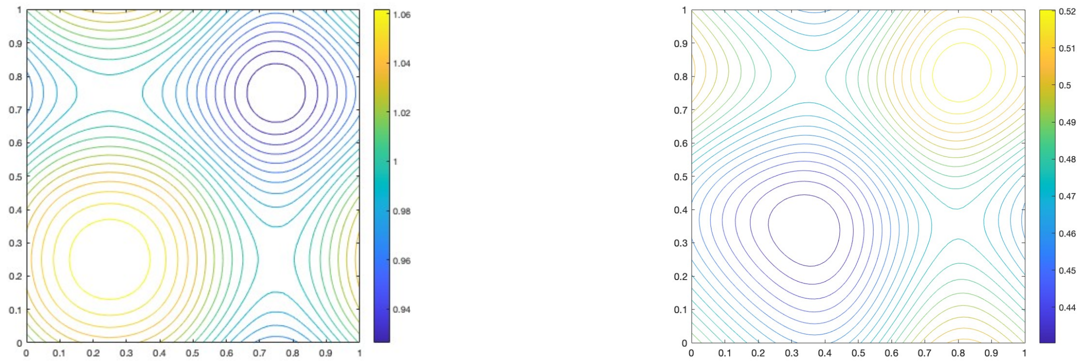
$$\begin{aligned} m_0(x_1, x_2) &= 1 + 12\cos(2\pi x_1) + 12\cos(2\pi x_2), \\ G(x_1, x_2) &= \cos(2\pi x_1) + \cos(2\pi x_2), \\ H(x_1, x_2, p) &= |p|^2 + V(x_1, x_2), \quad V(x_1, x_2) = \sin(2\pi x_1) + \sin(2\pi x_2) + \cos(4\pi x_1), \\ F(m) &= m^2. \end{aligned}$$

We set $\tau = 10^{-4}$, $N_h = 66$, and $\Delta t = h^2/2$. Figure 4.7 shows the errors (4.5.1) in a logarithmic scale on the y-axis with respect to number of Newton iterations on the x-axis, Figure 4.9 shows the contour level of the approximations to $u(T/2,\cdot)$ and $m(T/2,\cdot)$, respectively, where we can notice easily the effect of periodic boundary conditions. The solution m is shown in Figure 4.9 at different times. Figures 4.9b-4.9c show the stationary state reached by the density, at intermediate time, which can be interpreted as a turnpike effect. Turnpike phenomena for mean field games with local coupling has been discussed in [67].



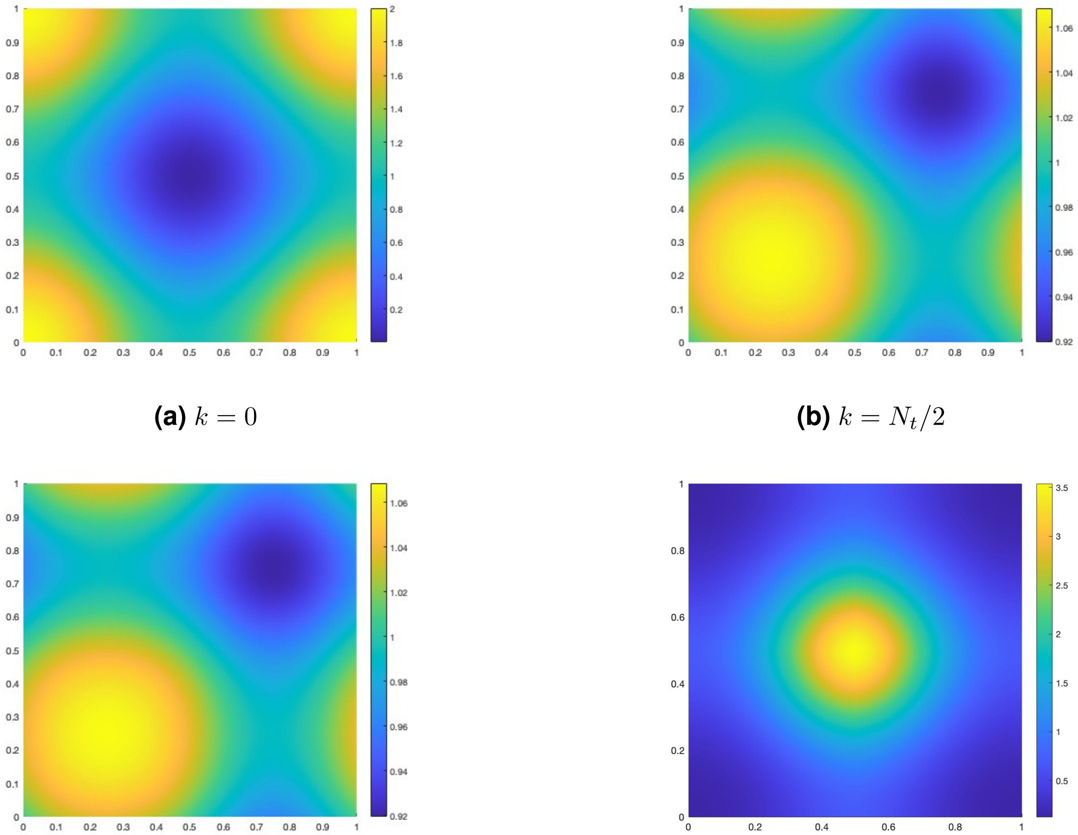
(a) $\|m^{n+1} - m^n\|_\infty$ **(b)** $\|u^{n+1} - u^n\|_\infty$

Figure 4.7: Newton-SL iterations error.



(a) $m(N_t/2, x)$ **(b)** $u(T/2, x)$

Figure 4.8: Contours level of $m(N_t/2, x)$ and $u(T/2, x)$.



(a) $k = 0$

(b) $k = N_t/2$

(c) $k = 3N_t/4$

(d) $k = N_t$

Figure 4.9: The approximated distribution m at times $t = 0, \frac{N_t \Delta t}{2}, \frac{3N_t \Delta t}{4}, T$.

5
Bibliography

[1] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta, “Mean field games: numerical methods for the planning problem,” *SIAM Journal on Control and Optimization*, vol. 50, Jan. 2012.

[2] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta, “Mean field games: convergence of a finite difference method,” *SIAM J. Numer. Anal.*, vol. 51, no. 5, 2013.

[3] Y. Achdou, F. Camilli, and L. Corrias, “On numerical approximation of the Hamilton-Jacobi-transport system arising in high frequency approximations,” *Discrete Contin. Dyn. Syst. Ser. B*, vol. 19, no. 3, 2014.

[4] Y. Achdou and I. Capuzzo-Dolcetta, “Mean field games: numerical methods,” *SIAM J. Numer. Anal.*, vol. 48, no. 3, 2010.

[5] Y. Achdou, P. Cardaliaguet, F. Delarue, A. Porretta, and F. Santambrogio, *Mean field games* (Lecture Notes in Mathematics). Springer, Cham; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 2020, vol. 2281.

[6] Y. Achdou and M. Laurière, “Mean field games and applications: numerical aspects,” in *Mean Field Games: Cetraro, Italy 2019*. Cham: Springer International Publishing, 2020, pp. 249–307.

[7] Y. Achdou and M. Laurière, “Mean field games and applications: numerical aspects,” in *Mean field games*, ser. Lecture Notes in Math. Vol. 2281, Springer, Cham, 2020.

[8] Y. Achdou, P. Mannucci, C. Marchi, and N. Tchou, “Deterministic mean field games with control on the acceleration and state constraints,” *SIAM Journal on Mathematical Analysis*, vol. 54, no. 3, 2022.

[9] Y. Achdou and A. Porretta, “Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games,” *SIAM J. Numer. Anal.*, vol. 54, no. 1, 2016.

[10] Y. Achdou, M. Bardi, and M. Cirant, “Mean field games models of segregation,” *Mathematical Models and Methods in Applied Sciences*, vol. 27, no. 01, 2017.

[11] Y. Achdou, J. Han, J. M. Lasry, P. L. Lions, and B. Moll, “Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach,” *The Review of Economic Studies*, vol. 89, no. 1, pp. 45–86, 2021.

[12] Y. Achdou and Z. Kobeissi, “Mean field games of controls: finite difference approximations,” *Mathematics in Engineering*, vol. 3, no. 3, 2021.

[13] Y. Achdou and V. Perez, “Iterative strategies for solving linearized discrete mean field games systems,” *Netw. Heterog. Media*, vol. 7, no. 2, 2012.

[14] C. Alasseur, I. Ben Taher, and A. Matoussi, “An extended mean field game for storage in smart grids,” *J. Optim. Theory Appl.*, vol. 184, no. 2, 2020.

[15] N. Almulla, R. Ferreira, and D. A. Gomes, “Two numerical approaches to stationary mean-field games,” *Dyn. Games Appl.*, vol. 7, no. 4, 2017.

[16] L. Ambrosio, “Transport equation and Cauchy problem for BV vector fields,” *Invent. Math.*, vol. 158, no. 2, 2004.

[17] L. Ambrosio, “Transport equation and Cauchy problem for non-smooth vector fields,” in *Calculus of variations and nonlinear partial differential equations*, ser. Lecture Notes in Math. Vol. 1927, Springer, Berlin, 2008.

[18] L. Ambrosio, N. Gigli, and G. Savare, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures* (Lectures in Mathematics. ETH Zürich). Birkhäuser Basel, 2008.

[19] A. Angiuli, J.-P. Fouque, and M. Laurière, “Unified reinforcement Q-learning for mean field game and control problems,” *Math. Control Signals Systems*, vol. 34, no. 2, 2022.

[20] Y. Ashrafyan and D. Gomes, *A fully-discrete semi-lagrangian scheme for a price formation mfg model*, 2024. arXiv: 2403.02785.

[21] J.-P. Aubin and A. Cellina, *Differential inclusions* (Grundlehren der mathematischen Wissenschaften). Springer-Verlag, Berlin, 1984, vol. 264.

[22] A. Aurell, R. Carmona, G. Dayanikli, and M. Laurière, “Optimal incentives to mitigate epidemics: a stackelberg mean field game approach,” *SIAM Journal on Control and Optimization*, vol. 60, no. 2, 2022.

[23] M. Bardi and I. C. Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Birkauser, 1996.

[24] M. Bardi and F. S. Priuli, “Linear-quadratic N-person and mean-field games with ergodic cost,” *SIAM J. Control Optim.*, vol. 52, no. 5, 2014.

[25] G. Barles and P. E. Souganidis, “Convergence of approximation schemes for fully nonlinear second order equations,” *Asymptotic Anal.*, vol. 4, no. 3, 1991.

[26] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi* (Math. Appl. (Berl.)). Paris: Springer-Verlag, 1994, vol. 17.

[27] J. D. Benamou and G. Carlier, “Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations,” *J. Optim. Theory Appl.*, vol. 167, no. 1, 2015.

[28] J.-D. Benamou, G. Carlier, and F. Santambrogio, “Variational mean field games,” in *Active particles. Vol. 1. Advances in theory, models, and applications*, ser. Model. Simul. Sci. Eng. Technol. Birkhäuser/Springer, Cham, 2017.

[29] A. Bensoussan and J. Frehse, “Nonlinear elliptic systems in stochastic game theory,” *J. Reine Angew. Math.*, vol. 350, 1984.

[30] A. Bensoussan and J. Frehse, “Ergodic bellman systems for stochastic games in arbitrary dimension,” *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, vol. 449, 1995.

[31] J. Berry, O. Ley, and F. J. Silva, *Approximation and perturbations of stable solutions to a stationary mean field game system*, 2024. arXiv: 2402.16377.

[32] V. I. Bogachev, *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[33] P. Bojan and T. Vladimir, “Central schemes for mean field games,” *Communications in Mathematical Sciences*, vol. 13, Jan. 2015.

[34] L. Bonaventura, E. Calzola, E. Carlini, and R. Ferretti, “Second order fully semi-Lagrangian discretizations of advection-diffusion-reaction systems,” *J. Sci. Comput.*, vol. 88, no. 1, 2021, Id/No 23.

[35] J. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.

[36] Bonnans, J. F., Liu, K., and Pfeiffer, L., “Error estimates of a theta-scheme for second-order mean field games,” *ESAIM: M2AN*, vol. 57, no. 4, 2023.

[37] L. Briceño-Arias, D. Kalise, Z. Kobeissi, M. Laurière, A. Mateos González, and F. J. Silva, “On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings,” *ESAIM: Proceedings and Surveys*, vol. 65, Jan. 2019.

[38] L. M. Briceño-Arias, D. Kalise, and F. J. Silva, “Proximal methods for stationary mean field games with local couplings,” *SIAM J. Control Optim.*, vol. 56, no. 2, 2018.

[39] G. W. Brown, “Iterative solution of games by fictitious play,” in *Activity Analysis of Production and Allocation*, ser. Cowles Commission Monographs, No. 13, John Wiley & Sons, Inc., New York, N.Y.; Chapman & Hall, Ltd., London, 1951.

[40] S. Cacace, F. Camilli, and A. Goffi, “A policy iteration method for mean field games,” *ESAIM, Control Optim. Calc. Var.*, vol. 27, 2021, Id/No 85.

[41] E. Calzola, E. Carlini, and F. J. Silva, “A high-order scheme for mean field games,” 2022, arXiv:2207.08463.

[42] F. Camilli and F. J. Silva, “A semi-discrete in time approximation for a first order-finite mean field game problem,” *Network and Heterogeneous Media*, vol. 7-2, 2012.

[43] F. Camilli and Q. Tang, “Rates of convergence for the policy iteration method for mean field games systems,” *J. Math. Anal. Appl.*, vol. 512, no. 1, 2022.

[44] F. Camilli and Q. Tang, “On the quadratic convergence of newton’s method for mean field games with non-separable hamiltonian,” *Dyn Games Appl*, 2024.

[45] P. Cannarsa and R. Capuani, “Existence and uniqueness for mean field games with state constraints,” in *PDE models for multi-agent phenomena*, ser. Springer INdAM Ser. Vol. 28, Springer, Cham, 2018, pp. 49–71.

[46] P. Cannarsa, R. Capuani, and P. Cardaliaguet, “Mean field games with state constraints: from mild to pointwise solutions of the pde system,” *Calculus of Variations and Partial Differential Equations*, vol. 60, no. 3, 2021.

[47] P. Cannarsa and C. Mendico, “Mild and weak solutions of mean field game problems for linear control systems,” *Minimax Theory Appl.*, vol. 5, no. 2, 2020.

[48] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control* (Progress in Nonlinear Differential Equations and Their Applications). Birkauser, 2004.

[49] P. Cardaliaguet, “Weak solutions for first order mean field games with local coupling,” *Springer INdAM Series*, vol. 11, May 2013.

[50] P. Cardaliaguet, F. Delarue, J. M. Lasry, and P. L. Lions, *The master equation and the convergence problem in mean field games* (Ann. Math. Stud.). Princeton, NJ: Princeton University Press, 2019, vol. 201.

[51] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon, “Second order mean field games with degenerate diffusion and local coupling,” *NoDEA, Nonlinear Differ. Equ. Appl.*, vol. 22, no. 5, 2015.

[52] P. Cardaliaguet and S. Hadikhanloo, “Learning in mean field games: the fictitious play,” *ESAIM, Control Optim. Calc. Var.*, vol. 23, no. 2, 2017.

[53] P. Cardaliaguet and C. A. Lehalle, “Mean field game of controls and an application to trade crowding,” *Mathematics and Financial Economics*, vol. 12, 2016.

[54] P. Cardaliaguet, A. R. Mészáros, and F. Santambrogio, “First order mean field games with density constraints: pressure equals price,” *SIAM J. Control Optim.*, vol. 54, no. 5, 2016.

[55] P. Cardaliaguet and A. Porretta, “An introduction to mean field game theory,” in *Mean field games*, ser. Lecture Notes in Math. Vol. 2281, Springer, Cham, 2020.

[56] E. Carlini, F. J. Silva, and A. Zorkot, “A lagrange–galerkin scheme for first order mean field game systems,” *SIAM Journal on Numerical Analysis*, vol. 62, 2024.

[57] E. Carlini and F. J. Silva, “A fully discrete semi-Lagrangian scheme for a first order mean field game problem,” *SIAM J. Numer. Anal.*, vol. 52, no. 1, 2014.

[58] E. Carlini and F. J. Silva, “On the discretization of some nonlinear Fokker-Planck-Kolmogorov equations and applications,” *SIAM J. Numer. Anal.*, vol. 56, no. 4, 2018.

[59] R. Carmona and F. Delarue, “Probabilistic analysis of mean-field games,” *SIAM Journal on Control and Optimization*, vol. 51, Oct. 2012.

[60] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications. I* (Probability Theory and Stochastic Modelling). Springer, Cham, 2018, vol. 83.

[61] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications. II* (Probability Theory and Stochastic Modelling). Springer, Cham, 2018, vol. 84.

[62] 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,” *SIAM J. Numer. Anal.*, vol. 59, no. 3, 2021.

[63] 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,” *Ann. Appl. Probab.*, vol. 32, no. 6, 2022.

[64] R. A. Carmona, “Applications of mean field games in financial engineering and economic theory,” *Proceedings of Symposia in Applied Mathematics*, 2020.

[65] C. Cheng and G. Knorr, “The integration of the vlasov equation in configuration space,” *Journal of Computational Physics*, vol. 22, no. 3, 1976.

[66] I. Chowdhury, O. Ersland, and E. Jakobsen, “On numerical approximations of fractional and nonlocal mean field games,” *Found Comput Math*, vol. 23, 2023.

[67] M. Cirant and A. Porretta, “Long time behavior and turnpike solutions in mildly non-monotone mean field games,” *ESAIM, Control Optim. Calc. Var.*, vol. 27, 2021, Id/No 86.

[68] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, “Electrical vehicles in the smart grid: a mean field game analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 30, Oct. 2012.

[69] M. G. Crandall, L. C. Evans, and P. L. Lions, “Some properties of viscosity solutions of hamilton-jacobi equations.,” *Transactions of the American Mathematical Society*, vol. 282, 1984.

[70] M. G. Crandall and P. L. Lions, “Viscosity solutions of hamilton-jacobi equations,” *Transactions of the American Mathematical Society*, vol. 277, 1983.

[71] F. Da Lio and O. Ley, “Convex Hamilton-Jacobi equations under superlinear growth conditions on data,” *Appl. Math. Optim.*, vol. 63, no. 3, 2011.

[72] B. Dacorogna, *Direct Methods in the Calculus of Variations*. Berlin, Heidelberg: Springer-Verlag, 1989.

[73] B. Djehiche, A. Tcheukam Siwe, and H. Tembine, “Mean-field-type games in engineering,” *AIMS Electronics and Electrical Engineering*, vol. 1, Nov. 2017.

[74] R. Elie, E. Hubert, T. Mastrolia, and D. Possamaï, “Mean–field moral hazard for optimal energy demand response management,” *Mathematical Finance*, vol. 31, Oct. 2020.

[75] R. Elie, T. Mastrolia, and D. Possamaï, “A tale of a principal and many, many agents,” *Mathematics of Operations Research*, vol. 44, Aug. 2016.

[76] M. Falcone and R. Ferretti, *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014.

[77] R. Ferretti, “Equivalence of semi-lagrangian and lagrange-galerkin schemes under constant advection speed,” *Journal of Computational Mathematics*, vol. 28, 2010.

[78] R. Ferretti, “On the relationship between semi-lagrangian and lagrange–galerkin schemes,” *Numerische Mathematik*, vol. 124, 2013.

[79] M. Fischer and F. J. Silva, “On the asymptotic nature of first order mean field games,” *Appl. Math. Optim.*, vol. 84, no. 2, 2021.

[80] J. Gianatti and F. J. Silva, “Approximation of deterministic mean field games with control-affine dynamics,” *Foundations of Computational Mathematics*, 2022.

[81] J. Gianatti, F. J. Silva, and A. Zorkot, *Approximation of deterministic mean field games under polynomial growth conditions on the data*, 2024.

[82] D. A. Gomes, J. Gutierrez, and M. Laurière, “Machine learning architectures for price formation models,” *Appl. Math. Optim.*, vol. 88, no. 1, 2023, Id/No 23.

[83] D. A. Gomes, J. Mohr, and R. Souza, “Discrete time, finite state space mean field games,” *Journal de Mathématiques Pures et Appliquées*, vol. 93, 2010.

[84] D. A. Gomes, J. Mohr, and R. R. Souza, “Discrete time, finite state space mean field games,” *J. Math. Pures Appl. (9)*, vol. 93, no. 3, 2010.

[85] D. A. Gomes, L. Nurbekyan, and E. Pimentel, *Economic models and mean-field games theory*. Jul. 2015.

[86] D. A. Gomes, E. A. Pimentel, and V. Voskanyan, *Regularity theory for mean-field game systems* (SpringerBriefs in Mathematics). Springer, Cham, 2016.

[87] D. A. Gomes and J. Saúde, “Mean field games models-A brief survey,” *Dyn. Games Appl.*, vol. 4, no. 2, 2014.

[88] D. A. Gomes and J. Saúde, “A mean-field game approach to price formation,” *Dynamic Games and Applications*, vol. 11, 2018.

[89] D. A. Gomes and J. Saúde, “Numerical methods for finite-state mean-field games satisfying a monotonicity condition,” *Appl. Math. Optim.*, vol. 83, no. 1, 2021.

[90] D. A. Gomes, R. Velho, and M. Wolfram, “Socio-economic applications of finite state mean field games,” *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 372, Mar. 2014.

[91] D. A. Gomes and V. K. Voskanyan, “Extended deterministic mean-field games,” *SIAM Journal on Control and Optimization*, vol. 54, no. 2, 2016.

[92] O. Guéant, “Mean field games equations with quadratic Hamiltonian: a specific approach,” *Math. Models Methods Appl. Sci.*, vol. 22, no. 9, 2012.

[93] S. Hadikhanloo, R. Laraki, P. Mertikopoulos, and S. Sorin, “Learning in nonatomic games. I: Finite action spaces and population games,” *J. Dyn. Games*, vol. 9, no. 4, 2022.

[94] S. Hadikhanloo and F. J. Silva, “Finite mean field games: fictitious play and convergence to a first order continuous mean field game,” *J. Math. Pures Appl. (9)*, vol. 132, 2019.

[95] K. Huang, X. Di, Q. Du, and X. Chen, *A game-theoretic framework for autonomous vehicles velocity control: bridging microscopic differential games and macroscopic mean field games*, 2020.

[96] M. Huang, P. Caines, and R. Malhame, “An invariance principle in large population stochastic dynamic games,” *Journal of Systems Science and Complexity*, vol. 20, Jun. 2007.

[97] M. Huang, P. Caines, and R. Malhame, “Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized ε-nash equilibria,” *IEEE Transactions on Automatic Control*, vol. 52, no. 9, 2007.

[98] M. Huang, P. Caines, and R. Malhame, “The nash certainty equivalence principle and mckean-vlasov systems: an invariance principle and entry adaptation,” Jan. 2008.

[99] 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," *Communications in Information and Systems*, vol. 6, no. 3, 2006.

[100] J. Jim and T. Russell, "Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures," *SIAM Journal on Numerical Analysis*, vol. 19, Oct. 1982.

[101] C. R. Johnson, "Positive definite matrices," *The American Mathematical Monthly*, vol. 77, no. 3, 1970.

[102] M. Kardar, *Statistical Physics of Fields*. Cambridge University Press, 2007.

[103] A. C. Kizilkale, R. Salhab, and R. P. Malhamé, "An integral control formulation of mean field game based large scale coordination of loads in smart grids," *Automatica*, vol. 100, 2019.

[104] P. E. Kloeden and E. Platen, "The numerical solution of stochastic differential equations," *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, vol. 20, 1977.

[105] V. Kolokoltsov and O. Malafeyev, "Corruption and botnet defense: a mean field game approach," *International Journal of Game Theory*, vol. 47, Sep. 2018.

[106] V. N. Kolokoltsov and M. S. Troeva, "Abstract McKean-Vlasov and Hamilton-Jacobi-Bellman equations, their fractional versions and related forward-backward systems on Riemannian manifolds," *Proc. Steklov Inst. Math.*, vol. 315, 2021.

[107] A. Lachapelle, J. M. Lasry, C. A. Lehalle, and P. L. Lions, "Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis," *Mathematics and Financial Economics*, vol. 10, 2013.

[108] D. Lacker, "On the convergence of closed-loop nash equilibria to the mean field game limit," *The Annals of Applied Probability*, 2018.

[109] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and quasi-linear equations of parabolic type*, Russian. Translations of Mathematical Monographs, 1968, vol. 23.

[110] L. Laguzet and G. Turinici, "Individual vaccination as nash equilibrium in a sir model with application to the 2009–2010 influenza a (h1n1) epidemic in france," *Bulletin of Mathematical Biology*, vol. 77, Jun. 2015.

[111] J. M. Lasry and P. L. Lions, "Jeux à champ moyen. II – Horizon fini et contrôle optimal," *Comptes Rendus. Mathématique*, vol. 343, no. 10, 2006.

[112] J. M. Lasry and P. L. Lions, “Mean field games,” *Japanese Journal of Mathematics*, vol. 2, Mar. 2007.

[113] J. M. Lasry and P. L. Lions, “Jeux à champ moyen. I – Le cas stationnaire,” *Comptes Rendus. Mathématique*, vol. 343, no. 9, 2006.

[114] M. Laurière, “Numerical methods for mean field games and mean field type control,” in *Mean field games*, ser. Proc. Sympos. Appl. Math. Vol. 78, Amer. Math. Soc., Providence, RI, 2021.

[115] M. Laurière, S. Perrin, M. Geist, and O. Pietquin, *Learning mean field games: a survey*, 2022. arXiv: `2205.12944 [cs.LG]`.

[116] M. Laurière, J. Song, and Q. Tang, “Policy iteration method for time-dependent mean field games systems with non-separable Hamiltonians,” *Appl. Math. Optim.*, vol. 87, no. 2, 2023, Id/No 17.

[117] P. Lavigne and L. Pfeiffer, “Generalized conditional gradient and learning in potential mean field games,” *Appl. Math. Optim.*, vol. 88, no. 3, 2023, Id/No 89.

[118] W. Lee, S. Liu, H. Tembine, W. Li, and S. Osher, “Controlling propagation of epidemics via mean-field control,” *SIAM Journal on Applied Mathematics*, vol. 81, no. 1, 2021.

[119] H. Li, Y. Fan, and L. Ying, “A simple multiscale method for mean field games,” *J. Comput. Phys.*, vol. 439, 2021, Id/No 110385.

[120] P.-L. Lions. “Cours au collège de france.” ().

[121] S. Liu, M. Jacobs, W. Li, L. Nurbekyan, and S. J. Osher, “Computational methods for first-order nonlocal mean field games with applications,” *SIAM J. Numer. Anal.*, vol. 59, no. 5, 2021.

[122] S. Liu and L. Nurbekyan, “Splitting methods for a class of non-potential mean field games,” *Journal of Dynamics and Games*, vol. 8, Jan. 2021.

[123] F. Mériaux, V. Varma, and S. Lasaulce, “Mean field energy games in wireless networks,” Jan. 2013.

[124] K. W. Morton, A. Priestley, and E. Süli, “Stability of the Lagrange-Galerkin method with non-exact integration,” *RAIRO, Modélisation Math. Anal. Numér.*, vol. 22, no. 4, 1988.

[125] J. R. Munkres, *Topology*, Second. Prentice Hall, Inc., Upper Saddle River, NJ, 2000.

[126] L. Nurbekyan and J. Saúde, “Fourier approximation methods for first-order nonlocal mean-field games,” *Port. Math.*, vol. 75, no. 3-4, 2018.

[127] B. Piccoli and A. Tosin, “Time-evolving measures and macroscopic modeling of pedestrian flow,” *Arch. Ration. Mech. Anal.*, vol. 199, no. 3, 2011.

[128] O. Pironneau, “On the transport-diffusion algorithm and its applications to the navierstokes equations.,” *Numerische Mathematik*, vol. 38, 1982.

[129] A. Quarteroni and A. Valli, *Numerical approximation of partial differential equations*. Springer Verlag, 1994.

[130] J. Robinson, “An iterative method of solving a game,” *Ann. of Math. (2)*, vol. 54, 1951.

[131] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis* (Grundlehren der mathematischen Wissenschaften). Springer-Verlag, Berlin, 1998, vol. 317.

[132] H. Shiri, J. Park, and M. Bennis, “Massive autonomous uav path planning: a neural network based mean-field game theoretic approach,” in *2019 IEEE Global Communications Conference*, 2019.

[133] A. Tosin and P. Frasca, “Existence and approximation of probability measure solutions to models of collective behaviors,” *Netw. Heterog. Media*, vol. 6, no. 3, 2011.

[134] J. Yong and X. Y. Zhou, *Stochastic controls. Hamiltonian systems and HJB equations* (Appl. Math. (N. Y.)). New York, NY: Springer, 1999, vol. 43.

Approximation de jeux à champ moyen

Résumé : L'objectif de la théorie des jeux à champ moyen est d'étudier une classe de jeux différentiels (déterministes ou stochastiques) comportant un grand nombre de joueurs. Étant donné que très peu de jeux à champ moyen admettent des solutions explicites, les méthodes numériques jouent un rôle essentiel dans la description quantitative, mais aussi qualitative, des équilibres de Nash associés. Cette thèse se concentrera sur des techniques numériques utilisées pour résoudre diverses classes de jeux à champ moyen.

Mots clés : Jeux à champ moyen, contrôle optimal, équations de Hamilton-Jacobi-Bellman et de Fokker-Planck, analyse numérique des équations aux dérivées partielles.

Approxiamtion to mean field games

Abstract : The purpose of the theory of mean field games is to study a class of differential games (deterministic or stochastic) with a large number of agents. Since very few mean field games admit explicit solutions, numerical methods play an essential role in describing quantitatively, and also qualitatively, the associated Nash equilibria. This thesis is focused on numerical techniques to solve several types of mean field game problems.

Keywords: Mean field games, optimal control, optimal control theory, Hamilton-Jacobi-Bellman and Fokker-Planck equations, numerical analysis of partial differential equations.

Univ. de Limoges, CNRS, XLIM, UMR 7252, F-87000 Limoges, France