INSTITUT POLYTECHNIQUE DE PARIS

NNT : 20XXIPPAXXXX

Thèse de doctorat

Numerical analysis and methods for mean-field-type optimization problems

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'Ecole Polytechnique

École doctorale n°574 Ecole Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat: Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 5 octobre 2023, par

KANG LIU

Composition du Jury :

Grégoire Allaire Professeur, Ecole Polytechnique	Examinateur
Frédéric Bonnans Directeur de recherche, CentraleSupélec et Inria	Co-directeur de thèse
Fabio Camilli Professeur, Université de Rome La Sapienza	Rapporteur
Nadia Oudjane Ingénieure de recherche, EDF-Lab	Invitée
Laurent Pfeiffer Chargé de recherche, CentraleSupélec et Inria	Directeur de thèse
Filippo Santambrogio Professeur, Université Claude Bernard - Lyon 1	Rapporteur
Hasnaa Zidani Professeure, INSA Rouen Normandie	Examinatrice

Contents

Chapter 1

Introduction

In today's increasingly interconnected industry and data-driven world, large-scale optimization problems have gained widespread prominence in modeling problems from various domains, such as in the field of energy management [SAB$^+$23], in the optimization of the social welfare [Wan17], in the training of neural networks [MMN18, MMM19]. A key characteristic of these optimization problems is the inclusion of an aggregate term in their objective function, which captures the interactions among individuals, supposed to be in great numbers. In general, resolving these optimization problems poses significant challenges due to their high dimensionality and non-convex nature.

Mean-field-type optimization problems constitute a powerful framework for addressing such problems. They offer an asymptotic perspective as the number of individuals tends to infinity. By solving mean-field-type optimization problems, we gain valuable insights into the solutions of the underlying aggregative problems with finitely many agents.

In the context of differential games, this asymptotic modeling approach yields equilibrium problems which are commonly referred to as *mean field games* (MFGs). They were introduced in 2006 independently by J.-M. Lasry and P.-L. Lions in [LL07] and M. Huang, R.P. Malhamé, and P.E. Caines in [HMC06]. MFGs have found significant applications in diverse areas, like crowd motion [LST10], sociology, biology, macroeconomics [ABL$^+$14], trade crowding [CL18a], and finance.

This thesis is dedicated to the numerical analysis and to methods for mean-field-type optimization problems. This introductory chapter is divided into two parts: In Section 1.1, we provide an overview of the general frameworks for constraint convex optimization and MFGs; in Section 1.2, we highlight the main results of the four chapters of the thesis.

- Chapter 2 and Section 1.2.1: **Large-scale nonconvex optimization: randomization, gap estimation, and numerical resolution**[1]. We address a large-scale and nonconvex optimization problem, involving an aggregative term. This term can be interpreted as the sum of the contributions of N agents to some common good, with N large. We investigate a relaxation of this problem, obtained by randomization. This relaxation lies in the framework of the MFO problem presented in Section 1.2.2. The relaxation gap is proved to have an order

[1] Chapter 2 corresponds to the article [BLO$^+$22] accepted for publication in *SIAM Journal on Optimization* and to the article [LOP22] accepted in *SIAM CT23*. The article [BLO$^+$22] is a joint work with J.F. Bonnans, N. Oudjane, L. Pfeiffer, and C. Wan. The article [LOP22] is a joint work with N. Oudjane and L. Pfeiffer.

$\mathcal{O}(1/N)$. We consider the Frank-Wolfe algorithm for the resolution of the randomized problem. Each iteration of the algorithm requires to solve a subproblem which can be decomposed into N independent optimization problems. A sublinear convergence rate is obtained for the FW algorithm. In order to handle the memory overflow problem possibly caused by the Frank-Wolfe algorithm, we propose a stochastic Frank-Wolfe algorithm, which ensures the convergence in both expectation and probability senses.

- Chapter 3 and Section 1.2.2: **Mean field optimization problems: stability results and Lagrangian discretization**[2]. We formulate and investigate a mean field optimization (MFO) problem involving probability distributions μ with a prescribed marginal m. The cost function depends on an aggregate term, which is the expectation of μ with respect to a contribution function. This problem is of particular interest in the context of Lagrangian MFGs and their discretization. We study the first-order optimality condition, prove strong duality, and investigate stability properties of the MFO problem from both primal and dual perspectives. In our stability analysis, we propose a feasible approach for recovering an approximate solution to an MFO problem with the help of an approximate solution to an MFO with a different marginal m, typically an empirical distribution. We combine this recovery method with the stochastic Frank-Wolfe algorithm stated in Section 1.2.1 to derive a complete resolution method.

- Chapter 4 and Section 1.2.3: **Error estimates of a theta-scheme for second-order mean field games**[3]. We introduce and analyze a new finite-difference scheme, relying on the theta-method, for solving monotone second-order MFGs. These games consist of a coupled system of the Fokker-Planck and the Hamilton-Jacobi-Bellman equation. The theta-method is used for discretizing the diffusion terms: we approximate them with a convex combination of an implicit and an explicit term. On contrast, we use an explicit centered scheme for the first-order terms. Assuming that the running cost is strongly convex and regular, we first prove the monotonicity and the stability of our theta-scheme, under a CFL condition. Taking advantage of the regularity of the solution of the continuous problem, we estimate the consistency error of the theta-scheme. Our main result is a convergence rate of order $\mathcal{O}(h^r)$ for the theta-scheme, where h is the step length of the space variable and $r \in (0,1)$ is related to the Hölder continuity of the solution of the continuous problem and some of its derivatives.

- Chapter 5 and Section 1.2.4: **A mesh-independent method for second-order potential mean field games**[4]. This part investigates the convergence of the generalized Frank-Wolfe algorithm for the resolution of potential and convex second-order MFGs. More specifically, the impact of the discretization of the MFG system on the effectiveness of the generalized Frank-Wolfe algorithm is analyzed. The article focuses on the theta-scheme introduced in Section 1.2.3. A sublinear and a linear rate of convergence are obtained, for two different choices of step sizes. These rates have the mesh-independence property: the underlying convergence constants are independent of the discretization parameters.

[2]Chapter 3 is a joint work with L. Pfeiffer.
[3]Chapter 4 corresponds to article [BLP22] accepted in *ESAIM: M2AN*, this is a joint work with J.F. Bonnans and L. Pfeiffer.
[4]Chapter 5 is a joint work with L. Pfeiffer.

1.1 Elements of convex analysis and mean-field-game models

Section 1.1.1 contains a recap of the duality theory for convex optimization problems and a presentation of the Frank-Wolfe (FW) algorithm, its generalization, and convergence results. We next present the different potential MFG models that we have investigated. For each of them, we explain how the FW algorithm can be applied and we highlight the interest of this method. Section 1.1.2 is dedicated to second-order MFGs. We present in Section 1.1.3 a general framework for discrete MFGs, which contains the finite difference scheme for second-order MFGs investigated in this thesis. In Section 1.1.4, we describe two potential problems associated with Lagrangian MFGs. They fit into a general class of optimization problems that we refer to as Mean Field Optimization (MFO) problems. We give a precise formulation of MFO problems and describe the associated optimality condition.

1.1.1 Convex optimization and the FW algorithm

We first introduce a general class of convex optimization problems and explore their relation with their dual problem. This provides us with a theoretical background for the application of various primal algorithms and primal-dual algorithms. Next, we focus on the Frank-Wolfe algorithm and its variants and present convergence results.

Constraint convex optimization and duality. Let $\mathcal{H}$ be a real Hilbert space and let $f\colon \mathcal{H} \to \mathbb{R}$ be a convex and lower semi-continuous (l.s.c.) function. Consider the following general optimization problem:

$$\inf_{x\in\mathcal{K}} \; f(x), \tag{1.1.1}$$

where $\mathcal{K}$ is a non-empty, bounded, closed, and convex subset of $\mathcal{H}$. The existence of a solution to problem (1.1.1) derives from [Bré11, Cor. 3.23]. It is convenient to consider the following equivalent formulation of (1.1.1):

$$\inf_{x\in\mathcal{H}} \; f(x) + \chi_{\mathcal{K}}(x), \tag{1.1.2}$$

where $\chi_{\mathcal{K}}$ denotes the indicator function of $\mathcal{K}$, which is 0 for $x \in \mathcal{K}$ and $+\infty$ for $x \notin \mathcal{K}$. The dual problem of (1.1.2) reads [BC11, Def. 15.19]:

$$\sup_{\lambda\in\mathcal{H}} \; \big(- f^*(\lambda) + \inf_{x\in\mathcal{K}} \langle \lambda, x\rangle \big), \tag{1.1.3}$$

where f^* is the Fenchel conjugate of f. Since f and $\chi_{\mathcal{K}}$ are convex and l.s.c., the domain of f is $\mathcal{H}$, and $\mathcal{K}$ is non-empty, we deduce from the Fenchel-Rockafellar theorem [Roc97][BC11, Thm. 15.23] that the dual problem (1.1.3) has a solution and that the strong duality holds for problem (1.1.1), i.e.,

$$\mathbf{val}(1.1.1) = \mathbf{val}(1.1.3).$$

Let us define the *Lagrangian* associated with (1.1.1),

$$\mathcal{L}\colon \mathcal{K} \times \mathcal{H} \to \mathbb{R},\ (x,\lambda) \mapsto -f^*(\lambda) + \langle \lambda, x\rangle.$$

We have obtained the existence of solutions for both the primal problem (1.1.1) and the dual problem (1.1.3). Additionally, given $(x^*, \lambda^*) \in \mathcal{K} \times \mathcal{H}$, the following two assertions are equivalent:

1. The point x^* is a solution of (1.1.1) and the point λ^* is a solution of (1.1.3).

2. The point (x^*, λ^*) is a saddle point of the Lagrangian $\mathcal{L}$, i.e.,

$$\inf_{x \in \mathcal{K}} \mathcal{L}(x, \lambda^*) = \mathcal{L}(x^*, \lambda^*) = \sup_{\lambda \in \mathcal{H}} \mathcal{L}(x^*, \lambda). \tag{1.1.4}$$

This equivalence provides us with two general perspectives for solving (1.1.1): (1) We can focus on the primal problem and compute a sequence of feasible candidates, letting the cost function decrease and converge to the optimum value; (2) We can find a saddle point (x^*, λ^*) of the Lagrangian $\mathcal{L}$, alternating between minimization and maximization phases for the primal and dual variables.

In the realm of numerical optimization, numerous algorithms have been developed to tackle problem (1.1.1). Among these, first-order algorithms have gained popularity thanks to their convergence guarantees and ease of implementation.

From the primal perspective, problem (1.1.1) can be solved using the FW algorithm [FW56] (also known as the conditional gradient algorithm [DH78]). Some algorithms having a forward-backward splitting structure can also be employed, such as the projected gradient descent [CW05] and its accelerated variant known as "FISTA" [BT09], the mirror descent [BT03] and its accelerated version of the entropic descent algorithm [KBB15].

From the primal-dual perspective (or "Arrow-Hurwicz" type), we can utilize the (inexact) Uzawa algorithm [EG94] to address the saddle point system (1.1.4). Recently, the Chambolle-Pock algorithm and its accelerated variant were investigated in [CP11, CP16] for resolving (1.1.4).

The successful application of the aforementioned first-order algorithms relies on the availability of certain "oracles". For instance, in the case of the FW algorithm, it is assumed that some linearized version of (1.1.1) (defined in (1.1.5)) is computationally tractable. Concerning primal-dual algorithms, knowledge of the prox operator of f and the ability to perform projections onto $\mathcal{K}$ are required.

The Frank-Wolfe algorithm. The FW algorithm was initially introduced in [FW56] for solving quadratic programming problems. Subsequent research in [DH78] demonstrated that the algorithm converges in a more general setting. Some variants of the FW algorithm can be found in [Jag13, LJJ15]. The FW algorithm, applied to problem (1.1.1), is described in Algorithm 1.1. It requires the following assumption:

- The function f is continuously differentiable and its gradient ∇f is Lipschitz continuous.

Let us mention that the subproblem (1.1.5) in Algorithm 1.1 is equivalent to minimize the first-order Taylor expansion of f at point x^k, i.e.,

$$\inf_{x \in \mathcal{K}} f(x^k) + \langle \nabla f(x^k), x - x^k \rangle.$$

Let us also mention that the resolution of the subproblem is equivalent to the evaluation of the dual criterion (in (1.1.3)), for $\lambda = \nabla f(x^k)$. The evaluation of $f^*(\lambda)$ is direct, since by the Fenchel-Young inequality, we have $f^*(\lambda) = \langle \lambda, x^k \rangle - f(x^k)$.

Algorithm 1.1: Frank-Wolfe Algorithm

Initialization: $x^0 \in \mathcal{K}$;
for $k = 0, 1, 2, \dots$ **do**
  Find a solution $\bar{x}^k$ to the subproblem

$$\inf_{x \in \mathcal{K}} \langle \nabla f(x^k), x \rangle; \tag{1.1.5}$$

  Choose $\omega_k \in [0, 1]$;
  Set $x^{k+1} = (1 - \omega_k) x^k + \omega_k \bar{x}^k$;
end

A commonly used choice for the learning rate ω_k in Algorithm 1.1 is $2/(k+2)$, as suggested in [DH78, Jag13]. Another approach for determining ω_k, known as the line-search method, consists in solving the following one-dimensional optimization problem:

$$\omega_k \in \underset{\omega \in [0,1]}{\operatorname{argmin}} \; f(x^k + \omega(\bar{x}^k - x^k)),$$

where $\bar{x}^k$ is the solution of the subproblem (1.1.5). Alternatively, if we know the Lipschitz constant L of ∇f, we can use the following one-dimensional quadratic programming approach to determine ω_k, in which the cost function is a quadratic majorization of $f(x^k + \omega(\bar{x}^k - x^k))$:

$$\omega_k \in \underset{\omega \in [0,1]}{\operatorname{argmin}} \; f(x^k) + \langle \nabla f(x^k), \bar{x}^k - x^k \rangle \omega + \frac{L \| \bar{x}^k - x^k \|^2}{2} \omega^2. \tag{1.1.6}$$

It has been proven (see [DH78, Jag13]) that for all the choices of ω_k mentioned above, Algorithm 1.1 exhibits a sub-linear convergence rate. This means that there exists a constant C, independent of k, such that for any $k \geq 1$,

$$f(x^k) - \mathbf{val}(1.1.1) \leq \frac{C}{k}. \tag{1.1.7}$$

Remark 1.1.1. The interest of the FW algorithm relies on the (fast) resolution of the subproblem (1.1.5) at each iteration. Let us consider the case where $\mathcal{K}$ is a convex polyhedron in $\mathbb{R}^d$. Then the subproblem (1.1.5) corresponds to a linear programming problem, which can be efficiently solved using the simplex algorithm [Dan63]. It is worth noting that the simplex algorithm returns a vertex of $\mathcal{K}$ as a solution [Dan63]. This provides us with the intuition that when applying the FW algorithm to a problem involving probability measures, the solutions of the subproblems are Dirac measures (which can be seen as extreme points of sets of probability measures).

Generalized FW algorithm. In some cases, the function f is decomposed into the sum of two convex functions, where one is differentiable and the other may be non-smooth, as for example in "Lasso" regression [Tib96]. More precisely, we replace our original assumption on f by the following:

- The function f is expressed as $f = f_1 + f_2$, where f_1 is convex and l.s.c., f_2 is convex, continuously differentiable, and its gradient ∇f_2 is Lipschitz continuous.

If f_1 is not differentiable, then the subproblem (1.1.5) is not well-defined. To overcome this limitation, Bredies et al. [BLM09] proposed the generalized Frank-Wolfe (GFW) algorithm, described in Algorithm 1.2. It is also referred to as the generalized conditional gradient algorithm. In their work, they established a connection between the GFW algorithm and the "ISTA" algorithm [DDDM04], which is widely utilized for solving inverse problems. Compared to Algorithm 1.1, Algorithm 1.2 introduces a modification: (1.1.5) is replaced by a new subproblem, which results from a partial linearization of f, denoted by $f_{\text{lin},y}$ and defined as follows: For any $y \in \mathcal{K}$,

$$f_{\text{lin},y} \colon \mathcal{K} \to \mathbb{R},\ x \mapsto f_1(x) + \langle \nabla f_2(y), x \rangle. \tag{1.1.8}$$

Algorithm 1.2: Generalized Frank-Wolfe Algorithm

... as Algorithm 1.1, except replacing (1.1.5) by

$$\inf_{x \in \mathcal{K}} \ f_{\text{lin},x^k}(x); \tag{1.1.9}$$

As in the original FW algorithm, if we choose the learning rate $\omega_k = 2/(k+2)$, then Algorithm 1.2 exhibits a sub-linear convergence rate, as shown in (1.1.7). In a recent study [KW22], several convergence results have been established for the GFW algortihm, under various assumptions. In particular, an improved convergence rate can be achieved when f_1 is assumed to be strongly convex. In this case, the learning rate ω_k is determined by solving a similar quadratic programming problem to (1.1.6),

$$\omega_k \in \underset{\omega \in [0,1]}{\operatorname{argmin}} \ -\beta_k \omega + \frac{L' \|\bar{x}^k - x^k\|^2}{2} \omega^2, \tag{1.1.10}$$

where $\bar{x}^k$ is a solution of (1.1.9), L' is the Lipschitz constant of ∇f_2, and

$$\beta_k = f_{\text{lin},x^k}(x^k) - f_{\text{lin},x^k}(\bar{x}^k), \tag{1.1.11}$$

which is positive by the definition of $\bar{x}^k$. Following the terminology in [LJ16], we call β_k the "Frank-Wolfe" gap at the point x^k. One can deduce from the convexity of f that x^k is a solution to the primal problem if and only if $\beta_k = 0$. Under the rule (1.1.10), Algorithm 1.2 exhibits a linear convergence rate. In other words, there exist positive constants C and $\delta \in (0,1)$ such that for any $k \geq 1$,

$$f(x^k) - \mathbf{val}(1.1.1) \leq C\delta^k. \tag{1.1.12}$$

1.1.2 Second-order MFGs

As mentioned at the beginning of this chapter, second-order MFGs describe the asymptotic behavior of Nash equilibria in N-player stochastic differential games, as the number of players N tends to infinity. In this section, we begin by considering a specific N-player differential game. From there, we introduce the MFG associated with this N-player game, which takes the form of a forward-backward partial differential equation (PDE) system. Next, we discuss the case of potential MFGs, for which the aforementioned PDE system can be interpreted as the first-order optimality condition associated with an optimal control problem of the Fokker-Planck equation.

An N-player differential game. We introduce an N-player differential game with a fixed time horizon $T > 0$. The notation $\mathbb{T}^d$ represents the d-dimensional torus, $\mathcal{P}(\mathbb{T}^d)$ denotes the set of probability distributions on $\mathbb{T}^d$, and $\mathcal{P}(T, \mathbb{T}^d)$ denotes the set of flows of distributions on $\mathbb{T}^d$, i.e., $\{m \mid m_t \in \mathcal{P}(\mathbb{T}^d), 0 \leq t \leq T\}$. Let Q denote $[0,T] \times \mathbb{T}^d$. The dynamic of each player i is governed by the following stochastic differential equation (SDE):

$$dx_t^i = v_t^i dt + \sqrt{2\sigma} dW_t^i, \quad \text{for } 0 \leq t \leq T, \qquad x_0^i \sim m_0^c. \tag{1.1.13}$$

Here, $\sigma > 0$, $m_0^c \in \mathcal{P}(\mathbb{T}^d)$ represents a fixed initial distribution, and $(W^i)_{i=1}^N$ are independent Brownian motions. The variable $(v_t^i)_{t\in[0,T]}$ is a stochastic process, adapted to the filtration generated by all Brownian motions and initial conditions of the agents.

Each player i has a cost function f^i defined as follows:

$$f^i(v^i, m^{-i}) := \mathbb{E}\left[\int_{t=0}^{T} \ell^c(t, x_t^i, v_t^i) + f^c(t, x_t^i, m_t^{-i}) dt + g^c(x_T^i)\right], \tag{1.1.14}$$

Here, $\ell^c\colon Q \times \mathbb{R}^d \to \mathbb{R}$ represents the running cost, $g^c\colon \mathbb{T}^d \to \mathbb{R}$ denotes the terminal cost, and $f^c\colon Q \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is a coupling function. Additionally, $m^{-i} \in \mathcal{P}(T, \mathbb{T}^d)$ is the flow of empirical distributions $m_t^{-i} = \frac{1}{N-1}\sum_{j\neq i} \delta_{x_t^j}$.

In general, addressing the N-player game system (1.1.13)-(1.1.14) becomes challenging as the number of players N increases. However, we can make the following two observations on the system (1.1.13)-(1.1.14):

- As N grows, the contribution from each individual player becomes negligible compared to the overall distribution m^{-i}.
- If we fix the second variable of f^i, i.e., the empirical distribution m^{-i}, then all players have a common best response strategy, which is the solution of (1.1.16) as stated later.

These observations lead to the MFG stated in the following paragraph, which captures the behavior of the entire population rather than focusing on each individual player.

Second-order MFGs. In the MFG associated with the N-player game (1.1.13)-(1.1.14), we fix the second variable of f^i for each i by a common flow in $\mathcal{P}(T, \mathbb{T}^d)$. The Nash equilibrium of this MFG is a pair (v^*, m^*) satisfying the following fixed-point system:

$$\begin{cases} v^* \in \mathbf{BR}^c(m^*), \\ m^* = \mathrm{Law}(x^{v^*}). \end{cases} \tag{1.1.15}$$

Here, $\mathbf{BR}^c$ is the best response mapping which returns the solution of the following individual optimal control problem, in which m^* is seen as a parameter:

$$\begin{cases} \inf_v \mathbb{E}\left[\int_{t=0}^{T} \ell^c(t, x_t, v_t) + f^c(t, x_t, m_t^*) dt + g^c(x_T)\right], \\ \text{s.t. } dx_t = v_t dt + \sqrt{2\sigma} dW_t, \quad \text{for } 0 \leq t \leq T, \qquad x_0 \sim m_0^c. \end{cases} \tag{1.1.16}$$

Simultaneously, $\text{Law}(x^{v^*})$ represents the flow of probability measures $\text{Law}(x_t^{v^*})_{t\in[0,T]}$, where x^{v^*} satisfies the following SDE:

$$dx_t^{v^*} = v_t^* dt + \sqrt{2\sigma} dW_t, \quad \text{for } 0 \le t \le T, \qquad x_0^{v^*} \sim m_0^c. \tag{1.1.17}$$

Hamilton-Jacobi-Bellman equation. A widely used technique for addressing the stochastic optimal control problem (1.1.16) relies on the Hamilton-Jacobi-Bellman (HJB) equation [FR75, BCD97]. To analyze (1.1.16), we introduce the value function $u^* \colon Q \to \mathbb{R}$, defined as follows:

$$u^*(t,x) = \begin{cases} \inf_v \mathbb{E}\left[\int_{\tau=t}^T \ell^c(t,x_t,v_t) + f^c(t,x_t,m_t^*)dt + g^c(x_T)\right], \\ \text{s.t. } dx_\tau = v_\tau d\tau + \sqrt{2\sigma} dW_\tau, \quad \text{for } t \le \tau \le T, \qquad x_t = x. \end{cases}$$

Under suitable assumptions, u^* is the viscosity solution [FR75] of the following HJB equation:

$$\begin{cases} -\partial_t u - \sigma\Delta u + H^c\left(t,x,\nabla u(t,x)\right) = f^c(t,x,m_t^*), & \forall (t,x) \in Q, \\ u^*(T,x) = g^c(x), & \forall x \in \mathbb{T}^d. \end{cases} \tag{1.1.18}$$

where the Hamiltonian H^c is related to the Fenchel conjugate of ℓ^c:

$$H^c(t,x,p) = \sup_{v\in\mathbb{R}^d} \langle -p, v\rangle - \ell^c(t,x,v).$$

Optimal control. In the case where H^c is differentiable with respect to its third variable, we can derive the following closed formula for the optimal control of (1.1.16):

$$v^*(t,x) = -H_p^c(t,x,\nabla u^*(t,x)), \quad \forall (t,x) \in Q. \tag{1.1.19}$$

Note that here the optimal control is expressed in feedback form, that is to say, as a function of time and state (and not as a stochastic process). When the optimal feedback is utilized, the state is the solution to the closed-loop equation:

$$dx_t^{v^*} = v_t^*(t, x_t^{v^*})dt + \sqrt{2\sigma} dW_t, \quad \text{for } 0 \le t \le T, \qquad x_0^{v^*} \sim m_0^c. \tag{1.1.20}$$

From now on, we only consider optimal controls described in a feedback form.

Fokker-Planck equation. Suppose that v^* is Lipschitz continuous. It is proved in [Car10] that m^* (the flow of the distribution of the solution of (1.1.20)) is a weak solution of the following Fokker-Planck (FP) equation:

$$\begin{cases} \partial_t m - \sigma\Delta m + \text{div}(v^* m) = 0 & \forall (t,x) \in Q, \\ m(0,x) = m_0^c(x), & \forall x \in \mathbb{T}^d. \end{cases} \tag{1.1.21}$$

MFG equations. Let us combine equations (1.1.18)-(1.1.21). The fixed point problem (1.1.15) is equivalent to the following coupled PDE system, including a backward HJB equation and a forward

FP equation:

$$\begin{cases} \text{(i)} & -\partial_t u - \sigma\Delta u + H^c(t,x,\nabla u(t,x)) = f^c(t,x,m(t)) & (t,x)\in Q, \\ \text{(ii)} & v(t,x) = -H^c_p(t,x,\nabla u(t,x)) & (t,x)\in Q, \\ \text{(iii)} & \partial_t m - \sigma\Delta m + \operatorname{div}(vm) = 0 & (t,x)\in Q, \\ \text{(iv)} & m(0,x) = m_0^c(x), \quad u(T,x) = g^c(x) & x\in\mathbb{T}^d. \end{cases} \tag{MFG}$$

The existence of classical solutions of (MFG) can be established under appropriate regularity assumptions, relying on Hölder estimates for parabolic equations and on fixed point arguments, see [LL07, Car10, BHP21, Kob22] for example. The existence of a weak solution for a variant of (MFG) with possibly degenerate coefficients and local coupling is proved in [CGPT15] by a duality method.

To ensure uniqueness of the solution to (MFG), a widely adopted assumption is the Lasry-Lions monotonicity assumption [LL07] for the coupling cost f^c: For any $t\in[0,T]$, for any m_1 and $m_2\in\mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d}\Big(f^c(t,x,m_1) - f^c(t,x,m_2)\Big)\big(m_1(x)-m_2(x)\big)dx \geq 0. \tag{1.1.22}$$

Potential MFGs. As introduced in [LL07] and studied in [CGPT15, BCS17, BHP21], the system (MFG) is said to be potential (or variational) if there exists a function $F^c\colon[0,T]\times\mathcal{P}(T,\mathbb{T}^d)\to\mathbb{R}^d$ such that for any $t\in[0,1]$ and $m_1,m_2\in\mathcal{P}(\mathbb{T}^d)$,

$$F^c(t,m_1) - F^c(t,m_2) = \int_0^1\int_{x\in\mathbb{T}^d} f^c(t,x,m_1+s(m_2-m_1))(m_2(x)-m_1(x))dxds. \tag{1.1.23}$$

In the presence of such F^c, system (MFG) can be interpreted as the first-order optimality condition of an optimal control problem driven by the FP equation,

$$\begin{cases} \inf\limits_{(m,v)} \displaystyle\int_Q \ell^c(t,x,v)m(t,x)dtdx + \int_0^T F^c(t,m(t))dt + \int_{\mathbb{T}^d} g^c(x)m(T,x)dx, \\ \text{such that } \begin{cases} \partial_t m - \sigma\Delta m + \operatorname{div}(vm) = 0, & \forall (t,x)\in Q, \\ m(0,x) = m_0^c(x), & \forall x\in\mathbb{T}^d. \end{cases} \end{cases} \tag{1.1.24}$$

Under the monotonicity condition (1.1.22), F^c is convex with respect to its second variable. However, problem (1.1.24) itself remains non-convex due to the lack of convexity of the term $\ell^c(v)m$ and the non-convexity of the admissible set. Fortunately, by employing the classical Benamou-Brenier transform, which maps (m,v) to $(m,w)=(m,mv)$ (see [BCS17] for more details), problem (1.1.24) can be transformed into an equivalent convex problem, given by:

$$\begin{cases} \inf\limits_{(m,w)} \displaystyle\int_Q \tilde{\ell}^c[m,w](t,x)dtdx + \int_0^T F^c(t,m(t))dt + \int_{\mathbb{T}^d} g^c(x)m(T,x)dx, \\ \text{such that } \begin{cases} \partial_t m - \sigma\Delta m + \operatorname{div}(w) = 0, & \forall (t,x)\in Q, \\ m(0,x) = m_0^c(x), & \forall x\in\mathbb{T}^d, \end{cases} \end{cases} \tag{1.1.25}$$

where the function $\tilde{\ell}^c[m,w]\colon Q \to \bar{\mathbb{R}}$ is the perspective function of ℓ^c and is defined by

$$\tilde{\ell}^c[m,w](t,x) = \begin{cases} \ell^c\Big(t,x,\frac{w(t,x)}{m(t,x)}\Big)m(t,x), & \text{if } m(t,x) \neq 0, \\ 0, & \text{if } m(t,x) = w(t,x) = 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{1.1.26}$$

If the MFG system (MFG) has a classical solution, then we can restrict the problem (1.1.25) to the Hilbert space $\mathbb{L}^2(Q) \times \mathbb{L}^2(Q,\mathbb{R}^d)$. In this scenario, problem (1.1.25) fits into a similar framework to the abstract optimization problem (1.1.1). We refer to [LL07, CGPT15, BCS17] for the dual problem of (1.1.25), which is an optimal control problem driven by the HJB equation.

Numerical methods exploiting the convexity of (1.1.25) can be applied to find a solution of (1.1.25), such as the fictitious play [CH17, HS19] and the GFW algorithm [LP22] from the primal perspective, the ADMM (Alternating direction method of multipliers) algorithm [BC15, And17] and the Chambolle-Pock algorithm [AL20] from the primal-dual perspective.

1.1.3 Discrete MFGs

We describe in this section a class of fully discrete MFGs (discrete in time and space), with a potential structure. This class is inspired from [BLP23] and the early reference [GMS10] about discrete MFG models. The main motivation behind our discrete MFG is to provide an abstract framework for the analysis and the resolution of the finite-difference method developed in Chapter 4.

Data of discrete MFG. Given a finite set A, we denote by $\mathbb{R}(A)$ (resp. $\mathbb{R}^d(A)$) the set of functions from A to $\mathbb{R}$ (resp. $\mathbb{R}^d$). Fixing $T \in \mathbb{N}_+$ and a finite subset S of $\mathbb{R}^d$, let us define:

$$\mathcal{T} = \{0,1,\ldots,T{-}1\}, \quad \tilde{\mathcal{T}} = \{0,1,\ldots,T\}, \quad \mathcal{P}(\tilde{\mathcal{T}},S) = \Big\{m \in \mathbb{R}(\tilde{\mathcal{T}} \times S) \;\big|\; \forall t \in \tilde{\mathcal{T}},\, m(t,\cdot) \in \mathcal{P}(S)\Big\}.$$

As in the continuous case, for the description of our discrete MFG model, we need a running cost ℓ, a coupling cost f, an initial condition m_0, and a terminal cost g, where

$$\ell\colon \mathcal{T} \times S \times \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}, \qquad f\colon \mathcal{T} \times S \times \mathbb{R}(S) \to \mathbb{R}, \qquad m_0 \in \mathcal{P}(S), \qquad g \in \mathbb{R}(S).$$

For a given $m \in \mathcal{P}(\tilde{\mathcal{T}},S)$, we denote by $\bar{\ell}_m$ the map defined by

$$\bar{\ell}_m\colon (t,x,\omega) \in \mathcal{T} \times S \mapsto \ell(t,x,\omega) + f(t,x,m(t)),$$

where $m(t) = (m(t,x))_{x\in S}$. To formulate the discrete MFG system, we also need a control bound $\bar{D} > 0$. The admissible control space, denoted by $\mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$, is the set of all elements v in $\mathbb{R}^d(\mathcal{T} \times S)$ such that $\|v(t,x)\| \leq \bar{D}$ for any (t,x). The probability of the motion from one state $x \in S$ to another state $y \in S$ at a time $t \in \mathcal{T}$ under some control $v \in \mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$ is given by

$$\pi[v](t,x,y) \coloneqq \pi(t,x,y,v(t,x)),$$

where π is a function from $\mathcal{T} \times S \times S \times \mathbb{R}^d$ to $\mathbb{R}$. We will assume that π is an affine function with respect to the last variable (the control variable).

Compared to the fixed-point problem (1.1.15) for continuous MFGs, the Nash equilibrium of our discrete MFG is a pair $(\bar{v}, \bar{m})$ satisfying the following:

$$\begin{cases} \bar{v} \in \mathbf{BR}^d(\bar{m}), \\ \bar{m} = \big(\mathrm{Law}(X_t^{\bar{v}})\big)_{t\in\tilde{\mathcal{T}}}. \end{cases} \tag{1.1.27}$$

Here $\mathbf{BR}^d$ is the best response mapping corresponding to a stochastic optimal control problem. For the sake of concision, we directly consider controls in feedback form, i.e. as deterministic functions of time and space. So $\mathbf{BR}^d(\bar{m})$ is the solution to

$$\inf_{v\in\mathbb{R}^d_{\bar{D}}(\mathcal{T}\times S)} \mathbb{E}\left[\sum_{t=0}^{T-1} \bar{\ell}_{\bar{m}}(t, X_t^v, v(t, X_t^v))\Delta t + g(X_T^v)\right], \tag{1.1.28}$$

where $(X_t^v)_{t\in\tilde{\mathcal{T}}}$ denotes a Markov chain satisfying

$$\mathbb{P}\big[X_{t+1}^v = y \mid X_t^v = x\big] = \pi[v](t,x,y), \text{ for } t \in \mathcal{T}, \quad X_0^v \sim m_0.$$

Formulation of discrete MFG. Similarly to the continuous case, the discrete fixed-point problem (1.1.27) is equivalent to a coupled system, involving the variables $u \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$, $v \in \mathbb{R}^d(\mathcal{T} \times S)$, and $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$:

$$\begin{cases} \text{(i)} \quad u = \mathbf{HJB}(m), \\ \text{(ii)} \quad v = \mathbf{V}(u), \\ \text{(iii)} \quad m = \mathbf{FP}(v), \end{cases} \tag{DMFG}$$

where the Hamilton-Jacobi-Bellman mapping **HJB**, the optimal control mapping **V**, and the Fokker-Planck mapping **FP** are defined as follows:

- Given $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, $u = \mathbf{HJB}(m) \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ is the solution to

$$\begin{cases} u(t,x) = \inf\limits_{\|\omega\|\le\bar{D}} \bar{\ell}_m(t,x,\omega)\Delta t + \sum_{y\in S} \pi(t,x,y,\omega)u(t+1,y), & \forall (t,x) \in \mathcal{T}\times S, \\ u(T,x) = g(x), & \forall x \in S. \end{cases} \tag{1.1.29}$$

- Given $u \in \mathbb{R}(\mathcal{T} \times S)$, $v = \mathbf{V}(u) \in \mathbb{R}^d(\mathcal{T} \times S)$ is defined by

$$v(t,x) = \underset{\|\omega\|\le\bar{D}}{\mathrm{argmin}}\ \ell(t,x,\omega)\Delta t + \sum_{y\in S} \pi(t,x,y,\omega)u(t+1,y), \quad \forall (t,x) \in \mathcal{T}\times S. \tag{1.1.30}$$

- Given $v \in \mathbb{R}^d(\mathcal{T} \times S)$, $m = \mathbf{FP}(v) \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ is defined as the solution to

$$\begin{cases} m(t+1,y) = \sum_{x\in S} \pi(t,x,y,v(t,x))m(t,x), & \forall (t,y) \in \mathcal{T}\times S, \\ m(0,x) = m_0(x), & \forall x \in S. \end{cases} \tag{1.1.31}$$

The existence of solutions of (DMFG) can be proved by applying the Brouwer fixed-point theorem to the mapping $\mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}$. The uniqueness of the solution of (DMFG) relies on a discrete version of the Lasry-Lions monotonicity condition (1.1.22): For any m_1 and $m_2 \in \mathcal{P}(S)$,

$$\sum_{x \in S} \big(f(t,x,m_1) - f(t,x,m_2)\big)\big(m_1(x) - m_2(x)\big) \geq 0. \tag{1.1.32}$$

Potential discrete MFG. Similarly to the potential system in the continuous case (1.1.23), we call (DMFG) a potential game if there exists $F \colon \mathcal{T} \times S \to \mathbb{R}$ such that for any $t \in \mathcal{T}$ and for any m_1 and m_2 in $\mathcal{P}(S)$, it holds

$$F(t,m_1) - F(t,m_2) = \int_0^1 \sum_{x \in S} f(t,x,m_1 + s(m_2 - m_1))(m_2(x) - m_1(x))ds. \tag{1.1.33}$$

This formulation allows us to consider a discrete time and finite state optimal control problem associated with (DMFG), similar to (1.1.24) in the continuous context. By applying the Benamou-Brenier transform to the aforementioned optimal control problem, we obtain the following equivalent convex problem (similar to (1.1.25) in the continuous context):

$$\inf_{\substack{m \in \mathcal{P}(\tilde{\mathcal{T}},S) \\ w \in \mathbb{R}^d(\mathcal{T} \times S)}} \tilde{J}(m,w), \quad \text{subject to: } (m,w) \in \tilde{\mathcal{A}}. \tag{1.1.34}$$

where the cost function $\tilde{J}$ and the set $\tilde{\mathcal{A}}$ are defined by

$$\begin{aligned}
&\tilde{J}(m,w) = \Delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \tilde{\ell}[m,w](t,x) + \Delta t \sum_{t \in \mathcal{T}} F(t,m(t)) + \sum_{x \in S} g(x) m(T,x); \\
&\tilde{\mathcal{A}} = \Big\{ (m,w) \in \mathcal{P}(\tilde{\mathcal{T}},S) \times \mathbb{R}^d(\mathcal{T} \times S) \,\Big|\, \exists\, v \in \mathbb{R}^d_D(\mathcal{T} \times S) \text{ such that } m = \mathbf{FP}(v),\ w = mv \Big\}.
\end{aligned}$$

Here, $\tilde{\ell}$ is the perspective function of ℓ, as the definition in (1.1.26). The equivalence between (DMFG) and (1.1.34) is discussed in the next paragraph.

A partially linearized problem. Noting that the coupling cost f corresponds to the directional derivative of F, one can easily express the linearization of the term $\sum_{t \in \mathcal{T}} F(t,m(t))$ in the objective function $\tilde{J}$. Exploiting the convexity of problem (1.1.34) along with this linearization, we are motivated to utilize the GFW algorithm 1.2 to solve (1.1.34). In this context, we consider the following partial linearization of $\tilde{J}$ at any point $m' \in \mathcal{P}(\tilde{\mathcal{T}},S)$ (which is similar to (1.1.8)): For any $(m,w) \in \mathcal{P}(\tilde{T},S) \times \mathbb{R}^d(\mathcal{T} \times S)$, let

$$\tilde{J}_{m'}(m,w) = \Delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \tilde{\ell}[m,w](t,x) + f(t,x,m'(t))m(t,x) + \sum_{x \in S} g(x) m(T,x).$$

The associated optimal control problem writes:

$$\inf_{(m,w) \in \tilde{\mathcal{A}}} \tilde{J}_{m'}(m,w). \tag{1.1.35}$$

Problem (1.1.35) plays the role of the partially linearized problem (1.1.9) of Algorithm 1.2 and will be solved at each iteration. Note that in (1.1.35), the third variable in f is fixed to m'. A major observation is that Problem (1.1.35) is equivalent to the stochastic optimal control problem (1.1.28). Therefore, (1.1.35) can be addressed by the dynamic programming principle. More precisely, the solution of (1.1.35) is given by the mapping $\mathbf{BR}\colon \mathcal{P}(\tilde{\mathcal{T}}, S) \to \tilde{\mathcal{A}}$ defined as follows: Given $m' \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, we obtain $(\tilde{m}, \tilde{w}) = \mathbf{BR}(m')$ by successively computing

$$\tilde{v} = \mathbf{V} \circ \mathbf{HJB}(m'), \quad \tilde{m} = \mathbf{FP}(\tilde{v}), \quad \text{and} \quad \tilde{w} = \tilde{m}\tilde{v}. \tag{1.1.36}$$

We refer to $\mathbf{BR}$ as the best-response mapping: Given a prediction m' of the equilibrium distribution of the agents, $\tilde{v}$ (as defined above) is the optimal feedback for the underlying optimal control problem and $\tilde{m}$ the resulting distribution.

If we assume furthermore that the running cost ℓ is α-convex with respect to its third variable, then for any $(m, w) \in \tilde{\mathcal{A}}$ and for any v such that $w = mv$, it holds that

$$\tilde{J}_{m'}(m, w) - \tilde{J}_{m'}(\tilde{m}, \tilde{w}) \geq \frac{\alpha}{2} \Delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \|(v - \tilde{v})(t, x)\|^2 m(t, x), \tag{1.1.37}$$

for m', $\tilde{m}$, $\tilde{v}$, and $\tilde{w}$ defined as above. Compared to (1.1.11), the term $\tilde{J}_{m'}(m, w) - \tilde{J}_{m'}(\tilde{m}, \tilde{w})$ plays the role of the "Frank-Wolfe" gap at point m' in the GFW algorithm applied to (1.1.34).

We mention that inequality (1.1.37) holds true not only in the potential case, but also in the non-potential case (when F is absent but the monotonicity assumption of f still holds). A similar inequality with the same nature for the continuous case has been investigated in [LP22, Lem. 29]. By combining (1.1.37) with the discrete monotonicity condition (1.1.32), we can establish the uniqueness of the solutions for both (DMFG) and (1.1.34), as well as their equivalence, as a consequence of 5.2.9:

1. System (DMFG) has a unique solution $(\bar{u}, \bar{v}, \bar{m})$.
2. The point $(\bar{m}, \bar{w})$ is the unique solution to (1.1.34), where $\bar{w} = \bar{m}\bar{v}$.

Inequality (1.1.37) also plays a crucial role in the proof of the fundamental inequality stated below, which is an essential part of the stability analysis of the finite-difference scheme presented in Section 1.2.3. Additionally, (1.1.37) serves as a key step in the convergence analysis of the GFW algorithm discussed in Section 1.2.4.

The fundamental inequality. The fundamental inequality, which is established for (DMFG), allows us to quantify the variation of the control variable v when the system (DMFG) is subject to perturbations. Let us consider a perturbed version of (DMFG) with additional terms $(\eta, \delta) \in \mathbb{R}^2(\mathcal{T} \times S)$ in the right-hand side:

$$\begin{cases} \text{(i)} \quad u = \mathbf{HJB}(m; \eta), \\ \text{(ii)} \quad v = \mathbf{V}(u), \\ \text{(iii)} \quad m = \mathbf{FP}(v; \delta), \end{cases} \tag{PDMFG}$$

where $\mathbf{HJB}(m;\eta)$ is defined by adding $\eta(t,x)$ to the right-hand-side of the first line in (1.1.29) for each $(t,x) \in \mathcal{T} \times S$. Similarly, $\mathbf{FP}(v;\delta)$ is defined by incorporating δ into (1.1.31) following the same approach.

Let (u,v,m) and $(\bar{u},\bar{v},\bar{m})$ be solutions of (PDMFG) and (DMFG), respectively. Assume that $m \geq 0$ and ℓ is α-convex with respect to its third variable. We can now state our fundamental inequality (see Proposition 4.3.7 for a rigorous statement):

$$\frac{\Delta t\alpha}{2}\sum_{t\in\mathcal{T}}\sum_{x\in S}\|(v-\bar{v})\|^2(m+\bar{m})(t,x) \leq \sum_{t\in\mathcal{T}}\sum_{x\in S}(u-\bar{u})(t+1,x)\delta(t,x)+(\bar{m}-m)\eta(t,x). \quad (1.1.38)$$

This inequality will be of key importance for the stability analysis of the theta-scheme.

1.1.4 MFO problems and Lagrangian MFGs

In this section, we introduce a class of convex optimization problems involving probability measures, which we call Mean-Field Optimization (MFO) problems. We first define MFO problems and provide a first-order optimality condition, which turns out to be equivalent to the Nash equilibrium conditions for a quite general class of games with non-atomic agents. Next we give two examples of potential Lagrangian MFGs which fit into this class.

MFO problems. Optimization problems involving probability measures have shown great promise in understanding and analyzing various aspects of neural networks (NN). One notable contribution in this regard is made in [MMN18], where the authors employ a mean-field approximation to investigate the asymptotic behavior of one-hidden layer NN as the number of neurons tends to infinity. In contrast, the article [CB18] focuses on the search for global minima of a convex function of a measure, through its many-particle limit. The authors explore the behavior of a non-convex particle gradient descent algorithm and establish its convergence properties.

Motivated by the previously mentioned articles, we introduce a general class of optimization problem that we refer to as Mean-Field Optimization problems. We consider two Polish spaces, denoted by X and Y (which are complete and separable metric spaces), and a closed subset Z of $X \times Y$. Let m be a probability measure on X and let g be a Borel measurable function mapping from Z to $\mathcal{H}$. Our objective is to solve the following problem parameterized by m:

$$\inf_{\mu\in\mathcal{P}_m(Z)} f\left(\int_Z g d\mu\right), \quad (\mathrm{P}_m)$$

where the integral $\int_Z g d\mu$ should be interpreted in the Bochner integration sense [Coh13, Appendix E]. The admissible set $\mathcal{P}_m(Z)$ is the set of all probability measures on Z whose marginal distribution on X is m. We assume that f is convex and differentiable, ∇f is L-Lipschitz, and g is bounded and continuous on Z. We call problem (P_m) an MFO problem, consistently with the terminology introduced in [CCRW23]. Problem (P_m) can be viewed as a social welfare optimization problem with considering nonatomic agents (the case with atomic agents is explored in [Wan17, BLO$^+$22]). In problem (P_m), the agents' positions are distributed according to a measure m, and the set Z represents the collection of feasible pairs of agent positions and strategies. The function $g(x,y)$

captures the contribution made by an agent located at position x and following strategy y to some common goods. The objective function f in (P_m) serves a similar purpose as in (P), representing a social cost evaluated at the aggregate term $\int_Z g d\mu$.

Remark 1.1.2. Problem (P_m) can be reformulated as a convex optimization problem of the form (1.1.1) by setting $\mathcal{K} = G_m \coloneqq \{\int_Z g d\mu, |, \mu \in \mathcal{P}_m(Z)\}$ in (1.1.1).

We next introduce a linearized version of (P_m) around a given measure $\hat{\mu}$, which appears in the application of the FW algorithm. It is given by

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \langle \lambda, g(x,y) \rangle d\mu(x,y) \tag{1.1.39}$$

where $\lambda = \nabla f \left(\int_Z g d\hat{\mu}\right)$. The solutions of this problem can be characterized through the best-response mapping $\mathbf{BR}_\lambda$, a set-valued mapping defined from X to subsets of Y as follows:

$$\mathbf{BR}_\lambda(x) = \operatorname*{argmin}_{y \in Z_x} \langle \lambda\, ,\, g(x,y) \rangle,$$

where Z_x is the set of $y \in Y$ such that $(x,y) \in Z$. Then, any $\mu \in \mathcal{P}_m(Z)$ is a solution to (1.1.39) if and only if

$$\operatorname{supp}(\mu_x) \subseteq \mathbf{BR}_\lambda(x), \quad \text{for } m\text{-a.e. } x \in X. \tag{1.1.40}$$

Here $(\mu_x)_{x \in X}$ denotes the disintegration of $\bar{\mu}$ (see Theorem 3.2.7) and $\operatorname{supp}(\mu_x)$ denotes the support of μ_x. This equivalence motivates the use of the FW algorithm: the linearized problem, originally posed on $\mathcal{P}_m(Z)$ can be solved through the evaluation of $\mathbf{BR}_\lambda$, that is to say, through the resolution of simpler problems, posed on Z.

We next discuss the first-order optimality condition associated with the problem (P_m). Given $\bar{\mu} \in \mathcal{P}_m(Z)$, we have that $\bar{\mu}$ is a solution to (P_m) if and only $\bar{\mu}$ is the solution to the linearized problem around $\bar{\mu}$ itself. A similar first-order optimality condition for the Lagrangian MFG can be found in [CH17, Eq. 3.13]. As a result, the MFO problem (P_m) is equivalent to the following equilibrium problem consisting in finding a pair (μ, λ) satisfying the following:

$$\begin{cases} \operatorname{supp}(\mu_x) \subseteq \mathbf{BR}_\lambda(x), & m\text{-a.e.} \\ \lambda = \nabla f \left(\int_Z g d\mu\right). \end{cases} \tag{1.1.41}$$

This coupled system can be interpreted as the Nash conditions for a game with non-atomic agents: an agent at position x must minimize $\langle \lambda, g(x, \cdot) \rangle$ and λ is a coupling variable, common to all agents, which results from their collective behavior μ. Several games fit into this framework, in particular, aggregative congestion games [LOW22] and nonatomic potential games [CL18b]. In the next paragraph, we give two examples of potential problems associated with Lagrangian MFGs.

Lagrangian MFGs. The concept of the Lagrangian MFGs follows from [BCS17, SS21, CC18, Sar22, MS19, CMS16]. In the context of the MFGs described with a coupled system of PDEs, we adopt an Eulerian point of view, which considers the evolution of the distribution of players over time. In contrast, the Lagrangian framework focuses on the distribution of the players' trajectories (which are not subject to Brownian perturbations).

Let us give a first example of a Lagrangian MFG, adapted from [Sar22]. Fix a domain $\Omega \subseteq \mathbb{R}^d$. Let $\mathrm{AC}([0,T],\mathbb{R}^d)$ be the set of all absolutely continuous functions from $[0,T]$ to $\mathbb{R}^d$. For any $x \in \Omega$, we denote,

$$\Gamma \coloneqq \{\gamma \in \mathrm{AC}([0,T],\mathbb{R}^d) \mid \gamma(t) \in \Omega, \forall t \in [0,T]\}, \qquad \Gamma_x \coloneqq \{\gamma \in \Gamma \mid \gamma(0) = x\}.$$

Let $Z = \{(x,\gamma) \mid x \in \Omega, \gamma \in \Gamma_x\}$. Let $m \in \mathcal{P}(\Omega)$ be the distribution of initial states of players. Define the admissible set of distribution of trajectories:

$$\mathcal{P}_m(Z) \coloneqq \{\mu \in \mathcal{P}(Z) \mid \pi_1 \# \mu = m\},$$

where $\pi_1 \colon Z \to \Omega, (x,\gamma) \mapsto x$. Define $e_t \colon \Gamma \to \Omega, \gamma \mapsto \gamma(t)$ and $\pi_2 \colon Z \to \Gamma, (x,\gamma) \mapsto \gamma$. Consider the following optimization problem:

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_0^T \int_Z \ell^c(t,\gamma(t),\dot{\gamma}(t)) dt d\mu(z) + \int_{t=0}^T F^c(t, e_t \# \pi_2 \# \mu) dt + \int_Z g^c(\gamma(T)) d\mu(z), \tag{1.1.42}$$

where ℓ^c and g^c follow the same definitions in Section 1.1.2 and F^c satisfies (1.1.23). The monotonicity condition on f^c implies that the function F^c is convex, as in the second-order case. It is easy to reformulate (1.1.42) as a particular case of an MFO problem. Moreover, utilizing the optimality condition (1.1.41), we can obtain an equivalent formulation of (1.1.42) as a Lagrangian MFG. This example is discussed more in detail in Chapter 3, Section 3.1.

We next present the potential formulation of another class of Lagrangian MFGs, involving a price variable, and taken from [GHS22]. For these MFGs, we take

$$Z = \Big\{ z = (x,\gamma) \in \Omega \times W^{1,\infty}(0,T;\mathbb{R}^d) \,\Big|\, x \in \Omega,\ \gamma(0) = x,\ \gamma(t) \in \Omega,\ \|\dot{\gamma}(t)\| \leq C, \text{ for a.e. } t \in (0,T) \Big\},$$

where $\Omega \subseteq \mathbb{R}^d$ and $C > 0$ are fixed. We next introduce a function $\Phi \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and consider the problem

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_0^T \int_Z \ell^c(t,\gamma(t),\dot{\gamma}(t)) dt d\mu(z) + \int_{t=0}^T \Phi\left(t, \int_Z \dot{\gamma}(t) d\mu(z)\right) dt + \int_Z g^c(\gamma(T)) d\mu(z), \tag{1.1.43}$$

where ℓ^c and g^c play the same as in second-order models. Again, this problem can be put in the form of an MFO problem and the associated optimality condition provides us with an interpretation of the problem as a Lagrangian MFG. More precisely, the variable λ (appearing in (1.1.41) plays the role of a price variable. We give more details about this class of problems in Section 3.6.

1.2 Contributions of the thesis

In this section, we provide a detailed description of the contributions of this thesis. Section 1.2.1 addresses a large-scale aggregative nonconvex optimization problem from both theoretical and algorithmic perspectives. Section 1.2.2 is dedicated to the study of the MFO problem (P_m), which generalizes the relaxed problem of Section 1.2.1. In Section 1.2.3, we explore a novel finite-difference discretization method (referred to as (Theta-mfg)) for the second-order system (MFG) and we analyze its convergence properties. Lastly, in Section 1.2.4, we focus on the "mesh-independent" convergence of the Generalized Frank-Wolfe algorithm for the problem (1.1.24), discretized with (Theta-mfg).

We mention that all constants C used in the rest of this chapter are independent from each other, unless explicitly indicated.

1.2.1 Large-scale nonconvex optimization: randomization, gap estimation, and numerical resolution

Framework and motivation. This section, associated with Chapter 2, is devoted to the theoretical analysis and the numerical resolution of the following large-scale, aggregative, and nonconvex optimization problem:

$$\inf_{x \in \mathcal{X}} J(x) \coloneqq f(G(x)), \qquad \text{where: } \begin{cases} G(x) = & \dfrac{1}{N} \displaystyle\sum_{i=1}^{N} g_i(x_i) \\ \mathcal{X} = & \prod_{i=1}^{N} \mathcal{X}_i. \end{cases} \tag{P}$$

Here, N can be seen as the number of agents and is assumed to be large. The key feature of this problem is the aggregative form of the function G, which is the average of N contribution mappings g_i. Each contribution mapping is defined on a set $\mathcal{X}_i$ and maps to a Hilbert space $\mathcal{H}$. The aggregate $G(x)$ represents the collective contribution of all agents. While very few structural assumptions are made on the sets $\mathcal{X}_i$ and the mappings g_i, we will assume that f is convex, with a Lipschitz-continuous gradientand that the image sets $g_i(\mathcal{X}_i)$ are all bounded. The central idea in Chapter 2 is that when N is large, the problem (P) can be well approximated by a convex problem.

Problem (P) finds applications in various domains including social welfare optimization [Wan17], power system management [SAB+23], and resource allocation problems [BBG+20]. It finds applications in supervised learning, such as training neural networks with one hidden layer [MMN18, CB18], sparse reconstruction problems [Mal09, MBP14], and the dual problem of linear support vector machines (SVM) [SST09, FR16].

Organization. We first propose a novel relaxation technique based on randomization to convexify problem (P). Then, we provide an upper bound of the relaxation gap, which is of order $\mathcal{O}(1/N)$. Next, we show that this gap estimate can be improved by studying a geometric relaxation of (P). From the algorithmic perspective, we introduce a method called the *selection method*, to reconstruct an approximate solution to the primal problem (P) out of an approximate solution to the relaxed problem. We derive from this method a general method, that we call *Stochastic Frank-Wolfe* (SFW) algorithm, for the resolution of (P). It combines the FW algorithm with the selection method. We provide a convergence result for the SFW algorithm. Finally, we apply the SFW algorithm to solve a mixed-integer quadratic programming problem and a high dimensional optimal control problem.

Randomized relaxation and a first gap. For any $i = 1, \ldots, N$, given a probability measure μ_i on $\mathcal{X}_i$, we denote by $E_{\mu_i}[g_i]$ the integral of g_i against the distribution μ_i, i.e., $\int_{\mathcal{X}_i} g_i d\mu_i$. The relaxed version of (P) is obtained by replacing the variables x_i by probability measures μ_i and by replacing the contribution mappings $g_i(x_i)$ by $E_{g_i}[\mu_i]$. The resulting randomized problem writes:

$$\inf_{\mu\in\mathcal{P}} \mathcal{J}(\mu) := f(E_\mu[G]), \qquad \text{where: } \begin{cases} E_\mu[G] = & \frac{1}{N}\sum_{i=1}^{N} E_{\mu_i}[g_i] \\ \mathcal{P} = & \prod_{i=1}^{N} \mathcal{P}(\mathcal{X}_i). \end{cases} \tag{PR}$$

Problem (PGR) is indeed a relaxation since for a tuple of Dirac measures $\mu = (\delta_{x_1}, \ldots, \delta_{x_N})$, we have $\mathcal{J}(\mu) = J(x_1, \ldots, x_N)$. One can see that the admissible set $\mathcal{P}$ is convex and that $E_\mu[G]$ is linear with respect to μ. Therefore, the relaxed problem (PR) is a convex optimization problem. We call relaxation gap the quantity

$$\mathbf{val}(\text{PG}) - \mathbf{val}(\text{PGR}).$$

It is easy to see that the relaxation gap is nonnegative.

Theorem 1.2.1 (Proposition 2.2.6 and Theorem 2.2.9)**.** *Let $\mu = (\mu_i)_{i=1,\ldots,N}$, where $\mu_i \in \mathcal{P}(\mathcal{X}_i)$ for any i. Let $(X_i)_{i=1,\ldots,N}$ be N independent random variables such that $X_i \sim \mu_i$. Then there exists a constant $C > 0$ independent of N such that for any $\epsilon > 0$,*

$$\mathbb{E}[J(X)] - \mathcal{J}(\mu) \leq \frac{C}{N}; \tag{1.2.1}$$

$$\mathbb{P}\left(J(X) - \mathcal{J}(\mu) \leq \frac{C}{N} + \epsilon\right) \geq 1 - \exp\left(-\frac{N\epsilon^2}{C}\right). \tag{1.2.2}$$

A consequence of the following result is that the relaxation gap is of order $\mathcal{O}(1/N)$. Here, we present a brief proof of (1.2.1). Let us consider $Y = \frac{1}{N}\sum_{i=1}^{N} g_i(X_i)$. Then, we have $\mathbb{E}[J(X)] = \mathbb{E}[f(Y)]$ and $\mathcal{J}(\mu) = f(\mathbb{E}[Y])$. By exploiting the independence of X_i, we can deduce that the variance of Y has an order of $\mathcal{O}(1/N)$. Using the Lipschitz continuity of ∇f, we obtain the inequality:

$$f(Y) \leq f\big(\mathbb{E}[Y]\big) + \big\langle \nabla f(\mathbb{E}[Y]), Y - \mathbb{E}[Y]\big\rangle + \frac{L}{2}\big\|Y - \mathbb{E}\big[Y\big]\big\|^2,$$

where L is the Lipschitz constant of ∇f. By taking the expectation on both sides of the inequality, we obtain (1.2.1), as the first-order term is zero and the quadratic term corresponds to the variance of Y. The proof of (1.2.2) relies on McDiarmid's inequality [McD89], a concentration inequality.

Geometric relaxation and a refined gap. It is convenient to write the primal problem (P) in an equivalent form:

$$\inf_{y\in\mathcal{H}} f(y), \quad \text{subject to: } y \in \mathcal{Y} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{Y}_i, \tag{PG}$$

where $\mathcal{Y}_i = g_i(\mathcal{X}_i)$ for any $i = 1, \ldots, N$. Indeed, by definition of $\mathcal{Y}$, any $y \in \mathcal{H}$ lies in $\mathcal{Y}$ if and only if there exists $x \in \mathcal{X}$ such that $y = \frac{1}{N}\sum_{i=1}^{N} g_i(x_i)$. For such an x, we have $f(y) = J(x)$. It is natural to consider the following relaxation:

$$\inf_{y\in\mathcal{H}} f(y), \quad \text{subject to: } y \in \text{conv}(\mathcal{Y}). \tag{PGR}$$

It is not difficult to observe that $\mathbf{val}(\mathrm{P}) = \mathbf{val}(\mathrm{PG})$ and $\mathbf{val}(\mathrm{PR}) = \mathbf{val}(\mathrm{PGR})$. Thus, the randomization gap is equal to $\mathbf{val}(\mathrm{PG}) - \mathbf{val}(\mathrm{PGR})$. Let q be the dimension of the aggregate space $\mathcal{H}$. We have the following refined gap estimate (in comparison to (1.2.1)).

Theorem 1.2.2 (Proposition 2.4.9)**.** *There exists a constant C, independent of N, such that*

$$\mathbf{val}(\mathrm{PG}) - \mathbf{val}(\mathrm{PGR}) \leq \frac{C \min\{q, N\}}{N^2}. \tag{1.2.3}$$

The proof of this sharper relaxation gap relies on a notion of measure of nonconvexity for sets, introduced in [Cas75]. Given a subset $\mathcal{K}$ of $\mathcal{H}$, we call nonconvexity measure of $\mathcal{K}$ the number $\rho(\mathcal{K})$ defined by

$$\rho(\mathcal{K}) = \left(\sup_{y \in \mathrm{conv}(\mathcal{K})} \inf_{\mathbb{E}[Y]=y} \mathrm{Var}(Y) \right)^{1/2},$$

where Y is a finitely supported random variable in $\mathcal{K}$ and $\mathrm{Var}(Y)$ is the variance of Y in the sense of [Vil03, Rem. 7.5], i.e. $\mathrm{Var}(Y) = \mathbb{E}\big[\|Y - \mathbb{E}[Y]\|^2\big]$. With the help of ρ, we can establish the following inequality, with a similar proof to the one of Theorem 1.2.1:

$$\mathbf{val}(\mathrm{PG}) - \mathbf{val}(\mathrm{PGR}) \leq \frac{L}{2}\rho(\mathcal{Y})^2 \leq \frac{L}{2N^2}\,\rho\left(\sum_{i=1}^N \mathcal{Y}_i\right)^2.$$

Next, we apply [Cas75, Thm. 2], which is derived from the Shapley-Folkman lemma, to obtain the following bound:

$$\rho\left(\sum_{i=1}^N \mathcal{Y}_i\right)^2 \leq \max_{\substack{Q \subseteq \{1,\dots,N\} \\ |Q| = \min\{q,N\}}} \sum_{i \in Q} \rho(\mathcal{Y}_i)^2.$$

Combining the preceding two inequalities, we obtain (1.2.3).

Stochastic Frank-Wolfe algorithm. From Theorem 1.2.1, we can conclude the following *selection method* to retrieve an approximate solution of (P) from an approximate solution of (PR):

- *Selection method.* Given μ_i a probability measure on $\mathcal{X}_i$ for any $i = 1, \dots, N$, we take N independent random variables X_i such that $X_i \sim \mu_i$.

Inequalities (1.2.1) and (1.2.2) indicate that the error resulting from the selection method decreases to 0 as N goes to infinity in both expectation and probability senses. The remaining question is to solve the randomized problem (PR). Let us consider the following linearization of the objective function $\mathcal{J}$ of (PR) at some point $\bar{\mu} \in \mathcal{P}$:

$$\langle \nabla f(E_{\bar{\mu}}[G])\,,\, E_\mu[G] \rangle = \frac{1}{N}\sum_{i=1}^N \left\langle \nabla f(E_{\bar{\mu}}[G])\,,\, \int_{\mathcal{X}_i} g_i d\mu_i \right\rangle.$$

As a consequence, if we want to minimize the above function on μ over $\mathcal{P}$, then it is equivalent to solve the following N subproblems: for $i = 1, \dots, N$,

$$\inf_{\mu_i \in \mathcal{P}(\mathcal{X}_i)} \left\langle \nabla f(E_{\bar{\mu}}[G])\,,\, \int_{\mathcal{X}_i} g_i d\mu_i \right\rangle. \tag{1.2.4}$$

The decomposition of the primal problem (P) into N smaller subproblems offers several advantages. Firstly, it allows for easier handling and enables parallel computation. Secondly, the subproblems (1.2.4) have solutions which are Dirac measures: if x_i minimizes $\langle f(E_{\bar{\mu}}[G]), g_i(\cdot)\langle$ over $\mathcal{X}_i$, then δ_{x_i} is a solution to (1.2.4). These properties motivate us to apply the FW algorithm (Algorithm 1.1) to solve problem (PR), in which the linearized problem (1.1.5) is equivalent to the N subproblems (1.2.4). We have a sublinear convergence rate for the FW algorithm 1.1 for problem (PR) by taking $\omega_k = 2/(k+2)$ as shown in (1.1.7).

A memory overflow problem can arise when applying the FW algorithm to problem (PR): The FW algorithm possibly requires storing N new points in the support of the solution at each iteration. This means that KN storage spaces are needed to store the result after K iterations, which can become prohibitive when the number of iterations K is large. This difficulty is mentioned in the related work [BSR17, CB18]. To overcome this problem, we propose a variant of the FW algorithm, in which the selection method is utilized n_k times each iteration k of the FW algorithm. We call this variant the *Stochastic Frank-Wolfe* (SFW) algorithm. It is presented in Algorithm 1.3.

Algorithm 1.3: Stochastic Frank-Wolfe Algorithm

```
Initialization: x^0 ∈ X;
for k = 0, 1, 2, ..., K do
    Step 1: Resolution of the subproblems.
    Compute y^k = (1/N) Σ_{i=1}^N g_i(x_i^k);
    for i = 1, 2, ..., N do
        Find x̄_i^k ∈ argmin_{x_i ∈ X_i} ⟨∇f(y^k), g_i(x_i)⟩;
    end

    Step 2: Update.
    Choose n_k ∈ ℕ*. Set ω_k = 2/(k + 2).;
    for j = 1, 2, ..., n_k do
        for i = 1, 2, ..., N do
            Simulate λ_i^{k,j} ~ Bern(ω_k), independently of all previously defined random variables;
            Set x̂_i^{k,j} = (1 − λ_i^{k,j}) x_i^k + λ_i^{k,j} x̄_i^k;
        end
        Set x̂^{k,j} = (x̂_i^{k,j})_{i=1,...,N};
    end
    Find x^{k+1} ∈ argmin{J(x) | x ∈ X^k}, where X^k = {x̂^{k,j}, j = 1, 2, ..., n_k} ∪ {x^k};
end
```

Starting from an initialization $x^0 \in \mathcal{X}$, Algorithm 1.3 generates a sequence $(x^k)_{k\in\mathbb{N}}$ in $\mathcal{X}$. For the analysis of the algorithm and for its description, it is convenient to introduce $\mu^k = (\delta_{x_1^k}, \dots, \delta_{x_N^k})$. With this notation at hand, we first observe that y^k, as defined in Step 1 of Algorithm 1.3, satisfies $y^k = \frac{1}{N}\sum_{i=1}^N E_{\mu_i^k}[g_i]$. Thus the Step 1 of Algorithm 1.3 plays the same role as (1.2.4) in the FW algorithm for (PR), which is decentralized and can be computed parallelly. Let us focus next on Step 2 of Algorithm 1.3 and let us define $\bar{\mu}^k = (\delta_{\bar{x}_i^1}, \dots, \delta_{\bar{x}_N^k})$ and $\hat{\mu}^k = (1-\omega_k)\mu^k + \omega_k\bar{\mu}^k$. In contrast with the learning procedure in the FW Algorithm, we do not directly use $\hat{\mu}^k$ at the next iteration but instead employ our selection method n_k times to generate n_k random variables $(\hat{x}^{k,j})_{j=1,\ldots,n_k}$. Finally, Step 2 selects a random variable $\hat{x}^{k,j}$ which minimizes J.

The following result concerns the convergence of Algorithm 1.3. As the outputs of Algorithm 1.3 are random variables, we provide convergence results in expectation and probability senses for the optimality gap

$$\gamma_k := J(x^k) - \mathbf{val}(\mathrm{PGR}).$$

Theorem 1.2.3 (Theorem 2.3.7 and Corollary 2.3.8)**.** *There exists a constant $C > 0$ such that*

$$\mathbb{E}[\gamma_K] \leq \frac{C}{K}, \qquad \text{for } K = 1, 2, \ldots, 2N. \tag{1.2.5}$$

Moreover, there exists $C' > 0$ such that for any $A > 0$, if we take $n_k > \max(\frac{Ak^2}{N}, 1)$ for any k, then

$$\mathbb{P}\Big[\gamma_K < \frac{C + C'}{K}\Big] \geq 1 - \exp\left(-\frac{A}{12}\right), \qquad \text{for } K = 1, 2, \ldots, 2N. \tag{1.2.6}$$

The proof of the theorem relies on a similar approach to the classical analysis of the FW algorithm. In a first step, we prove the following inequality:

$$\gamma_K \leq \frac{C}{K} + S_K,$$

where S_K is a random variable that accumulates the errors arising from the selection method at each iteration. Then (1.2.5) follows from the fact that $\mathbb{E}[S_K] = 0$. The proof of (1.2.6) is considerably more intricate, it requires the computation of an upper bound of $\mathbb{P}\big[S_K \geq \epsilon\big]$, obtained with an extension of McDiarmid's concentration inequality provided in [Del15, Thm. 7].

A more precise convergence result has been obtained in Theorem 2.3.7. In particular, we have a more precise formula for the probability in (1.2.6), which outlines the benefit of increasing the number of simulations n_k.

Numerical simulations. In Sections 2.6 and 2.7, we show the effectiveness of the SFW algorithm by applying it to two different problems: a mixed integer quadratic program (MIQP) and an aggregative battery charging problem discussed in [LOP22], which involves a high-dimensional optimal control setting. For the MIQP problem, our numerical results indicate that the SFW algorithm exhibits a superior convergence rate compared to the sublinear rate predicted by the theory. Moreover, in terms of computation time, the SFW algorithm outperforms popular solvers such as SCIP and GUROBI, for sufficiently large values of N. In the case of the battery charging problem, the SFW algorithm avoids the curse of dimensionality, which often plagues high-dimensional optimal control problems. By using the SFW algorithm, we can compute an approximate optimal control in a tractable manner.

Literature comparison and perspectives. If we assume that (P) is convex, then classical Lagrangian relaxation (Chapter XII of [HUL93]) methods can be applied. Thanks to the aggregative form of G, the dual problem of (P) has a separable structure, as explained in [SAB$^+$23, Pac18] in related contexts. As a result, the primal-dual algorithms mentioned in Section 1.1.1 can be employed to address (P). Furthermore, if we make additional assumptions regarding the regularity of g_i and $\mathcal{X}_i$, the block coordinate descent (BCD) algorithm (and its extensions, see [BT13, FR16]) could be utilized to tackle (P). Despite the nonconvexity of problem (P), numerical approaches leveraging the separability of the cost of the dual problem are particularly appealing, since they allow for a decomposed resolution of the problem. Yet they raise two difficulties: the potential large duality gap and the reconstruction of a primal solution from the dual optimal solution.

These two difficulties were already addressed by Wang in [Wan17]. She proposed a convex relaxation of the problem, based on the same geometrical approach of (PGR), that allows to obtain an estimate of the duality gap decreasing with N. Her main tool was the Shapley-Folkman lemma [Sta69], which allows to show that the image of G is close to a convex set.

In comparison to the theoretical results in [Wan17], we provide a sharper estimate of the relaxation gap. According to Theorem 1.2.2, the refined gap is of the order $\mathcal{O}(\min(q, N)/N^2)$, where q represents the dimension of $\mathcal{H}$. This estimate is tighter to the one obtained by applying [Wan17, Thm. 3.5], which yields a gap of the order $\mathcal{O}(q^2/N^2)$.

Let us compare our algorithmic approaches with the one in [Wan17]. Both approaches leverage the decomposability of the problem into N problems and require that the subproblems can be easily solved. The method in [Wan17, Algorithm 2] requires to compute the full set of ξ-optimal solutions, which is more demanding. Contrary to [Wan17], Algorithm 1.3 does not require to perform Shapley-Folkman decompositions. This is a major advantage when the dimension of the aggregate q is very large. As a counterpart, we are only able to find $\mathcal{O}(1/N)$-optimal solutions, while the algorithm of [Wan17] can find $\mathcal{O}(q^2/N^2)$-optimal solutions.

The design of a method for the computation of $\mathcal{O}(\min\{q, N\}/N^2)$-solutions will be the topic of future research. We also aim at working on more complex problems, involving for example convex constraints on the aggregate, as for example the resource allocation problems investigated in [BBG$^+$20]. Such constraints could be handled with extensions of the Frank-Wolfe algorithm for non-smooth costs as those proposed in [SFMF20, YFC19]. Finally, we intend to apply our method to large-scale optimal control problems, which have more complex structure than the battery example presented in this work, such as nonconvex variants of the problem investigated in [SAB$^+$23].

1.2.2 Mean field optimization problems: stability results and Lagrangian discretization

Framework and motivation. In Chapter 3 of this thesis, our focus is on the MFO problem (P_m), which is motivated by the price model of Lagrangian MFG (1.1.43). Recall the formulation of (P_m):

$$\inf_{\mu \in \mathcal{P}_m(Z)} f\left(\int_Z g d\mu\right). \tag{P_m}$$

We refer to Section 1.1.4 for a description of the applications of this class of problems and their connection with some potential Lagrangian MFG models. Let us emphasize that the relaxed problem associated with the aggregative problems of the previous section, problem (PR), fits into the general framework of MFO problems.

Organization. We start with the derivation of the first-order optimality condition of problem (P_m). Then, we study the stability of the primal problem (P_m) with respect to its parameter m. Additionally, we formulate the dual problem of (P_m), demonstrate the strong duality, and investigate the stability of the dual solution. At a numerical level, we propose to discretize the marginal m and to solve the resulting problem with the SFW algorithm (Algorithm 1.3). We study the convergence of this method. Finally, we perform some numerical simulations for a Lagrangian MFG model taken from [GHS22].

First-order optimality condition. Recall the definition of the individual best-response mapping $\mathbf{BR}_\lambda$ from X to subsets of Y:

$$\mathbf{BR}_\lambda(x) = \operatorname*{argmin}_{y \in Z_x} \langle \lambda \, , \, g(x,y) \rangle.$$

Theorem 1.2.4 (Corollary 3.3.5)**.** *Any $\bar{\mu} \in \mathcal{P}_m(Z)$ is a solution to (P_m) if and only if the following equilibrium conditions are satisfied:*

$$\begin{cases} \bar{\lambda} = \nabla f \left(\int_Z g d\bar{\mu} \right), \\ \operatorname{supp}(\bar{\mu}_x) \subseteq \mathbf{BR}_{\bar{\lambda}}(x), \quad m\text{-a.e.} \end{cases} \tag{1.2.7}$$

The proof follows the main steps described previously in Section 1.1.4. We first prove that $\bar{\mu}$ is a solution to (P_m) if and only $\bar{\mu}$ is the solution to a linearized problem around $\bar{\mu}$. Then, in a second step, we provide a characterization of the set of solutions of the linearized problems: at a technical level, the main difficulty of this step is related to the application of the measurable selection theorem.

Additionally, we provide an existence result for the solution under a “tightness” assumption on the minimizing sequence of (P_m).

Stability analysis. Let m_0 and m_1 be two probability measures in $\mathcal{P}(X)$. We consider two instances of problem (P_m) with $m = m_0$ and $m = m_1$, denoted as (P_{m_0}) and (P_{m_1}) respectively. We are first concerned with the variation of the optimal cost, in relation with the Kantorovich-Rubinstein distance d_1 between m_0 and m_1.

We perform the stability analysis under the following additional assumption: There exists $L_g > 0$ such that for any x_1 and x_2 in X, for any $y_1 \in Z_{x_1}$, there exists $y_2 \in Z_{x_2}$ such that

$$\|g(x_1, y_1) - g(x_2, y_2)\| \leq L_g d_X(x_1, x_2). \tag{1.2.8}$$

Theorem 1.2.5 (Theorem 3.3.16)**.** *There exists a constant $C > 0$, independent of m_0 and m_1, such that*

$$\left| \mathbf{val}(P_{m_0}) - \mathbf{val}(P_{m_1}) \right| \leq C d_1(m_0, m_1).$$

The proof of Theorem 1.2.5 relies on a recovery method, described in Algorithm 3.1, which bridges approximate solutions for problems (P_{m_0}) and (P_{m_1}). More precisely, given an approximate solution $\bar{\mu}_0$ for (P_{m_0}), the recovery method provides a way to derive an approximate solution μ_1 for (P_{m_1}). To provide an intuitive understanding of our recovery method, we consider the following particular case where $m_0 = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, $m_1 = \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{x}_i}$, and $\bar{\mu}_0 = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i,y_i)}$. Here $(x_i)_{i=1}^N, (\tilde{x}_i)_{i=1}^N \in X^N$ and $(y_i)_{i=1}^N \in \prod_{i=1}^N Z_{x_i}$. From [PC19, Prop. 2.1], there exists a permutation of $\{\tilde{x}_1, \dots, \tilde{x}_N\}$, denoted by $\{x'_1, \dots, x'_N\}$, such that $\rho^* = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i,x'_i)}$ is a solution of the following optimal transport problem:

$$\inf_{\rho\in\Pi(m_0,m_1)} \int_{X\times X} d_X(x,x')d\rho(x,x'), \tag{OT}$$

where $\Pi(m_0, m_1)$ is the set of all transference plans between m_0 to m_1. As a consequence,

$$d_1(m_0, m_1) = \int_{X\times X} d_X(x,x')d\rho^*(x,x') = \frac{1}{N}\sum_{i=1}^N d_X(x_i, x'_i).$$

By (1.2.8), for any i, there exists $y'_i \in Z_{x'_i}$ such that $\|g(x'_i, y'_i) - g(x_i, y_i)\| \leq L_g d_X(x_i, x'_i)$. In our recovery method, each x_i is transported to x'_i while simultaneously y_i is transported to the point $y'_i \in Z_{x'_i}$ for $i = 1, \dots, N$. This can be expressed as follows:

$$\bar{\mu}_0 = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i,y_i)} \longrightarrow \mu_1 = \frac{1}{N}\sum_{i=1}^N \delta_{(x'_i,y'_i)}. \tag{1.2.9}$$

The distribution μ_1 belongs to $\mathcal{P}_{m_1}(Z)$, moreover,

$$\left\|\int_Z g d\bar{\mu}_0 - \int_Z g d\mu_1\right\| = \left\|\frac{1}{N}\sum_{i=1}^N \big(g(x_i,y_i) - g(x'_i,y'_i)\big)\right\| \leq \frac{L_g}{N}\sum_{i=1}^N d_X(x_i,x'_i) = L_g d_1(m_0,m_1). \tag{1.2.10}$$

Therefore, by utilizing the special recovery method (1.2.9) for this particular example, we can obtain a feasible solution for (P_{m_1}) while effectively controlling the distance between the aggregate terms by the distance $d_1(m_0, m_1)$.

In the general case, where an optimal-assignment-type solution is not available for the optimal transport problem (OT), the Gluing lemma [Vil09, p. 11] is employed in the recovery method. This allows us to construct a probability measure $\nu \in \mathcal{P}(X \times Y \times X)$ such that its marginal distribution with respect to the first two variables is equal to $\bar{\mu}_0$ and such that the marginal distribution of the first and third variables is ρ^*, where ρ^* is a solution of (OT). Then, we introduce the set-valued function $S\colon Z \times X \rightsquigarrow Z$,

$$S(x,y,x') = \left\{(x',y') \in Z \mid \|g(x',y') - g(x,y)\| \leq L_g d_X(x,x')\right\}.$$

We prove that S admits a measurable selection s under suitable assumptions. In our general recovery method, the resulting approximate solution of problem (P_{m_1}) is given by $\mu_1 = s\#\rho^*$. We prove similar results as in (1.2.10) in this general case, from which Theorem 1.2.5 follows.

Dual problem. As mentioned in Remark 1.1.2, we can reformulate problem (P_m) as a convex optimization problem of the form (1.1.1) by setting $\mathcal{K} = G_m := \{\int_Z g d\mu, |, \mu \in \mathcal{P}_m(Z)\}$ in (1.1.1). Similarly to (1.1.3), the resulting dual problem writes:

$$\sup_{\lambda\in\mathcal{H}} -f^*(\lambda) - \chi^*_{G_m}(-\lambda).$$

The previous dual problem is equivalent to

$$-\inf_{\lambda\in\mathrm{dom}(f^*)} \mathcal{D}_m(\lambda) \coloneqq f^*(\lambda) - \int_X \inf_{y\in Z_x} \langle \lambda, g(x,y)\rangle \, dm(x). \tag{D_m}$$

We first show the existence and the uniqueness of the solution to the dual problem (D_m). This can be established based on two observations: (1) the function f^* is strongly convex, (2) the second term in $\mathcal{D}_m(\lambda)$ is convex with respect to λ. Next, assuming that G_m is a closed set, we establish the strong duality relation between problems (P_m) and (D_m) with the Fenchel-Rockafellar theorem. Let $\lambda^*(m)$ be the solution of the dual problem (D_m). Define the following function:

$$v\colon \mathcal{P}(X)\times X \to \mathbb{R},\ (m,x)\mapsto \inf_{y\in Z_x} \langle \lambda^*(m), g(x,y)\rangle.$$

Thanks to the strong duality, we can prove the stability of the dual problem and characterize the directional derivative of the optimal cost of (P_m) with respect to m.

Theorem 1.2.6 (Lemma 3.4.4 and Proposition 3.4.5)**.** *There exists a constant $C>0$ independent of m_0 and m_1 such that*

$$\max\left\{ \left|\mathcal{D}_{m_0}(\lambda^*(m_0)) - \mathcal{D}_{m_1}(\lambda^*(m_1))\right|,\ \|\lambda^*(m_0)-\lambda^*(m_1)\|^2 \right\} \leq C d_1(m_0,m_1).$$

The function v is the directional derivative of the value function of (P_m)*, i.e.,*

$$\mathbf{val}(\mathrm{P}_{m_1}) - \mathbf{val}(\mathrm{P}_{m_0}) = \int_{t=0}^1 \int_X v(m_0 + t(m_1-m_0),x) d(m_1-m_0)(x) dt.$$

Discretization. We present an original resolution method for the MFO problem (P_m). Our approach relies first on a discretization of the marginal m, as proposed in [Sar22] for Lagrangian MFGs. We approximate the common marginal distribution m in (P_m) by an empirical distribution $m_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, where $x_i \in X$ for $i=1,\dots,N$ and N is a sufficiently large integer. This allows us to write the associated discretized problem for (P_m) as

$$\inf_{\mu\in\mathcal{P}_{m_N}(Z)} f\left(\int_Z g d\mu\right). \tag{P_{m_N}}$$

We can establish a connection between problem (P_{m_N}) and the relaxed problem (PR) introduced in Section 1.2.1. First, we observe that a probability distribution μ belongs to $\mathcal{P}_{m_N}(Z)$ if and only if there exist $\mu_i \in \mathcal{P}(Z_{x_i})$ for each $i=1,\dots,N$ such that $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}\otimes \mu_i$. This leads to the following equivalent formulation of the discretized problem:

$$\inf_{\mu_i\in\mathcal{P}(Z_{x_i})} f\left(\frac{1}{N}\sum_{i=1}^N \int_{Z_{x_i}} g(x_i,y_i) d\mu_i(y_i)\right). \tag{1.2.11}$$

We can observe that (1.2.11) is a special case of (PR) by choosing $\mathcal{X}_i = Z_{x_i}$ and $g_i(\cdot) = g(x_i,\cdot)$. This equivalence allows us to apply the SFW algorithm (Algorithm 1.3 of Section 1.2.1) to (1.2.11).

In this context, if $y \in \prod_{i=1}^{N} Z_{x_i}$ is the result after some iterations of Algorithm 1.3, then the subproblems for the next iteration are

$$\bar{y}_i \in \mathbf{BR}_\lambda(x_i), \qquad \text{where } \lambda = \frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i). \tag{1.2.12}$$

Finally, the combination of Algorithm 1.3 with the recovery method (Algorithm 3.1) allows to obtain an approximate solution of (P_m) whose quality improves as the discretization parameter N increases. Here, we describe this combination:

1. Let $y^K \in \prod_{i=1}^{N} Z_{x_i}$ be the output of Algorithm 1.3 applied to (1.2.11), for $1 \leq K \leq 2N$ and for arbitrary numbers $n_k \geq 1$ of simulations. Let $\mu_N^K = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_i^K)}$.
2. Move on to the recovery method (Algorithm 3.1) with the following inputs: $m_0 = m_N$, $m_1 = m$, and $\bar{\mu}_0 = \mu_N^K$. The output is denoted as $\tilde{\mu}^K$, which is an element of the set $\mathcal{P}_m(Z)$.

Combining the convergence result of the SFW algorithm (Theorem 1.2.3) with the stability of the primal problem (Theorem 1.2.5), we have the following convergence result.

Theorem 1.2.7 (Theorem 3.5.8)**.** *There exists a constant $C > 0$, independent of m_N and m, such that*

$$\mathbb{E}\left[f\left(\int_Z g d\tilde{\mu}^K\right)\right] - \mathbf{val}(\mathrm{P}_m) \leq C\left(\frac{1}{K} + d_1(m_N, m)\right).$$

Remark 1.2.8. According to Theorem 1.2.7, in order to achieve better convergence, it is desirable to find an empirical distribution $m_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$ that is as close as possible to m in terms of the d_1-distance. This problem is commonly referred to as the optimal quantization problem. A precise estimate on $d_2(m_N, m)$ is given in [MM16, Prop. 12], and we slightly modify this estimate for $d_1(m_N, m)$ in Lemma 3.5.1. For more detailed information concerning optimal quantization, we refer to [GG12].

Numerical simulations. We apply in Section 3.6 our discretization method and the SFW algorithm to a Lagrangian MFG taken from [GHS22], in which the agents exploit their own stock of an exhaustible resource. In this model, the agents have different initial stock which follows a distribution m. The agents make decisions regarding their selling speed. The price of this resource at a time t is a decreasing function of the aggreagate selling speed at t. Following [GHS22], the main purpose of this model is to find the Nash equilibrium as defined in (3.6.2).

Literature comparison and perspectives. Our first-order necessary and sufficient optimality condition (1.2.7) is similar to the Nash equilibrium conditions studied in [CL18b, Sec. 3] for nonatomic potential games and in [CC18, Def. 3.1, Eq. 3.32] and [SS21, Def. 2.2] for Lagrangian MFGs. The authors of [CC18] also establish the existence of solutions to the associated Nash equilibrium problem using Kakutani's fixed point argument.

The article [CH17, Sec. 3] proposes a resolution of the potential Lagrangian MFGs with the fictitious play algorithm, which can be seen as a specific case of the FW algorithm with the learning rate $\omega_k = 1/(k+1)$. The possible memory overflow problem and the need to discretize the distribution of the initial conditions of the agents are not discussed.

The article [Sar22] investigates the discretization of a Lagrangian MFG with local congestion terms. The author proposes a regularization technique for the congestion term, a discretization in time of the trajectories, and a discretization of the initial distribution of the agents, leading to a nonlinear program which is solved with the quasi-Newton method. Convergence properties for the discrete solutions have been obtained.

Let us discuss on some possible extensions of this work. An easy way to construct m_N, in comparison with an optimal quantization approach, consists in drawing N samples of the distribution m and to construct the corresponding empirical distribution. One could modify the convergence analysis done for Theorem 1.2.6 to take into account the randomness of m_N, utilizing concentration inequalities such as [FG15, Thm. 1, Thm. 2]. They indeed allow to estimate $d_p(m_N, m)$ in both expectation and probability senses, where d_p is the Wasserstein distance of order p.

Another perspective deals with the study of a fully discretized version of the MFO problem (P_m). A crucial step in the SFW algorithm is the resolution of the subproblems (1.2.12) (involved in the definition of the best response mapping). These subproblems could be posed over an infinite dimensional space, as for example, in the price model (1.1.43). In general, one has to discretize those problems (as is done in [Sar22, Sec. 5]). One could extend the convergence analysis and look for a quantitave result that takes into account the effect of the discretization of the subproblems.

1.2.3 Error estimates of a theta-scheme for second-order mean field games

Framework and motivation. This section is associated with Chapter 4. It is dedicated to a novel finite-difference scheme for the second order MFGs system (MFG), relying on the theta-method. Let us denote by ∇_h, div_h and Δ_h the discrete gradient, divergence and Laplace operators of the centered finite-difference scheme. Let $\theta \in [0,1]$. Let us first describe the discretization of the FP equation. At any time t, the theta-scheme of the FP equation consists of two steps:

1. An explicit scheme for an intermediate FP equation, with a weight $(1-\theta)$ for the Laplacian term:
$$\frac{m(t+1/2) - m(t)}{\Delta t} - (1-\theta)\sigma \Delta_h m(t) + \mathrm{div}_h(mv(t)) = 0. \tag{S1}$$
2. An implicit scheme for an intermediate heat equation (without divergence term):
$$\frac{m(t+1) - m(t+1/2)}{\Delta t} - \theta\sigma\Delta_h(m(t+1)) = 0. \tag{S2}$$

Notice that when there is no divergence term ($v = 0$), the above scheme (S1)-(S2) coincides with the classical theta-scheme for the heat equation [All07]. For the HJB equation, we propose an adjoint scheme; at each time t, two steps are performed: (1) an implicit scheme for an intermediate heat equation (without the Hamiltonian term) and (2) an explicit scheme for an intermediate HJB equation.

Let us describe the main properties of the theta-scheme, which justify our interest for it. If $\theta = 0$, our scheme is an explicit scheme which has a natural interpretation as a discrete mean field game. However, it is not clear whether the explicit scheme for the FP equation, when $\theta = 0$, enjoys stability properties for some ℓ^2-norm. To ensure stability, a natural idea consists in taking an implicit scheme for the second-order term, i.e. $\theta = 1$. This yields a mixed scheme (implicit for the Laplacian term and explicit for the divergence term). We emphasize that the divergence term should remain explicit, in order to guarantee that the discrete system has a structure of (DMFG). When $\theta = 1$, we see that (S1) is an explicit scheme of a continuity equation (without diffusion term). To ensure the monotonicity of (S1), an upwind discretization for the divergence term should be employed, instead of centered scheme. In comparison with a centered discretization, the upwind discretization has the following disadvantages: (1) the consistency error is of a lower order, (2) we need then to construct a numerical Hamiltonian (see [ACD10, ACCD13]) to preserve the adjoint structure. Finally, we propose to take $\theta \in (1/2, 1)$ in (S1)-(S2) and to keep the centered scheme for the first-order term. The ℓ^2-stability and the monotonicity property hold in this case by some discrete energy estimate and a CFL condition detailed later. We end up with a discrete system which has a structure of (DMFG), has a higher order for the consistency error, and which does not require the construction of a numerical Hamiltonian.

Organization. We first introduce the theta-scheme associated with (MFG) and state our main result on the error estimate of this theta-scheme. This theta-scheme lies in the framework of (DMFG). The existence and uniqueness of the solution of the theta-scheme is deduced from condition (CFL) and (1.1.38). Subsequently, we present several stability properties of the theta-scheme. Next, we discuss the consistency error of the theta-scheme. Finally, we give the sketch of the proof of the main result, which is based on the stability and the consistency analysis.

Formulation and the main result. The main assumptions on the data of (MFG), which hold true in the subsequent sections, are as follows:

- The functions $\ell^c(\cdot,\cdot,v)$, $\ell^c_v(\cdot,x,v)$, $g^c(\cdot)$, $f^c(\cdot,\cdot,\cdot)$ and $m^c_0(\cdot)$ are L^c-Lipschitz continuous, where the Lipschitz continuity of f^c with respect to m is for the $\|\cdot\|_{\mathbb{L}^2}$-norm;
- The running cost ℓ^c is α^c-strongly convex with respect to the control variable;
- The monotonicity condition (1.1.22) holds;
- The continuous system (MFG) has a unique solution (u^*, v^*, m^*), with $u^*, m^* \in \mathcal{C}^{1+r/2,2+r}(Q)$ and $v^* \in \mathcal{C}^r(Q) \cap \mathbb{L}^\infty([0,1]; \mathcal{C}^{1+r}(\mathbb{T}^d))$, where $r \in (0,1)$.

The terminal time in (MFG) is fixed to 1 in this section. We discretize the time horizon $[0,1]$ into T equal steps, resulting in a time step size of $\Delta t = 1/T$. Similarly, we discretize the state space $\mathbb{T}^d$ into a grid with N points in each dimension, leading to a spatial step size of $h = 1/N$. The discretized time horizon is denoted by $\mathcal{T}$ (or $\tilde{\mathcal{T}}$ if the terminal time is included) and the discretized state space is denoted by S. For any $x \in S$, we denote $B_h(x) = \prod_{i=1}^d [x - he_i/2, x + he_i/2)$, where $(e_i)_{i=1,\dots,d}$ is the canonical basis of $\mathbb{R}^d$. For any function $\varphi \in \mathbb{R}^n(\tilde{\mathcal{T}} \times S)$ and any $(p,q) \in [1,\infty]^2$, we define the norm $\|\varphi\|_{p,q} = \| \left(\|\varphi(t,\cdot)\|_{\ell^q(S)} \right)_{t \in \tilde{\mathcal{T}}} \|_{\ell^p(\tilde{\mathcal{T}})}$.

Let ℓ, H, f, m_0, and g be the discretization of ℓ^c, H^c, f^c, m^c_0, and g^c, respectively, as defined in (4.2.9)-(4.2.10). We note that the discretization of H^c in this context does not require the construction of a numerical Hamiltonian, in contrast to [ACCD13]. Taking any $\theta \in (1/2, 1)$, we can now introduce the theta-scheme of (MFG): Find $(u,v,m) \in \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}^d(\mathcal{T} \times S) \times \mathbb{R}(\bar{\mathcal{T}} \times S)$ such that $\forall (t,x) \in \mathcal{T} \times S$,

$$\begin{cases} \text{(i)} & u = \mathbf{HJB}_\theta(m), \\ \text{(ii)} & v = \mathbf{V}_\theta(u), \\ \text{(iii)} & m = \mathbf{FP}_\theta(v), \end{cases} \tag{Theta-mfg}$$

where the Hamilton-Jacobi-Bellman mapping $\mathbf{HJB}_\theta \colon \mathcal{P}(\tilde{\mathcal{T}}, S) \to \mathbb{R}(\tilde{\mathcal{T}} \times S)$, $m \mapsto u$, is defined by

$$\begin{cases} -\frac{u(t+1,x)-u(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h u(t+1/2,x) = 0, \\ -\frac{u(t+1/2,x)-u(t,x)}{\Delta t} - (1-\theta)\sigma\Delta_h u(t+1/2,x) + H[\nabla_h u(\cdot+1/2,\cdot)](t,x) = f(t,x,m(t)), \\ u(T,x) = g(x), \end{cases}$$

the optimal control mapping $\mathbf{V}_\theta \colon \mathbb{R}(\tilde{\mathcal{T}} \times S) \to \mathbb{R}^d(\mathcal{T} \times S)$, $u \mapsto v$, is defined by

$$\begin{cases} -\frac{u(t+1,x)-u(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h u(t+1/2,x) = 0, \\ v(t,x) = -H_p(t,x,\nabla_h u(t+1/2,x)), \end{cases}$$

and the Fokker-Planck mapping $\mathbf{FP}_\theta \colon \mathbb{R}^d(\mathcal{T} \times S) \to \mathbb{R}(\tilde{\mathcal{T}} \times S)$, $v \mapsto m$, is defined by

$$\begin{cases} \frac{m(t+1/2,x)-m(t,x)}{\Delta t} - (1-\theta)\Delta_h m(t,x) + \mathrm{div}_h(vm)(t,x) = 0, \\ \frac{m(t+1,x)-m(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h m(t+1,x) = 0, \\ m(0,x) = m_0(x). \end{cases}$$

From the definitions of $\mathbf{HJB}_\theta$ and $\mathbf{FP}_\theta$, we see that the adjoint structure of the coupled system (MFG) is preserved in the resulting discretized system, which is an important property for the stability analysis in this section and for the potential case discussed in Section 1.2.4. We fix now a constant M, defined as follows:

$$M = \frac{1}{\alpha^c}\Big(2 \max_{(t,x)\in Q} \|\ell_v^c(t,x,0)\| + \sqrt{d}(L_\ell^c + L_f^c + L_g^c)\Big). \tag{1.2.13}$$

We prove in Theorem 4.4.4 that the constant M is an upper bound of $\|v\|_{\infty,\infty}$. We consider the following condition on $(\Delta t, h)$:

$$\Delta t \leq \frac{h^2}{2d(1-\theta)\sigma}, \qquad h \leq \frac{2(1-\theta)\sigma}{M}. \tag{CFL}$$

The main result of this section is the following error estimate on (Theta-mfg).

Theorem 1.2.9 (Theorem 4.2.10)**.** *Let $\theta \in (1/2,1)$ and let $(\Delta t, h)$ satisfy the condition* (CFL)*. Then* (Theta-mfg) *has a unique solution (u_h, v_h, m_h). Moreover, there exists a constant $C > 0$, independent of Δt and h, such that*

$$\|u_h - u_h^*\|_{\infty,\infty} + \|m_h - m_h^*\|_{\infty,1} \leq Ch^r,$$

where $u_h^*, m_h^* \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ are defined by $u_h^*(t,x) = u^*(t\Delta t, x)$ and $m_h^*(t,x) = \int_{B_h(x)} m^*(t\Delta t, y)dy$.

Existence and uniqueness. Recall our general framework for discrete MFGs, introduced in Section 1.1.3. We prove in Lemma 4.4.1 that (Theta-mfg) is equivalent to the system (DMFG). We provide an explicit formula for the mapping π describing the probability transitions under a given control. A key point is that $\pi(t, x, \cdot, \omega)$ must be a probability distribution over S, which can be established under the CFL condition. The uniqueness of the solution of (Theta-mfg) is a consequence of the fundamental inequality (1.1.38).

Stability of the theta-scheme. Following the general perturbation system (PDMFG), we introduce the perturbed version of (Theta-mfg) with additional terms $(\eta, \delta) \in \mathbb{R}^2(\tilde{\mathcal{T}} \times S)$:

$$\begin{cases} \text{(i)} \quad u = \mathbf{HJB}_\theta(m; \eta), \\ \text{(ii)} \quad v = \mathbf{V}_\theta(u), \\ \text{(iii)} \quad m = \mathbf{FP}_\theta(v; \delta). \end{cases} \tag{1.2.14}$$

Let (u, v, m) and (u_h, v_h, m_h) be solutions of (1.2.14) and (Theta-mfg) respectively. Suppose that $m \geq 0$. Our stability analysis consists of the following three properties, related to each of the mappings $\mathbf{HJB}_\theta$, $\mathbf{V}_\theta$, and $\mathbf{FP}_\theta$:

1. *Stability of the HJB equation.* From the dynamical programming principle and the regularity of f^c, we can derive the following inequality:
$$\|u - u_h\|_{\infty,\infty} \leq \frac{L^c}{h^{d/2}} \|m - m_h\|_{\infty,2} + \|\eta\|_{1,\infty}, \tag{1.2.15}$$
where L^c denotes the Lipschitz continuity modulus of f^c.

2. *Stability of the optimal control.* We deduce from the fundamental inequality (1.1.38) that
$$\frac{\Delta t \alpha^c}{2} \left\| \|v - v_h\|^2 (m + m_h) \right\|_{1,1} \leq \sum_{t \in \mathcal{T}} \sum_{x \in S} (u - u_h)(t+1, x)\delta(t, x) + (m_h - m)(t, x)\eta(t, x). \tag{1.2.16}$$

3. *Stability of the FP equation.* To analyze the stability of the FP equation, we consider the energy estimate for the discrete parabolic equation given by:
$$\begin{cases} \frac{\mu(t+1) - \mu(t)}{\Delta t} - \theta \Delta_h \mu(t+1) - (1-\theta)\Delta_h \mu(t) + \text{div}_h \mu v(t) = \text{div}_h \delta_1(t) + \delta_2(t), \\ \mu(0) = \mu_0, \end{cases}$$
where $\delta_1, \delta_2 \in \mathbb{R}(\mathcal{T} \times S)$ represent perturbations in the form of discrete divergence and other forms, respectively. Assume that $\|v\|_{\infty,\infty} \leq M$. Then, there exists a constant C independent of h and Δt such that
$$\max_{t \in \tilde{\mathcal{T}}} \left\|\mu(t)\right\|_2^2 \leq C \left(\left\|\mu_0\right\|_2^2 + \left\|\nabla_h^+ \mu_0\right\|_2^2 + \sum_{\tau \in \mathcal{T}} \Delta t \left(\left\|\delta_1(\tau)\right\|_2^2 + \left\|\delta_2(\tau)\right\|_2^2 \right) \right). \tag{1.2.17}$$
The notation ∇_h^+ denotes the forward finite-difference operator. The proof of the energy estimate (1.2.17) is similar to the one for parabolic PDEs, see [Lio71, LSU88], and the one for the implicit scheme, see [ACCD13]. Let us emphasize the fact that for establishing (1.2.17), we need to take $\theta > 1/2$.

Consistency error. Recall the definition of (u_h^*, m_h^*) in Theorem 1.2.9. For $(t,x) \in \mathcal{T} \times S$, let

$$v_h^*(t,x) = -H_p(t,x,\nabla_h u_h^*(t+1/2,x)),$$

where $u_h^*(t+1/2)$ and $u_h^*(t+1)$ satisfy the first line of $\mathbf{HJB}_\theta$ for any $t \in \mathcal{T}$. Define the invertible matrix $B_1 = \mathrm{Id} - \theta\sigma\Delta t\Delta_h$. Using the regularity of the continuous solution, we have the following consistency result.

Theorem 1.2.10 (Lemma 4.5.5)**.** *The triplet* (u_h^*, v_h^*, m_h^*) *satisfies the perturbed system* (1.2.14) *with perturbation terms* η *and* δ *such that*

$$\eta = \mathcal{O}(\Delta t h^r), \qquad B_1\delta = \Delta t \mathrm{div}_h(\delta_1) + \mathcal{O}(\Delta t h^{r+d}), \quad \textit{where } \delta_1 = \mathcal{O}(h^{2r+d}).$$

Sketch of the proof of Theorem 1.2.9. Let us define $\epsilon = \Delta t \left\| |v_h^* - v_h|^2 m_h^* \right\|_{1,1}$.

- By combining the stability property for the HJB equation (1.2.15), the fundamental inequality (1.2.16), and the consistency error from Theorem 1.2.10, we can derive an upper bound of ϵ as a function of $\|m_h^* - m_h\|_{\infty,2}$ and h.
- Using the stability of the FP equation (1.2.17) and considering the consistency error from Theorem 1.2.10, we can derive an upper bound of $\|m_h^* - m_h\|_{\infty,2}$ depending on ϵ and h.

Combining the previous two estimates, we can deduce that $\|m_h^* - m_h\|_{\infty,2} \leq Ch^{r+d/2}$ for some constant C independent of Δt and h. The conclusion of Theorem 1.2.9 follows from Hölder's inequality for discrete norms (see Lemma 4.2.2).

Literature comparison and perspectives. In 2010, a first result concerning the convergence of a finite-difference scheme for stationary MFGs was obtained in [ACD10]. In this article, the authors also proposed an implicit scheme for time-dependent MFGs and proved the existence and uniqueness of the solution of this scheme. In 2013, a convergence result was obtained for the same implicit scheme in [ACCD13] when the Hamiltonian has a monomial form. The two cited works assume the existence of a classical solution of (MFG). In 2016, in the absence of this existence assumption, [AP16] proved that the solution of the implicit scheme converges to a weak solution of (MFG) when the grid steps tend to zero.

We mention the articles [CS14, CS15, HS19] investigating semi-Lagrangian discretizations of MFG systems. We also mention the article [BC22] which contains an explicit rate of convergence for a semi-discretization in space, obtained with finite differences. To the best of our knowledge, the work presented in this section is the first one, in the context of MFGs, to give a precise convergence order for a fully discrete numerical scheme.

We will discuss in the next section the resolution of (Theta-mfg) with the Generalized Frank Wolfe algorithm. As will be explained, the application of this algorithm is made possible by the fact that the potential structure of continuous MFGs is preserved at the discrete level by (Theta-mfg). The GFW algorithm essentially relies on successive resolutions of the discrete HJB and the Fokker-Planck equations, which do not require to solve nonlinear equations (as would be the case with the fully implicit scheme of [ACCD13]). Future work could focus on the numerical analysis for an extension of the theta-scheme relying on splitting methods [Tho95, Sec. 4.4], applied to the discrete Laplacian involved in the implicit linear equations. This would allow to facilitate the resolution of the discrete HJB and Fokker-Planck and thus to reduce the numerical complexity of the GFW algorithm.

Another perspective concerns the extension of (Theta-mfg) to more complex MFG models, such as the MFG models with a coupling through a price variable [BHP21]. They fall into the class of MFGs of controls (abbreviated MFGCs), since the price depends on the distribution and the control of the agents. Note that the price appears in the Hamiltonian of the HJB equation of the coupled system. One would have to propose a suitable discretization of the additional coupling relation (involving the price variable) and to establish a fundamental inequality for the obtained discrete setting. This would allow to conduct a stability analysis. Let us mention that the article [LP22] already contains stability results for such MFGCs, in the continuous framework. Concerning consistency, the analysis would exploit the regularity properties of the continuous solution obtained in [BHP21], in particular, the Hölder continuity of the price variable.

1.2.4 A mesh-independent method for second-order potential mean field games

Framework and motivation. In this section, we investigate the resolution of a second-order potential MFG, which is equivalent to the optimal control problem (1.1.25). As already mentioned in Section 1.2.3, the potential structure is preserved by (Theta-mfg). As a result, the discrete MFG can be addressed with the GFW algorithm, described in Section 1.1.3.

The general objective is to show that the performance of the GFW algorithm is not impacted by a refinement of the discretization grid. The main results of this section are two mesh-independence properties for the resolution of (MFG) with (Theta-mfg) and the GFW algorithm, as stated in Theorems 1.2.12 and 1.2.13. The terminology mesh-independence was coined in the article [ABPR86]. It is said that an algorithm satisfies a mesh-independence property when approximately the same number of iterations is required to satisfy a stopping criterion, when comparing an infinite-dimensional problem and its discrete counterpart.

Organization. We begin by presenting the GFW algorithm for the resolution of general potential and discrete MFGs. We provide two results concerning the rate of convergence of the potential cost. One rate is sublinear and the other is linear, depending on the choice of the learning rate (1.2.18)-(1.2.19). Next, we prove that the GFW algorithm exhibits mesh-independent sublinear and linear convergence rates when applied to (Theta-mfg) . These convergence results rely on the study of the semi-concavity of $\mathbf{HJB}_\theta$, the ℓ^2-stability of $\mathbf{FP}_\theta$, and the ℓ^∞-stability of $\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta$.

GFW algorithm for potential discrete MFG. Recall the data of (DMFG): The time space $\mathcal{T}$, the state space S, the running cost ℓ, the coupling cost f, the initial condition m_0, and the terminal cost g, where

$$\ell \colon \mathcal{T} \times S \times \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}, \qquad f \colon \mathcal{T} \times S \times \mathbb{R}(S) \to \mathbb{R}, \qquad m_0 \in \mathcal{P}(S), \qquad g \in \mathbb{R}(S).$$

We assume that ℓ is α-convex with respect to its third variable, f is L_f-Lipschitz with respect to its third variable, and the discrete potential condition (1.1.33) is satisfied.

We present the GFW algorithm for solving the potential discrete MFG in Algorithm 1.4. As discussed in Section 1.1.3, at each iteration of the GFW algorithm, we are required to solve the partially linearized problem (1.1.35). The solution to this problem is obtained through the best-response mapping $\mathbf{BR}\colon \mathcal{P}(\tilde{\mathcal{T}}, S) \to \tilde{\mathcal{A}}$, which is defined in (1.1.36). For the learning rate λ_k in Algorithm 1.4, we propose the following two rules:

1. an open update rule as in Algorithm 1.1:
$$\lambda_k = \frac{2}{k+2}; \tag{1.2.18}$$
2. a closed update rule similar to the line search (1.1.10) used in Algorithm 1.2:
$$\lambda_k = \min\left\{ \frac{\tilde{J}_{m^k}(m^k, w^k) - \tilde{J}_{m^k}(\bar{m}^k, \bar{w}^k)}{L_f |S|^{1/2} \|m^k - \bar{m}^k\|^2_{\infty,2}}\, , 1 \right\}. \tag{1.2.19}$$

Algorithm 1.4: GFW for potential discrete MFG

Initialization: $(m^0, w^0) \in \tilde{\mathcal{A}}$;
First iteration: $(m^1, w^1) = (\bar{m}^0, \bar{w}^0) = \mathbf{BR}(m^0)$;
Choose an update rule from (1.2.18) and (1.2.19) ;
for $k = 1, 2, \dots$ **do**

  Step 1: Resolution of the partial linearized problem.
  Set $(\bar{m}^k, \bar{w}^k) = \mathbf{BR}(m^k)$;

  Step 2: Update.
  Set $\lambda_k \in [0,1]$ by the chosen update rule;
  Set $(m^{k+1}, w^{k+1}) = (1-\lambda_k)(m^k, w^k) + \lambda_k(\bar{m}^k, \bar{w}^k)$;

end

To state the convergence results of Algorithm 1.4, we need to introduce three key constants. The first two are defined by
$$C_1 = \sup_{v \in \mathbb{R}^d_{\tilde{D}}(\mathcal{T}\times S)} \|\mathbf{FP}(v)\|^2_{\infty,2} \quad \text{and} \quad C_2 = \sup_{m \in \mathcal{P}(\tilde{\mathcal{T}}, S)} \|\mathbf{FP}\circ \mathbf{V} \circ \mathbf{HJB}(m)\|_{\infty,\infty}.$$
The third constant C_3 is such that for any $v_1, v_2 \in \mathbb{R}^d_{\tilde{D}}(\mathcal{T}\times S)$, we have:
$$\|\mathbf{FP}(v_1) - \mathbf{FP}(v_2)\|^2_{\infty,2} \leq C_3 \Delta t \sum_{t\in\mathcal{T}} \sum_{x \in S} \|(v_1 - v_2)(t,x)\mathbf{FP}(v_1)(t,x)\|^2.$$
We will see later that for the theta-scheme (which is a particular case of (DMFG)), estimates of the constants C_1 and C_3 derive from an energy estimate (1.2.17) and an estimate of the constant C_3 can be deduced from a semi-concavity property of the discrete HJB mapping and the ℓ^∞-stability of $\mathbf{FP}\circ\mathbf{V}\circ\mathbf{HJB}$.

Recall that $\tilde{J}$ is the objective function in the convex optimal control problem (1.1.34) associated with the potential (DMFG). Let $(\bar{m}, \bar{w})$ be the solution of (1.1.34).

Theorem 1.2.11 (Proposition 5.3.1)**.** *We consider the sequence* $(m_k, w_k)_{k\geq 1}$ *generated by Algorithm* (1.4)*. Let* $D = C_1 L_f |S|^{1/2}$ *and* $c = \max\left\{1 - \frac{\alpha}{4C_2C_3L_f|S|^{1/2}}, \frac{1}{2}\right\} \in (0,1)$*.*

1. Sublinear rate. *If the chosen update rule is* (1.2.18)*, then,*

$$\tilde{J}(m^k, w^k) - \tilde{J}(\bar{m}, \bar{w}) \leq \frac{8D}{k}, \qquad \forall k \geq 1. \tag{1.2.20}$$

2. Linear rate. *If the chosen update rule is* (1.2.19)*, then,*

$$\tilde{J}(m^k, w^k) - \tilde{J}(\bar{m}, \bar{w}) \leq 4Dc^k, \qquad \forall k \geq 1. \tag{1.2.21}$$

The proof of Theorem 1.2.11 is similar to the convergence analysis of the GFW algorithm in Hilbert space [KW22] and the one for continuous potential MFGs [LP22].

We can see from Theorem 1.2.11 that the convergence constants D and c rely on various factors, including the strong convexity constant α, the product $L_f|S|^{1/2}$, and the three key constants C_1, C_2, and C_3 defined earlier. In order to investigate the mesh-independent property of Algorithm (1.4) when applied to some numerical discretization of the potential (MFG), the crucial point is to establish the independence of these factors with respect to the discretization parameters Δt and h.

Now let us apply Algorithm 1.4 to the theta-scheme associated with the potential (MFG). We make the same assumptions for this continuous system as in Section 1.2.3. From Lemma 5.4.4, the associated theta-scheme (Theta-mfg) preserves the potential structure in the sense of (1.1.33).

A mesh-independent sublinear convergence result. In this paragraph, we will see that the constant D is independent of Δt and h. Combining with (1.2.20), a first mesh-independent sublinear convergence result is obtained in Theorem 1.2.12. We have the following two a priori estimates:

- Lemma 5.4.2 shows that $\alpha = \alpha^c$, $L_f = L^c h^{-d/2}$, and $|S| = 1/h^d$, where α^c is the strong convexity constant of ℓ^c and L^c is the Lipschitz constant of f^c;
- The energy estimate (1.2.17) implies that C_1 and C_3 are mesh-independent, since the constant C in (1.2.17) is independent of Δt and h.

From the definition of the constant D and the above two points, we derive the mesh-independence of D. Let (u_h, v_h, m_h) be the solution of (Theta-mfg). Let us set $\gamma_k = \tilde{J}(m^k, w^k) - \tilde{J}(\bar{m}, \bar{w})$. We have the following convergence result.

Theorem 1.2.12 (Sublinear rate, Theorem 5.4.5)**.** *In Algorithm 1.4, apply the update rule* (1.2.18)*. Then there exists a constant* C_θ*, independent of* Δt *and* h*, such that for any* $k \geq 1$*,*

$$\gamma_k \leq \frac{C_\theta}{k}.$$

A mesh-independent linear convergence result. To explore the mesh-independent linear convergence rate of Algorithm 1.4 for (Theta-mfg), it is necessary to establish the independence of the constant C_2 with respect to the discretization parameters Δt and h. To accomplish this, we need to study the ℓ^∞-stability of $\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta$. Two additional assumptions are required:

- The functions $\ell^c(t,x,v)$, $f^c(t,x,m)$ and $g^c(x)$ are L-semi-concave with respect to x, for any $t \in [0,1]$, $v \in \mathbb{R}^d$, and $m \in \mathcal{P}(\mathbb{T}^d)$;
- The function ℓ^c has a separable form with respect to v, i.e. $\ell^c(t,x,v) = \sum_{i=1}^d \ell_i^c(t,x,v_i)$, where v_i is the i-th coordinate of v.

We refer to [CS04, Def. 1.1.1] for the definition of the semi-concavity. Under the previous two additional assumptions, we have a second convergence result.

Theorem 1.2.13 (Linear rate, Theorem 5.4.9)**.** *In Algorithm 1.4, apply the update rule* (1.2.19)*. Then there exist two constants $c_\theta \in (0,1)$ and $C_\theta > 0$, both independent of Δt and h, such that for any $k \geq 1$,*

$$\gamma_k \leq C_\theta c_\theta^k.$$

The proof of Theorem 1.2.13 relies on the ℓ^∞-stability of $\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta$. The proof of the ℓ^∞-stability of $\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta$ is inspired from the continuous case (see for example [CL18a, Lem. 5.3]). Let us outline the proof in the continuous case:

1. The semi-concavity of the solution u of the HJB equation implies that $D^2 u \leq L \cdot \mathrm{Id}$;
2. Combining with the convexity of H^c, we deduce that there exists some constant $C > 0$ such that $-\mathrm{div}\ v = \mathrm{Tr}(H^c_{pp}[\nabla u]\ D^2 u) \leq C$ uniformly;
3. Since div v is the coefficient preceding m in the FP equation, the maximum principle of parabolic equations implies that $\|m\|_{\mathbb{L}^\infty} \leq \exp(C)\|m_0^c\|_{\mathbb{L}^\infty}$ (the terminal time is 1).

In our discrete context, we first prove that for any $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, the function $u = \mathbf{HJB}_\theta(m)$ is $3L$-semi-concave with respect to x. Let $v = \mathbf{V}_\theta(u)$ and $\tilde{m} = \mathbf{FP}_\theta(v)$. Thanks to the semi-concavity of u, there exists a constant C independent of Δt and h such that $-\mathrm{div}_h v(t,x) \leq C$. As a consequence, we can deduce that $\|\tilde{m}(t+1,\cdot)\|_\infty \leq (1 + C\Delta t)\|\tilde{m}(t,\cdot)\|_\infty$ for any $t \in \mathcal{T}$. It follows that $\|\tilde{m}\|_{\infty,\infty} \leq \exp(C)\|m_0\|_\infty$. Since the right-hand-side of the previous inequality is independent of the input m, it gives an upper bound of C_2, which is mesh-independent. Combining with (1.2.21), Theorem 1.2.13 follows.

Literature comparison and perspectives. Various methods have been proposed in the literature for the resolution of potential MFGs, besides the GFW algorithm that was analyzed in the continuous setting in [LP22]. Note that the sublinear and linear convergence rates have obtained for the continuous model in this reference. Note also that the GFW algorithm can be seen as a generalization of the fictitious play method introduced in [CH17, HS19]. In the convex potential case, the ADMM algorithm was utilized in [BC15, And17] and the Chambolle-Pock algorithm was utilized in [AL20]. Some articles propose to discretize the optimal control problem (1.1.25), see for example [LST10, And17]. In this context, it is very desirable that the potential structure of the continuous MFG is preserved at the level of the discretized coupled system, so that one can apply in a direct fashion suitable optimization methods to the discrete system. This is in particular the case for the implicit scheme proposed in [ACCD13].

To the best of our knowledge, in the context of mean field games, the mesh-independence property has never been established so far for any other method. Though it seems a natural property, it may not hold in general. In particular, it might not hold for primal-dual methods, whose application relies on a saddle-point formulation of the convex problem (1.1.25) of the form, in which the FP equation is "dualized". This saddle-point formulation involves a linear operator, encoding the (discrete) FP equation (see for example [AL20, Sec. 3.2]). As the discretization parameters decrease, the operator norm of these operators (for the Euclidean norm) increases, which has an impact on the convergence properties of methods such as the Chambolle-Pock algorithm. In contrast, the discrete Fokker-Planck equation remains satisfied at each iteration of the GFW equation.

Let us discuss some possible extensions of this work. We can study the convergence of the GFW applied to the non-monotone discrete MFG, where the monotonicity condition (1.1.32) is not satisfied. This results in the non-convexity of the associated potential function F. Consequently, the convergence of the GFW algorithm to a global minimum of the objective function is not guaranteed in general. However, we can show that a point $(\bar{m}, \bar{v})$ is a Nash-equilibrium of the potential MFG in the sense of (1.1.15) if the "Frank-Wolfe" gap (the left-hand-side of (1.1.37)) at the point $\bar{m}$ is 0. Fortunately, similarly to the approach used in [LJ16], we can prove that the GFW algorithm converges to a stationary point of (1.1.34) such that the "Frank-Wolfe" gap at this point vanishes. This implies that the GFW algorithm can still converge to a Nash equilibrium despite the presence of non-monotonicity. Consequently, we can further study the mesh-independent property of the theta-scheme in this scenario.

Another perspective deals with the discretization and the application of the GFW algorithm to the planning problem of [ACCD12], referred to as MFGP. In MFGP, instead of having a terminal condition g^c for the HJB equation, we consider a fixed terminal condition for the FP equation, denoted as $m_T^c \in \mathcal{P}(\mathbb{T}^d)$. This problem can be seen as a generalized optimal transport problem. The article [ACCD12] also considers an approximation, called MFGPP, obtained by regularization of the final-time constraint on the distribution. This opens the path to a numerical resolution of the planning problem through an application of the GFW algorithm to MFGPP. Yet the application of the GFW algorithm may be difficult because the Lipschitz constant of the coupling term in MFGPP would increase with the penalty parameter.

Chapter 2

Large-scale nonconvex optimization: randomization, gap estimation, and numerical resolution

2.1 Introduction

Problem formulation This article is devoted to the theoretical analysis and the numerical resolution of the following large-scale, aggregative, and nonconvex optimization problem:

$$\inf_{x \in \mathcal{X}} J(x) \coloneqq f(G(x)), \qquad \text{where: } \begin{cases} G(x) = & \dfrac{1}{N} \displaystyle\sum_{i=1}^{N} g_i(x_i) \\ \mathcal{X} = & \prod_{i=1}^{N} \mathcal{X}_i. \end{cases} \tag{P}$$

Here, N can be seen as the number of agents and is assumed to be large. The mappings $g_i \colon \mathcal{X}_i \to \mathcal{E}$ are given and referred to as the contribution mappings. The space $\mathcal{E}$ is a real Hilbert space. The main feature of this problem is the aggregative form of the function $G \colon \mathcal{X} \to \mathcal{E}$, which is defined as the average of the N mappings g_i. We will call $G(x)$ the aggregate. Let us emphasize that the dimension q of the aggregate space $\mathcal{E}$ can be arbitrarily large and possibly infinite. While very few structural assumptions are made on the sets $\mathcal{X}_i$ and the mappings g_i, we will assume that f is convex, with a Lipschitz-continuous gradient and that the image sets $g_i(\mathcal{X}_i)$ are all bounded. A central idea in this work is that the problem can be well approximated by a convex problem when N is large.

In various examples of interest, the function f has a separable structure as defined below. It turns out that taking into account the separability of f, when possible, allows us to refine our theoretical results (more precisely, to reduce some of the constants of interest, see Remark 2.2.7). From now on, we suppose that $\mathcal{E}$ is the Cartesian product of M separable Hilbert spaces denoted $\mathcal{E}_j$, for $j = 1, \ldots, M$. We assume that f is additively separable, that is to say, we assume that

$$f(y) = \sum_{j=1}^{M} f_j(y_j), \quad \forall (y_1, \ldots, y_M) \in \prod_{i=1}^{M} \mathcal{E}_j,$$

where $f_j\colon \mathcal{E}_j \to \mathbb{R}$. Note that when f is not separable, one can take $M = 1$ and $\mathcal{E}_1 = \mathcal{E}$. We assume that the contribution mappings are of the form

$$g_i(x_i) = \big(g_{ij}(x_i)\big)_{j=1,\dots,M}, \quad \text{where } g_{ij}\colon \mathcal{X}_i \to \mathcal{E}_j.$$

Hence the criterion J of problem (P) writes

$$J(x) = f(G(x)) = \sum_{j=1}^{M} f_j\Big(\frac{1}{N}\sum_{i=1}^{N} g_{ij}(x_i)\Big). \tag{2.1.1}$$

We present and discuss some motivating examples in Section 2.5, arising from social welfare problems, optimal control problems and supervised learning.

Related works and methods Let us return to the general problem (P). Classical Lagrangian relaxation (Chapter XII of [HUL93]) methods can be relevant here because the dual problem is separable in the sense below, thanks to the aggregative form of G. To see this, let us reformulate (P) as: $\inf_{(x,v)\in\mathcal{X}\times\mathcal{E}} f(v)$, subject to the constraint that $v = G(x)$. Its dual problem is:

$$\sup_{\lambda\in\mathcal{E}} \big(-f^*(\lambda) + \Phi(\lambda)\big), \tag{2.1.2}$$

where f^* is the Fenchel conjugate function of f, and $\Phi(\lambda)$ is defined by

$$\Phi(\lambda) \coloneqq \inf_{x\in\mathcal{X}} \langle \lambda, G(x)\rangle = \frac{1}{N}\sum_{i=1}^{N} \inf_{x_i\in\mathcal{X}_i} \langle \lambda, g_i(x_i)\rangle. \tag{2.1.3}$$

One sees that $\Phi(\lambda)$ can be evaluated by solving N independent sub-problems, one for each i in $\{1,\dots,N\}$. Solving these sub-problems can be much easier than addressing frontally the original problem with N coupled variables. This approach has been extensively employed in convex settings [SAB$^+$23, Pac18]. However, the nonconvexity of the problem raises two major difficulties: the potentially large duality gap and the reconstruction of a primal solution from the dual optimal solution.

These two difficulties are addressed by Wang in [Wan17]. She proposed a convex relaxation of the problem, based on a geometrical approach, that allows to obtain an estimate of the duality gap of order $\mathcal{O}(q^2/N^2)$. Her main tool was the Shapley-Folkman lemma [Sta69], which allows to show that the image of G is close to a convex set. This idea was already present in the seminal work of Aubin and Ekeland in [AE76], dealing with a different setting involving a coupling constraint. We refer the reader to [KCD22] for the most recent improvements dealing with this class of problems. We also refer to [Wan17] for a more exhaustive of mathematical works dedicated to the estimation of the duality gap, where a kind of convexification occurs. After having solved the dual problem by a cutting plane method and then found an approximate solution to the relaxed primal problem via a projection problem, Wang's method recovers an approximate solution to the original nonconvex problem, by computing a Shapley-Folkman decomposition of the aggregate with a standard linear programming approach.

There exist another important class of methods for large-scale optimization problems which are the block coordinate descent algorithm and its variants [BT13, FR16]. These methods may not be applicable without additional assumptions on the sets $\mathcal{X}_i$ and the maps g_i (in the current framework, the sets $\mathcal{X}_i$ could be discrete). Even if we make additional regularity assumptions, they may be inefficient, in particular because the cost function J is not convex in general.

Contributions and organization of the paper We first introduce in Section 2.2 a convex relaxation of the original problem (P). The relaxed problem is obtained by randomization, that is to say, we replace the variables x_i by probability measures μ_i on $\mathcal{X}_i$. The contribution mappings $g_i(x_i)$ are replaced by $\int_{\mathcal{X}_i} g_i(x_i) d\mu_i(x_i)$; these terms are linear with respect to μ_i. The resulting randomized cost function, denoted $\mathcal{J}$, is convex, and so is the randomized problem. We give a first upper bound of the relaxation gap of order $\mathcal{O}(1/N)$. The randomized problem has a stochastic interpretation: it amounts to replace the variables x_i by independent random variables X_i of probability distribution μ_i, and to replace $g_i(x_i)$ by the expectation of $g_i(X_i)$. To derive a good candidate (for (P)), given an approximate solution to the randomized problem $\mu = (\mu_1, ..., \mu_N)$, we propose to simulate random variables X_i with probability distribution μ_i. We will call this technique the *selection method*. We give a sharp estimate of the probability of error for the selection method. More precisely, we estimate the probability that $J(X_1, \ldots, X_N) \geq \mathcal{J}(\mu) + \left(\frac{C}{N} + \epsilon\right)$, given $\epsilon > 0$. The proof relies on McDiarmid's inequality, a concentration inequality [McD89].

From a numerical point of view, our main contribution is a method which is parallelizable, which benefits from the convexity of the randomized problem, but avoids the difficulty of the manipulation of probability measures (arising in the formulation of the randomized problem). This could be achieved by combining the Frank-Wolfe (FW) algorithm [DH78, Jag13], applied to the randomized problem, and the selection method described previously. The resulting algorithm, called stochastic Frank-Wolfe (SFW) algorithm, is described and analyzed in Section 2.3. Each iteration of the algorithm requires to solve a subproblem of the form (2.1.3), which is decomposable into N subproblems. Resorting to the selection method, we avoid to manipulate explicitely probability measures on the sets $\mathcal{X}_i$, which may otherwise cause memory issues. The SFW method is able to find an $\mathcal{O}(1/N)$-solution to problem P. In addition, we estimate the probability that the iterate x_k is $\left(\frac{C}{k} + \epsilon\right)$-optimal, for $k \leq 2N$, where k is the iteration counter. This result relies on concentration inequalities for martingales [Del15] which generalize McDiarmid's inequality.

Let us note that many articles in the literature are dedicated to stochastic variants of the Frank-Wolfe algorithm. These variants are concerned with the situation where the cost function is in the form of the expectation of a random cost and where its gradient is evaluated by sampling. See for example [DU18, HKMS20, HL16, MHK20, YSC19], see also [FMSF21, LYFC19] and the references therein. Let us emphasize that the stochasticity of our algorithm has another origin, namely the selection method. In all these articles, convergence is established in expectation; to our knowledge, only the article [TBL22] quantifies the probability of success of some stochastic method based on the Frank-Wolfe algorithm.

Our last theoretical contribution is a sharp estimate of the relaxation gap, of order $\mathcal{O}(q \wedge N/N^2)$, where q is the (potentially infinite) dimension of the aggregate space $\mathcal{E}$. It is proved in Section 2.4.

It relies on a geometrical relaxation of problem (P), shown to be equivalent to the relaxation by randomization. The relaxation gap is estimated with the help of a measure of nonconvexity for sets (introduced in [Cas75]) and with the help of the Shapley-Folkman lemma [Sta69]. We also give an estimate of the price of decentralization (as defined by Wang in [Wan17]). We conclude the section with a detailed comparison of our approach and the one of [Wan17].

Section 2.5 is dedicated to examples and discussions on numerical aspects. We provide in sections 2.6 and 2.7 numerical results for a mixed-integer linear-quadratic program and a discrete aggregative optimal control problem.

2.1.1 Notations

On sets For two sets $\mathcal{A}$ and $\mathcal{B}$ in a normed vector space $\mathcal{X}$, we denote by $d(\mathcal{A})\colon = \sup_{x,y\in\mathcal{A}} \|x - y\|_{\mathcal{X}}$ the diameter of $\mathcal{A}$, by $\mathcal{A}+\mathcal{B} = \{x+y \mid x\in\mathcal{A}, y\in\mathcal{B}\}$ the Minkowski sum of A and B, by $\lambda\mathcal{A} = \{\lambda x \mid x \in \mathcal{A}\}$ the scalar multiplication of A with $\lambda\in\mathbb{R}$ and by $\text{conv}(\mathcal{A})$ the convex hull of $\mathcal{A}$. Note that $\text{conv}(\mathcal{A}+\mathcal{B}) = \text{conv}(\mathcal{A}) + \text{conv}(\mathcal{B})$.

For all $i \in \{1,\dots,N\}$, we denote $\mathcal{X}_{-i} = \left(\prod_{i'=1}^{i-1}\mathcal{X}_{i'}\right) \times \left(\prod_{i'=i+1}^{N}\mathcal{X}_{i'}\right)$. Given $x\in\mathcal{X}$, we denote $x_{-i} = (x_1,\dots,x_{i-1},x_{i+1},\dots,x_N) \in \mathcal{X}_{-i}$. From time to time, we represent x by the pair (x_i, x_{-i}).

On functions Let $\mathcal{H}$ be a real Hilbert space. Let $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}}$ denote the corresponding scalar product and norm. Let $F\colon \mathcal{H}\to\mathbb{R}\cup\{+\infty\}$. The domain of F, denoted by $\text{dom}(F)$, is defined by $\text{dom}(F) = \{x \mid F(x) \neq +\infty\}$. When F is differentiable, we denote its gradient by ∇F. The gradient is defined as a function from $\mathcal{H}$ to itself. We say that ∇F is L-Lipschitz on a subset $\mathcal{A}$ of $\mathcal{H}$ if for any $x,y\in\mathcal{A}$, we have

$$\|\nabla F(x) - \nabla F(y)\|_{\mathcal{H}} \leq L\|x-y\|_{\mathcal{H}}. \tag{2.1.4}$$

The subgradient of F at some point $x \in \text{dom}(F)$ is denoted by $\partial F(x)$ and defined by

$$\partial F(x) = \{p\in\mathcal{H} \mid F(y) \geq F(x) + \langle p, y-x\rangle, \forall\, y\in\mathcal{H}\}.$$

The Fenchel conjugate of F is denoted by $F^*\colon H\to\mathbb{R}$ and defined by $F^*(p) = \sup_{x\in\mathcal{H}} \langle p,x\rangle - F(x)$.

On measures Given a set Ω, we denote by δ_x the Dirac distribution at some point $x\in\Omega$. We denote by $\mathcal{P}_\delta(\Omega)$ the set of finitely supported probability distributions, defined by

$$\mathcal{P}_\delta(\Omega) := \left\{ \sum_{k=1}^{K} \lambda_k \delta_{x_k} \,\middle|\, K\in\mathbb{N},\, (\lambda_k)_{k=1}^K \in (\mathbb{R}_+)^K,\, (x_k)_{k=1}^K \in \Omega^K,\, \sum_{k=1}^{K}\lambda_k = 1 \right\}.$$

Let $\mu = \sum_{k=1}^K \lambda_k\delta_{x_k} \in \mathcal{P}_\delta(\Omega)$. Given a Hilbert space $\mathcal{H}$ and a mapping $F\colon\Omega\to\mathcal{H}$, we denote

$$E_\mu\big[F\big] = \sum_{k=1}^{K}\lambda_k F(x_k), \qquad \sigma_\mu^2\big[F\big] = \sum_{k=1}^{K}\lambda_k \big\|F(x_k) - E_\mu\big[F\big]\big\|_{\mathcal{H}}^2.$$

In other words, $E_\mu\big[F\big]$ is the integral of F with respect to the measure μ and $\sigma_\mu^2\big[F\big]$ is the variance of the probability measure $\sum_{j=1}^J \lambda_j\delta_{F(x_j)}$, in the sense of [Vil03, Remark 7.5]. Finally, the Bernoulli distribution with parameter $\omega\in[0,1]$ is denoted by $\text{Bern}(\omega)$.

On numbers and real-valued random variables We denote by $m \wedge n$ the minimum of the numbers m and n in $\mathbb{R} \cup \{+\infty\}$. Let X be a real-valued random variable. The expectation of X is denoted by $\mathbb{E}[X]$, the variance of X is denoted by $\mathrm{Var}(X)$ and the conditional expectation of X w.r.t. some $\sigma-$algebra $\mathcal{F}$ is denoted by $\mathbb{E}[X \mid \mathcal{F}]$. Given $\mu \in \mathcal{P}_\delta(\Omega)$ and a random variable X in Ω, the notation $X \sim \mu$ indicates that μ is the probability distribution of X.

2.2 Relaxation by randomization and gap estimation

In this section we first make a structural assumption on the general problem of interest, problem (P). Next we introduce a relaxation of the problem, obtained by randomization. We give an upper bound of the randomization gap in Proposition 2.2.6. Finally we propose a method to recover an approximate solution to (P), given an approximate solution to the randomized problem. Its performance is investigated in Theorem 2.2.9.

2.2.1 Assumptions and constants

We recall that $\mathcal{E}$ is the Cartesian product of M separable real Hilbert spaces $\mathcal{E}_j$. We denote by $\langle \cdot, \cdot \rangle_{\mathcal{E}_j}$ the associated scalar products and by $\| \cdot \|_{\mathcal{E}_j}$ the corresponding norms. Let us emphasize that we will not consider any other norm in the spaces $\mathcal{E}_j$. We equip $\mathcal{E}$ with the scalar product $\langle \cdot, \cdot \rangle$, defined by $\langle (y_1, \ldots, y_M), (y'_1, \ldots, y'_M) \rangle = \sum_{j=1}^{M} \langle y_i, y'_i \rangle_{\mathcal{E}_j}$ and we denote by $\| \cdot \|$ the corresponding norm.

For any $i = 1, \ldots, N$ and for any $j = 1, \ldots, M$, we denote

$$S_{ij} := \big\{ g_{ij}(x_i) \mid x_i \in \mathcal{X}_i \big\} \qquad \text{and} \qquad S_j := \frac{1}{N} \sum_{i=1}^{N} S_{ij}.$$

The following regularity assumption will be in force all along the article.

Assumption A. For $i = 1, 2, \ldots, N$ and $j = 1, 2 \ldots, M$:

1. The range set S_{ij} in $\mathcal{E}_j$ has finite diameter $d_{ij} := d(S_{ij})$.
2. The function f_j is L_j-Lipschitz on $\mathrm{conv}\,(S_j)$.
3. The function f_j is continuously differentiable on a neighborhood of $\mathrm{conv}\,(S_j)$, and ∇f_j is $\tilde{L}_j-$Lipschitz on $\mathrm{conv}\,(S_j)$, in the sense of (2.1.4).

We next define two constants $C_0 > 0$ and $C_1 > 0$ by

$$C_0 = \sum_{j=1}^{M} \Big(L_j \max_{1 \le i \le N} \{d_{ij}\} \Big), \quad \text{and} \quad C_1 = \frac{1}{N} \sum_{j=1}^{M} \Big(\tilde{L}_j \sum_{i=1}^{N} d_{ij}^2 \Big).$$

Remark 2.2.1. We will regularly employ notations of the form $O(h(N, q, k))$, where h is an explicit function of N, q (the dimension of $\mathcal{E}$), and k (some iteration counter). We use it to express the fact that some variable is bounded by $C\, h(N, q, k)$, where the constant C only depends on

$\big(\max_{1\leq i\leq N} d_{ij}\big)_{j=1,\dots,M}$ and the Lipschitz moduli $(L_j)_{j=1,\dots,M}$ and $(\tilde{L}_j)_{j=1,\dots,M}$. With this convention in mind, we have

$$C_0 = O(1) \quad \text{and} \quad C_1 = O(1).$$

Remark 2.2.2. Our results can be applied to aggregative problems of the form

$$\inf_{x\in\mathcal{X}} \sum_{j=1}^{M} f_j\Big(\sum_{i=1}^{N} \hat{g}_{ij}(x_i)\Big),$$

i.e. of the same form as in (P), but without the coefficient $\frac{1}{N}$. Indeed, it suffices to define $g_{ij} = N\hat{g}_{ij}$ to come down to the formulation (P) and to use the fact that $d(g_{ij}(\mathcal{X}_i)) = Nd(\hat{g}_{ij}(\mathcal{X}_i))$. The introduction of the coefficient $\frac{1}{N}$ induces a natural scaling of the problem as N increases. It also enables to us to highlight the convexification of the problem as N becomes large, assuming that the coefficients d_{ij} are uniformly bounded.

We state in the following lemma a straightforward inequality, exhibiting the role of the constant C_0. Note that the role of the constant C_1 will be revealed in Lemma 2.2.6.

Lemma 2.2.3. *Let Assumption A be satisfied. For all $i \in \{1,\dots,N\}$, for all $x_{-i} \in \mathcal{X}_{-i}$, x_i and x'_i in $\mathcal{X}_i$, it holds:*

$$|J(x'_i, x_{-i}) - J(x_i, x_{-i})| \leq \frac{C_0}{N}.$$

2.2.2 The randomized problem

The *randomized problem* is obtained by replacing each optimization variable x_i by a probability measure $\mu_i \in \mathcal{P}_\delta(\mathcal{X}_i)$. The contribution mappings $g_i(x_i)$ are replaced by their integral with respect to μ_i, $E_{\mu_i}\big[g_i\big]$. Denoting $\mathcal{P}_\delta = \prod_{i=1}^{N} \mathcal{P}_\delta(\mathcal{X}_i)$, we obtain

$$\inf_{\mu\in\mathcal{P}_\delta} \mathcal{J}(\mu) := f\Big(\frac{1}{N}\sum_{i=1}^{N} E_{\mu_i}\big[g_i\big]\Big) = \sum_{j=1}^{M} f_j\Big(\frac{1}{N}\sum_{i=1}^{N} E_{\mu_i}\big[g_{ij}\big]\Big). \tag{PR}$$

The following equality justifies the denomination of the relaxed problem: given $\mu \in \mathcal{P}_\delta$ and given N random variables X_i in $\mathcal{X}_i$ such that $X_i \sim \mu_i$, we have

$$\mathcal{J}(\mu) = f\Big(\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\big[g_i(X_i)\big]\Big). \tag{2.2.1}$$

Remark 2.2.4. Working with probability measures with finite support, we do not need to equip the sets $\mathcal{X}_i$ with a topology and to consider regularity assumptions on the mappings g_i. Note that the original problem and the randomized one do not necessarily have a solution under the standing assumptions of the article.

Let J^* and $\mathcal{J}^*$ denote the values of the primal problem (P) and the randomized problem (PR) respectively. One is interested in comparing J^* and $\mathcal{J}^*$. The next lemma gives a direct result for one direction of this comparison.

Lemma 2.2.5. *Let Assumption A hold true. Then $-\infty < \mathcal{J}^* \leq J^*$.*

Proof. By the definitions of $E_{\mu_i}[g_{ij}]$ and S_j, we have that $\frac{1}{N}\sum_{i=1}^N E_{\mu_i}[g_{ij}] \in \text{conv}(S_j)$. Since f_j is Lipschitz-continuous over the bounded set $\text{conv}(S_j)$, we deduce that $\mathcal{J}^* > -\infty$. Let $x \in \mathcal{X}$. Define $\mu = (\delta_{x_1}, \dots, \delta_{x_N}) \in \mathcal{P}_\delta$. Then $\mathcal{J}(\mu) = J(x)$. As a consequence, inequality $\mathcal{J}^* \leq J^*$ follows. □

The *randomization gap* is then defined as

$$\text{randomization gap} = J^* - \mathcal{J}^* \geq 0.$$

Next we prove a first upper bound of the randomization gap, of order $O(\frac{1}{N})$.

Proposition 2.2.6. *Let Assumption A hold true. Let $\mu \in \mathcal{P}_\delta$ and let $(X_i)_{i=1,\dots,N}$ denote N independent random variables such that $X_i \sim \mu_i$. Then,*

$$\mathbb{E}[J(X)] - \mathcal{J}(\mu) \leq \frac{1}{2N^2}\sum_{j=1}^M \Big(\tilde{L}_j \sum_{i=1}^N \sigma^2_{\mu_i}[g_{ij}]\Big) \leq \frac{C_1}{2N}, \tag{2.2.2}$$

where $X = (X_1, \dots, X_N)$. As a consequence, $J^* - \mathcal{J}^* \leq \frac{C_1}{2N}$.

Proof. Let us define $Y_j = \frac{1}{N}\big(\sum_{i=1}^N g_{ij}(X_i)\big)$, for $j = 1, \dots, M$. Let us set $Y = (Y_j)_{j=1,\dots,M}$. We have

$$\mathbb{E}\big[J(X)\big] = \mathbb{E}\big[f(Y)\big] \qquad \text{and} \qquad \mathcal{J}(\mu) = f\big(\mathbb{E}[Y]\big).$$

Since the variables X_i are independent, the random variables $g_{ij}(X_i)$ are also independent (for fixed j). It follows that

$$\mathbb{E}\Big[\big\|Y_j - \mathbb{E}\big[Y_j\big]\big\|^2_{\mathcal{E}_j}\Big] = \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\Big[\big\|g_{ij}(X_i) - \mathbb{E}[g_{ij}(X_i)]\big\|^2_{\mathcal{E}_j}\Big] = \frac{1}{N^2}\sum_{i=1}^N \sigma^2_{\mu_i}\big[g_{ij}\big].$$

By Assumption A, we have

$$f(Y) \leq f\big(\mathbb{E}[Y]\big) + \big\langle \nabla f(\mathbb{E}[Y]), Y - \mathbb{E}[Y]\big\rangle_{\mathcal{E}_j} + \frac{1}{2}\sum_{j=1}^M \Big(\tilde{L}_j \big\|Y_j - \mathbb{E}\big[Y_j\big]\big\|^2_{\mathcal{E}_j}\Big).$$

Taking the expectation of the above inequality and recalling the definition of C_1, we deduce (2.2.2). □

Remark 2.2.7. As we explained in the introduction, our analysis covers the case of a non-separable cost f (when $M = 1$), however, when f is separable, it is useful to take this property into account. The aim of this rem is to justify this fact. Let us assume (in this remark only) that f is indeed separable, i.e. $M > 1$. Let us treat f as a non-separable function. It is easy to verify that the mapping ∇f is Lipschitz continuous with modulus $\big(\max_{j=1,\dots,M} \tilde{L}_j\big)$; this estimate is tight. If we do not take into account the additive structure of f in the proof of Proposition 2.2.6, we end up with the following estimate:

$$\mathbb{E}\big[J(X)\big] \leq \mathcal{J}(\mu) + \frac{1}{2N^2}\Big(\max_{j=1,\dots,M} \tilde{L}_j\Big)\sum_{i=1}^N \sum_{j=1}^M \sigma^2_{\mu_i}\big[g_{ij}\big],$$

which is less precise than inequality (2.2.2). The same kind of comment could be made for the constants appearing afterwards in the convergence results of our numerical method.

We finish this subsection with an equivalent relaxed problem in the situation when the sets $\mathcal{X}_i$ (resp. the contribution functions g_i) are identical. We refer to this situation as the *symmetric case.*

Lemma 2.2.8. *Suppose that there exists a set $\mathcal{X}$ and a function $g\colon \mathcal{X} \to \mathcal{E}$ such that $\mathcal{X}_i = \mathcal{X}$ and $g_i = g$, for all i. Then,*

$$\mathcal{J}^* = \inf_{\nu \in \mathcal{P}_\delta(\mathcal{X})} f\big(E_\nu[g]\big). \tag{2.2.3}$$

Proof. Let $\nu \in \mathcal{P}_\delta(\mathcal{X})$. Take $\mu = (\nu, \dots, \nu) \in \mathcal{P}_\delta$. It follows that $f\left(E_\nu[g]\right) = \mathcal{J}(\mu)$. As a consequence, $\inf_{\nu \in \mathcal{P}_\delta(\mathcal{X})} f\left(E_\nu[g]\right) \leq \inf_{\mu \in \mathcal{P}_\delta} \mathcal{J}(\mu)$. On the other hand, let $\bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_N) \in \mathcal{P}_\delta$. Take $\bar{\nu} = \sum_{i=1}^N \bar{\mu}_i / N \in \mathcal{P}_\delta(\mathcal{X})$. Then, we deduce that $\mathcal{J}(\bar{\mu}) = f\left(E_{\bar{\nu}}[g]\right)$. The conclusion follows. □

The relaxed problem in (2.2.3) has a natural interpretation as a mean field relaxation: instead of considering an optimization problem with N symmetric agents, we consider an arbitrarily large number of agents and optimize their distribution ν.

2.2.3 Selection method

Suppose that a minimizer or an approximate minimizer μ of the randomized problem (PR) has been obtained. We address in this subsection the issue of recovering an approximate minimizer of the original problem (P) from μ.

A naive approach would consist in *averaging* the measures μ_i, assuming that the sets $\mathcal{X}_i$ are convex. In such a case, one can define the point $x_i = E_{\mu_i}[\mathrm{Id}]$. Another approach, motivated by Proposition 2.2.6, consists in sampling μ, that is, in simulating N independent random variables $(X_1, \dots, X_N)$, with distributions $X_i \sim \mu_i$. This can be done without additional structural assumption on the sets $\mathcal{X}_i$, moreover, Proposition 2.2.6 ensures that for any $\varepsilon > 0$,

$$\mathbb{P}\Big[J(X_1, \dots, X_N) < \mathcal{J}(\mu) + \frac{C_1}{2N} + \varepsilon\Big] > 0. \tag{2.2.4}$$

Of course, one can realize several samplings of μ to increase the probability of finding a good candidate for the original problem. We will refer to this approach as the *selection method.*

Example. Consider the following instance of the problem (P), where N is a large even number:

$$\begin{cases} \text{minimize} \left\{ J(x_1, x_2, \dots, x_N) = -\frac{1}{N} \sum_{i=1}^N x_i^2 + \Big(\frac{1}{N} \sum_{i=1}^N x_i\Big)^2 \right\}; \\ \text{subject to } x_i \in [-1, 1], \quad i = 1, \dots, N. \end{cases} \tag{2.2.5}$$

It is easy to see that x^* is a minimizer of (2.2.5) if and only if x^* has $N/2$ coordinates equal to 1 and the others equal to -1. In this example, the original and the relaxed problem have the same value, $J^* = \mathcal{J}^* = -1$. The relaxed problem does not have a unique solution. One of them is $\tilde{\mu}_i = \frac{1}{2}\big(\delta_{-1} + \delta_1\big)$. Averaging $\tilde{\mu}$ as suggested above yields $\tilde{x} = (0, \dots, 0)$ and $J(\tilde{x}) = 0$. Thus in this example, the averaging method yields a poor candidate, whatever the value of N.

On the other hand, the selection method yields good candidates when N is large. Indeed, assume that $\mathbb{P}[X_i = -1] = \mathbb{P}[X_i = 1] = 1/2$. When N is large, by the law of large numbers [Tsy09], nearly half of the random variables X_i are equal to 1 while the others are equal to -1, with probability close to 1. Then in such a case X is almost a minimizer of (2.2.5).

The next theorem provides a sharp estimate of the probability in (2.2.4) and confirms the interest of the selection method for large values of N. It relies on a concentration inequality, *McDiarmid's inequality* [McD89], and its variant [Del15] (cf. Corollary 2.9.2) of "variance type". It is quite intuitive that if the probability measures μ_i have a small variance (in a sense to be specified), then the selection method will be more efficient. The interest of taking into account the variances of the probability distributions will be revealed in the analysis of the stochastic Frank-Wolfe algorithm in Subsection 2.3.3.

Theorem 2.2.9. *Let Assumption A be satisfied. Let $\mu \in \mathcal{P}_\delta$ and let $X_1, \ldots, X_N$ be N independent random variables such that $X_i \sim \mu_i$. Let $X = (X_1, \ldots, X_N)$. Then, for all $\epsilon > 0$,*

$$\mathbb{P}\left[J(X) < \mathcal{J}(\mu) + \frac{C_1}{2N} + \epsilon\right] \geq 1 - \exp\left(-\frac{2N\epsilon^2}{C_0^2}\right). \tag{2.2.6}$$

Assume further that for all $i = 1, \ldots, N$, there exists a constant v_i such that

$$\sigma^2_{\mu_i}\big[J(\cdot, x_{-i})\big] \leq v_i^2, \tag{2.2.7}$$

for all $x_{-i} \in X_{-i}$. Then (2.2.6) can be strengthened as:

$$\mathbb{P}\left[J(X) < \mathcal{J}(\mu) + \sum_{j=1}^{M}\sum_{i=1}^{N} \frac{\tilde{L}_j}{2N^2}\sigma^2_{\mu_i}\big[g_{ij}\big] + \epsilon\right] \geq 1 - \exp\left(-\frac{N\epsilon^2}{2\left(\sum_{i=1}^N N v_i^2 + \frac{C_0\epsilon}{3}\right)}\right). \tag{2.2.8}$$

Proof. Combining Lemma 2.2.3 and McDiarmid's inequality [McD89], we obtain

$$\mathbb{P}\Big[J(X) < \mathbb{E}\big[J(X)\big] + \epsilon\Big] \geq 1 - \exp\left(-\frac{2N\epsilon^2}{C_0^2}\right).$$

Combining this estimate with the second inequality of Proposition 2.2.6, we obtain (2.2.6).

Estimate (2.2.8) is proved similarly, combining McDiarmid's inequality of "variance type" proved in Corollary 2.9.2 and the first inequality of Proposition 2.2.6. □

We provide in the next lemma an explicit candidate for (2.2.7).

Lemma 2.2.10. *Inequality (2.2.7) is satisfied with $v_i^2 = \frac{1}{N^2}\big(\sum_{j=1}^M L_j^2\big)\sigma^2_{\mu_i}(g_i)$.*

Proof. We first state a general following property: given a probability measure μ and two maps h_1 and h_2 suitably defined, we have the inequality

$$\sigma^2_\mu\big[h_1 \circ h_2\big] \leq L^2 \sigma^2_\mu\big[h_2\big], \tag{2.2.9}$$

assuming that h_1 is L-Lipschitz continuous. Let us prove this property. For any x, we have

$$\begin{aligned}\big\|h_1 \circ h_2(x) - E_\mu[h_1 \circ h_2]\big\|^2 = {} & \big\|h_1 \circ h_2(x) - h_1(E_\mu[h_2])\big\|^2 \\ & + 2\big\langle h_1 \circ h_2(x) - h_1(E_\mu[h_2]), h_1(E_\mu[h_2]) - E_\mu[h_1 \circ h_2]\big\rangle \\ & + \big\|h_1(E_\mu[h_2]) - E_\mu[h_1 \circ h_2]\big\|^2.\end{aligned}$$

Taking the expectation, we obtain that

$$\sigma_\mu^2[h_1 \circ h_2] = E_\mu\Big[\big\|h_1 \circ h_2 - h_1(E_\mu[h_2])\big\|^2\Big] - \big\|h_1(E_\mu[h_2]) - E_\mu[h_1 \circ h_2]\big\|^2.$$

Since h_1 is L-Lipschitz continuous, we have $E_\mu\big[\|h_1 \circ h_2 - h_1(E_\mu[h_2])\|^2\big] \leq L^2 \sigma_\mu[h_2]^2$. Inequality (2.2.9) follows immediately. Next, it is easy to verify that the function f is L-Lipschitz continuous, with $L = \big(\sum_{j=1}^M L_j^2\big)^{1/2}$. Using (2.2.9), we conclude that

$$\sigma_{\mu_i}^2\big[J(\cdot, x_{-i})\big] \leq L^2 \sigma_{\mu_i}^2\Big[\frac{1}{N} g_i(\cdot) + C\Big] = \frac{L^2}{N^2}\sigma_{\mu_i}^2\big[g_i\big],$$

where $C = \frac{1}{N}\sum_{i' \neq i} g_{i'}(x_{i'})$ is regarded as a constant. The estimate follows. □

2.3 Stochastic Frank-Wolfe algorithm

2.3.1 Assumptions

We introduce two new assumptions, which will be in force until the end of the article.

Assumption B. For all $j = 1, \ldots, M$, the function $f_j \colon \mathcal{E}_j \to \mathbb{R}$ is convex over $\mathrm{conv}(S_j)$.

Let μ^1 and μ^2 lie in $\mathcal{P}_\delta$. Take $\omega \in [0,1]$. Let $\mu = (\mu_1, \ldots, \mu_N)$ be defined, for any $i = 1, \ldots, N$, by $\mu_i = (1-\omega)\mu_i^1 + \omega\mu_i^2$. Here, the addition and the multiplication by a scalar are understood as usual in the set of signed measures. In the sequel, we simply denote $\mu = (1-\omega)\mu^1 + \omega\mu^2$. We have $\mu \in \mathcal{P}_\delta$; moreover, $E_{\mu_i}[g_i] = (1-\omega)E_{\mu_i^1}[g_i] + \omega E_{\mu_i^2}[g_i]$, for any $i = 1, \ldots, N$. Then, Assumption B implies that $\mathcal{J}(\mu) \leq (1-\omega)\mathcal{J}(\mu^1) + \omega\mathcal{J}(\mu^2)$. In words, the randomized problem (PR) is convex.

In this section, we address the numerical resolution of the randomized problem (and the original problem) under Assumption B. Let us mention that this convexity assumption is natural for the application problems described in the introduction. It allows the application of the Frank-Wolfe algorithm (also called conditional gradient algorithm) [DH78], for which convergence can be established. The Frank-Wolfe algorithm requires to solve at each iteration a subproblem. Here, the subproblems can be decomposed in N optimization problems, which can be solved in parallel. This property is particularly interesting, since we aim at solving instances of (P) with large values of N. We do not detail here the practical resolution of the subproblems, which can only be investigated case by case. Instead, we make the following assumption. Let us set $\mathcal{A} := \{\nabla f(y) \mid y \in \mathrm{conv}(G(\mathcal{X}))\} \subset \mathcal{E}$.

Assumption C. For all $i = 1, \ldots, N$, for all $\lambda \in \mathcal{A}$, the problem

$$\inf_{x_i \in \mathcal{X}_i} \langle \lambda, g_i(x_i) \rangle \tag{2.3.1}$$

has at least a solution. For all $i = 1, \ldots, N$, we fix a map $\mathbb{S}_i \colon \mathcal{A} \mapsto \mathcal{X}_i$ such that for any $\lambda \in \mathcal{A}$, $\mathbb{S}_i(\lambda)$ is a solution to (2.3.1).

The map $\mathbb{S}_i$ can be understood as a best-response function corresponding to agent i. The involved cost function is a linear combination of the contribution mappings g_{ij}, with $j = 1, \ldots, M$. In problem (2.3.1), λ can be interpreted as a price variable associated with $g_i(x_i)$.

Remark 2.3.1. It is easy to find assumptions which ensure the existence of the map $\mathbb{S}_i$. For example, one can assume that $\mathcal{X}_i$ is a compact set in a topological vector space and that g_i is continuous. Let us emphasize that Assumption C is essentially an assumption of numerical nature: $\mathbb{S}_i$ should be understood as the output of an (efficient) numerical procedure for the resolution of (2.3.1). The algorithms described afterwards largely rely on evaluations of $\mathbb{S}_i$.

2.3.2 Basic Frank-Wolfe algorithm

We first describe a rather direct application of the Frank-Wolfe algorithm, which is referred to as the basic Frank-Wolfe algorithm. The starting point of our numerical approach is the following lemma, the proof of which is straightforward.

Lemma 2.3.2. *Let $\lambda \in \mathcal{A}$ and let $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_N) \in \mathcal{P}_\delta$. Then, $\bar{\mu}$ is a solution to*

$$\inf_{\mu \in \mathcal{P}_\delta} \Big\langle \lambda, \frac{1}{N} \sum_{i=1}^{N} E_{\mu_i}(g_i) \Big\rangle. \tag{2.3.2}$$

if and only if for all $i = 1, \ldots, N$, $\bar{\mu}_i$ is supported in $\operatorname{argmin}_{x_i \in \mathcal{X}_i} \langle \lambda, g_i(x_i) \rangle$.

The cost function in (2.3.2) should be regarded as a linearization of $\mathcal{J}$, as needed in the abstract formulation of the Frank-Wolfe algorithm in [DH78]. An immediate consequence of Lemma 2.3.2 is that $(\delta_{\mathbb{S}_1(\lambda)}, \ldots, \delta_{\mathbb{S}_N(\lambda)})$ is a solution to (2.3.2). The resolution of problem (2.3.2) is a key step in the numerical procedures developed afterwards; let us emphasize that the maps $\mathbb{S}_i(y)$ can be evaluated independently from each other, i.e. the resolution of (2.3.2) can be parallelized.

Algorithm 2.1: Frank-Wolfe Algorithm

Initialization: $\mu^0 \in \mathcal{P}_\delta$;
for $k = 0, 1, \ldots, K$ **do**
  Step 1: Resolution of the subproblems.
  Set $y^k = \frac{1}{N} \sum_{i=1}^{N} E_{\mu_i^k}\big[g_i\big]$ and set $\lambda^k = \nabla f(y^k)$;
  for $i = 1, \ldots, N$ **do**
    Compute $\bar{x}_i^k = \mathbb{S}_i(\lambda^k)$;
  end
  Step 2: Update.
  Set $\bar{\mu}^k = (\delta_{\bar{x}_1^k}, \ldots, \delta_{\bar{x}_N^k})$;
  Set $\mu^{k+1} = (1 - \omega_k)\mu^k + \omega_k \bar{\mu}^k$.
end

The convergence analysis performed afterwards relies on standard arguments (compare our proof with [Jag13]). We introduce the primal gap γ_k and the primal-dual gap β_k, defined by

$$\gamma_k = \mathcal{J}(\mu^k) - \mathcal{J}^*, \quad \beta_k = \langle \nabla f(y^k), y^k - \bar{y}^k \rangle, \quad \text{where:} \quad \bar{y}^k = \frac{1}{N} \sum_{i=1}^{N} g_i(\bar{x}_i^k). \tag{2.3.3}$$

Note that β_k can be evaluated numerically. The following lemma shows that β_k is an upper bound of the primal gap γ_k.

Lemma 2.3.3. *For all $k \in \mathbb{N}$, $\gamma_k \leq \beta_k$.*

Proof. Let $k \in \mathbb{N}$. Let $\mu \in \mathcal{P}_\delta$ and let $y = \frac{1}{N}\sum_{i=1}^N E_{\mu_i}[g_i]$. By Lemma 2.3.2, we have $\langle \nabla f(y^k), \bar{y}^k \rangle \leq \langle \nabla f(y^k), y \rangle$. Thus, using the convexity of f, we obtain

$$\beta_k = \langle \nabla f(y^k), y^k - \bar{y}^k \rangle \geq \langle \nabla f(y^k), y^k - y \rangle \geq f(y^k) - f(y) = \mathcal{J}(\mu^k) - \mathcal{J}(\mu). \tag{2.3.4}$$

Since μ is arbitrary, we deduce that $\beta_k \geq \mathcal{J}(\mu^k) - \mathcal{J}^* = \gamma_k$. □

We have the following convergence result.

Proposition 2.3.4. *Let Assumptions A, B, and C hold. Then, in Algorithm 2.1, for any $K \in \mathbb{N}^*$,*

$$\gamma_K \leq \frac{2C_1}{K}.$$

Proof. As we will see, the result is a consequence of Lemma 2.9.3, with $C = \frac{C_1}{2}$ and $u_k = 0$. By Assumption A,

$$f(y^{k+1}) \leq f(y^k) + \langle \nabla f(y^k), y^{k+1} - y^k \rangle + \sum_{j=1}^M \frac{\tilde{L}_j}{2} \|y_j^{k+1} - y_j^k\|^2.$$

We have $y^{k+1} - y^k = \omega_k(\bar{y}^k - y^k)$. Therefore, by definition of β_k,

$$f(y^{k+1}) \leq f(y^k) - \omega_k \beta_k + \omega_k^2 \sum_{j=1}^M \frac{\tilde{L}_j}{2} \|\bar{y}_j^k - y_j^k\|^2. \tag{2.3.5}$$

By definition, $\|\bar{y}_j^k - y_j^k\|^2 = \frac{1}{N^2}\big\| \sum_{i=1}^N E_{\mu_i^k}\big[g_{ij}(\bar{x}_i^k) - g_{ij}(\cdot)\big]\big\|^2$, thus by Cauchy-Schwarz inequality,

$$\|\bar{y}_j^k - y_j^k\|^2 \leq \frac{1}{N}\sum_{i=1}^N \big\|E_{\mu_i^k}\big[g_{ij}(\bar{x}_i^k) - g_{ij}(\cdot)\big]\big\|^2 \leq \frac{1}{N}\sum_{i=1}^N d_{ij}^2.$$

Combining the above estimate with (2.3.5) and using the inequality $\gamma_k \leq \beta_k$ proved in Lemma 2.3.3, we obtain that $\gamma_{k+1} \leq (1-\omega_k)\gamma_k + \frac{C_1}{2}\omega_k^2$. Thus Lemma 2.9.3 applies, which concludes the proof. □

In the following remark, we give an alternative value of ω_k in Step 2 of Algorithm 2.1, while preserving the convergence rate from the previous proposition.

Remark 2.3.5. For any $k \in \mathbb{N}$, denote $h_k(\omega) = -\omega\beta_k + \frac{C_k}{2}\omega^2$, where the constant C_k is defined by $C_k = \sum_{j=1}^M \tilde{L}_j \|\bar{y}_j^k - y_j^k\|^2$. In view of inequality (2.3.5), the result of Proposition 2.3.4 remains true if the sequence $(\omega_k)_{k\in\mathbb{N}}$ is chosen such that for any $k \in \mathbb{N}$, $h(\omega_k) \leq h(\bar{\omega}_k)$. The result remains in particular true for

$$\omega_k = \operatorname*{argmin}_{\omega\in[0,1]} h(\omega) = \min\Big(\frac{\beta_k}{C_k}, 1\Big). \tag{2.3.6}$$

The above proposition shows the convergence of the Frank-Wolfe algorithm. Yet the algorithm only provides a relaxed solution. In order to get a solution to the original problem, one can use the selection method introduced in Subsection 2.2.3. A first direct application of Proposition 2.2.6 yields the following. Let $(X_1, \dots, X_N)$ be N independent random variables such that $X_i \sim \mu_i^k$, for all i. Then,

$$\mathbb{E}\big[J(X)\big] \leq J^* + \frac{2C_1}{k} + \frac{C_1}{2N}.$$

Therefore, from a theoretical point of view, there is no guaranty of improvements when $k \gg N$ since, then, the error term $\frac{2C_1}{k}$ becomes negligible in comparison with $\frac{C_1}{2N}$. The following lemma provides a convergence result (in probability) for the combination of the Frank-Wolfe algorithm and the selection method, for a number of iterations $k \leq N$.

Lemma 2.3.6. *Let $(\mu_k)_{k\in\mathbb{N}}$ be the output of Algorithm 2.1. Let $k \leq N$. Let $\zeta \in (0,1)$. Let $n \in \mathbb{N}^*$ and let $(X_i^j)_{i=1,\dots,N}^{j=1,\dots,n}$ be Nn independent random variables such that $X_i^j \sim \mu_i^k$. Let $X^j = (X_1^j, \dots, X_N^j)$. Then,*

$$\mathbb{P}\left[\min_{j=1,\dots,n} J(X^j) < \mathcal{J}^* + \frac{3C_1}{k}\right] \geq 1 - \zeta, \quad \textit{if} \quad n \geq \frac{2C_0^2}{C_1^2}\,\frac{k^2}{N}\,\ln\Big(\frac{1}{\zeta}\Big). \tag{2.3.7}$$

Proof. Since $k \leq N$, we have $\frac{C_1}{2N} \leq \frac{C_1}{2k}$. Therefore, by Theorem 2.2.9,

$$\mathbb{P}\left[\min_{j=1,\dots,n} J(X^j) < \mathcal{J}^* + \frac{2C_1}{k} + \frac{C_1}{2k} + \epsilon\right] \geq 1 - \exp\Big(-\frac{2N\epsilon^2 n}{C_0^2}\Big),$$

for any $\epsilon > 0$. Take $\epsilon = \frac{C_1}{2k}$. If n satisfies (2.3.7), then $\exp\big(-\frac{2N\epsilon^2 n}{C_0^2}\big) \leq \zeta$. □

2.3.3 Stochastic Frank-Wolfe algorithm

At each iteration of Algorithm 2.1, a new point $\bar{x}_i^k$ is added to the support of each distribution μ_i^k. Therefore, if at iteration K, the points $(\bar{x}_i^k)_{k=0,\dots,K-1}$ are distinct from each other (for each i), then KN places are needed to store the iterate μ^K, which can be prohibitive as K becomes large. We propose in this subsection a variant of Algorithm 2.1 which significantly mitigates the risk of memory overflow. We call it the *Stochastic Frank-Wolfe* (SFW) algorithm, it is given in Algorithm 2.2 below.

Starting from an initialization $x^0 \in \mathcal{X}$, Algorithm 2.2 generates a sequence $(x^k)_{k\in\mathbb{N}}$ in $\mathcal{X}$. Let us emphasize that there is no probability distribution involved in the practical implementation of Algorithm 2.2. However, for the analysis of the algorithm and for its description, it is convenient to introduce $\mu^k = (\delta_{x_1^k}, \dots, \delta_{x_N^k})$. With this notation at hand, we first observe that y^k, as defined in Step 1 of Algorithm 2.2, satisfies $y^k = \frac{1}{N}\sum_{i=1}^N E_{\mu_i^k}[g_i]$. Thus the Steps 1 of Algorithms 2.1 and 2.2 play exactly the same role. Let us focus next on Step 2 of Algorithm 2.2 and let us define $\bar{\mu}^k = (\delta_{\bar{x}_i^1}, \dots, \delta_{\bar{x}_N^k})$ and $\hat{\mu}^k = (1-\omega_k)\mu^k + \omega_k\bar{\mu}^k$. In contrast with Algorithm 2.1, we do not directly use $\hat{\mu}^k$ at the next iteration but instead employ our selection method so that $\hat{\mu}^k$ is reduced to an N-uplet of Dirac measures. The application of the selection method is here simple since $\hat{\mu}_i^k = (1-\omega_k)\delta_{x_i^k} + \omega_k\delta_{\bar{x}_i^k}$. Thus, to simulate a random variable with distribution $\hat{\mu}_i^k$, it suffices to

Algorithm 2.2: Stochastic Frank-Wolfe Algorithm

Initialization: $x^0 \in \mathcal{X}$;
for $k = 0, 1, 2, \ldots, K$ **do**

  Step 1: Resolution of the subproblems.
  Compute $y^k = \frac{1}{N} \sum_{i=1}^N g_i(x_i^k)$;
  for $i = 1, 2, \ldots, N$ **do**
    Find $\bar{x}_i^k \in \operatorname{argmin}_{x_i \in \mathcal{X}_i} \langle \nabla f(y^k), g_i(x_i) \rangle$;
  end

  Step 2: Update.
  Choose $n_k \in \mathbb{N}^*$. Set $\omega_k = 2/(k+2)$.;
  for $j = 1, 2, \ldots, n_k$ **do**
    for $i = 1, 2, \ldots, N$ **do**
      Simulate $P_i^{k,j} \sim \text{Bern}(\omega_k)$, independently of all previously defined random variables;
      Set $\hat{x}_i^{k,j} = (1 - P_i^{k,j}) x_i^k + P_i^{k,j} \bar{x}_i^k$;
    end
    Set $\hat{x}^{k,j} = (\hat{x}_i^{k,j})_{i=1,\ldots,N}$;
  end
  Find $x^{k+1} \in \operatorname{argmin}\{J(x) \,|\, x \in X^k\}$, where $X^k = \{\hat{x}^{k,j},\ j = 1, 2, \ldots, n_k\}$;

end

simulate a random variable P with Bernoulli distribution $\text{Bern}(\omega_k)$ and to consider $(1-P)x_i^k + P\bar{x}_i^k$. Using this method, Step 2 consists in simulating n_k random variables $(\hat{x}^{k,j})_{j=1,\ldots,n_k}$ such that their probability distribution is equal to $\hat{\mu}^k$ (to be rigorous, their probability distribution conditionally to x^k). Finally, Step 2 selects a random variable $\hat{x}^{k,j}$ which minimizes J.

It is important to keep in mind that all variables involved in the algorithm (x^k, $\bar{x}^k$, $\hat{x}^{k,j}$) and all variables defined above (μ^k, $\bar{\mu}^k$, $\hat{\mu}^k$) are themselves random variables, since they depend on the Bernoulli random variables $P_i^{k,j}$. For the analysis of the algorithm, we need to consider the filtration generated by the Bernoulli random variables. We introduce the set of indices $\mathcal{I}$ defined by

$$\mathcal{I} = \Big\{ (k, j, i) \,|\, k \in \mathbb{N},\ j \in \{1, \ldots n_k\},\ i \in \{1, \ldots, N\} \Big\} \cup \big\{ (0, 0, 0) \big\}.$$

We equip the set $\mathcal{I}$ with the lexicographic order: given (k_1, j_1, i_1) and (k_2, j_2, i_2) in $\mathcal{I}$, we write $(k_1, j_1, i_1) < (k_2, j_2, i_2)$ if and only if

$$[k_1 < k_2] \quad \text{or} \quad [k_1 = k_2 \text{ and } j_1 < j_2] \quad \text{or} \quad [(k_1, j_1) = (k_2, j_2) \text{ and } i_1 < i_2].$$

We further write $(k_1, j_1, i_1) \leq (k_2, j_2, i_2)$ if and only if $(k_1, j_1, i_1) < (k_2, j_2, i_2)$ or $(k_1, j_1, i_1) = (k_2, j_2, i_2)$. Note that this order coincides with the simulation order of the random variables $P_i^{k,j}$ in the algorithm. The relation $\leq$ defines a total order with minimal element $(0, 0, 0)$. For any $(k, j, i) \neq (0, 0, 0)$, we denote by $(k, j, i) - 1$ the maximal element of the set $\{(k', j', i') \in \mathcal{I} \,|\, (k', j', i') < (k, j, i)\}$.

Finally, we consider the filtration $(\mathcal{G})_{(k,j,i)\in\mathcal{I}}$ defined by

$$\mathcal{G}_{(k,j,i)} = \begin{cases} \text{trivial } \sigma\text{−algebra}, & \text{if } (k,j,i) = (0,0,0), \\ \sigma\big(\mathcal{G}_{(k,j,i)-1}, P_i^{k,j}\big), & \text{otherwise}, \end{cases}$$

where $\sigma\big(\mathcal{G}_{(k,j,i)-1}, P_i^{k,j}\big)$ denotes the σ-algebra generated by $\mathcal{G}_{(k,j,i)-1}$ and $P_i^{k,j}$. Note that $\hat{x}_i^{k,j}$ is $\mathcal{G}_{(k,j,i)}$-adapted and that x^k and $\bar{x}^k$ are $\mathcal{G}_{(k,1,1)-1}$-adapted.

Theorem 2.3.7. *Let Assumptions A, B, and C hold true. Then, for all $K = 1, \dots, 2N$,*

$$\mathbb{E}[\gamma_K] \leq \frac{4C_1}{K}, \quad \textit{where } \gamma_K = J(x^K) - \mathcal{J}^*.$$

Moreover, for all $\epsilon > 0$,

$$\mathbb{P}\Big[\gamma_K < \frac{4C_1}{K} + \epsilon\Big] \geq 1 - \exp\left(\frac{-\epsilon^2 N}{2(v_K + \epsilon m_K/3)}\right), \tag{2.3.8}$$

where $v_K = \frac{2C_0^2}{K^2(K+1)^2}\left(\sum_{k=1}^{K-1} \frac{k(k+1)^2}{n_k}\right)$ and $m_K = \frac{C_0}{K(K+1)}\left(\max_{k=1,\dots,K-1} \frac{(k+1)(k+2)}{n_k}\right)$. Finally, the following estimates quantify the variability of γ_K:

$$\mathrm{Var}\big[\gamma_K\big] \leq \frac{16C_1^2}{K^2} + \frac{v_K}{N} \quad \textit{and} \quad \mathbb{E}\Big[\Big(\max\Big(\gamma_K - \frac{4C_1}{K}, 0\Big)\Big)^2\Big] \leq \frac{v_K}{N}. \tag{2.3.9}$$

The proof is postponed to Section 2.3.4. Let us note that the constants m_K and v_K involved in the theorem depend on the sequence $(n_k)_{k=0,1,\dots}$ but do not depend on N.

Corollary 2.3.8. *Let $A > 0$. Assume that $n_k \geq \max\big(\frac{Ak^2}{N}, 1\big)$, for any k. Then, for all $K = 1, \dots, 2N$,*

$$\mathbb{P}\Big[\gamma_K < \frac{4C_1 + C_0}{K}\Big] \geq 1 - \exp\left(-\frac{A}{12}\right).$$

Proof. Using $k + 1 \leq 2k$, we obtain

$$\begin{aligned} v_K &\leq \frac{2C_0^2}{K^2(K+1)^2}\left(\sum_{k=1}^{K-1} \frac{Nk(k+1)^2}{Ak^2}\right) \leq \frac{8NC_0^2}{AK^2(K+1)^2}\left(\sum_{k=1}^{K-1} k\right) \\ &= \frac{4NC_0^2(K-1)K}{AK^2(K+1)^2} \leq \frac{4NC_0^2}{AK^2} \end{aligned}$$

and $m_K \leq \frac{C_0}{K(K+1)}\Big(\max_{k=1,\dots,K-1} \frac{N(k+1)(k+2)}{Ak^2}\Big) \leq \frac{6NC_0}{AK^2}$. Applying Theorem 2.3.7 with $\epsilon = \frac{C_0}{K}$, we obtain that $\mathbb{P}\Big[\gamma_K < \frac{4C_1+C_0}{K}\Big] \geq 1 - p$, with

$$p \leq \exp\left(\frac{-(C_0/K)^2 N}{2\Big(\frac{4NC_0^2}{AK^2} + \frac{6NC_0^2}{3AK^3}\Big)}\right) \leq \exp\left(\frac{-A}{12}\right),$$

as was to be proved. ☐

Remark 2.3.9. A variant of Algorithm 2.2 consists in setting $x^{k+1} = x^k$ if $J(\hat{x}^{k,j}) \geq J(x^k)$ for all $j = 1, \dots, n_k$. Theorem 2.3.7 is still satisfied under this modification.

2.3.4 Proof of Theorem 2.3.7 and comments

Step 1: proof of the convergence in expectation We make use of the notations μ^k, $\bar{\mu}^k$, and $\hat{\mu}^k$, introduced right after Algorithm 2.2. We also introduce $\beta_k = \langle \nabla f(y^k), y^k - \bar{y}^k \rangle$, where $\bar{y}^k = \frac{1}{N} \sum_{i=1}^N g_i(\bar{x}_i^k)$. By construction, we have

$$J(x^{k+1}) = \min_{j=1,\dots,n_k} J(\hat{x}^{k,j}) \leq \frac{1}{n_k} \sum_{j=1}^{n_k} J(\hat{x}^{k,j}).$$

Recalling that $\mathcal{J}(\mu^k) = J(x^k)$, we deduce that $\gamma_{k+1} \leq \gamma_k + a_k + b_k + c_k$, where

$$\begin{aligned} a_k &= \frac{1}{n_k} \sum_{j=1}^{n_k} \Big(J(\hat{x}^{k,j}) - \mathbb{E}\big[J(\hat{x}^{k,j}) \,|\, \mathcal{G}_{(k,1,1)-1} \big] \Big), \\ b_k &= \frac{1}{n_k} \sum_{j=1}^{n_k} \Big(\mathbb{E}\big[J(\hat{x}^{k,j}) \,|\, \mathcal{G}_{(k,1,1)-1} \big] - \mathcal{J}(\hat{\mu}^k) \Big), \\ c_k &= \mathcal{J}(\hat{\mu}^k) - \mathcal{J}(\mu^k) = \mathcal{J}(\hat{\mu}^k) - J(x^k). \end{aligned}$$

The term a_k does not play a significant role at the moment since its expectation is null. The term b_k must be understood as a relaxation cost, induced by the use of the selection method. The term c_k is estimated exactly as in Proposition 2.3.4: as was seen in its proof, we have $c_k \leq -\omega_k \beta_k + \omega_k^2 \frac{C_1}{2}$. A direct adaptation of Proposition 2.2.6 shows that

$$b_k \leq \frac{1}{2N^2} \sum_{j=1}^M \sum_{i=1}^N \tilde{L}_j \sigma^2_{\hat{\mu}_i^k}[g_{ij}] \leq \frac{1}{2N^2} \sum_{j=1}^M \sum_{i=1}^N \tilde{L}_j \omega_k (1-\omega_k) d_{ij}^2 = \omega_k (1-\omega_k) \frac{C_1}{2N}.$$

Combining the above estimates, we obtain

$$\gamma_{k+1} \leq \gamma_k + a_k + \Big(-\omega_k \beta_k + \omega_k^2 \frac{C_1}{2} \Big) + \omega_k (1-\omega_k) \frac{C_1}{2N}. \tag{2.3.10}$$

For the choice $\omega_k = \bar{\omega}_k$, we have $(1-\omega_k)/N = k/(N(k+2)) \leq \omega_k$, since $k \leq 2N$. It follows that

$$\omega_k (1-\omega_k) \frac{C_1}{2N} \leq \omega_k^2 \frac{C_1}{2}$$

and finally, since $\gamma_k \leq \beta_k$, we have $\gamma_{k+1} \leq (1-\omega_k)\gamma_k + \omega_k^2 C_1 + a_k$. Next by Lemma 2.9.3,

$$\gamma_K \leq \frac{4C_1}{K} + S_K, \quad \text{where: } S_K = \sum_{k=0}^{K-1} \frac{(k+1)(k+2)}{K(K+1)} a_k. \tag{2.3.11}$$

We have $\mathbb{E}[a_k] = 0$, thus $\mathbb{E}[S_K] = 0$ and finally $\mathbb{E}\big[\gamma_K\big] \leq \frac{4C_1}{K}$.

Step 2: proof of the probability and variance estimates We next need to find an estimate of $\mathbb{P}[S_K \geq \epsilon]$. For this purpose, we need to further decompose the term a_k as a sum of random variables. A first observation is the following equality: $\mathbb{E}\big[J(\hat{x}^{k,j}) \mid \mathcal{G}_{(k,1,1)-1}\big] = \mathbb{E}\big[J(\hat{x}^{k,j}) \mid \mathcal{G}_{(k,j,1)-1}\big]$, which easily follows from Lemma 2.9.5. As a consequence,

$$J(\hat{x}^{k,j}) - \mathbb{E}\big[J(\hat{x}^{k,j}) \mid \mathcal{G}_{(k,1,1)-1}\big] = \sum_{i=1}^N U_{(k,j,i)},$$

where

$$U_{(k,j,i)} = \mathbb{E}\big[J(\hat{x}^{k,j}) \mid \mathcal{G}_{(k,j,i)}\big] - \mathbb{E}\big[J(\hat{x}^{k,j}) \mid \mathcal{G}_{(k,j,i)-1}\big].$$

We obtain the following decomposition of S_K:

$$S_K = \sum_{k=1}^{K-1} \sum_{j=1}^{n_k} \sum_{i=1}^{N} \frac{(k+1)(k+2)}{n_k K(K+1)} U_{(k,j,i)}.$$

Note that the index k starts at 1. Indeed, $\omega_0 = 1$, thus $\hat{x}^{0,j} = \bar{x}^0$ and then $a_0 = 0$. Let us apply Proposition 2.9.1 to S_K. We have $\mathbb{E}\big[U_{(k,j,i)} \mid \mathcal{G}_{(k,j,i)-1}\big] = 0$. Viewing the term $J(\hat{x}^{k,j})$ as a function F of the random variables $A := (P_{i'}^{k',j'})_{(k',j',i')<(k,j,i)}$, $B := P_i^{k,j}$, and $C := (P_{i'}^{k',j'})_{(k,j,i)<(k',j',i')}$, we can apply Lemma 2.9.4 to $U_{(k,j,i)}$, with $\delta = C_0/N$ (by Lemma 2.2.3). This yields

$$U_{(k,j,i)} \leq \frac{C_0}{N} \qquad \text{and} \qquad \mathbb{E}\big[U_{(k,j,i)}^2 \mid \mathcal{G}_{(k,j,i)-1}\big] \leq \frac{\omega_k(1-\omega_k)C_0^2}{N^2}.$$

Therefore, Proposition 2.9.1 applies to $\mathbb{P}[S_K \geq \epsilon]$, where the constants m and v are given by

$$m = \max_{k=1,\ldots,K-1} \frac{(k+1)(k+2)}{n_k K(K+1)} \frac{C_0}{N} = \frac{m_K}{N},$$

$$v = \sum_{k=1}^{K-1} \sum_{j=1}^{n_k} \sum_{i=1}^{N} \Big(\frac{(k+1)(k+2)}{n_k K(K+1)}\Big)^2 \frac{2kC_0^2}{(k+2)^2 N^2} = \frac{v_K}{N}.$$

This proves estimate (2.3.8). Recalling that $\gamma_K \leq \frac{4C_1}{K} + S_K$ a.s., we obtain

$$\mathrm{Var}\big[\gamma_K\big] \leq \mathbb{E}\big[\gamma_K^2\big] \leq \mathbb{E}\Big[\Big(\frac{4C_1}{K} + S_K\Big)^2\Big] = \frac{16C_1^2}{K^2} + \mathbb{E}\big[S_K^2\big].$$

Next by Proposition 2.9.1, $\mathbb{E}\big[S_K^2\big] \leq v_K/N$. The first inequality in (2.3.9) follows. The second inequality follows from the inequality: $\max\big(\gamma_K - \frac{4C_1}{K}, 0\big)^2 \leq S_K^2$.

Remark 2.3.10. Let us set $h_k(\omega) = -\omega\beta_k + \omega^2 \frac{C_1}{2} + \omega(1-\omega)\frac{C_1}{2N}$. If for all $k \in \mathbb{N}$, we have $h_k(\omega_k) \leq h_k(2/(k+2))$, then the convergence in expectation of Theorem 2.3.7 still holds, i.e. $\mathbb{E}\big[\gamma_K\big] \leq 4C_1/K$, in view of inequality (2.3.10). In particular, one can take

$$\omega_k = \operatorname*{argmin}_{\omega\in[0,1]} h_k(\omega) = \max\left(\min\left(\frac{\beta_k - C_1/2N}{C_1(1-1/N)}, 1\right), 0\right). \tag{2.3.12}$$

2.3.5 A speed-up of the SFW algorithm

Step 1 of Algorithm 2.2 requires to solve N independent subproblems. It turns out that only a subset of those subproblems need to be solved for the implementation of Step 2. At iteration k consider the following set:

$$I_k = \bigcup_{j=1,2,\ldots,n_k} \Big\{ i \in \{1,\ldots,N\} \mid P_i^{k,j} = 1 \Big\}.$$

If $i \notin I_k$, then $\hat{x}_i^{k,j} = x_i^k$, in other words, for such an index i, it is not necessary to evaluate $\mathbb{S}_i(\lambda^k)$. A speed-up of the SFW algorithm can therefore be obtained by simulating the Bernoulli random variables before Step 1, next by evaluating $\mathbb{S}_i(\lambda^k)$ only for the indices i in I_k, and finally by computing $\hat{x}^{k,j}$ and x^{k+1} as before. The expectation of the number of subproblems to be solved at iteration k is given by

$$\begin{aligned}\mathbb{E}\big[|I_k|\big] &= \sum_{i=1}^{N} \mathbb{P}\big[i \in I_k\big] = N\big(1 - \mathbb{P}\big[1 \notin I_k\big]\big) = N\Big(1 - \mathbb{P}\big[P_1^{k,j} = 0,\ \forall j = 1, \dots, n_k\big]\Big) \\ &= N\Big(1 - \Big(\frac{k}{k+2}\Big)^{n_k}\Big).\end{aligned}$$

Note that this speed-up technique cannot be applied if ω_k is chosen according to formula (2.3.12). Indeed, this formula requires to evaluate β_k, which implies that the N subproblems must all be solved.

2.3.6 Stopping time strategy

In Algorithm 2.2, the number of samplings n_k is chosen at the beginning of Step 2. We consider here a variant: we generate a sequence of random variables $\hat{x}^{k,j}$ with probability distribution equal to $\hat{\mu}_k$ (conditionally to $\mathcal{G}_{(k,1,1)-1}$); the variables are constructed via Bernoulli variables independent from each other. We define n_k as the first index j such that

$$J(\hat{x}^{k,j}) \leq \mathcal{J}(\hat{\mu}^k) + \Big(\frac{C_1}{2} + C_0\Big)\omega_k^2, \tag{2.3.13}$$

or, equivalently,

$$f(\hat{y}^{k,j}) \leq f\big((1-\omega_k)y^k + \omega_k \bar{y}^k\big) + \Big(\frac{C_1}{2} + C_0\Big)\omega_k^2, \tag{2.3.14}$$

where $\bar{y}^k = \frac{1}{N}\sum_{i=1}^N g_i(\bar{x}_i^k)$ and $\hat{y}^{k,j} = \frac{1}{N}\sum_{i=1}^N g_i(\hat{x}_i^{k,j})$. The next iterate is defined by $x^{k+1} = \hat{x}^{k,n_k}$.

Lemma 2.3.11. *Let $(x^k)_{k\in\mathbb{N}}$ denote the sequence obtained with the stopping rule (2.3.13). Then*

$$J(x^{K+1}) - \mathcal{J}^* \leq \frac{4(C_1 + C_0)}{K}, \quad \forall K = 1, \dots 2N, \quad a.s.$$

Moreover,

$$\mathbb{E}\big[n_k\big] \leq \Big(1 - \exp\Big(-\frac{4N}{(k+2)^3}\Big)\Big)^{-2}, \quad \forall k = 1, \dots, K.$$

Proof. Let $\hat{x}$ be a random variable with probability distribution equal to $\hat{\mu}^k$, conditionally to $\mathcal{G}_{(k,1,1)-1}$. Then, for all $\epsilon > 0$, estimate (2.2.8) of Theorem 2.2.9 yields:

$$\mathbb{P}\Big[J(\hat{x}) \geq \mathcal{J}(\hat{\mu}^k) + \frac{C_1}{2N}\omega_k(1-\omega_k) + \epsilon \,\Big|\, \mathcal{G}_{(k,1,1)-1}\Big] \leq p_\epsilon \tag{2.3.15}$$

where $p_\epsilon = \exp\Big(\frac{-N\epsilon^2}{2(\omega_k(1-\omega_k)C_0^2 + \frac{C_0}{3}\epsilon)}\Big)$. For $\epsilon = C_0\omega_k^2$, we have

$$p_\epsilon = \exp\left(\frac{-NC_0^2\omega_k^4}{2\big(\omega_k C_0^2 - \frac{2}{3}\omega_k^2 C_0^2\big)}\right) \leq p := \exp\left(\frac{-N\omega_k^3}{2}\right) = \exp\left(\frac{-4N}{(k+2)^3}\right).$$

Recalling that $\frac{C_1}{2N}\omega_k(1-\omega_k) \leq \frac{C_1}{2}\omega_k^2$, we deduce that

$$\mathbb{P}\Big[J(\hat{x}) \geq \mathcal{J}(\hat{\mu}^k) + \Big(\frac{C_1}{2} + C_0\Big)\omega_k^2 \,\Big|\, \mathcal{G}_{(k,1,1)-1}\Big] \leq p.$$

Now, let us consider a sequence of independent random variables $(\hat{x}^{k,j})_{j=1,\ldots}$ (conditionally to $\mathcal{G}_{(k,1,1)-1}$), with conditional probability distribution $\hat{\mu}^k$. By estimate (2.3.15),

$$\mathbb{P}\big[n_k = j\big] \leq \mathbb{P}\Big[J(\hat{x}^{k,j'}) \geq \mathcal{J}(\hat{\mu}^k) + \Big(\frac{C_1}{2} + C_0\Big)\omega_k^2,\ \forall j' \,\Big|\, \mathcal{G}_{(k,1,1)-1}\Big] \leq p^{j-1}.$$

We finally deduce that $\mathbb{E}\big[n_k\big] \leq \sum_{n=1}^{\infty} jp^{j-1} = \frac{1}{(1-p)^2}$, which proves the second part of the lemma. For the first part of the lemma, it suffices to observe that

$$J(x^{k+1}) \leq \mathcal{J}(\hat{\mu}^k) + \Big(\frac{C_1}{2} + C_0\Big)\omega_k^2 \leq J(x^k) - \beta_k\omega_k + (C_1 + C_0)\omega_k^2,$$

and to conclude with Lemma 2.9.3. □

2.3.7 Distributed algorithm

In this subsection we present a privacy-preserving implementation of Algorithm 2.2. The Algorithm 2.3 is equivalent to Algorithm 2.2; the instructions are distributed over an **operator**, N **agents**, a **simulator**, and an **aggregator**, who communicate with each other. Roughly speaking, the operator sets up prices that are sent to the agents, which compute independently from each other their best-response. The aggregator computes in a confidential fashion the aggregate associated with a given value of $(x_i)_{i=1,\ldots,N}$. The simulator implements the random variables $P_i^{j,k}$ of the Stochastic Frank-Wolfe algorithm.

More precisely, at the beginning of iterattion k of Algorithm 2.3, the operator sends a price λ_k to the agents, who calculate their best-response. The aggregator sends the corresponding aggregate $\bar{y}_k$ to the operator, who can compute the primal-dual gap β^k and can fix the value of the stepsize ω_k. Next the simulator realizes stochastic simulations, communicated to the agents. Only the aggregate associated with each simulation, $\hat{y}^{k,j}$, is communicated to the operator. The operator decides when to stop the simulation phase through the logical variable *test*. For example, *test* can be set to true as long as $j < n_k$, for predefined values of n_k. The variable test can also be designed so as to implement the stopping rule (2.3.14) of Subsection 2.3.6. Finally, the operator identifies the number j^* of the simulation that has yielded the best aggregate and communicates it to the agents.

The key point in this algorithm is that the operator never receives information that is specific to a given agent: it only collects aggregates (the variables $\bar{y}^k$, $\hat{y}^{k,j}$, and y^k). Similarly, the agents have only access to the prices λ_k and to j^*. We do not detail here algorithms used by the aggregator to compute the aggregate and refers the reader to [BBG$^+$20], which investigates a similar approach for preserving privacy, with an operator that only has access to aggregates (note that the underlying mathematical method is different from ours). It is proposed in that reference to use a cryptographic protocol called *secure multiparty computation* for the non-intrusive computation of aggregates, taken from [SMZ$^+$16] and [ABL$^+$04].

Algorithm 2.3: Distributed SFW Algorithm

[**Agents**] Initialization: $x^0 \in \mathcal{X}$.
[**Aggregator**] Compute and send $y^0 = \frac{1}{N}\sum_{i=1}^N g_i(x_i^0)$ to **Operator**.
for $k = 0, 1, 2, \dots, K$ **do**
  [**Operator**] Compute and send $\lambda^k = \nabla f(y^k)$ to the **Agents**.
  for $i = 1, 2, \dots, N$ **do**
    [**Agent** i] Compute $\bar{x}_i^k \in \mathbb{S}_i(\lambda^k)$.
  end for
  [**Aggregator**] Compute and send $\bar{y}^k = \frac{1}{N}\sum_{i=1}^N g_i(\bar{x}_i^k)$ to **Operator**.
  [**Operator**] Compute $\beta^k = \langle \lambda^k, y^k - \bar{y}^k \rangle$.
  [**Operator**] Compute, send ω_k with (2.3.12) or with $\omega_k = \frac{2}{k+2}$ to **Simulator**.
  [**Operator**] Set $j = 0$ and send $test$ = true to **Simulator**.
  while $test$ **do**
    [**Operator**] Increment j.
    for $i = 1, 2, \dots, N$ **do**
      [**Simulator**] Simulate and send $P_i^{k,j} \sim \mathrm{Bern}(\omega_k)$ to **Agent** i.
      [**Agent** i] Set $\hat{x}_i^{k,j} = (1 - P_i^{k,j})x_i^k + P_i^{k,j}\bar{x}_i^k$.
    end for
    [**Aggregator**] Compute, send $\hat{y}^{k,j} = \frac{1}{N}\sum_{i=1}^N g_i(\hat{x}_i^{k,j})$ to **Operator**.
    [**Operator**] Update and send $test$ to **Simulator**.
  end while
  [**Operator**] Find $j^* \in \underset{j'=1,\dots,j}{\operatorname{argmin}} f(\hat{y}^{k,j'})$. Set $y^{k+1} = \hat{y}^{k,j^*}$.
  [**Operator**] Send j^* to the **Agents**.
  for $i = 1, 2, \dots, N$ **do**
    [**Agent** i] Set $x_i^{k+1} = \hat{x}_i^{k,j^*}$.
  end for
end for

2.4 Refined gap estimates

2.4.1 Nonconvexity measure and gap estimate

We give in this subsection a refinement of the randomization gap obtained in Proposition 2.2.6. Our analysis relies on the concept of nonconvexity measure, introduced in [Cas75].

Definition 2.4.1. Given a subset $\mathcal{K}$ of $\mathcal{E}$, we call nonconvexity measure of $\mathcal{K}$ the number $\rho(\mathcal{K})$ defined by

$$\rho(\mathcal{K}) = \left(\sup_{y \in \operatorname{conv}(\mathcal{K})} \inf_{\substack{\mu \in \mathcal{P}_\delta, \\ E_\mu[\mathrm{Id}]=y}} \sigma_\mu \big[\mathrm{Id} \big]^2 \right)^{1/2},$$

where $\mathrm{Id} \colon \mathcal{E} \to \mathcal{E}$ denotes the identity mapping.

The "nonconvexity measure" terminology is motivated by the following: if $\mathcal{K}$ is convex, then obviously $\rho(\mathcal{K}) = 0$ and conversely, if $\rho(\mathcal{K}) = 0$, then $\mathcal{K}$ is dense into $\operatorname{conv}(\mathcal{K})$. We have the following two properties, easily verified. The map ρ is homogeneous in the following sense: given $a \in \mathbb{R}$, we have $\rho(a\mathcal{K}) = |a|\rho(\mathcal{K})$. Moreover $\rho(\mathcal{K}) \leq d(\mathcal{K})$, where $d(\mathcal{K})$ is the diameter of $\mathcal{K}$. Another particularly interesting property for our aggregative problem is the sub-additivity of $\rho(\cdot)^2$: given two subsets $\mathcal{K}_1$ and $\mathcal{K}_2$, we have $\rho(\mathcal{K}_1 + \mathcal{K}_2)^2 \leq \rho(\mathcal{K}_1)^2 + \rho(\mathcal{K}_2)^2$, see [Cas75, Theorem 1]. We will use an improvement of this inequality in the proof of Theorem 2.4.4, based on the Shapley-Folkman theorem.

The next lemma provides a general relaxation estimate based on the nonconvexity measure of the feasible set. Let us emphasize that the central idea behind this result is the same as the one in the proof of Proposition 2.2.6. The only difference is the point of view, which is here geometric while it was previously probabilistic.

Lemma 2.4.2. *Let $\mathcal{K}$ be a subset of $\mathcal{E}$. Let F be a differentiable real-valued function defined on some neighborhood of* $\operatorname{conv}(\mathcal{K})$*. Assume that ∇F is $\tilde{L}$-Lipschitz continuous over* $\operatorname{conv}(\mathcal{K})$*. Then,*

$$\inf_{y \in \mathcal{K}} F(y) \leq \Big(\inf_{y \in \operatorname{conv}(\mathcal{K})} F(y) \Big) + \frac{\tilde{L}}{2} \rho(\mathcal{K})^2.$$

Proof. Let $y \in \operatorname{conv}(\mathcal{K})$. Let $\mu \in \mathcal{P}_\delta(\mathcal{K})$ be such that $E_\mu[\mathrm{Id}] = y$. Then, since ∇F is $\tilde{L}$-Lipschitz continuous, we have

$$\inf_{y' \in \mathcal{K}} F(y') \leq E_\mu[F] \leq F(y) + \frac{\tilde{L}}{2} \sigma_\mu^2 \big[\mathrm{Id} \big].$$

Minimizing the right-hand side with respect to μ, we obtain that

$$\inf_{y' \in \mathcal{K}} F(y') \leq F(y) + \frac{\tilde{L}}{2} \rho(\mathcal{K})^2.$$

Minimizing the result with respect to y yields the announced estimate. □

Some notations are needed for the application of Lemma 2.4.2 to (P). We set

$$\tilde{g}_{ij}(x_i) = \sqrt{\tilde{L}_j}\, g_{ij}(x_i), \quad \tilde{g}_i(x_i) = \big(\tilde{g}_{ij}(x_i)\big)_{j=1,\dots,M}$$

$$\tilde{f}_j(y_j) = f_j\Big(\tfrac{y_j}{\sqrt{\tilde{L}_j}}\Big), \qquad \tilde{f}(y) = \sum_{j=1}^{M} \tilde{f}_j(y_j).$$

Obviously, $J(x) = \tilde{f}\big(\frac{1}{N}\sum_{i=1}^N \tilde{g}_i(x_i)\big) = \sum_{j=1}^M \tilde{f}_j\big(\frac{1}{N}\sum_{i=1}^N \tilde{g}_{ij}(x_i)\big)$. Finally we denote

$$\mathcal{Y}_i = \tilde{g}_i(\mathcal{X}_i) \quad \text{and} \quad \mathcal{Y} = \frac{1}{N}\sum_{i=1}^N \mathcal{Y}_i.$$

We give next two new formulations of problems (P) and (PR), revealing the geometric nature of the relaxation technique employed so far.

Lemma 2.4.3. *We have*

$$J^* = \inf_{y\in\mathcal{Y}} \tilde{f}(y), \tag{PG}$$

$$\mathcal{J}^* = \inf_{y\in\operatorname{conv}(\mathcal{Y})} \tilde{f}(y). \tag{PGR}$$

Proof. The first equality is straightforward. For the second one, it suffices to observe that $\operatorname{conv}(\mathcal{Y}) = \frac{1}{N}\sum_{i=1}^N \operatorname{conv}(\mathcal{Y}_i)$ and that $\operatorname{conv}(\mathcal{Y}_i) = \big\{E_{\mu_i}[\tilde{g}_i] \mid \mu_i \in \mathcal{P}_\delta(\mathcal{X}_i)\big\}$. □

We introduce the following constants:

$$D_i = \sum_{j=1}^M \tilde{L}_j d_{ij}^2, \qquad D[k] = \max_{\substack{K\subseteq\{1,\dots,N\}\\|K|=k}} \sum_{i\in K} D_i. \tag{2.4.1}$$

Theorem 2.4.4. *Let Assumption A hold true. It holds:*

$$J^* - \mathcal{J}^* \leq \frac{1}{2N^2}\left(\max_{\substack{Q\subseteq\{1,\dots,N\}\\|Q|=q\wedge N}} \sum_{i\in Q} \rho(\mathcal{Y}_i)^2\right) \leq \frac{D[q\wedge N]}{2N^2}. \tag{2.4.2}$$

Note that $D[N] = NC_1$, thus the new gap estimate is the same as the one obtained in Proposition 2.2.6 when $q \geq N$ and it is strictly better when $q < N$.

Proof of Theorem 2.4.4. We let the reader verify that $\nabla\tilde{f}$ is 1-Lipschitz. Then Lemma 2.4.3 and the homogeneity of ρ yield

$$J^* - \mathcal{J}^* \leq \frac{1}{2}\rho(\mathcal{Y})^2 \leq \frac{1}{2N^2}\,\rho\left(\sum_{i=1}^N \mathcal{Y}_i\right)^2.$$

Applying [Cas75, Theorem 2], we obtain that

$$\rho\left(\sum_{i=1}^N \mathcal{Y}_i\right)^2 \leq \max_{\substack{Q\subseteq\{1,\dots,N\}\\|Q|=q\wedge N}} \sum_{i\in Q} \rho(\mathcal{Y}_i)^2,$$

which proves the first inequality. Observing that

$$\rho(\mathcal{Y}_i)^2 \leq d(\mathcal{Y}_i)^2 \leq \sum_{j=1}^{M} d\big(\tilde{g}_{ij}(\mathcal{X}_i)\big)^2 = \sum_{j=1}^{M} \tilde{L}_j d\big(g_{ij}(\mathcal{X}_i)\big)^2 = D_i,$$

we obtain the second inequality. □

2.4.2 Duality and price of decentralization

In this subsection we introduce a dual problem (we work again under Assumption B) and investigate its connection with the geometric relaxed problem (PGR). This allows us to obtain a last refinement of the randomization gap. For all $i = 1, \dots, N$ and for all $\lambda \in \mathcal{E}$, we introduce

$$\Phi_i(\lambda) = \inf_{x_i \in \mathcal{X}_i} \langle \lambda, \tilde{g}_i(x_i) \rangle, \qquad \mathcal{Y}_i(\lambda) = \operatorname*{argmin}_{y_i \in \mathcal{Y}_i} \langle \lambda, y_i \rangle, \qquad \mathcal{X}_i(\lambda) = \operatorname*{argmin}_{x_i \in \mathcal{X}_i} \langle \lambda, \tilde{g}_i(x_i) \rangle.$$

We refer to the following problem as the dual problem:

$$\sup_{\lambda \in \mathcal{E}} \Big(-\tilde{f}^*(\lambda) + \frac{1}{N} \sum_{i=1}^{N} \Phi_i(\lambda) \Big). \tag{D}$$

Let $\mathcal{D}^*$ denote the value of Problem (D).

Assumption D. The function $f \colon \mathcal{E} \to \mathbb{R}$ is lower semi-continuous and convex, and the set $\operatorname{conv}(\mathcal{Y})$ is closed.

Remark 2.4.5. Assume that $\mathcal{E}$ is finite-dimensional. If the sets $\mathcal{X}_i$ are compact and the maps $\tilde{g}_i$ continuous, then the sets $\mathcal{Y}_i = \tilde{g}_i(\mathcal{X}_i)$ are also compact. It is then easy to verify with Carathéodory's theorem that $\operatorname{conv}(\mathcal{Y}_i)$ is also compact, thus closed, which finally implies Assumption D.

Lemma 2.4.6. *The problem* (PGR) *has a solution.*

Proof. This is a direct application of [BC11, Theorem 11.9]. □

The next lemma provides a duality result and a characterization of optimal solutions for problem (PGR).

Lemma 2.4.7. *Let Assumptions A, B, C, and D hold true. Then, $\mathcal{J}^* = \mathcal{D}^*$ and the dual problem* (D) *has at least one solution. Fix a solution λ to Problem* (D)*. Let $y \in \mathcal{E}$. Then, y is a solution to* (PGR) *if and only if $y \in \partial \tilde{f}^*(\lambda)$ and $y \in \frac{1}{N} \sum_{i=1}^{N} \operatorname{conv}\big(\mathcal{Y}_i(\lambda)\big)$.*

Proof. Let h denote the indicatrix function of $\operatorname{conv}(\mathcal{Y})$. By Assumption A, the domain of $\tilde{f}$ contains a neighborhood of $\operatorname{conv}(\mathcal{Y})$. By Assumption D, h is lower semi-continuous. Therefore, the Fenchel-Rockafellar theorem [Roc97] applies and yields

$$\mathcal{J}^* = \inf_{y \in \mathcal{E}} \Big(f(y) + h(y) \Big) = \sup_{\lambda \in \mathcal{E}} \Big(-\tilde{f}^*(\lambda) - h^*(-\lambda) \Big).$$

Moreover, the supremum in the right-hand side is a maximum. We have

$$-h^*(-\lambda) = \inf_{y \in \text{conv}(\mathcal{Y})} \langle \lambda, y \rangle = \inf_{y \in \mathcal{Y}} \langle \lambda, y \rangle = \frac{1}{N} \sum_{i=1}^{N} \Phi_i(\lambda).$$

As a consequence, $\mathcal{J}^* = \mathcal{D}^*$ and problem (D) has at least one solution.

Now let us fix a solution λ to the dual problem (D). Let $y \in \mathcal{E}$. Then y is a solution if and only if (i) $\tilde{f}(y) + \tilde{f}^*(\lambda) = \langle \lambda, y \rangle$ and (ii) $h(y) + h^*(-\lambda) = -\langle \lambda, y \rangle$. The condition (i) is equivalent to $y \in \partial \tilde{f}(\lambda)$. The condition (ii) is equivalent to

$$y \in \text{conv}(\mathcal{Y}) \text{ and } \langle \lambda, y \rangle = -h^*(-\lambda) = \inf_{y' \in \mathcal{Y}} \langle \lambda, y' \rangle.$$

Thus (ii) $\iff y \in Y$, where $Y = \operatorname*{argmin}_{y' \in \text{conv}(\mathcal{Y})} \langle \lambda, y' \rangle$. We further have

$$Y = \text{conv}\left(\operatorname*{argmin}_{y' \in \mathcal{Y}} \langle \lambda, y' \rangle \right) = \text{conv}\Big(\frac{1}{N} \sum_{i=1}^{N} \mathcal{Y}_i(\lambda) \Big) = \frac{1}{N} \sum_{i=1}^{N} \text{conv}\big(\mathcal{Y}_i(\lambda)\big),$$

which concludes the proof. □

Remark 2.4.8. If $\tilde{f}$ is differentiable on $\mathcal{E}$, with a Lipschitz-continous gradient, then $\tilde{f}^*$ is strongly convex (see [BC11, Theorem 18.15]), which implies that (D) has a unique solution.

Let us fix a solution λ to the dual problem until the end of the subsection. Let us consider

$$J_{\text{dec}} = \inf_{x \in \mathcal{X}} J(x), \quad \text{subject to: } x_i \in \mathcal{X}_i(\lambda), \quad \forall i = 1, \ldots, N.$$

In words, we restrict $\mathcal{X}_i$ to the best-responses corresponding to the dual variable λ. Following the terminology of [Wan17], we call price of decentralization the real number $p = J_{\text{dec}} - J^*$.

Proposition 2.4.9. *Let Assumptions A, B, C, and D hold true. It holds:*

$$p \leq J_{\text{dec}} - \mathcal{J}^* \leq \frac{1}{2N^2} \left(\max_{\substack{Q \subseteq \{1,\ldots,N\} \\ |Q| = q \wedge N}} \sum_{i \in Q} \rho\big(\mathcal{Y}_i(\lambda)\big)^2 \right).$$

Proof. The definition of J_{dec} and Lemma 2.4.7 respectively yield:

$$J_{\text{dec}} = \inf_{y \in \frac{1}{N} \sum_{i=1}^{N} \mathcal{Y}_i(\lambda)} \tilde{f}(y) \qquad \text{and} \qquad \mathcal{J}^* = \inf_{y \in \frac{1}{N} \sum_{i=1}^{N} \text{conv}(\mathcal{Y}_i(\lambda))} \tilde{f}(y).$$

The announced estimate follows then from Lemma 2.4.2 and [Cas75, Theorem 2], as in the proof of Theorem 2.4.4. □

Remark 2.4.10. The randomization gap is bounded from above by $J_{\text{dec}} - \mathcal{J}^*$. Moreover, one can show that $\rho(\mathcal{Y}_i(\lambda)) \leq \rho(\mathcal{Y}_i)$. Thus Proposition 2.4.9 provides a last refinement of the gap estimate (2.4.2).

2.5 Comments on numerical aspects and examples

2.5.1 Literature comparison

Let us compare our results and our method with the work of Wang [Wan17]. Our gap estimate, as well as our estimate of the price of decentralization, are of order $\mathcal{O}(\min(q, N)/N^2)$, while the estimates obtained by applying [Wan17, Theorem 3.5] are of order $\mathcal{O}(q^2/N^2)$. We emphasize that our first gap estimate, of order $\mathcal{O}(1/N)$, already improves [Wan17] when $q \gg \sqrt{N}$. Note that the geometric relaxation employed in Section 2.4.1 is the same as the one used in [Wan17].

Let us compare our algorithmic approaches. At a general level, one can observe that we have a primal approach, while Wang solves the dual problem to the relaxed problem. Our approach is restricted to the case where f is differentiable, while the dual approach allows to tackle the case of hard constraints (for example when f is the indicator function of some convex set). Both approaches leverage the decomposability of the problem into N problems and require that the subproblems can be easily solved. Let us emphasize however that we only need to be able to compute a single solution for those problems, while [Wan17, Algorithm 2] requires to compute the full set of ξ-optimal solutions, which may be much more difficult. Our algorithm does not require to perform Shapley-Folkman decompositions, contrary to [Wan17]. This is a major advantage when the dimension of the aggregate q is very large. Also, we do not need to evaluate f^*. As a counterpart, we are only able to find $\mathcal{O}(1/N)$-optimal solutions, while the algorithm of [Wan17] can find $\mathcal{O}(q^2/N^2)$-optimal solutions. The design of a method for the computation of $\mathcal{O}(q \wedge N/N^2)$-solutions will be the topic of future research.

2.5.2 Social welfare example

A particularly interesting instance of (P) is the social welfare optimization problem investigated in a closely related paper by Mengdi Wang [Wan17]. The cost function is the following:

$$\inf_{x_i \in \mathcal{X}_i} f_0 \left(\frac{1}{N} \sum_{i=1}^{N} h_i(x_i) \right) + \frac{1}{N} \sum_{i=1}^{N} l_i(x_i). \tag{2.5.1}$$

Following her terminology, the function h_i is the contribution of agent i to some common goods, f_0 is a social cost function of the common goods, and l_i describes the individual preference of agent i. There are various applications fitting into the framework of (2.5.1), see [Wan17]. In particular, some power system management problems can be modeled as (2.5.1). Such a problem is investigated in [SAB$^+$23]: x_i represents the production profile of the generator i, $l_i(x_i)$ is its individual production cost, f_0 denotes the demand elasticity or, equivalently, a penalty function that depends on the difference between the average production and some inflexible demand D (e.g. $f_0 := \| \cdot - D\|^2$) so as to penalize the deviation of the overall production from the inflexible demand.

Let us also mention the *resource allocation problems*, investigated in [BBG$^+$20], for example. These problems are of the form (2.5.1), where f_0 is the indicator function (as defined in [BC11, Example 1.25]) of a given point $y \in \mathcal{E}$, modelling the resource to be allocated over the agents. These problems find applications in energy management (see for example [GPA21] and [JBGO18]).

They do not fit to the current framework, since the indicator function is not differentiable, but can be reasonably well approximated, replacing the indicator by a penalty function.

2.5.3 Discussion on the case of finite feasible sets

The stochastic Frank-Wolfe algorithm investigated in the previous sections was motivated by the difficulty of manipulating probability measures, from a numerical point of view. However, when the sets $\mathcal{X}_i$ are finite, with relatively low cardinality, it is possible to store probability measures with possibly full support and some other numerical methods can be used to solve the randomized problem. Let us assume (in this subsection only) that the sets $\mathcal{X}_i$ are of cardinality $n_i \in \mathbb{N}$ and that $\mathcal{X}_i = \{\mathbf{x}_i^1, \dots, \mathbf{x}_i^{n_i}\}$. Then the randomized problem reads:

$$\min_{\nu=(\nu_1,\dots,\nu_N)} \quad f\Big(\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{n_i} \nu_i^\ell g_i(\mathbf{x}_i^\ell)\Big), \quad \text{subject to: } \nu_i \in \Delta(n_i), \tag{2.5.2}$$

where $\Delta(n_i)$ denotes the (n_i-1)-simplex, i.e.

$$\Delta(n_i) = \Big\{\nu \in \mathbb{R}^{n_i} \,\Big|\, \sum_{\ell=1}^{n_i} \nu^\ell = 1 \text{ and } \nu^\ell \geq 0, \, \forall \ell = 1,\dots,n_i\Big\}.$$

The problem is a convex program on a Cartesian product of N simplices. Let us first note that in this framework, Assumption C is trivially verified, since problem (2.3.1) is just a minimization problem over $\mathcal{X}_i$ which can be solved by enumeration. Moreover any variant of the Frank-Wolfe algorithm can be implemented, in order to solve the randomized problem in a faster way. We refer the reader to [Jag13, LJJ15]. Some other methods could also be implemented. The problem could be solved with the projected gradient descent algorithm, but the projection on the simplices is expensive (see [Con16]). Instead, the problem can be naturally addressed with the mirror descent algorithm [BT03] (see in particular the entropic descent algorithm in Section 5), and with accelerated versions of the entropic descent algorithm [KBB15].

Let us observe that if we require ν to have integer entries in the problem (2.5.2), then we are back to the original problem. Indeed, the elements of the simplex with integer entries are its vertices, that is, the vectors of the form $(0,\dots,0,1,0,\dots,0)$. Therefore the original problem can be viewed as a mixed-integer convex program (MICP) and can be addressed numerically with combinatorial techniques, see [BKL12, CLV20] and the references therein.

2.5.4 Aggregative optimal control

We describe here a large-scale optimal control problem of the form of problem (P), with an infinite-dimensional aggregate space. We verify Assumptions A, B, and C and we discuss the applicability of the Stochastic Frank-Wolfe algorithm.

Let us first fix the data of the problem. For any $i = 1,\dots,N$, we consider: an initial condition $z_i^0 \in \mathbb{R}^n$, a control set $U_i \subseteq \mathbb{R}^m$, a dynamics $F_i \colon (z_i,u_i) \in \mathbb{R}^n \times U_i \mapsto F_i(z_i,u_i) \in \mathbb{R}^n$, and a contribution function $\phi_i \colon \mathbb{R}^n \times U_i \to \mathbb{R}^k$. We also consider a social cost $\ell \colon \mathbb{R}^k \to \mathbb{R}$. We make the following assumptions:

1. *Regularity and boundedness.* For any $i = 1, \dots, N$,

   - U_i is non-empty and compact
   - F_i is continuous, Lipschitz continuous with respect to z_i, uniformly with respect to u_i; moreover, there exists a constant K_i such that $\|F_i(z_i, u_i)\| \leq K_i(1+\|z_i\|)$, for any $(z_i, u_i) \in \mathbb{R}^n \times U_i$
   - ϕ_i is continuous; moreover, there exists a function $R_i \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\|\phi_i(z_i, u_i)\| \leq R_i\big(\|z_i\| + \|u_i\|\big)$, for any $(z_i, u_i) \in \mathbb{R}^n \times U_i$.

2. *Regularity of the social cost.* The function ℓ is continuously differentiable, moreover, ℓ and $\nabla\ell$ are Lipschitz continuous with moduli L_ℓ and $L_{\nabla\ell}$, respectively.

3. *Convexity assumption.* For any $i = 1, \dots, N$, for any $y \in \mathbb{R}^k$, for any $z_i \in \mathbb{R}^n$, we define $\mathcal{Z}_i(y, z_i)$ the set of all elements $(\bar{z}_1, \bar{z}_2)$ in $\mathbb{R}^{n+1}$, where there exists $u_i \in U_i$, such that $\bar{z}_1 = F_i(z_i, u_i)$ and $\bar{z}_2 \geq \langle \nabla\ell(y), \phi_i(z_i, u_i)\rangle$. The set $\mathcal{Z}_i(y, z_i)$ is convex.

Let us mention a particular case in which the above convexity assumption is true: for any $i = 1, \dots, N$, for any $y \in \mathbb{R}^k$, for any $z_i \in \mathbb{R}^n$,

- For any z_i, the map $u_i \mapsto F_i(z_i, u_i)$ is affine.
- The set U_i is convex and the function $u_i \in U_i \mapsto \langle \nabla\ell(y), \phi_i(z_i, u_i)\rangle$ is convex.

For any $i = 1, \dots, N$, consider the set $\mathcal{X}_i$ of pairs $(z_i, u_i) \in W^{1,\infty}(0, T; \mathbb{R}^n) \times L^\infty(0, T; \mathbb{R}^m)$ satisfying

$$\dot{z}_i(t) = F_i(z_i(t), u_i(t)), \quad z_i(0) = z_i^0, \quad u_i(t) \in U_i, \quad \text{for a.e. } t \in (0, T).$$

A direct application of Gronwall's lemma shows that for any $(z_i, u_i) \in \mathcal{X}_i$, we have $\|z_i\|_{L^\infty(0,T;\mathbb{R}^n)} \leq \tilde{K}_i$, where $\tilde{K}_i = (1 + \|y_0^i\|)\exp(K_i T) - 1$.

The aggregative optimal control problem of interest is defined as follows:

$$\inf_{(z_i,u_i)_{i=1}^N \in \prod_{i=1}^N \mathcal{X}_i} \int_0^T \ell\Big(\frac{1}{N}\sum_{i=1}^N \phi_i\big(z_i(t), u_i(t)\big)\Big)\,\mathrm{d}t. \tag{2.5.3}$$

It is a special case of problem (P) with $m = 1$, $\mathcal{E}_1 = \mathcal{E} = L^2(0, T; \mathbb{R}^k)$, and

$$\begin{aligned} g_i \colon \quad (z_i, u_i) \in \mathcal{X}_i \quad &\mapsto \quad \big(t \in (0, T) \mapsto \phi_i(z_i(t), u_i(t))\big) \in L^2(0, T; \mathbb{R}^k) \\ f \colon y \in L^2(0, T; \mathbb{R}^k) \quad &\mapsto \quad \textstyle\int_0^T \ell(y(t))\,\mathrm{d}t. \end{aligned}$$

Problem (2.5.3) can be seen as a nonconvex optimal control problem with state variable $(z_i)_{i=1}^N$. It finds application in energy management, in the situations mentioned in the introduction and in particular those involving storage devices, for which the dynamics of the state-of-charge must be taken into account. Once again we refer the reader to [SAB+23], which considers a convex stochastic aggregative optimal control problem. In general, only dynamic-programming-based methods can provide global solutions to nonlinear optimal control problems. They are not applicable here because of the high dimension of the state variable, equal to Nn.

It is easy to verify that ∇f is continuously differentiable and that f and ∇f are Lipschitz-continuous with moduli $\sqrt{T} L_\ell$ and $L_{\nabla \ell}$, respectively. Let $\hat{K}_i$ be an upper bound of $\sup_{u_i \in U_i} \|u_i\|$, for all $i \in 1, \ldots, N$. Then $g_i(\mathcal{X}_i)$ is bounded in $L^2(0,T;\mathbb{R}^k)$, with diameter bounded by $2\sqrt{T} R_i(\tilde{K}_i + \hat{K}_i)$. Therefore, Assumption A is satisfied. If ℓ is convex, then f is also convex and then Assumption B holds true. Let us verify Assumption C. Given $y \in G(\mathcal{X})$, the problem (2.3.1) to be solved at each iteration of the SFW algorithm reads

$$\inf_{(z_i,u_i)\in\mathcal{X}_i} \int_0^T \big\langle \nabla \ell(y(t)), \phi_i(z_i(t), u_i(t)) \big\rangle \, \mathrm{d}t. \tag{2.5.4}$$

This is an optimal control with state variable z_i, which falls into the class of problems introduced in [FR75, Chapter III, Theorem 4.1] and therefore possesses a solution. It the dimension of the state variable, n, is small, then it can be solved by dynamic programming. We refer the reader to [FF13].

2.5.5 Supervised learning problems

We describe and discuss here two applications of problem (P) in the context of supervised learning.

Neural networks with one hidden layer We refer the reader to [CB18, MMN18, MMM19]. Consider a neural network of the form $\frac{1}{N}\sum_{i=1}^N \sigma_*(\mathbf{a}, x_i)$, where $\mathbf{a} \in \mathbb{R}^d$ is the feature vector, $x = (x_i)_{i=1}^N \in (\mathbb{R}^D)^N$ are the network parameters (to be optimized), and $\sigma_* \colon \mathbb{R}^d \times \mathbb{R}^D \to \mathbb{R}$ an activation function. We consider a loss function $\varphi \colon \mathbb{R} \to \mathbb{R}_+$. Given a data set $(\mathbf{a}_j, b_j)_{j=1}^M \in (\mathbb{R}^d \times \mathbb{R})^M$, the learning problem of interest writes

$$\inf_{(x_i)_{i=1}^N \in (\mathbb{R}^D)^N} \frac{1}{M} \sum_{j=1}^M \varphi\Big(b_j - \frac{1}{N} \sum_{i=1}^N \sigma_*(\mathbf{a}_j, x_i) \Big). \tag{2.5.5}$$

It is of the form (P), with $\mathcal{E} = \mathbb{R}^M$, $\mathcal{E}_j = \mathbb{R}$, $f_j(y_j) = \varphi(b_j - y_j)/M$, $g_{ij}(x_i) = \sigma_*(\mathbf{a}_j, x_i)$. Assume that the set $\{\sigma_*(\mathbf{a}_j, x) \,|\, x \in \mathbb{R}^D,\, j \in \{1, \ldots, M\}\}$ has a bounded diameter $\bar{d}$. Assume moreover that φ is continuously differentiable and that $\nabla\varphi$ is $L_{\nabla\varphi}$-Lipschitz continuous. Then Assumption A is satisfied and we have $D_i = L_{\nabla\varphi}\bar{d}^2$, for the coefficients D_i introduced in (2.4.1). Therefore, by Theorem 2.4.4, the optimality gap is bounded by

$$\frac{(M \wedge N) L_{\nabla\varphi} \bar{d}^2}{2N^2}.$$

Moreover, if φ is convex, then Assumption B holds true. The resolution of the subproblems (2.3.1) is not easy in general, we refer the reader to [dP20] where the linearized problems are shown to be solvable by second-order cone programming in the case of ReLu activation functions.

Note that we are here in the symmetric case, as defined at the end of Section 2.2.2. The mean-field relaxation proposed in Lemma 2.2.8 was also utilized in [MMM19, MMN18] for learning problems of the form (2.5.5). A gap estimate of order $\mathcal{O}(1/N)$ is demonstrated, in the case of a quadratic loss function φ, see [MMN18, Prop. 1]. Our gap estimate is more general since $\nabla\varphi$ is only supposed to be Lipschitz; moreover, it is more precise in the case of an overparametrized network (i.e. when $M < N$), since then it is of order $\mathcal{O}(M/N^2)$.

Sparse reconstruction Another important learning example is the sparse reconstruction with the ℓ_0-penalty, see [Mal09, MBP14]. Let D be a M by N dictionary matrix. The objective is to approximate the observed vector $x \in \mathbb{R}^M$ by a sparse linear combination of the columns of D. Following [MBP14, Eq. 5.6], we are interested in the following least square problem with the ℓ_0-penalty:

$$\inf_{\alpha \in \mathbb{R}^N} \frac{1}{2}\left\|x - D\alpha\right\|^2 + \beta\|\alpha\|_{\ell_0} = \inf_{\alpha \in \mathbb{R}^N} \frac{1}{2}\sum_{j=1}^{M}\Big(x_j - \sum_{i=1}^{N} D_{ji}\alpha_i\Big)^2 + \beta\sum_{i=1}^{N}\mathbf{1}_{\mathbb{R}\setminus\{0\}}(\alpha_i),$$

where β is a constant and $\|\alpha\|_{\ell_0}$ counts the number of non-zero entries in a vector α. Adding constraints of the form $\alpha_i \in [u_i, v_i]$ to the problem, it is easy to see that Assumptions A and B are satisfied. The subproblems (2.3.1) are here of the form

$$\inf_{\alpha_i \in [u_i, v_i]} z\alpha_i + \mathbf{1}_{\mathbb{R}\setminus\{0\}}(\alpha_i)$$

for some real number z. One can show that there is a solution that necessarily lies in $\{u_i, v_i, 0\}$, thus it is easy to compute.

Finally, let us mention other applications of the problem (P) in a convex framework, for instance, the "sharing problem" in [BPC11], Lasso regression in [FR16] and the dual problem of a linear support vector machine (SVM) in [SST09, FR16].

2.6 Numerical test for MIQP

In this section we provide numerical results for a mixed-integer linear quadratic problem of the form (P). Let A be a real $M \times N$ matrix and let $\bar{y} \in \mathbb{R}^M$. Consider the following problem:

$$\min_{x \in \{0,1\}^N} J(x) := \frac{1}{N^2}\|Ax - \bar{y}\|^2_{\mathbb{R}^M} = \sum_{j=1}^{M}\left(\frac{1}{N}\sum_{i=1}^{N} A_{ji}x_i - \frac{\bar{y}_j}{N}\right)^2. \tag{MIQP}$$

Problem (MIQP) has the form (P), with $f_j(y_j) = \big(y_j - \frac{\bar{y}_j}{N}\big)^2$ for $1 \le j \le M$, and $g_{ij}(x_i) = A_{ji}x_i$ for $1 \le i \le N$, $1 \le j \le M$. Moreover, Assumption A is satisfied with $\tilde{L}_j = 2$ and $d_{ij} = |A_{ji}|$. Thus $C_1 = \frac{2}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}|A_{ji}|$. Due to the linearity of g_{ij}, the randomized problem coincides with the minimization problem of J on $[0,1]^N$, which is a convex linear-quadratic program that can be solved with independent methods; thus it is easy here to obtain a precise estimate of $\mathcal{J}^*$.

In the numerical simulation, we draw the parameters A_{ji} according to the uniform distribution on the interval $[0,1]$ while y_j is drawn according to the uniform distribution on $[0, N/2]$. Thus, $C_1 \approx M$ and the gap estimate is given by $\frac{C_1}{2N} \approx 0.5$. We perform our numerical experiments on a laptop with one Intel Core i5-8250U processor (4 cores) at 1.60 GHz and 8 GB RAM.

The first experiment is a comparison of Algorithm 2.2 with an open source solver, SCIP, [BBC$^+$21] and a commercial solver, GUROBI, [GO18]. As mentioned before, the dual (randomized) problem is a convex linear-quadratic program. We can compute $\mathcal{J}^*$ easily by solver GUROBI. Table 2.1 shows the value $\mathcal{J}^*$ and results of (MIQP) obtained from SCIP, GUROBI and Algorithm 2.2, for different values of M, N ranging from 100 to 3200. In Table 2.1, "Nan" indicates that the solver has failed to return a result or that computation time has exceeded one hour. Denote by v_s the result of Algorithm 2.2. The indicated gap is a relative gap, in percent, defined by $(v_s - \mathcal{J}^*)/\mathcal{J}^*$. We can observe that the relative gap decreases as N increases, which is consistent with the randomized gap (2.2.2). The last three columns of Table 2.1 show that Algorithm 2.2 is competitive in terms of execution time, in comparison with SCIP and GUROBI. Finally, observe that for $N = M = 3200$, none of the two solvers could solve the problems while Algorithm 2.2 has provided a solutions in approximately 6 minutes.

$N = M$	$\mathcal{J}^*$	SCIP	GUR.	SFW		SCIP	GUR.	SFW
		value	value	value	gap in %	time in seconds		
100	2.077	2.077	2.077	2.136	2.870	0.88	0.20	0.03
200	4.120	4.120	4.120	4.159	0.956	5.99	0.69	0.09
400	7.871	7.871	7.871	7.904	0.430	87.78	7.90	0.91
800	15.953	Nan	15.954	15.966	0.079	Nan	10.63	6.18
1600	32.045	Nan	32.048	32.0585	0.042	Nan	81.41	42.51
3200	64.717	Nan	Nan	64.724	0.012	Nan	Nan	330.95

Table 2.1: Comparison of the approximate values and execution times obtained with SCIP, GUROBI and Algorithm 2.2 for problem (MIQP) with $M = N = 100$, 200, 400, 800, 1600 and 3200. In Algorithm 2.2, we take $n_k = 1$ and $K = 2N$ iterations.

The second experiment is on the basic Frank-Wolfe algorithm 2.1 and its stochastic version 2.2. In this experiment, we fix $M = N = 1000$. Figure 2.1 shows the outcome of the basic Frank-Wolfe algorithm 2.1 with 200 iterations. The left sub-figure shows the evolution of γ_k for $\omega_k = 2/(k+2)$ (green curve) and for ω_k determined by line search (2.3.6) (red curve). A sub-linear rate of convergence is observed (note that logarithmic scales are employed for both axes). The right sub-figure represents the evolution of $J(X^k) - \mathcal{J}^*$, where X^k is a random variable with distribution μ^k. For both choices of ω_k, approximate solutions to the problems are simulated, with a gap smaller than 10^{-3}, significantly smaller than the gap estimate $\frac{C_1}{2N}$. The line search approach is quicker than the approach with $\omega_k = \frac{2}{k+2}$.

Figure 2.2 shows the outcome of Algorithm 2.2 (with the modification suggested in Remark 2.3.9), for different (constant) choices of n_k with 200 iterations, for two different stepsize rules ($\omega_k = 2/(k+2)$ on the left, line search on the right). Since the algorithm is stochastic, we have tested it 50 times to evaluate its efficiency; the curves represent the average value of γ_k. The standard deviation (for these 8 instances of the SFW method) is displayed on Figure 2.3. In all cases, an average value of the gap significantly smaller than $\frac{C_1}{2N}$ can be reached; the standard deviation is also significantly smaller than $\frac{C_1}{2N}$ at the last iterations. There is a benefit (both in expectation and standard deviation) in increasing the number of simulations n_k (note that the choice $n_k = 1000$ is much smaller the rule suggested by Corollary 2.3.8). Yet the convergence is slower in comparison with the basic Franck-Wolfe algorithm, which can be explained by the use of the selection method at each iteration.

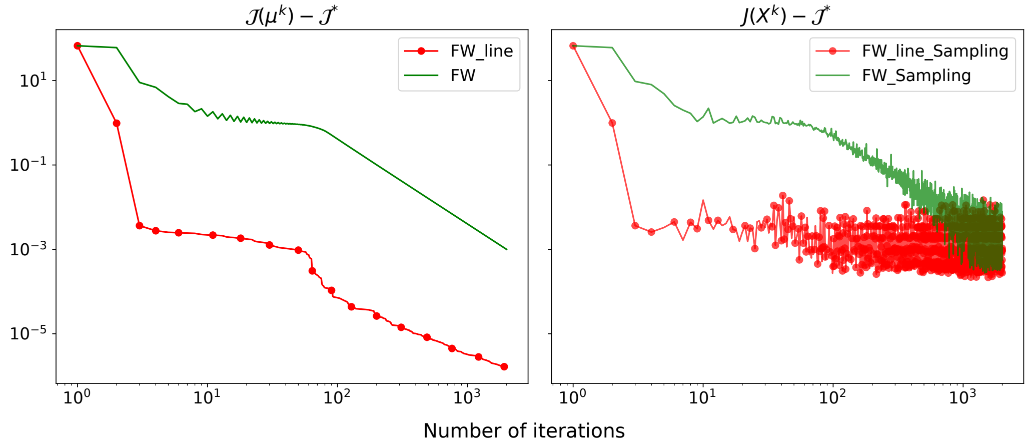


Figure 2.1: MIQP by Algorithm 2.1, 2000 iterations, with $\omega_k = 2/(k+2)$ and line search (2.3.6).

2.7 Numerical test for discrete aggregative optimal control problems

This section is dedicated to a class of aggregative optimal control problems in discrete time and discrete state space (the continuous version is presented in section 2.5.4). These problems involve a large number N of agents, indexed by $i = 1, \ldots, N$, and time steps ranging over $t = 0, 1, \ldots, T$. For any agent i, we fix a finite *state set* S_i and a finite *control set* U_i. The evolution of agent i is described by *transition functions* $\pi_i^t \colon S_i \times U_i \to S_i$, where $t = 0, \ldots, T$. We also fix mappings $U_i^t \colon S_i \to 2^{U_i}$ describing the feasible controls of the agents: at time t, if the agent i is in state s_i^t, he can make use of all controls in $U_i^t(s_i^t)$. The initial state of each agent i is constrained to be in S_i^0, a subset of S_i. The problem also involves some functions $f_t \colon \mathbb{R} \to \mathbb{R}$ which we call *social* cost (at time t) and some functions $h_i^t \colon S_i^t \times U_i \to \mathbb{R}$, which we call *contribution* functions. We also make use of functions $\ell_i^t \colon S_i^t \times U_i \to \mathbb{R}$, which we call *individual* costs. The optimal control problem of interest reads:

$$\left\{\begin{array}{ll} \inf\limits_{(s,u)} & J(s,u) := \sum_{t=0}^{T} f_t\Big(\frac{1}{N}\sum_{i=1}^{N} h_i^t(s_i^t, u_i^t)\Big) \\ & \qquad\qquad + \frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T} \ell_i^t(s_i^t, u_i^t), \\ \\ \text{s.t.} & s_i^{t+1} = \pi_i^t(s_i^t, u_i^t),\ u_i^t \in U_i^t(s_i^t),\ s_i^0 \in S_i^0, \\ & \forall t = 0, 1, \ldots, T-1,\ i = 1, 2, \ldots, N, \end{array}\right. \tag{2.7.1}$$

where $(s, u) = (s_i^t, u_i^t)_{i=1,\ldots,N}^{t=0,\ldots,T}$.

A standard approach to deal with problem (2.7.1) relies on the dynamic programming principle (see [Ber12]), in which a key step is to compute the value function $V \colon \{0, 1, \ldots, T\} \times S \to \mathbb{R}$, where the state space S is defined by $\prod_{i=1}^{N} S_i$. For our problem, Bellman's equation reads as follows: for

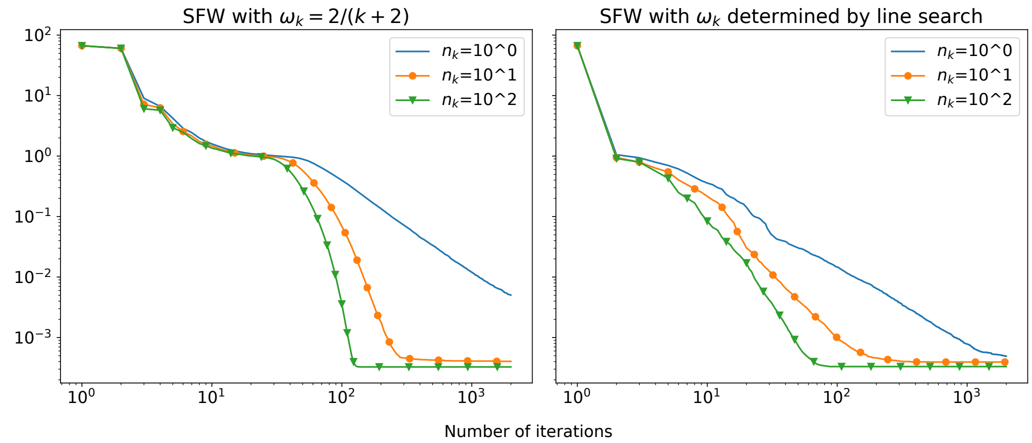


Figure 2.2: MIQP by Algorithm 2.2 with 2000 iterations, expectation of the gap.

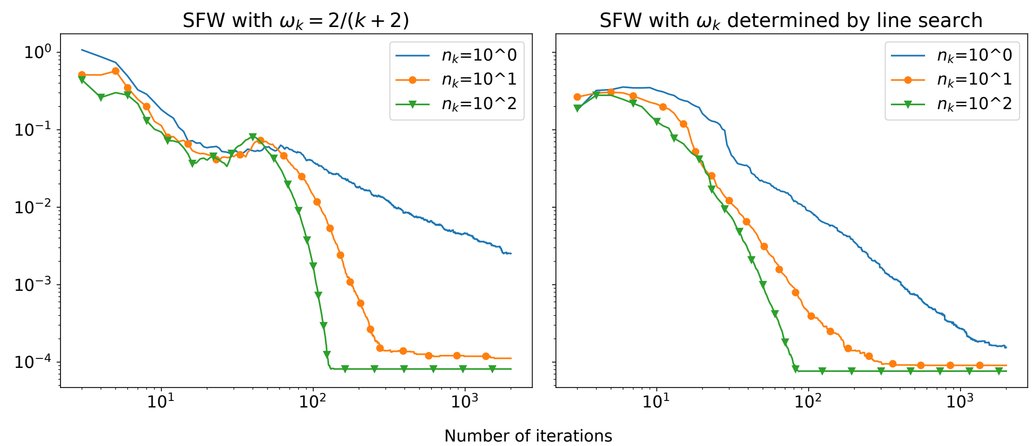


Figure 2.3: MIQP by Algorithm 2.2 with 2000 iterations, standard deviation of the gap.

any $t \in \{0, \dots T\}$, for any $s \in S$,

$$V^t(s) = \min_{u \in U^t(s)} f_t\Big(\frac{1}{N}\sum_{i=1}^N h_i^t(s_i, u_i)\Big) + \frac{1}{N}\sum_{i=1}^N \ell^t(s_i, u_i) + V^{t+1}(\pi^t(s,u)),$$

where $U^t(s) = \prod_{i=1}^N U_i^t(s_i)$ and where $\pi^t(s,u) = (\pi_i^t(s_i,u_i))_{i=1}^N$. We observe that the complexity of Bellman's equation increases exponentially with N; this phenomenon is the well-known curse of dimensionality. As a consequence, the dynamic programming approach is not tractable for problem (2.7.1) when the number of agents N is large.

Let us reformulate optimal control problem (2.7.1) as a problem of the form (P). Next, we address its resolution with the SFW algorithm.

2.7.1 Reformulation

Let us consider an agent i and let us describe its state-control feasible set $\mathcal{X}_i$. Recall that a state S_i, a control set U_i, mappings $U_i^t\colon S_i \to 2^{U_i}$, and transition mapping $\pi_i^t\colon S_i \times U_i \to S_i$ are given. We call feasible *state-control trajectory* an element $x_i = (s_i, u_i)$, where $s_i = (s_i^0, \dots, s_i^T) \in (S_i)^{T+1}$ and $u_i = (u_i^0, \dots, u_i^T) \in (U_i)^{T+1}$, such that

$$s_i^0 \in S_i^0, \quad u_i^t \in U_i^t(s_i^t), \quad s_i^{\theta+1} = \pi_i^\theta(s_i^\theta, u_i^\theta),$$

for any $t = 0, \dots, T$ and any $\theta = 0, \dots, T-1$. We denote by $\mathcal{X}_i$ the set of feasible state-control trajectories. We set $\mathcal{X} = \prod_{i=1}^N \mathcal{X}_i$. The non-emptyness of $\mathcal{X}_i$ is a straightforward consequence of the following assumption.

Assumption 2.1. The set S_i^0 is non-empty. For all $t = 0, \dots, T$, for all $s_i^t \in S_i$, the set $U_i^t(s_i^t)$ is non-empty.

With each agent i are associated $(T+1)$ contribution functions h_i^t, $t = 0, \dots, T$ and $(T+1)$ individual costs ℓ_i^t, $t = 0, \dots, T$. We set $\mathcal{E}_0 = \dots \mathcal{E}_{T+1} = \mathbb{R}$ and we define $T+2$ functions $g_{it}\colon \mathcal{X}_i \to \mathcal{E}_t$ by

$$g_{it}(x_i) = \begin{cases} h_i^t(s_i^t, u_i^t) & \text{if } t \leq T \\ \sum_{t'=0}^T \ell_i^{t'}(s_i^{t'}, u_i^{t'}) & \text{if } t = T+1. \end{cases}$$

The social costs $f_0, \dots, f_T$ are the same as in the original problem (2.7.1). The social cost $f_{T+1}\colon \mathcal{E}_{T+1} \to \mathbb{R}$ is the identity function. With these definitions, problem (2.7.1) is equivalent to

$$\inf_{(x_i)_{i=1}^N \in \prod_{i=1}^N \mathcal{X}_i} \quad \sum_{t=0}^{T+1} f_t\Big(\frac{1}{N}\sum_{i=1}^N g_{it}(x_i)\Big). \tag{2.7.2}$$

2.7.2 Assumptions

As before, we denote $g_i(x_i) = (g_{it}(x_i))_{t=0,\dots,T+1}$, $\mathcal{E} = \prod_{t=0}^{T+1} \mathcal{E}_t = \mathbb{R}^{T+2}$ and for $y \in \mathcal{E}$, $f(y) = \sum_{t=0}^{T+1} f_t(y_t)$. For any $i = 1, \dots, N$ and for any $t = 0, \dots, T+2$, we denote

$$Y_{it} = \big\{g_{it}(x_i) \mid x_i \in \mathcal{X}_i\big\} \quad \text{and} \quad Y_t = \frac{1}{N}\sum_{i=1}^N Y_{it}.$$

Assumption 2.2. For $i = 1, 2, \dots, N$ and for $t = 0, 1, \dots, T$,

- f_t is L_t-Lipschitz on $\operatorname{conv}(Y_t)$,
- f_t is continuously differentiable on a neighborhood of $\operatorname{conv}(Y_t)$, ∇f_t is $\tilde{L}_t$-Lipschitz on $\operatorname{conv}(Y_t)$
- f_t is convex on $\operatorname{conv}(Y_t)$.

Assumptions 2.1 and 2.2 imply Assumptions A and B for problem (2.7.2). Assumption C is trivially satisfied since $\mathcal{X}_i$ is a finite set.

2.7.3 Resolution of the sub-problems

We explain now how to solve the sub-problems (2.3.1) associated with the aggregative optimal control problem (2.7.2). Let $y \in \mathcal{E}$. Let $\mu \in \mathcal{E}$ be defined by $\mu^t = \nabla f_t(y^t)$. By definition of f_{T+1}, $\mu^{T+1} = 1$. The sub-problem (2.3.1) reads:

$$\inf_{x_i \in \mathcal{X}_i} \sum_{t=0}^{T} \Big(\ell_i^t(s_i^t, u_i^t) + \langle \mu^t, h_i^t(s_i^t, u_i^t) \rangle \Big). \tag{2.7.3}$$

The sub-problem (2.7.3) can be solved by dynamic programming. Algorithm 2.4 yields a solution to (2.7.3). For convenience, we denote

$$\ell_i^t[\mu^t](s_i^t, u_i^t) = \ell_i^t(s_i^t, u_i^t) + \langle \mu^t, h_i^t(s_i^t, u_i^t) \rangle$$

in the algorithm. The algorithm consists of two steps: first in a backward pass, a sequence of value functions $(V_i^t)_{t=0,\dots,T+1}$ is computed, where $V_i^t \colon S_i \to \mathbb{R}$. A globally optimal solution is obtained in a forward pass. Note that the value of the optimization problem of Step 1 is finite as a consequence of Assumption 2.1.

Algorithm 2.4: Dynamic programming algorithm

Step 1: Backward pass.
Set $V_i^{T+1}(s_i^{T+1}) = 0$, for any $s_i^{T+1} \in S_i$.
for $t = T, T-1, \dots, 0$ **do**
  for $s_i^t \in S_i$ **do**
    Define $V_i^t(s_i^t)$ as

$$\min_{u_i^t \in U_i^t(s_i^t)} \ \ell_i^t[\mu^t](s_i^t, \cdot) + V_i^{t+1}\big(\pi_i^t(s_i^t, \cdot)\big).$$

  end
end

Step 2: Forward pass.
Find $\bar{s}_i^0 \in \operatorname{argmin}_{s_i^0 \in S_i^0} V_i^0(s_i^0)$;
for $t = 0, \dots, T$ **do**
  Find a solution $\bar{u}_i^t$ to the problem

$$\min_{U_i^t(\bar{s}_i^t, u_i^t)} \ \ell_i^t[\mu^t](\bar{s}_i^t, \cdot) + V_i^{t+1}\big(\pi_i^t(\bar{s}_i^t, \cdot)\big).$$

  If $t < T$, set $\bar{s}_i^{t+1} = \pi_i^t(\bar{s}_i^t, \bar{u}_i^t)$.
end

Remark 2.7.1. As mentioned in the introduction, problem (2.7.2) could be addressed by dynamic programming. This would allow the computation of an exact solution. However, this would require to compute a value function of the form $V^t(s^t)$, where $s^t = (s_1^t, \dots, s_N^t) \in \prod_{i=1}^N S_i$. The resulting complexity, of order $T \prod_{i=1}^N |S_i|$, is prohibitive even for moderate values of N. In contrast, the complexity of each iteration of the SFW algorithm is linear with respect to N, while the accuracy of the algorithm improves as N increases.

2.7.4 Numerical simulations on a battery charging problem

Let us now turn to the problem of the charging of a fleet of batteries. We propose a very simple model which is essentially illustrative, rather than realistic. However, it is emphasised that the proposed approach can easily incorporate more realistic constraints on battery operation (e.g. taking into account limits on cycles numbers). Indeed, these refinements remain localized at the sub-problem level (impacting only the dynamic programming Algorithm 2.4). They consist either in adding a state variable or in modifying the local costs in order to penalise undesired behaviour. Suppose that there are N batteries to be charged. Let s_i^t be the state of charge (SoC) for the battery i at the time t.

Dynamics. The dynamics of each battery is characterized by three parameters: an initial state of charge $s_i^{\text{in}} \in \mathbb{N}$, a maximal state of charge $s_i^{\max} \in \mathbb{N}$, a maximal load speed $u_i^{\max} \in \mathbb{N}$. We define:

$$\begin{aligned} &S_i = \{s_i^{\text{in}}, \dots, s_i^{\max}\},\ S_i^0 = \{s_i^{\text{in}}\},\ U_i = \{0, \dots, u_i^{\max}\}, \\ &U_i^t(s_i^t) = \{0, \dots, \min(u_i^{\max}, s_i^{\max} - s_i^t)\}, \\ &\pi_i^t(s_i^t, u_i^t) = s_i^t + u_i^t. \end{aligned}$$

In words: the initial condition s_i^{in} is given, the charging of the battery is additive, the charging speed is bounded by $u_i^{\max}$ and is such that s_i^t can never exceed $s_i^{\max}$.

Cost functions. Some positive coefficients $(\beta_i)_{i=1,\dots,N}$, $(\alpha_t)_{t=0,\dots,T-1}$, and $(c_t)_{t=0,\dots,T-1}$ are given. The individual costs are

$$\begin{aligned} \ell_i^t(s_i^t, u_i^t) &= 0, \quad \forall t = 0, \dots, T-1, \\ \ell_i^T(s_i^T, u_i^T) &= \beta_i (s_i^{\max} - s_i^T)^2. \end{aligned}$$

The contributions are defined by $h_i^T(s_i^T, u_i^T) = 0$ and

$$h_i^t(s_i^t, u_i^t) = u_i^t, \quad \forall t = 0, \dots, T-1.$$

The social costs f_t are defined by $f_T(y_T) = 0$ and

$$f^t(y_t) = \alpha^t (y_t - c_t)^2, \quad \forall t = 0, \dots, T-1.$$

Therefore, the cost function J reads

$$\sum_{t=0}^{T-1} \alpha^t \left(\left(\frac{1}{N} \sum_{i=1}^{N} u_i^t \right) - c^t \right)^2 + \frac{1}{N} \sum_{i=1}^{N} \beta_i \left(s_i^T - s_i^{\max} \right)^2.$$

The cost function has two contributions, one depends on the average of charging levels of all the batteries, the other one depends on the individual final SoC of each battery. To be more precise, for $t \leq T-1$, the average charging level needs to approach some target power c_t. For $t = T$, the batteries expect to approach their maximum SoCs.

Numerical simulations. The parameters are chosen as follows:

- $N = 100$, $T = 24$

- s_i^{in} (resp. $s_i^{\max}$) is chosen randomly and uniformly in $\{0, 1, \ldots, 20\}$ (resp. $\{20, 21, \ldots, 40\}$), $u_i^{\max} = 4$
- α^t is chosen randomly and uniformly in $[1, 2]$, β_i is chosen randomly and uniformly in $[0, 1]$
- $c^t = 1.5\lfloor \sin(\pi t/12) + 1 \rfloor$.

Thus, for $t = 0, 1, \ldots, 23$, the diameter of the range set Y_{it} is less than $u_i^{\max} = 4$, and the Lipschitz constant $\tilde{L}_t$ is $2\alpha^t$, which is less than 4. Then, we have the following upper bound for the relaxation gap $C_1/2N$:

$$\frac{C_1}{2N} \leq \frac{1}{200} \cdot \frac{1}{100} \cdot \sum_{t=0}^{23} \left(4 \cdot \sum_{i=1}^{100} 4^2 \right) = 7,68.$$

Fig. 2.4 shows the outcome of Algorithm 2.1 with 500 iterations to get an approximation of the minimum $\mathcal{J}^*$ of the relaxed problem. The curve represents the relaxed cost. Fig. 2.5 shows the outcome of Algorithm 2.2, for different choices of n_k with 100 iterations. Since the algorithm is stochastic, we ran it 50 times independently to evaluate its efficiency; the curves represent the average value of $\gamma_k = J(x^k) - \mathcal{J}^*$. The standard deviation is displayed in the right part of Fig. 2.6. In all cases, an average value of the gap significantly smaller than $7,68$ can be reached; the standard deviation is also significantly smaller than $7,68$ at the last iterations. We have initialized the algorithm with values of x_i^0 such that $u_i^t = 0$, for any $t = 0, \ldots, T-1$.

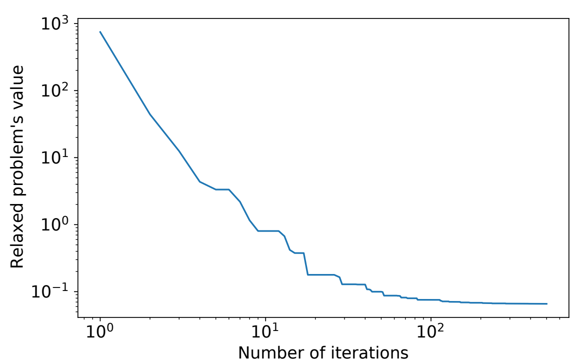


Figure 2.4: Frank-Wolfe Algorithm with 500 iterations for the relaxed problem.

2.8 Conclusion

We have investigated a large-scale and aggregative optimization problem and its relaxation. New error bounds for the relaxation gap have been obtained. We have proposed a tractable algorithm for its resolution with a detailed convergence analysis relying on concentration inequalities. Assuming that an efficient method for the resolution of the subproblems is available, the implementation of our stochastic Frank-Wolfe method is easy.

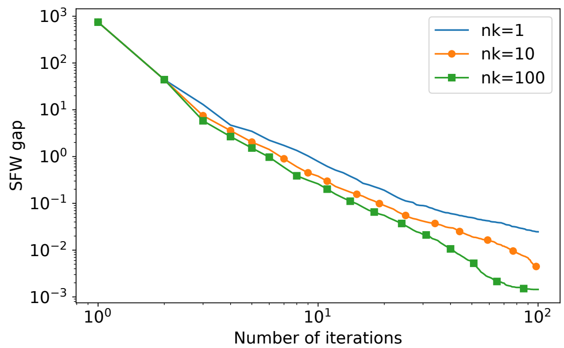


Figure 2.5: Algorithm 2.2 with 100 iterations, expectation of the gap.

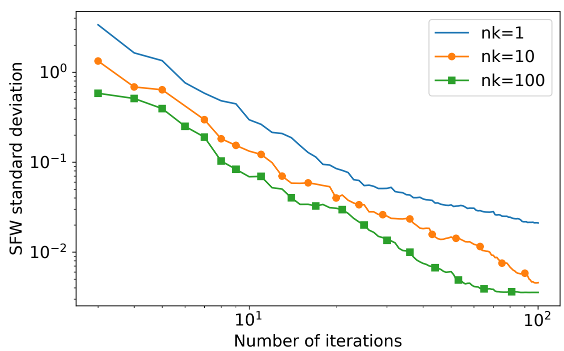


Figure 2.6: Algorithm 2.2 with 100 iterations, standard deviation of the gap.

Future research will focus on refinements of the selection method, allowing the computation of $\mathcal{O}(q \wedge N/N^2)$-solutions. We also aim at working on more complex problems, involving for example convex constraints on the aggregate, as for example the resource allocation problems mentioned in section 2.5.2. Such constraints could be handled with extensions of the Frank-Wolfe algorithm for non-smooth costs as those proposed in [SFMF20, YFC19]. Finally, we intend to apply our method to large-scale optimal control problems, such as nonconvex variants of the problem investigated in [SAB$^+$23].

2.9 Appendix

2.9.1 Concentration inequalities and other technical lemmas

Proposition 2.9.1. *Consider T real-valued random variables $(Y_t)_{t=1,\dots,T}$. Let $(\mathcal{F}_t)_{t=1,\dots,T}$ denote the associated filtration ($\mathcal{F}_0$ is the trivial σ-algebra). Let $Z_t = \mathbb{E}[Y_t^2|\mathcal{F}_{t-1}]$ and let $S_T = \sum_{t=1}^T Y_t$. Assume the following:*

$$(i) \quad \mathbb{E}[Y_t \mid \mathcal{F}_{t-1}] = 0, \qquad (ii) \quad Y_t \leq m, \qquad (iii) \quad \sum_{t'=1}^{T} Z_{t'} \leq v, \quad a.s. \tag{2.9.1}$$

for all $t = 1, ..., T$ and for some constants m and v. Then, $\mathbb{E}[S_T^2] \leq v$. Moreover, for any $\epsilon > 0$,

$$\mathbb{P}\big[S_T \geq \epsilon\big] \leq \exp\Big(- \frac{\epsilon^2}{2\,(v + \epsilon m/3)}\Big). \tag{2.9.2}$$

Proof. The estimate of $\mathbb{E}[S_T^2]$ can be easily obtained by induction. For the estimate of $\mathbb{P}\big[S_T \geq \epsilon\big]$, see [Del15, Theorem 7]. □

As a corollary, we obtain the following McDiarmid's inequality of "variance type".

Corollary 2.9.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\Omega_i)_{i=1,...,n}$ be n measurable subsets of Ω. Let $X = (X_i)_{i=1,...,n}$ be n independent random variables valued respectively in $(\Omega_i)_{i=1,...,n}$. Consider a measurable function $f\colon \prod_{i=1}^n \Omega_i \to \mathbb{R}$ and real constants $m \in \mathbb{R}$ and $(v_i)_{i=1,...,n}$ such that*

$$\mathrm{Var}\big[f(X_i, x_{-i})\big] \leq v_i^2, \quad a.s., \qquad \big|f(X_i, x_{-i}) - \mathbb{E}\big[f(X_i, x_{-i})\big]\big| \leq m, \qquad a.s.,$$

for all $i = 1, ..., n$ and for all $x_{-i} \in \Big(\prod_{j=1}^{i-1} \Omega_i\Big) \times \Big(\prod_{j=i+1}^{n} \Omega_j\Big)$. Then, for any $\epsilon > 0$,

$$\mathbb{P}\Big[f(\mathbf{x}) - \mathbb{E}\big[f(\mathbf{x})\big] \geq \epsilon\Big] \leq \exp\Big(- \frac{\epsilon^2}{2\,\big(\sum_{i=1}^n v_i^2 + \frac{m\epsilon}{3}\big)}\Big). \tag{2.9.3}$$

Proof. Define $Y_t = \mathbb{E}\left[f(X) \mid X_1, \ldots, X_t\right] - \mathbb{E}\left[f(X) \mid X_1, \ldots, X_{t-1}\right]$ and apply Proposition 2.9.1. □

Lemma 2.9.3. *For all $k \in \mathbb{N}$, denote $\omega_k = \frac{2}{k+2}$. Let $(u_k)_{k\in\mathbb{N}}$ and $(\gamma_k)_{k\in\mathbb{N}}$ be two sequences of real numbers. Assume that there exists a positive number C such that*

$$\gamma_{k+1} \leq (1 - \omega_k)\gamma_k + C\omega_k^2 + u_k, \tag{2.9.4}$$

for all $k \in \mathbb{N}$. Then, for all $K \in \mathbb{N}^*$,

$$\gamma_K \leq \frac{4C}{K} + \sum_{k=0}^{K-1} \frac{(k+1)(k+2)}{K(K+1)} u_k. \tag{2.9.5}$$

Proof. We proof this lemma by induction on K. We have $\omega_0 = 1$, thus taking $k = 0$ in (2.9.4), we obtain that $\gamma_1 \leq C + u_0$, which proves the claim for $K = 1$. Let us assume that the claim holds true for some $K \in \mathbb{N}^*$. We deduce from (2.9.4) that

$$\begin{aligned}
\gamma_{K+1} &\leq \Big(\frac{1}{K+2} + \frac{1}{(K+2)^2}\Big)4C + \frac{K}{K+2}\Big(\sum_{k=0}^{K-1} \frac{(k+1)(k+2)}{K(K+1)} u_k\Big) + u_K \\
&\leq \frac{4C}{K+1} + \sum_{k=0}^{K} \frac{(k+1)(k+2)}{(K+1)(K+2)} u_k.
\end{aligned}$$

Therefore the claim holds for $K + 1$. This concludes the proof. □

Lemma 2.9.4. *Let A, B, and C be three random variables. Assume that B is independent of (A, C) and that $B \sim \mathrm{Bern}(\omega)$ for some $\omega \in [0, 1]$. Let F be a real-valued function of (A, B, C). Assume that $|F(A, 1, C) - F(A, 0, C)| \leq \delta$, a.s. Finally, define $U = \mathbb{E}[F(A, B, C) \mid A, B] - \mathbb{E}[F(A, B, C) \mid A]$. Then,*

$$\mathbb{E}[U \mid A] = 0, \qquad U \leq \delta, \qquad \mathbb{E}[U^2 \mid A] \leq \omega(1 - \omega)\delta^2, \quad a.s.$$

Proof. The equality $\mathbb{E}[U \mid A] = 0$ is trivial. We have $U = \mathbb{E}[Z \mid A, B]$, where

$$Z = F(A, B, C) - \mathbb{E}[F(A, B, C) \mid A, C].$$

It is easy to verify that $Z \leq \delta$, a.s., which implies that $\mathbb{E}[U \mid A] = \mathbb{E}[Z \mid A] \leq \delta$. The first inequality is proved. For the second inequality, we first note that

$$\mathbb{E}[Z^2 \mid A, C] = \omega(1-\omega)(F(A, 1, C) - F(A, 0, C))^2,$$

as can be easily verified. Thus $\mathbb{E}[Z \mid A] \leq \omega(1-\omega)\delta^2$. Next by Jensen's inequality, we have $U^2 \leq \mathbb{E}[Z^2 \mid A, B]$. Therefore,

$$\mathbb{E}[U^2 \mid A] \leq \mathbb{E}\big[\mathbb{E}[Z^2 \mid A, B] \mid A\big] = \mathbb{E}[Z^2 \mid A] \leq \omega(1-\omega)\delta^2,$$

as was to be proved. □

The following lemma is an elementary property of the conditional expectation. For the sake of simplicity, we only state it (and prove it) with discrete random variables.

Lemma 2.9.5. *Let X, Y, and Z be three random variables. Assume that Y and Z are discrete and that Z is independent of (X, Y). Then, $\mathbb{E}\big[X \mid Y, Z\big] = \mathbb{E}\big[X \mid Y\big]$.*

Proof. By definition, $\mathbb{E}\big[X \mid Y, Z\big] = \phi(Y, Z)$, where ϕ is defined as follows: for any pair (y, z) such that $\mathbb{P}\big[Y = y \text{ and } Z = z\big] \neq 0$,

$$\phi(y, z) = \frac{\mathbb{E}\big[X \mathbf{1}_{Y=y} \mathbf{1}_{Z=z}\big]}{\mathbb{P}\big[Y = y \text{ and } Z = z\big]} = \frac{\mathbb{E}\big[X \mathbf{1}_{Y=y}\big]}{\mathbb{P}\big[Y = y\big]},$$

since Z is independent of (X, Y). Thus ϕ does not depend on Z and the result follows. □

Chapter 3

Mean field optimization problems: stability results and Lagrangian discretization

3.1 Introduction

Mean field optimization (MFO) problems have recently gained significant interest in various domains, including training problems with neural networks with one hidden layer [CB18, MMN18], sparse inverse problems with differentiable observation models [BSR17], etc. The abstract mean field optimization problem of interest writes:

$$\inf_{\mu \in \mathcal{P}} \mathcal{F}(\mu),$$

where $\mathcal{P}$ is a set of probability measures and $\mathcal{F}$ is a function from $\mathcal{P}$ to $\mathbb{R}$. This article is concerned with the case where $\mathcal{P}$ is the set of probability measures sharing the same marginal distribution and $\mathcal{F}(\mu)$ depends on the expectation of μ with respect to a certain function. To be more concrete, let X and Y be two complete and separable metric spaces and let $\mathcal{H}$ be a separable Hilbert space. Let Z be a closed subset of $X \times Y$ and let m be a probability measure on X. We focus on the following problem, parameterized by m:

$$\inf_{\mu \in \mathcal{P}_m(Z)} f\left(\int_Z g d\mu\right), \tag{P_m}$$

where $g\colon Z \to \mathcal{H}$ is a Borel measurable function and $f\colon \mathcal{H} \to \mathbb{R}$ is a convex function. The admissible set $\mathcal{P}_m(Z)$ is the set of all probability measures on Z whose marginal distribution on X is m. We mention that the rigorous setting for (P_m) is presented in Sec. 3.2.3.

Motivation. Problem (P_m) can be viewed as a social welfare optimization problem, where we consider nonatomic agents (the case with atomic agents is explored in [Wan17, BLO$^+$22]). Here, the agents' positions are distributed according to m, and the set Z consists of all feasible pairs of agent positions and strategies. The function $g(x, y)$ represents the contribution to some common goods made by an agent situated at some position x and following a strategy y. The objective function of

(P_m) involves a function f, called social cost f, evaluated at the aggregate term $\int_Z g d\mu$. It is worth mentioning that problem (P_m) has the general form of the potential formulation of several games with variational structure, including aggregative congestion games [LOW22], nonatomic potential games [CL18b], and potential Lagrangian mean field games (MFGs) [BCS17] described in detail in the following paragraph. This is a consequence of the first-order necessary and sufficient optimality condition: As will be shown, any $\bar{\mu} \in \mathcal{P}_m(Z)$ is a solution to (P_m) if and only if for $\bar{\lambda} = \nabla f(\int_Z g d\bar{\mu})$, it holds that

$$\mathrm{supp}(\bar{\mu}_x) \subseteq \mathbf{BR}_{\bar{\lambda}}(x) \coloneqq \operatorname*{argmin}_{y \in Y} \left\{ \langle \bar{\lambda}, g(x,y) \rangle \mid (x,y) \in Z \right\}, \quad \text{for a.e. } x \text{ in } m. \tag{3.1.1}$$

Here $(\bar{\mu}_x)_{x \in X}$ denotes the disintegration of $\bar{\mu}$ (see Theorem 3.2.7) and $\mathrm{supp}(\bar{\mu}_x)$ denotes the support of $\bar{\mu}_x$ (see (3.2.1)). This result is precisely stated in Corollary 3.3.5. In other words, (3.1.1) defines the conditions for a Nash equilibrium: an agent at position x must minimize $\langle \bar{\lambda}, g(x, \cdot) \rangle$ and $\bar{\lambda}$ is a coupling variable, common to all agents, which results from their collective behavior through the relation $\bar{\lambda} = \nabla f(\int_Z g d\bar{\mu})$.

Lagrangian MFGs. The framework of the potential Lagrangian mean field games (MFGs) follows from [BCS17, SS21, CC18, Sar22]. Let us fix a domain $\Omega \subseteq \mathbb{R}^d$ and a final time $T > 0$. Let $\mathrm{AC}([0,T], \mathbb{R}^d)$ be the set of all absolutely continuous functions from $[0,1]$ to $\mathbb{R}^d$. For any $x \in \Omega$, we denote,

$$\Gamma \coloneqq \{\gamma \in \mathrm{AC}([0,T], \mathbb{R}^d) \mid \gamma(t) \in \Omega, \, \forall t \in [0,T]\}, \qquad \Gamma_x \coloneqq \{\gamma \in \Gamma \mid \gamma(0) = x\}.$$

Let $Z = \{(x, \gamma) \mid x \in \Omega, \, \gamma \in \Gamma_x\}$. Let $m \in \mathcal{P}(\Omega)$ be the distribution of initial states of players. Following [Sar22], a general potential Lagrangian MFG writes:

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \int_0^T \tilde{L}(\dot{\gamma}(t)) dt d\mu(z) + \int_{t=0}^T \tilde{\mathcal{F}}(e_t \# \pi_2 \# \mu) dt, \tag{3.1.2}$$

where $\tilde{L} \colon \mathbb{R}^d \to \mathbb{R}$ is a running cost, $\tilde{\mathcal{F}} \colon \mathcal{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is a congestion cost, $e_t \colon \Gamma \to \Omega, \gamma \mapsto \gamma(t)$, and $\pi_2 \colon Z \to \Gamma, (x,\gamma) \mapsto \gamma$. We are interested in the case where there exist two functions $\tilde{g} \colon \mathbb{R}^d \to \mathbb{R}$ and $\tilde{f} \colon \mathbb{R} \to \mathbb{R}$ such that for any $\nu \in \mathcal{P}(\Omega)$,

$$\tilde{\mathcal{F}}(\nu) = \tilde{f}\left(\int \tilde{g}(x) d\nu(x) \right).$$

In this case, the previous Lagrangian MFG (3.1.2) writes:

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \int_0^T \tilde{L}(\dot{\gamma}(t)) dt d\mu(x,\gamma) + \int_{t=0}^T \tilde{f}\left(\int_Z \tilde{g}(\gamma(t)) d\mu(x,\gamma) \right) dt, \tag{3.1.3}$$

which follows the structure of problem (P_m). We comment more in details on this example in Remark 3.3.6.

Theoretical results. We study the MFO problem (P_m) from both primal and dual perspectives. In the primal sense, we establish a first-order optimality condition for (P_m) under mild assumptions. As we already explained, this condition turns out to be equivalent to a Nash equilibrium condition for Lagrangian MFGs and nonatomic potential games, as discussed in Remarks 3.3.6-3.3.7. Additionally, we provide an existence result for the solution under a "tightness" assumption on the minimizing sequence of (P_m).

The first contribution of this article lies in the stability analysis of the primal problem (P_m), see Theorem 3.3.16. Firstly, we derive an upper bound for the variation of the optimal cost of (P_m) for two different parameters m. The obtained bound is linear with respect to the Kantorovich-Rubinstein distance of the parameters. Secondly, we introduce a recovery method (Algorithm 3.1), which bridges approximate solutions for problems with different parameters. This recovery method is important for the discretization of the problem, since it permits to construct an approximate minimizer of the original problem (P_m) based on an approximate solution of the discretized problem (P_{m_N}) stated later.

From the dual perspective, we prove the strong duality for (P_m) under certain qualification assumptions. We also conduct a stability analysis for dual problems with different parameters, obtaining upper bounds for both the gap of the dual values and the distance of the dual solutions. At the end of this section, we provide a formula for the directional derivative of the value function of (P_m) using the strong duality result.

Discretization and algorithms. We present an original resolution method for the MFO problem (P_m). Our approach relies first on a discretization of the marginal m, as proposed in [Sar22] for Lagrangian MFGs. We approximate the common marginal distribution m in (P_m) by an empirical distribution $m_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$, where $x_i \in X$ for $i = 1, \dots, N$ and N is a sufficiently large integer. This allows us to write the associated discretized problem for (P_m) as

$$\inf_{\mu \in \mathcal{P}_{m_N}(Z)} f\left(\int_Z g d\mu\right). \tag{P_{m_N}}$$

To solve (P_{m_N}) numerically, a first approach is the Frank-Wolfe algorithm [Jag13], in which one needs to solve a linearized problem (3.5.2) at each iteration. This approach is similar to the fictitious play algorithm applied to the Lagrangian MFG in [CH17]. As mentioned in [BSR17, CB18, BLO$^+$22], one major issue of the Frank-Wolfe algorithm is the increase of the cardinality of the support of the approximate solution along the iterations, which may cause a memory overflow problem.

We suggest to leverage the fact that the discretized problem (P_{m_N}) can be equivalently represented as the relaxation of an aggregative optimization problem involving N agents, as defined by the authors in [BLO$^+$22]. This allows to apply to (P_{m_N}) the Stochastic Frank-Wolfe algorithm proposed and investigated in [BLO$^+$22]. The output of Algorithm 3.3 has a support of fixed size N.

Finally, the combination of the Stochastic Frank-Wolfe algorithm (Algorithm 3.3) with the recovery method (Algorithm 3.1) allows to obtain an approximate solution of (P_m) whose quality improves as the discretization parameter N increases, as demonstrated in Lemma 3.5.1 and Theorem 3.5.8. To illustrate the effectiveness of our approach, we apply it to a numerical example taken from [GHS22], involving the optimal exploitation of exhaustible resources.

Organization. In Section 3.2, we present some notations and results in measure theory and set-valued functions, as well as the rigorous description of the data of problem (P_m). Section 3.3 is dedicated to the primal problem: We provide a first-order optimality condition, an existence result and a stability analysis for the primal problem. In Section 3.4, we formulate the dual problem of (P_m), we prove strong duality, and we prove the stability of the dual solution. We provide our numerical method in Section 3.5. We perform in Section 3.6 some numerical simulations for a Lagrangian MFG model taken from [GHS22].

3.2 Preliminaries

3.2.1 Results in measure theory

A metric space is called a *Polish* space if it is complete and separable. Let X be a Polish space equipped with a metric d_X, and let $\mathcal{X}$ be a σ-algebra on X. The Borel σ-algebra on X is denoted by $\mathcal{B}^X$. Given any measure m on $\mathcal{X}$, we refer to the triplet $(X, \mathcal{X}, m)$ as a measure space. Measure spaces are said to be complete if for any $A \in \mathcal{X}$ with $m(A) = 0$ and for any subset B of A, we have $B \in \mathcal{X}$. We define

$$\mathcal{P}(X) := \big\{ m \text{ is a positive Borel measure on } X, \text{ and } m(X) = 1 \big\};$$

$$\mathcal{P}^1(X) := \left\{ m \in \mathcal{P}(X) \,\middle|\, \exists x_0 \in X \text{ such that } \int_X d_X(x, x_0) dm < +\infty \right\}.$$

Let δ_x denote the Dirac measure at point x. We denote by $\mathcal{P}_\delta(\Omega)$ the set of finitely supported probability measures, defined by

$$\mathcal{P}_\delta(X) := \left\{ \sum_{k=1}^K \omega_k \delta_{x_k} \,\middle|\, K \in \mathbb{N},\, (\omega_k)_{k=1}^K \in (\mathbb{R}_+)^K,\, (x_k)_{k=1}^K \in X^K,\, \sum_{k=1}^K \omega_k = 1 \right\}.$$

In particular, we call $m \in \mathcal{P}_\delta(X)$ an *empirical distribution* if $\lambda_k = 1/K$ for $k = 1, \ldots, K$.

The set $\mathcal{P}(X)$ is endowed with the narrow topology. We say that a sequence $(m_n)_{n \geq 1}$ in $\mathcal{P}(Z)$ narrowly converges to some $m \in \mathcal{P}(X)$ if for any bounded and continuous function $F \colon X \to \mathbb{R}$,

$$\lim_{n \to +\infty} \int_X F dm_n = \int_X F dm.$$

The space $\mathcal{P}^1(X)$ is endowed with the *Kantorovich–Rubinstein Distance*,

$$d_1(m_0, m_1) := \sup_{F \in \mathrm{Lip}_1(X)} \int_\Omega F d(m_0 - m_1),$$

where $\mathrm{Lip}_1(X)$ is the set of all 1-Lipschitz continuous functions on X. For any $m \in \mathcal{P}(X)$, the support of m is defined by

$$\mathrm{supp}(m) := \big\{ x \in X \mid m(V) > 0 \text{ for all open set } V \text{ such that } x \in V \big\}. \tag{3.2.1}$$

Lemma 3.2.1. *Let $m \in \mathcal{P}(X)$. Let $F \colon X \to \mathbb{R}_+$ be a Borel measurable function. Assume that*

$$\int_X F dm = 0.$$

Then $F = 0$, m-a.e. Moreover, if $F^{-1}(\{0\})$ is closed, then $\mathrm{supp}(m) \subseteq F^{-1}(\{0\})$.

Proof. The fact that $F = 0$, m-a.e., is from [Rud87, Thm. 1.39(a)]. Now, let $F^{-1}(\{0\})$ be closed. Suppose that there exists $x \in \mathrm{supp}(m)$ such that $x \notin F^{-1}(\{0\})$. Since $F^{-1}(\{0\})$ is closed, there exists an open neighborhood V of x such that $F(x) > 0$, for all $x \in V$. By the definition of the support of a probability measure, we have $m(V) > 0$. Therefore, $\int_X F dm \geq \int_V F dm > 0$, contradiction. □

3.2.2 Results about set-valued functions

In this subsection, we consider a metric space X equipped with a metric d_X, a σ-algebra $\mathcal{X}$ on X, and a measure m on $\mathcal{X}$. Additionally, we fix a Polish space Y with a metric d_Y, and we denote the Borel σ-algebra on Y by $\mathcal{B}^Y$. We call F a set-valued function from X to Y if $F(x) \subseteq Y$ for all $x \in X$, denoted by $X \rightsquigarrow Y$ for short. The graph of F is defined by

$$\mathrm{Graph}(F) \coloneqq \{(x, y) \in X \times Y \mid y \in F(x)\}.$$

We say that F has closed (non-empty) images, if for any $x \in X$, $F(x)$ is closed (non-empty) in Y.

Let us give some definitions concerning regularity properties of set-valued functions, which are from [AF09, Def. 1.4.1, Def. 1.4.2, Def. 1.4.5, and Def. 8.1.1].

Definition 3.2.2. Let $F \colon X \rightsquigarrow Y$ be a set-valued function with non-empty images.

1. (Lower semi-continuity). The set-valued function F is *lower semi-continuous* at point $x \in X$ if for any $y \in F(x)$ and any sequence $(x_n \in X)_{n \geq 1}$ converging to x, there exists $y_n \in F(x_n)$ converging to y. The set-valued function F is said to be lower semi-continuous if it is lower semi-continuous at each point $x \in X$.

2. (Upper semi-continuity). The set-valued function F is *upper semi-continuous* at point $x \in X$ if for any neighborhood $\mathcal{U}$ of $F(x)$, there exists $\eta > 0$ such that for any $x' \in B_X(x, \eta)$, we have

   $$F(x') \subseteq \mathcal{U}.$$

   The set-valued function F is said to be upper semi-continuous if it is upper semi-continuous at each point $x \in X$.

3. (Lipschitz continuity). When X and Y are normed vector spaces, we say that F is L-Lipschitz continuous on X, for some $L > 0$, if for any $x_1, x_2 \in X$,

   $$F(x_1) \subseteq F(x_2) + B_Y(0, L d_X(x_1, x_2)).$$

   Here $B_Y(0, r)$ denotes the closed ball in Y centered at 0 with radius $r > 0$.

4. (Measurability). The set-valued function F is *measurable* if the inverse image of any open subset $\mathcal{O}$ of Y is measurable, i.e.,

   $$F^{-1}(\mathcal{O}) \coloneqq \{x \in X \mid F(x) \cap \mathcal{O} \neq \emptyset\} \in \mathcal{X}.$$

An important property of measurable set-valued functions is the existence of measurable selections.

Theorem 3.2.3 (Measurable selection)**.** *Let $F\colon X \rightsquigarrow Y$ be a measurable set-valued function with non-empty images. Then F has a measurable selection f, i.e., $f\colon X \to Y$ is $(\mathcal{X}, \mathcal{B}^Y)$-measurable and $f(x) \in F(x)$ for any $x \in X$.*

Proof. See [AF09, Thm. 8.1.3]. □

The following two lemmas will allow us to prove the measurability of some set-valued functions.

Lemma 3.2.4. *If $F\colon X \rightsquigarrow Y$ is a set-valued function such that $F^{-1}(\mathcal{C}) \in \mathcal{X}$ for any closed subset $\mathcal{C}$ of Y, then F is measurable.*

Proof. See [CV06, Prop. III.11]. □

Lemma 3.2.5. *Let $(X, \mathcal{X}, m)$ be a complete measure space, with m a positive measure such that $m(X) = 1$. Then any set-valued mapping $F\colon X \rightsquigarrow Y$ is measurable if and only if Graph(F) belongs to $\mathcal{X} \otimes \mathcal{B}^Y$.*

Proof. See [AF09, Thm. 8.1.4]. □

3.2.3 Data setting and technical lemmas

Recall the MFO problem (P_m). We consider the following setting:

- Two Polish spaces and their Borel σ-algebras: $(X, \mathcal{B}^X)$ and $(Y, \mathcal{B}^Y)$.
- A probability distribution on X: $m \in \mathcal{P}(X)$.
- A set-valued function $F\colon X \rightsquigarrow Y$ with a closed graph and non-empty images. Let
$$Z \coloneqq \mathrm{Graph}(F), \qquad Z_x \coloneqq F(x),\ \forall x \in X.$$
- The admissible set of probability measures:
$$\mathcal{P}_m(Z) \coloneqq \{\mu \in \mathcal{P}(Z) \mid \pi_1 \# \mu = m\},$$
where $\pi_1\colon Z \to X$, $(x, y) \mapsto x$.
- A separable Hilbert space: $\mathcal{H}$.
- Two Borel measurable functions: $g\colon Z \to \mathcal{H}$ and $f\colon \mathcal{H} \to \mathbb{R}$.

The integral $\int_Z g d\mu$ in (P_m) should be interpreted in the Bochner integration sense. We refer to [Coh13, Appx. E] for Bochner integrable functions.

Lemma 3.2.6. *If there exists a constant $M > 0$ such that $\|g(z)\| \leq M$ for any $z \in Z$, then the function g is Bochner integrable with respect to any $\mu \in \mathcal{P}(Z)$, i.e., $\int_Z g d\mu$ exists. Moreover, for any $\lambda \in \mathcal{H}$, we have*
$$\left\langle \lambda, \int_Z g d\mu \right\rangle = \int_Z \langle \lambda, g \rangle d\mu.$$
As a consequence, for any $\mu_1, \mu_2 \in \mathcal{P}(Z)$, we have
$$\left\langle \int_Z g d\mu_1\,, \int_Z g d\mu_2 \right\rangle = \int_Z \int_Z \langle g(x), g(y) \rangle d\mu_1(x) d\mu_2(y).$$

Proof. As $\mathcal{H}$ is separable, the function g is strongly measurable. Moreover, as the constant function M is Bochner integrable with respect to any $\mu \in \mathcal{P}(Z)$, and $|g(z)| \leq M$ for any $z \in Z$, it follows from [Coh13, Prop. E.2, Thm. E.6] that g is Bochner integrable with respect to any $\mu \in \mathcal{P}(Z)$. Therefore, we can apply [Coh13, Prop. E.11] to obtain the first equality of this lemma. The second equality is obtained by applying twice the first one. □

Theorem 3.2.7 (Disintegration theorem)**.** *For any $\mu \in \mathcal{P}_m(Z)$, there exists a family of probability measures $\{\mu_{x\in X} \in \mathcal{P}(Y)\}_x$ such that for any Borel measurable function $f\colon Z \to \mathbb{R}_+$, we have*

$$\int_Z f d\mu = \int_X \int_{Z_x} f(x,y) d\mu_x(y) dm(x).$$

Moreover, for a.e. $x \in X$, μ_x is uniquely determined.

Proof. See [AGS05, Thm. 5.3.1]. □

Remark 3.2.8. It is not difficult to generalize Theorems 3.2.7 to functions f bounded from below, by adding to f a sufficient large positive constant.

3.3 Primal mean field optimization problem

3.3.1 Assumptions and constants

To simplify the presentation of the assumptions and the results of the article, we introduce the following (set-valued) functions, parameterized by $\lambda \in \mathcal{H}$:

- $g_\lambda\colon Z \to \mathbb{R}$ and $u_\lambda\colon X \to \mathbb{R}$,

$$g_\lambda(x,y) = \langle \lambda\,,\, g(x,y)\rangle\,, \qquad u_\lambda(x) = \inf_{y\in Z_x} g_\lambda(x,y);$$

- $G_\lambda\colon X \rightsquigarrow \mathbb{R}$ and $\mathbf{BR}_\lambda\colon X \rightsquigarrow Y$,

$$G_\lambda(x) = \{g_\lambda(x,y) \mid y \in Z_x\}\,, \qquad \mathbf{BR}_\lambda(x) = \operatorname*{argmin}_{y\in Z_x} g_\lambda(x,y).$$

Assumption A. The following holds:

1. The function g is bounded. The function f is convex and differentiable, and ∇f is Lipschitz continuous with modulus L.
2. Let $\mathcal{H}_f \coloneqq \nabla f(\mathcal{H})$. Fixing any $\lambda \in \mathcal{H}_f$, we have:
   - the function g_λ is lower semi-continuous;
   - the set-valued function G_λ is lower semi-continuous;
   - the set-valued function $\mathbf{BR}_\lambda$ has non-empty images.

Three useful constants below are defined, following Assumption A:

$$M := \sup_{z \in Z} \|g(z)\|, \qquad D := \sup_{z_1, z_2 \in Z} \|g(z_1) - g(z_2)\|^2, \qquad C := \sup_{\mu \in \mathcal{P}(Z)} \left\| \nabla f \left(\int_Z g d\mu \right) \right\|.$$

We present here a lemma following Assumption A. A similar result for the Lagrangian MFG is presented in [CC18, Lem. 3.4].

Lemma 3.3.1. *Under Assumption A, for any $\lambda \in \mathcal{H}_f$, the set-valued function $\mathbf{BR}_\lambda$ has a closed graph.*

Proof. Let $x_k \in X$ converge to some $\bar{x} \in X$, and let $y_k \in \mathbf{BR}_\lambda(x_k)$ converge to some $\bar{y} \in Y$. We have to prove that $\bar{y} \in \mathbf{BR}_\lambda(\bar{x})$. First, we have $\bar{y} \in Z_{\bar{x}}$, since Z is closed. Fix any $y \in Z_{\bar{x}}$. Since G_λ is lower semi-continuous, there exists a sequence $(\hat{y}_k)_{k \in \mathbb{N}}$ in Z_{x_k} such that

$$g_\lambda(\bar{x}, y) = \lim_{k \to \infty} g_\lambda(x_k, \hat{y}_k).$$

By the lower semi-continuity of g_λ, we have

$$g_\lambda(\bar{x}, \bar{y}) \leq \liminf_{k \to \infty} g_\lambda(x_k, y_k).$$

Since $y_k \in \mathbf{BR}_\lambda(x_k)$ and $\hat{y}^k \in Z_{x_k}$, we have $g_\lambda(x_k, y_k) \leq g_\lambda(x_k, \hat{y}_k)$ for any k. Passing to the limit in this inequality (using the above inequalities), we deduce that $g_\lambda(\bar{x}, \bar{y}) \leq g_\lambda(\bar{x}, y)$. Thus, $\mathbf{BR}_\lambda$ has a closed graph. □

In Section 3.4, we will consider the dual problem of (P_m). For the analysis of Section 3.4, Assumption A needs to be strengthened as follows:

Assumption A*. Assumption A(1) holds true and Assumption A(2) holds true for all $\lambda \in \mathrm{dom}(f^*)$, where f^* is the Fenchel conjugate of f.

Remark 3.3.2. Assumption $\mathbf{A}^*$ is indeed stronger than Assumption A since $\mathcal{H}_f = \nabla f(\mathcal{H}) \subseteq \mathrm{dom}(f^*)$. This inclusion is deduced from Fenchel's relation: $y = \nabla f(x) \Leftrightarrow f^*(y) = \langle x, y \rangle - f(x)$.

3.3.2 First-order-optimality condition

The following lemma plays a key role in proving the first-order optimality condition for (P_m).

Lemma 3.3.3. *Let Assumption A hold true. For any $\lambda \in \mathcal{H}_f$, we have*

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z g_\lambda d\mu = \int_X u_\lambda dm.$$

Here we present a proof of Lemma 3.3.3 for the case where m has finite support, that is, $m \in \mathcal{P}_\delta(X)$. This particular case provides us with insight into the general proof, and proves beneficial for resolving the discretized problem introduced in Section 3.5.

Proof of Lemma 3.3.3 when $m \in \mathcal{P}_\delta(X)$. Fix any $\mu \in \mathcal{P}_m(Z)$. Since g is bounded over Z, the function g_λ is bounded from below. By Lemma 3.2.7 and Remark 3.2.8, we have

$$\int_Z g_\lambda d\mu = \int_X \int_{Z_x} g_\lambda(x,y) d\mu_x(y) dm(x) \geq \int_X u_\lambda dm,$$

where the second inequality follows from the definition of u_λ.

Let us prove the converse inequality. Let us fix $m \in \mathcal{P}_\delta(X)$. Let $K \in \mathbb{N}$, let $(x_k)_{k=1,\dots,K} \in X^k$ and let $(\omega_k)_{k=1,\dots,K} \in \mathbb{R}_+^K$ be such that $\sum_{k=1}^K \omega_k = 1$ and $m = \sum_{k=1}^K \omega_k \delta_{x_k}$. For any $k = 1,\dots,K$, let $y_k \in \mathbf{BR}_\lambda(x_k)$. Let us define

$$\tilde{\mu} = \sum_{k=1}^K \omega_k \delta_{(x_k,y_k)}.$$

Clearly $\tilde{\mu} \in \mathcal{P}_m(Z)$. Moreover,

$$\int_Z g_\lambda d\tilde{\mu} = \sum_{k=1}^K \omega_k g_\lambda(x_k,y_k) = \sum_{k=1}^K \omega_k u_\lambda(x_k) = \int_X u_\lambda dm.$$

The conclusion follows, moreover, $\tilde{\mu}$ minimizes $\int_Z g_\lambda d\mu$ over $\mathcal{P}_m(Z)$. □

In the general case, one has to find a measurable selection of $\mathbf{BR}_\lambda$, which requires us to prove the measurability of $\mathbf{BR}_\lambda$, which cannot be done in a direct fashion. The complete proof is given in Appendix 3.7.1.

Theorem 3.3.4 (First-order optimality condition)**.** *Let Assumption A(1) hold true. Let $\bar{\mu} \in \mathcal{P}_m(Z)$ and $\bar{\lambda} = \nabla f\left(\int_Z g d\bar{\mu}\right)$. Consider the following three assertions:*

1. *The measure $\bar{\mu}$ is a solution of problem* (P_m)*;*
2. $\int_Z g_{\bar{\lambda}} d\bar{\mu} = \inf_{\mu \in \mathcal{P}_m(Z)} \int_Z g_{\bar{\lambda}} d\mu$*;*
3. $\mathrm{supp}(\bar{\mu}_x) \subseteq \mathbf{BR}_{\bar{\lambda}}(x)$*, m-a.e., where $\bar{\mu}_x$ is defined by the disintegration theorem.*

Then, assertions (1) and (2) are equivalent. Moreover, under Assumption A(2), assertions (1), (2), and (3) are equivalent.

Proof. **Step 1**. (Equivalence between (1) and (2)). We first prove that (1) ⇒ (2). Suppose that $\bar{\mu}$ is a solution of problem (P_m). Take an arbitrary $\mu \in \mathcal{P}_m(Z)$. Then, for any $\alpha \in [0,1]$, we have

$$\begin{aligned} f\left(\int_Z g d\bar{\mu}\right) &\leq f\left(\int_Z g d(\bar{\mu} + \alpha(\mu - \bar{\mu}))\right) \\ &\leq f\left(\int_Z g d\bar{\mu}\right) + \alpha \left\langle \bar{\lambda}, \int_Z g d(\mu - \bar{\mu}) \right\rangle + \frac{\alpha^2 LD}{2}, \end{aligned}$$

where the second inequality follows from the Lipschitz-continuity of ∇f and the definition of D. Therefore

$$0 \leq \left\langle \bar{\lambda}, \int_Z g d(\mu - \bar{\mu}) \right\rangle + \frac{\alpha LD}{2}$$

Let α go to 0. We obtain that

$$\left\langle \bar{\lambda}, \int_Z g d\bar{\mu} \right\rangle = \inf_{\mu \in \mathcal{P}_m(Z)} \left\langle \bar{\lambda}, \int_Z g d\mu \right\rangle. \tag{3.3.1}$$

This implies (2) by the definition of $g_{\bar{\lambda}}$.

We now prove (2) $\Rightarrow$ (1). Let (2) hold true. We obtain (3.3.1) by the definition of $g_{\bar{\lambda}}$. The convexity of f implies that for any $\mu \in \mathcal{P}_m(Z)$,

$$f\left(\int_Z g d\mu\right) \geq f\left(\int_Z g d\bar{\mu}\right) + \left\langle \bar{\lambda}, \int_Z g d\mu - \int_Z g d\bar{\mu} \right\rangle \geq f\left(\int_Z g d\bar{\mu}\right).$$

Therefore, $\bar{\mu}$ is a solution of problem (P_m).

Step 2. (Equivalence between (2) and (3)). By Theorem 3.2.7, we have

$$\int_Z g_{\bar{\lambda}} d\bar{\mu} = \int_X \int_{Z_x} g_{\bar{\lambda}}(x, y) d\bar{\mu}_x(y) dm(x).$$

By Lemma 3.3.3, we have

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z g_{\bar{\lambda}} d\mu = \int_X u_{\bar{\lambda}} dm.$$

Therefore, assertion (2) is equivalent to

$$\int_X \int_{Z_x} g_{\bar{\lambda}}(x, y) d\bar{\mu}_x(y) dm(x) = \int_X u_{\bar{\lambda}} dm. \tag{3.3.2}$$

Let (3) hold true. It follows that $\int_{Z_x} g_{\bar{\lambda}}(x, y) d\bar{\mu}_x(y) = u_{\bar{\lambda}}(x)$, m-a.e., which implies (3.3.2).

Let (2) hold true. We obtain (3.3.2). The function $x \mapsto \left(\int_{Z_x} g_{\bar{\lambda}}(x, y) d\bar{\mu}_x(y)\right) - u_{\bar{\lambda}}(x)$ is nonnegative, for a.e. $x \in X$, by the definition of $u_{\bar{\lambda}}$. By (3.3.2), its integral is null, thus, as a consequence of Lemma 3.2.1, we have

$$\int_{Z_x} g_{\bar{\lambda}}(x, y) d\bar{\mu}_x(y) = u_{\bar{\lambda}}(x) = \inf_{y \in Z_x} g_{\bar{\lambda}}(x, y), \qquad m\text{-a.e.} \tag{3.3.3}$$

Fix $x \in X$ such that equality holds in (3.3.3). Consider the map $y \in Z_x \mapsto g_{\bar{\lambda}}(x, y) - u_{\bar{\lambda}}(x)$. It is nonnegative, with a null integral, and $\mathbf{BR}_{\bar{\lambda}}(x)$ is non-empty and closed. Then assertion (3) follows with Lemma 3.2.1. □

Corollary 3.3.5. *Under Assumption A, $\bar{\mu}$ is a solution of* (P_m) *if and only if the following equilibrium equation is satisfied:*

$$\begin{cases} \bar{\lambda} = \nabla f\left(\int_Z g d\bar{\mu}\right), \\ \operatorname{supp}(\bar{\mu}_x) \subseteq \mathbf{BR}_{\bar{\lambda}}(x), \quad m\text{-a.e.} \end{cases} \tag{3.3.4}$$

Proof. This is a consequence of Theorem 3.3.4. □

The equilibrium equation (3.3.4) shares similarities with the conditions that characterize Nash equilibria in Lagrangian MFGs [CC18, SS21] and nonatomic potential games [CL18b], as illustrated in the following two remarks.

Remark 3.3.6. In the Lagrangian MFG (3.1.3), if $(\bar{\lambda}, \bar{\mu})$ satisfies (3.3.4), then $\bar{\lambda} = (1, p(\bar{\mu}))$, where $p(\bar{\mu}) = \nabla \tilde{f}\left(\int_Z \tilde{g}(\gamma(t)) d\bar{\mu}(x,\gamma)\right)$. The variable $p(\bar{\mu})$ can be interpreted as a social price determined by the population distribution $\bar{\mu}$. The second line of (3.3.4) indicates that for almost every x with respect to the measure m, the elements in $\mathrm{supp}(\bar{\mu}_x)$ are solutions of an optimal control problem, similarly to [CC18, Def. 3.1, Eq. 3.32] and [SS21, Def. 2.2]. In the framework of (3.1.3), this problem writes:

$$\inf_{\gamma \in \Gamma_x} \int_0^T \tilde{L}(\dot{\gamma}(t)) + \langle p(\bar{\mu})(t), \tilde{g}(\gamma(t)) \rangle dt.$$

Remark 3.3.7. To investigate the connection with the nonatomic potential games presented in [CL18b], let us consider a specific case of (P_m) where $m = \delta_{x_0}$ for some $x_0 \in X$. In this scenario, (P_m) can be rewritten as:

$$\inf_{\nu \in \mathcal{P}(Z_{x_0})} f\left(\int_{Z_{x_0}} g(x_0, y) d\nu(y)\right).$$

Let $\bar{\nu}$ be a solution to the aforementioned problem, and define $\bar{\lambda} = \nabla f\left(\int_{Z_{x_0}} g(x_0, y) d\bar{\nu}(y)\right)$. The second line of (3.3.4) can be expressed as:

$$\langle \bar{\lambda}, g(x_0, \bar{y}) \rangle \leq \langle \bar{\lambda}, g(x_0, y) \rangle, \quad \forall \bar{y} \in \mathrm{supp}(\bar{\nu}),\ \forall y \in Z_{x_0},$$

which is consistent with the definition of Nash equilibrium in nonatomic potential games presented in [CL18b, Sec. 3].

3.3.3 Existence of a solution under tightness assumptions

We denote by $\mathbf{val}(\mathrm{P}_m)$ the value of problem (P_m). We can easily deduce from Assumption A that $\mathbf{val}(\mathrm{P}_m) > -\infty$. The following proposition demonstrates the existence of a solution to problem (P_m) under some additional assumptions.

Proposition 3.3.8 (Existence)**.** *Let Assumption A hold true. Let $(\mu_n)_{n\geq 1}$ be a minimizing sequence for problem (P_m). Suppose that $\{\mu_n\}_{n\geq 1}$ is tight in $\mathcal{P}(Z)$, i.e. for any $\epsilon > 0$, there exists a compact subset K_ϵ of Z such that*

$$\mu_n(K_\epsilon) \geq 1 - \epsilon, \qquad \forall n \geq 1.$$

Then every accumulation point of $\{\mu_n\}_{n\geq 1}$ for the narrow topology (there exists at least one) is a solution of (P_m).

Proof. By *Prokhorov's* theorem [Vil09, p. 43], the set $\{\mu_n\}_{n\geq 1}$ is relatively compact with respect to the narrow topology. Without loss of generality, suppose that μ_n narrowly converges to some $\bar{\mu} \in \mathcal{P}(Z)$. The set $\mathcal{P}_m(Z)$ is closed with respect to narrow topology by [SS21, Prop. 2.4]. This implies that $\bar{\mu} \in \mathcal{P}_m(Z)$. Let $\bar{\lambda} = \nabla f(\int_Z g d\bar{\mu})$. Since f is convex, we have

$$f\left(\int_Z g d\mu_n\right) \geq f\left(\int_Z g d\bar{\mu}\right) + \int_Z g_{\bar{\lambda}} d(\mu_n - \bar{\mu}). \tag{3.3.5}$$

Since $g_{\bar{\lambda}} \colon Z \to \mathbb{R}$ is lower semi-continuous and bounded from below by Assumption A, we deduce the following inequality from [Vil09, Lem. 4.3]:

$$\liminf_{n \to +\infty} \int_Z g_{\bar{\lambda}} d(\mu_n - \bar{\mu}) \geq 0.$$

In inequality (3.3.5), letting n go to infinity, by the definition of μ_n, we have

$$\mathbf{val}(\mathrm{P}_m) = \liminf_{n \to +\infty} f\left(\int_Z g d\mu_n\right) \geq f\left(\int_Z g d\bar{\mu}\right) \geq \mathbf{val}(\mathrm{P}_m).$$

Therefore, $\bar{\mu}$ is a solution of problem (P_m). □

Remark 3.3.9. Let us make some comments on the assumption of tightness of $\{\mu_n\}_{n\geq 1}$ in Proposition 3.3.8. One simple example is that where Z is compact in $X \times Y$. Let us consider another example inspired from the Lagrangian MFGs (3.1.3). We assume that $\tilde{f}$ in (3.1.3) is positive. Let $(\mu_n)_{n\geq 1}$ be a minimizing sequence for (3.1.3). Since $\mathbf{val}(3.1.3) < +\infty$, there exists $0 < \bar{M} < \infty$ such that $f(\int_Z g d\mu_n) \leq \bar{M}$ for all n. Then, we construct

$$K_\epsilon = \left\{ (x, \gamma) \in Z \,\middle|\, \int_0^T L(\dot{\gamma}(t))dt \leq \frac{\bar{M}}{\epsilon} \right\}. \tag{3.3.6}$$

By Markov's inequality, for any $n \geq 1$,

$$\mu_n(K_\epsilon) = 1 - \mu_n\left(\left\{ (x, \gamma) \in Z \,\middle|\, \int_0^T L(\dot{\gamma}(t))dt > \frac{\bar{M}}{\epsilon} \right\}\right) \geq 1 - \epsilon \frac{\int_Z \int_0^T L(\dot{\gamma}(t))dt d\mu_n}{\bar{M}}.$$

Since $\tilde{f} \geq 0$, we have $\int_Z \int_0^T L(\dot{\gamma}(t))dt d\mu_n \leq f(\int_Z g d\mu_n) \leq \bar{M}$. Then, $\mu_n(K_\epsilon) \geq 1 - \epsilon$, for any $n \geq 1$. Furthermore, if we assume that K_ϵ defined by (3.3.6) is compact in Z, then the tightness of $\{\mu_n\}_{n\geq 1}$ follows. This approach is employed in [SS21], in the proof of existence of a solution of a Lagrangian MFG.

3.3.4 Stability of primal problem

In this subsection, we study the stability of the primal problem (P_m) with respect to its parameter m. We need the following assumptions (recall the data setting introduced in Sec. 3.2.3).

Assumption B. The following holds:

1. The space X is a closed subset of a separable Banach space;
2. The function $g \colon Z \to \mathcal{H}$ is continuous;
3. The set Z_x is compact for any $x \in X$ and the set-valued function $F \colon X \rightsquigarrow Y$ is upper semi-continuous;
4. There exists $L_g \geq 0$ such that the set-valued function

$$\mathcal{Z} \colon X \rightsquigarrow \mathcal{H},\ x \mapsto \{g(x, y) \mid y \in Z_x\} \tag{3.3.7}$$

is L_g-Lipschitz on X.

Let m_0 and m_1 lie in $\mathcal{P}(X)$. We consider the following two instances of (P_m) with $m = m_0$ and $m = m_1$ respectively:

$$\inf_{\mu \in \mathcal{P}_{m_0}(Z)} f\left(\int_Z g d\mu\right); \tag{P_{m_0}}$$

$$\inf_{\mu \in \mathcal{P}_{m_1}(Z)} f\left(\int_Z g d\mu\right). \tag{P_{m_1}}$$

Suppose that we have an (approximate) solution of problem (P_{m_0}), denoted by $\bar{\mu}_0$. Our goal is to propose a feasible approach for recovering an approximate solution of problem (P_{m_1}) from $\bar{\mu}_0$ and to study the performance of this approximation. Our recovery approach relies on $\bar{\mu}_0$ and the solution of the optimal transport problem (OT1) stated later. To introduce (OT1), we need to define some projection operators

Recall that $\pi_1 \colon Z \to X$, $(x, y) \mapsto x$, and $\pi_2 \colon Z \to X$, $(x, y) \mapsto y$. The other projection operators used in this subsection are defined as:

$$\tilde{\pi}_1 : X \times X \to X,\, (x, x') \mapsto x, \quad \tilde{\pi}_2 : X \times X \to X,\, (x, x') \mapsto x', \quad \pi_3 \colon Z \times X \to X, (x, y, x') \mapsto x',$$
$$\pi_{12} \colon Z \times X \to Z,\, (x, y, x') \mapsto (x, y), \qquad \pi_{13} \colon Z \times X \to Z,\, (x, y, x') \mapsto (x, x').$$

It directly follows from the above definitions that

$$\tilde{\pi}_1 \circ \pi_{13} = \pi_1 \circ \pi_{12} \quad \text{and} \quad \tilde{\pi}_2 \circ \pi_{13} = \pi_3.$$

Now we consider the following optimal transport problem:

$$\inf_{\rho \in \Pi(m_0, m_1)} \int_{X \times X} d_X(x, x') d\rho(x, x'), \tag{OT1}$$

where $\Pi(m_0, m_1) = \{\rho \in \mathcal{P}(X \times X) \mid \tilde{\pi}_1 \# \rho = m_0,\, \tilde{\pi}_2 \# \rho = m_1\}$. It follows from [Vil09, Rem. 6.5] that if m_0 and m_1 lie in $\mathcal{P}^1(X)$, then $d_1(m_0, m_1) = \textbf{val}(\text{OT1})$.

The following particular example will provide an intuitive understanding of our recovery method.

A particular case. Let us assume that the distributions m_0, m_1, and $\bar{\mu}_0$ are empirical distributions with supports of size N, i.e., there exists $(x_i)_{i=1}^N, (\tilde{x}_i)_{i=1}^N \in X^N$ and $(y_i)_{i=1}^N \in \prod_{i=1}^N Z_{x_i}$ such that

$$m_0 = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}, \qquad m_1 = \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{x}_i}, \qquad \bar{\mu}_0 = \frac{1}{N} \sum_{i=1}^N \delta_{(x_i, y_i)}. \tag{3.3.8}$$

Lemma 3.3.10. *Let m_0 and m_1 be defined by* (3.3.8)*. Then problem* (OT1) *has a solution*

$$\rho = \frac{1}{N} \sum_{i=1}^N \delta_{(x_i, x_i')}, \tag{3.3.9}$$

where $\{x_1', \dots, x_N'\}$ is a permutation of $\{\tilde{x}_1, \dots, \tilde{x}_N\}$.

Proof. This is a consequence of [PC19, Prop. 2.1]. □

Let ρ be given by Lemma 3.3.10. By Assumption B(4), for any i, there exists $y'_i \in Z_{x'_i}$ such that

$$\|g(x'_i, y'_i) - g(x_i, y_i)\| \leq L_g d_X(x_i, x'_i). \tag{3.3.10}$$

In our recovery method, each x_i is transformed to x'_i while simultaneously moving y_i to the point $y'_i \in Z_{x'_i}$ for $i = 1, \ldots, N$. This can be expressed as follows:

$$\bar{\mu}_0 = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_i)} \longrightarrow \mu_1 = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x'_i, y'_i)}. \tag{3.3.11}$$

To provide a clearer formula of the construction of μ_1, we introduce the empirical distribution $\nu_N \in \mathcal{P}(Z \times X)$ and the mapping $s_N \colon \{(x_i, y_i, x'_i)\}_{i=1}^N \to \{(x_i, y'_i)\}_{i=1}^N$, defined as:

$$\begin{aligned}
&\nu_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_i, x'_i)}, \\
&s_N(x_i, y_i, x'_i) = (x'_i, y'_i), \quad \forall i = 1, 2, \ldots, N.
\end{aligned}$$

It can be observed that $\pi_{12}\#\nu_N = \bar{\mu}_0$ and $\pi_3\#\nu_N = m_1$. Furthermore, we will demonstrate later in Lemma 3.3.11 that ν_N is a solution of another optimal transport problem (OT2). Then the approximate solution μ_1 of problem (P_{m_1}) can be written as:

$$\mu_1 = s_N \# \nu_N.$$

The distribution μ_1 belongs to $\mathcal{P}_{m_1}(Z)$, and furthermore,

$$\begin{aligned}
\left\| \int_Z g d\bar{\mu}_0 - \int_Z g d\mu_1 \right\| &= \left\| \frac{1}{N}\sum_{i=1}^{N} \big(g(x_i, y_i) - g(x'_i, y'_i)\big) \right\| \\
&\leq \frac{L_g}{N}\sum_{i=1}^{N} \|g(x_i, y_i) - g(x'_i, y'_i)\| \\
&\leq \frac{L_g}{N}\sum_{i=1}^{N} d_X(x_i, x'_i) = L_g d_1(m_0, m_1),
\end{aligned}$$

where the second line follows from the triangle inequality and the third line follows from (3.3.10) and Lemma 3.3.10. The above inequality demonstrates that the distance between the aggregates associated with $\bar{\mu}_0$ and μ_1 is controlled by the d_1-distance of m_0 and m_1.

The general case. To investigate the stability and present the recovery method in the general case, we draw inspiration from the constructions of ν_N and s_N in the previous particular case and introduce the following:

- the auxiliary optimal transport problem:

$$\inf_{\nu \in \Pi(\bar{\mu}_0, m_1)} \int_{Z \times X} d_X(x, x') d\nu(x, y, x'), \tag{OT2}$$

  where $\Pi(\bar{\mu}_0, m_1) \coloneqq \{\nu \in \mathcal{P}(Z \times X) \mid \pi_{12}\#\nu = \bar{\mu}_0,\ \pi_3\#\nu = m_1\}$;

- the set-valued function $S\colon Z \times X \rightsquigarrow Z$,

$$S(x,y,x') = \left\{(x',y') \in Z \ \mid \ \|g(x',y') - g(x,y)\| \leq L_g d_X(x,x')\right\}. \tag{3.3.12}$$

Note that problems (OT2) and (OT1) are similar in so far as the integrand of the cost function is the same, moreover, the second marginal of ν in (OT2) (resp. ρ in (OT1)) must be equal to m_1. The following lemma shows the equivalence between problems (OT1) and (OT2). We will see that the solution of (OT2) will play the role of ν_N in the particular case mentioned earlier.

Lemma 3.3.11. *If m_0 and m_1 lie in $\mathcal{P}^1(X)$, then both problems* (OT1) *and* (OT2) *have solutions, moreover,*

$$\mathbf{val}(\text{OT1}) = \mathbf{val}(\text{OT2}) = d_1(m_0, m_1).$$

Proof. Since $m_0, m_1 \in \mathcal{P}^1(X)$, by [Vil09, Rem. 6.5], we have $d_1(m_0,m_1) = \mathbf{val}(\text{OT1})$. The existence of solutions of problems (OT2) and (OT1) is from [Vil09, Thm. 4.1].

Let ν be a solution to (OT2) and let $\rho = \pi_{13}\#\nu$, which is clearly an element of $\mathcal{P}(X\times X)$. By the basic properties of push-forward measures, we have that $\tilde{\pi}_1\#\rho = \tilde{\pi}_1\#(\pi_{13}\#\nu) = (\tilde{\pi}_1 \circ \pi_{13})\#\nu$. Using the relation $\tilde{\pi}_1 \circ \pi_{13} = \pi_1 \circ \pi_{12}$, we obtain that $(\tilde{\pi}_1\circ\pi_{13})\#\rho = (\pi_1\circ\pi_{12})\#\rho = \pi_1\#(\pi_{12}\#\nu) = \pi_1\#\bar{\mu}_0 = m_0$. It follows that $\tilde{\pi}_1\#\rho = m_0$. By similar arguments, we deduce that $\tilde{\pi}_2\#\rho = m_1$ from the relation $\tilde{\pi}_2\circ\pi_{13} = \pi_3$. Therefore, $\mu \in \Pi(m_0,m_1)$, moreover,

$$\int_{X\times X} d_X(x,x')d\pi_{13}\#\nu(x,x') = \int_{Z\times X} d_X(\pi_{13}(x,y,x'))d\nu(x,y,x') = \mathbf{val}(\text{OT2}).$$

It follows that $d_1(m_0,m_1) \leq \mathbf{val}(\text{OT2})$.

On the other hand, let ρ be a solution of (OT1). Since $\bar{\mu}_0$ and ρ have the same marginal distribution m_0 with respect to their first variable, by the Gluing lemma [Vil09, p. 11], there exists a probability measure $\nu \in \mathcal{P}(X\times Y\times X)$ such that

$$\pi_{12}\#\nu = \bar{\mu}_0, \qquad \pi_{13}\#\nu = \rho.$$

Since $\bar{\mu}_0 \in \mathcal{P}(Z)$, we have $\nu \in \mathcal{P}(Z\times X)$. From the relation $\pi_3 = \tilde{\pi}_2\circ\pi_{13}$, we deduce that $\pi_3\#\nu = \tilde{\pi}_2\#(\pi_{13}\#\nu) = \tilde{\pi}_2\#\rho = m_1$. Thus, $\nu\in\Pi(\bar{\mu}_0,m_1)$, moreover,

$$\begin{aligned}\int_{Z\times X} d_X(x,x')d\nu(x,y,x') = \int_{Z\times X} d_X(\pi_{13}(x,y,x'))d\nu(x,y,x') = \int_{X\times X} d_X(x,x')d\pi_{13}\#\nu(x,x')\\ = \int_{X\times X} d_X(x,x')d\rho(x,x') = d_1(m_0,m_1).\end{aligned}$$

It follows that $d_1(m_0,m_1) \geq \mathbf{val}(\text{OT2})$. □

The following two lemmas demonstrate that the set-valued function S has a measurable selection, which will be denoted by s. We will see that the function s will fulfill the role of s_N in the particular case discussed earlier.

Lemma 3.3.12. *Under Assumption B(3), let $(x_n)_{n\geq 1}$ be a sequence in X converging to some $x_0 \in X$. Then any sequence $(y_n \in Z_{x_n})_{n\geq 1}$ has a convergent sub-sequence with its limit in Z_{x_0}.*

Proof. Since $F\colon X \rightsquigarrow Y$ is upper semi-continuous, for any $k \geq 1$, there exists $\eta_k > 0$ such that for any $x \in B_X(x_0, \eta_k)$, we have $Z_x \subseteq B_Y(Z_{x_0}, 1/k) \coloneqq \cup_{y \in Z_{x_0}} B_Y(y, 1/k)$. For $k = 1$, there exists $\varphi(1) \in \mathbb{N}_+$ such that $Z_{x_{\varphi(1)}} \subseteq B_Y(Z_{x_0}, 1)$, i.e., there exists $\bar{y}_1 \in Z_{x_0}$ such that

$$d_Y(y_{\varphi(1)}, \bar{y}_1) \leq 1.$$

Assume now we have $\varphi(k) \in \mathbb{N}_+$ and $\bar{y}_k \in Z_{x_0}$ for $k = 1, \ldots, K$ such that

$$d_Y(y_{\varphi(k)}, \bar{y}_k) \leq \frac{1}{k}, \qquad k = 1, \ldots, K.$$

Since $x_n \to x_0$, we can find $\varphi(K+1) > \varphi(K)$ such that $x_{\varphi(K+1)} \in B_X(x_0, \eta_{K+1})$. As a consequence, there exists $\bar{y}_{K+1} \in Z_{x_0}$ such that

$$d_Y(y_{\varphi(K+1)}, \bar{y}_{K+1}) \leq \frac{1}{K+1}.$$

Since Z_{x_0} is compact, $(\bar{y}_k)_{k \geq 1}$ has a convergent sub-sequence $(\bar{y}_{\phi(k)})_{k \geq 1}$ with a limit $\bar{y} \in Z_{x_0}$. By the triangle inequality,

$$d_Y(y_{\phi(\varphi(k))}, \bar{y}) \leq d_Y(y_{\phi(\varphi(k))}, \bar{y}_{\phi(\varphi(k))}) + d_Y(\bar{y}_{\phi(\varphi(k))}, \bar{y}) \leq \frac{1}{\phi(\varphi(k))} + d_Y(\bar{y}_{\phi(\varphi(k))}, \bar{y}).$$

Since ϕ and φ are strictly increasing functions going to $+\infty$, we have $\lim_{k \to +\infty} d_Y(y_{\phi(\varphi(k))}, \bar{y}) = 0$. Therefore, $(y_{\phi(\varphi(k))})_{k \geq 1}$ is a convergent sub-sequence of $(y_n)_{n \geq 1}$ with its limit $\bar{y} \in Z_{x_0}$. ☐

Lemma 3.3.13. *Under Assumption B, the set-valued function S has a Borel measurable selection function $s\colon Z \times X \to Z$. Furthermore, we have $\|g(s(x, y, x')) - g(x, y)\| \leq L_g d_X(x, x')$.*

Proof. We will apply Theorem 3.2.3 and Lemma 3.2.4 to prove the result. The images of S are non-empty since the set-valued mapping $\mathcal{Z}$ (defined in (3.3.7)) is supposed to be L_g-Lipschitz. Let us first verify that S has non-empty closed images. Fix any $(x, y, x') \in Z \times X$, and assume that $(x', z_n) \in S(x, y, x')$ converges to some $(x', z) \in Z$. It suffices to prove that $(x', z) \in S(x, y, x')$, i.e., $\|g(x', z) - g(x, y)\| \leq L_g d_X(x, x')$. This is true since g is continuous and $z_n \to z$.

Then, let us show that $S^{-1}(\mathcal{C})$ is closed for any closed subset $\mathcal{C}$ in Z. By (3.3.12), we have

$$\begin{aligned} S^{-1}(\mathcal{C}) &= \left\{(x, y, x') \in Z \times X \;\middle|\; S(x, y, x') \cap \mathcal{C} \neq \emptyset\right\} \\ &= \left\{(x, y, x') \in Z \times X \;\middle|\; \exists\, y' \in Z_{x'} \text{ such that } \begin{cases} (x', y') \in \mathcal{C}, \\ \|g(x', y') - g(x, y)\| \leq L_g d_X(x, x') \end{cases} \right\}. \end{aligned}$$

If $S^{-1}(\mathcal{C}) = \emptyset$, then the conclusion is obvious. Assume that $S^{-1}(\mathcal{C}) \neq \emptyset$ and let $(x_n, y_n, x'_n)_{n \geq 1} \in S^{-1}(\mathcal{C})$ be a convergent sequence with its limit point $(x_0, y_0, x'_0) \in Z \times X$. Then, it suffices to prove that $(x_0, y_0, x'_0) \in S^{-1}(\mathcal{C})$. Since $(x_n, y_n, x'_n) \in S^{-1}(\mathcal{C})$, there exists $y'_n \in Z_{x'_n}$, for any n, such that

$$(x'_n, y'_n) \in \mathcal{C}, \qquad \|g(x'_n, y'_n) - g(x_n, y_n)\| \leq L_g d_X(x_n, x'_n).$$

By Lemma 3.3.12, the sequence $(y'_n)_{n \geq 1}$ has a convergent sub-sequence $(y'_{\varphi(n)})_{n \geq 1}$ with its limit $y'_0 \in Z_{x'_0}$. Hence, $\lim_{n \to \infty}(x'_{\varphi(n)}, y'_{\varphi(n)}) = (x'_0, y'_0)$. Since $\mathcal{C}$ is closed, we have $(x'_0, y'_0) \in \mathcal{C}$. By the triangle inequality,

$$\|g(x_0', y_0') - g(x_0, y_0)\| \leq \|g(x_0', y_0') - g(x'_{\varphi(n)}, y'_{\varphi(n)})\| + \|g(x'_{\varphi(n)}, y'_{\varphi(n)}) - g(x_{\varphi(n)}, y_{\varphi(n)})\| \\ + \|g(x_{\varphi(n)}, y_{\varphi(n)}) - g(x_0, y_0)\|.$$

By the continuity of g, we have

$$\begin{aligned} \|g(x_0', y_0') - g(x_0, y_0)\| &\leq \limsup_{n\to\infty} \|g(x'_{\varphi(n)}, y'_{\varphi(n)}) - g(x_{\varphi(n)}, y_{\varphi(n)})\| \\ &\leq \limsup_{n\to\infty} L_g d_X(x_{\varphi(n)}, x'_{\varphi(n)}) = L_g d_X(x_0, x_0'). \end{aligned}$$

Therefore, $y_0' \in Z_{x_0'}$, $(x_0', y_0') \in \mathcal{C}$ and $\|g(x_0', y_0') - g(x_0, y_0)\| \leq L_g d_X(x_0, x_0')$. It follows that $(x_0, y_0, x_0') \in S^{-1}(\mathcal{C})$, which implies that $S^{-1}(\mathcal{C})$ is closed, thus a Borel set.

Lemma 3.2.4 shows that the set-valued function S is Borel measurable, and Theorem 3.2.3 shows the existence of a Borel measurable selection s of S. Since $s(x, y, x') \in S(x, y, x')$, the inequality $\|g(s(x, y, x')) - g(x, y)\| \leq L_g d_X(x, x')$ is a direct consequence of (3.3.12). □

Lemma 3.3.14. *Let Assumptions A-B hold true. Let $\nu \in \Pi(\bar{\mu}_0, m_1)$ and s be the Borel measurable selection of S obtained Lemma 3.3.13. Let $\mu_1 = s\#\nu$. Then $\mu_1 \in \mathcal{P}_{m_1}(Z)$ and*

$$f\left(\int_Z g d\mu_1\right) - f\left(\int_Z g d\bar{\mu}_0\right) \leq L_g(C + LM) \int_{Z\times X} d_X(x, x') d\nu(x, y, x').$$

Proof. **Step 1**. (Properties of μ_1). Since $\nu \in \mathcal{P}(Z \times S)$ and $s \colon Z \times S \to Z$ is a Borel measurable function, we have $\mu = s\#\nu \in \mathcal{P}(Z)$. Observing that $\pi_1 \circ s = \pi_3$, it follows that $\pi_1\#\mu = (\pi_1 \circ s)\#\nu = \pi_3\#\nu = m_1$. Thus, $\mu_1 \in \mathcal{P}_{m_1}(Z)$. For any $\lambda \in \mathcal{H}$, $g_\lambda \colon Z \to \mathbb{R}$ is bounded. Then,

$$\int_Z g_\lambda d\mu_1 = \int_{Z\times X} g_\lambda \circ s \, d\nu. \tag{3.3.13}$$

Step 2. (Quadratic upper bound). By the Lipschitz continuity of ∇f, we have

$$f\left(\int_Z g d\mu_1\right) - f\left(\int_Z g d\bar{\mu}_0\right) \leq \int_Z g_{\lambda_0}(d\mu_1 - d\bar{\mu}_0) + \frac{L}{2}\left\|\int_Z g d\mu_1 - \int_Z g d\bar{\mu}_0\right\|^2, \tag{3.3.14}$$

where $\lambda_0 = \nabla f(\int_Z g d\bar{\mu}_0)$.

Step 3. (First-order estimate). Let us study the first-order term in (3.3.14). By (3.3.13), we have

$$\int_Z g_{\lambda_0} d\mu_1 = \int_{Z\times X} g_{\lambda_0} \circ s d\nu = \int_{Z\times X} \langle \lambda_0 \, , \, g(s(x, y, x'))\rangle d\nu(x, y, x').$$

From the relation $\pi_{12}\#\nu = \bar{\mu}_0$, we deduce that

$$\int_Z g_{\lambda_0} d\bar{\mu}_0 = \int_{Z\times X} \langle \lambda_0 \, , \, g(x, y)\rangle d\nu(x, y, x').$$

Using the previous two equalities, Lemma 3.3.13, and the Cauchy–Schwarz inequality, we obtain that

$$\int_Z g_{\lambda_0}(d\mu_1 - d\bar{\mu}_0) \leq L_g \|\lambda_0\| \int_{Z\times X} d_X(x, x') d\nu(x, y, x').$$

Step 4. (Second-order estimate). Let us study the second-order term in (3.3.14). Developing it and using Lemma 3.2.6, we obtain that

$$\left\| \int_Z g d\mu_1 - \int_Z g d\bar{\mu}_0 \right\|^2 = \gamma_1 + \gamma_2, \tag{3.3.15}$$

where

$$\gamma_1 = \int_Z \int_Z \langle g(z_1), g(z_2) \rangle \, d(\mu_1 - \bar{\mu}_0)(z_1) \, d\mu_1(z_2);$$
$$\gamma_2 = \int_Z \int_Z \langle g(z_1), g(z_2) \rangle \, d(\bar{\mu}_0 - \mu_1)(z_1) \, d\bar{\mu}_0(z_2).$$

Fix any $z_2 \in Z$. Following the same argument as in step 3, we have

$$\left\| \int_Z \langle g(z_1), g(z_2) \rangle \, d(\mu_1 - \bar{\mu}_0)(z_1) \right\| \leq L_g \|g(z_2)\| \int_{Z \times X} d_X(x, x') d\nu(x, y, x').$$

It follows that

$$\gamma_1 + \gamma_2 \leq 2 L_g M \int_{Z \times X} d_X(x, x') d\nu(x, y, x').$$

Step 5. As a consequence of Steps 2-4, we deduce that

$$f\left(\int_Z g d\mu_1 \right) - f\left(\int_Z g d\bar{\mu}_0 \right) \leq L_g(\|\lambda_0\| + LM) \int_{Z \times X} d_X(x, x') d\nu(x, y, x').$$

By the definition of the constant C, we have $C \geq \|\lambda_0\|$. The conclusion follows. □

We can now state the recovery algorithm that enables us to obtain an approximate solution of (P_{m_1}), given an approximate solution $\bar{\mu}_0$ of (P_{m_0}).

Algorithm 3.1: Recovery method

Input: $m_0, m_1 \in \mathcal{P}^1(Z)$, and $\bar{\mu}_0 \in \mathcal{P}_{m_0}(Z)$.
Step 1. Find a solution ρ of the optimal transport problem (OT1).
Step 2. Find $\nu \in \mathcal{P}(Z \times X)$ such that $\pi_{12}\#\nu = \bar{\mu}_0$ and $\pi_{13}\#\nu = \rho$.
Step 3. Set $\mu_1 = s\#\nu \in \mathcal{P}_{m_1}(Z)$, where s is constructed in Lemma 3.3.13.
Output: μ_1.

Remark 3.3.15. We have already discussed the case where m_0, m_1, and $\bar{\mu}_0$ are empirical distributions. We discuss now the slightly more general case where only m_0 and $\bar{\mu}_0$ are empirical distributions:

$$m_0 = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}, \qquad \bar{\mu}_0 = \frac{1}{N} \sum_{i=1}^N \delta_{(x_i, y_i)}.$$

This situation corresponds to the algorithm presented in Section 3.5. Since $\rho \in \mathcal{P}_{m_0}(X \times X)$, by Lemma 3.5.5, we have $\rho = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \otimes \rho_{x_i}$, where ρ_{x_i} is defined in Theorem 3.2.7. Then the probability distribution μ, obtained in general with the Gluing lemma, is given here in an explicit form:

$$\nu = \frac{1}{N} \sum_{i=1}^N \delta_{(x_i, y_i)} \otimes \rho_{x_i}.$$

Theorem 3.3.16. *Let Assumptions A-B hold true. Assume that $m_0, m_1 \in \mathcal{P}^1(Z)$ and that $\bar{\mu}_0$ is an ϵ_0-minimizer of problem (P_{m_0}) for some $\epsilon_0 \geq 0$. The following holds true.*

1. $|\mathbf{val}(\mathrm{P}_{m_0}) - \mathbf{val}(\mathrm{P}_{m_1})| \leq L_g(C + LM)d_1(m_0, m_1)$*;*
2. *If $\mu_1 \in \mathcal{P}_{m_1}(Z)$ is the output of Algorithm 3.1, then μ_1 is an η-minimizer of problem (P_{m_1}), where*
$$\eta = \epsilon_0 + 2L_g(C + LM)d_1(m_0, m_1).$$

Proof. We prove (1). Fix any $\epsilon > 0$. Let μ_0^ϵ be an ϵ-minimizer of problem (P_{m_0}). By Lemma 3.3.11, there exists $\nu^\epsilon \in \Pi(\mu_0^\epsilon, m_1)$ such that

$$\int_{Z \times X} d_X(x, x') d\nu(x, y, x') = d_1(m_0, m_1).$$

We deduce from Lemma 3.3.14 that there exists $\mu_1^\epsilon \in \mathcal{P}_{m_1}(Z)$ such that

$$f\left(\int_z g\mu_1^\epsilon\right) - f\left(\int_Z g d\mu_0^\epsilon\right) \leq L_g(C + LM)d_1(m_0, m_1).$$

Since $\mu_1^\epsilon \in \mathcal{P}_{m_1}(Z)$, $\mathbf{val}(\mathrm{P}_{m_1}) \leq f\left(\int_z g\mu_1\right)$. Combining this with the fact $\mathbf{val}(\mathrm{P}_{m_0}) \geq f\left(\int_Z g d\mu_0^\epsilon\right) - \epsilon$, we obtain that

$$\mathbf{val}(\mathrm{P}_{m_1}) - \mathbf{val}(\mathrm{P}_{m_0}) \leq f\left(\int_z g\mu_1^\epsilon\right) - f\left(\int_Z g d\mu_0^\epsilon\right) + \epsilon.$$

Therefore, $\mathbf{val}(\mathrm{P}_{m_1}) - \mathbf{val}(\mathrm{P}_{m_0}) \leq L_g(C + LM)d_1(m_0, m_1)$ by the arbitrariness of ϵ. We conclude the first part of the proof by exchanging the positions of m_0 and m_1.

Let μ_1 be the output of Algorithm 3.1. To prove (2), we do the following decomposition:

$$f\left(\int g d\mu_1\right) - \mathbf{val}(\mathrm{P}_{m_1}) = \gamma_1 + \gamma_2 + \gamma_3,$$

where

$$\gamma_1 = f\left(\int g d\mu_1\right) - f\left(\int g d\bar{\mu}_0\right), \quad \gamma_2 = f\left(\int g d\bar{\mu}_0\right) - \mathbf{val}(\mathrm{P}_{m_0}), \quad \gamma_3 = \mathbf{val}(\mathrm{P}_{m_0}) - \mathbf{val}(\mathrm{P}_{m_1}).$$

From the proof of Lemma 3.3.11, we know that ν (the result of step 2 in Algorithm 3.1) is a solution of (OT2). Then, Lemma 3.3.14 shows that $\gamma_1 \leq L_g(C + LM)d_1(m_0, m_1)$. Since $\bar{\mu}_0$ is an ϵ-minimizer, $\gamma_2 \leq \epsilon$. By point (1), $\gamma_3 \leq L_g(C + LM)d_1(m_0, m_1)$. The conclusion follows. □

3.4 Dual problem

This section is dedicated to the duality analysis of the primal problem (P_m). In the sequel of this section, let Assumptions $\mathbf{A}^*$ and B hold true. Consider the equivalent formulation of problem of (P_m),

$$\inf_{\mu \in \mathcal{P}_m(Z),\, \beta \in \mathcal{H}} f(\beta), \qquad \text{s.t. } \beta = \int_Z g d\mu. \tag{$\tilde{\mathrm{P}}_m$}$$

The *Lagrangian* $\mathcal{L}\colon \mathcal{H}^2 \times \mathcal{P}_m(Z) \to \mathbb{R}$ associated with $(\tilde{\mathrm{P}}_m)$ writes,

$$\mathcal{L}(\lambda, \beta, \mu) = f(\beta) + \left\langle \lambda, \int_Z g d\mu - \beta \right\rangle.$$

Then, the dual problem of $(\tilde{\mathrm{P}}_m)$ is,

$$\sup_{\lambda \in \mathcal{H}} \inf_{\beta \in \mathcal{H}\,,\, \mu \in \mathcal{P}_m(Z)} \mathcal{L}(\lambda, \beta, \mu) = \sup_{\lambda \in \mathcal{H}} \Big(- f^*(\lambda) + \inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \langle \lambda\,,\, g(z) \rangle \, d\mu(z) \Big), \tag{3.4.1}$$

where f^* is the Fenchel conjugate of f. For any $\lambda \in \mathcal{H}$, since g is bounded over Z, the second term $\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \langle \lambda\,,\, g(z) \rangle \, d\mu(z)$ is finite. Therefore, it suffices to study (3.4.1) for $\lambda \in \mathrm{dom}(f^*)$, i.e.,

$$\sup_{\lambda \in \mathrm{dom}(f^*)} \Big(- f^*(\lambda) + \inf_{\mu \in \mathcal{P}_m(Z)} \int_Z g_\lambda d\mu \Big).$$

The result of Lemma 3.3.3 holds true for all $\lambda \in \mathrm{dom}(f^*)$ under Assumption $\mathbf{A}^*$. Applying it to the previous problem, we obtain the following equivalent dual problem:

$$- \inf_{\lambda \in \mathrm{dom}(f^*)} \mathcal{D}_m(\lambda) \coloneqq f^*(\lambda) - \int_X u_\lambda dm. \tag{D_m}$$

Lemma 3.4.1. *The function $\mathcal{D}_m$ is strongly convex with modulus $1/L$. As a consequence, problem (D_m) has a unique solution, denoted by $\lambda^*(m)$. Moreover, there exists a constant C^* independent of m such that*

$$\|\lambda^*(m)\| \leq C^*.$$

Proof. Since ∇f is L-Lipschitz continuous, we know that f^* is strongly convex with modulus $1/(2L)$ (i.e. $f^* - 1/L\|\cdot\|^2$ is convex) (see [BC11, Thm. 18.15]). Let us consider $u_\lambda(x)$ as a function of λ while fixing any $x \in X$. By definition, $\lambda \mapsto u_\lambda(x)$ is the infimum of a family of affine functions (with respect to λ), thus it is concave with respect to λ. Consequently, $-\int_X u_\lambda dm$ is convex with respect to λ. Therefore, $\mathcal{D}_m$ is $1/L$-strongly convex. Additionally, $\mathrm{dom}(f^*)$ is both convex and closed. These properties guarantee the existence and uniqueness of the minimizer $\lambda^*(m)$.

Since M is an upper bound of $\|g(z)\|$, it follows that for all $\lambda \in \mathcal{H}$:

$$-M\|\lambda\| \leq \inf_{y \in Z_x} -\|\lambda\|\|g(x,y)\| \leq u_\lambda(x) \leq \sup_{y \in Z_x} \|\lambda\|\|g(x,y)\| \leq M\|\lambda\|.$$

Let $\lambda_0 \in \mathrm{dom}(f^*)$. As $\mathcal{D}_m(\lambda^*(m)) \leq \mathcal{D}_m(\lambda_0)$, we can derive the following inequalities:

$$f^*(\lambda_0) + M\|\lambda_0\| \geq \mathcal{D}_m(\lambda_0) \geq \mathcal{D}_m(\lambda^*(m)) \geq f^*(\lambda^*(m)) - M\|\lambda^*(m)\|.$$

The strong convexity of f^* yields that

$$\frac{1}{2L}\|\lambda^*(m) - \lambda_0\|^2 + \langle p_0, \lambda^*(m) - \lambda_0 \rangle \leq f^*(\lambda^*(m)) - f^*(\lambda_0),$$

where $p_0 \in \partial f^*(\lambda_0)$. Combining the two above inequalities, we obtain:

$$\frac{1}{2L}\|\lambda^*(m) - \lambda_0\|^2 + \langle p_0, \lambda^*(m) - \lambda_0 \rangle \leq M(\|\lambda^*(m)\| + \|\lambda_0\|).$$

where $p_0 \in \partial f^*(\lambda_0)$. The announced result follows, with $C^* = 3\|\lambda_0\| + 2L(M + \|p_0\|)$. $\square$

3.4.1 Strong duality

Let us now prove the strong duality principle between (P_m) and (D_m), i.e., $\mathbf{val}(\mathrm{P}_m) = \mathbf{val}(\mathrm{D}_m)$. We will apply the Fenchel-Rockafellar theorem [Roc97] to prove this relation.

Proposition 3.4.2. *Assume that the set* $G_m := \{\int_\mu g d\mu \mid \mu \in \mathcal{P}_m(Z)\} \subseteq \mathcal{H}$ *is closed. Then,*

1. $\mathbf{val}(\mathrm{P}_m) = \mathbf{val}(\mathrm{D}_m)$*;*
2. *the primal problem* (P_m) *has a solution;*
3. *let* $\lambda^*(m)$ *be the solution of* (D_m) *and let* μ *be a solution of* (P_m)*, then*

$$\lambda^*(m) = \nabla f\left(\int_Z g d\mu\right).$$

Proof. Let us consider the following optimization problem with variable in $\mathcal{H}$:

$$\inf_{z\in\mathcal{H}} f(z) + \chi_{G_m}(z). \tag{3.4.2}$$

It is obvious that $\mathbf{val}(\mathrm{P}_m) = \mathbf{val}(\tilde{\mathrm{P}}_m) = \mathbf{val}(3.4.2)$. The dual problem of (3.4.2) writes

$$\sup_{\lambda\in\mathcal{H}} -f^*(\lambda) - \chi^*_{G_m}(-\lambda). \tag{3.4.3}$$

By the definition of the Fenchel conjugate and the definition of G_m, we have

$$-\chi^*_{G_m}(-\lambda) = \inf_{z\in G_m} \langle \lambda, z\rangle = \inf_{\mu\in\mathcal{P}_m(Z)} \left\langle \lambda, \int_Z g d\mu \right\rangle.$$

Therefore, $\mathbf{val}(\mathrm{D}_m) = \mathbf{val}(3.4.1) = \mathbf{val}(3.4.3)$. Let us apply the Fenchel-Rockafellar theorem to (3.4.2). The function f is convex and continuous. The function χ_{G_m} is convex and lower semi-continuous from the fact that G_m is convex and closed. It is obvious that G_m is non-empty. Therefore, $0 \in \mathrm{int}(\mathcal{H} - G_m) = \mathrm{int}(\mathrm{dom} f - \mathrm{dom}\chi_{G_m})$. By the Fenchel-Rockafellar theorem [Roc97], $\mathbf{val}(3.4.3) = \mathbf{val}(3.4.2)$, thus, $\mathbf{val}(\mathrm{P}_m) = \mathbf{val}(\mathrm{D}_m)$.

Since G_m is non-empty, bounded, convex, and closed, and since f is continuous and convex, we deduce from [Bré11, Cor. 3.23] that problem (3.4.2) has a solution. Therefore, (P_m) has a solution, denoted by μ. Since $\lambda^*(m)$ is the solution of (D_m), by the strong duality,

$$-f^*(\lambda^*(m)) + \inf_{z\in G_m} \langle \lambda^*(m), z\rangle = f\left(\int_Z g d\mu\right).$$

On the other hand, by the definition of Fenchel's conjugate,

$$f\left(\int_Z g d\mu\right) + f^*(\lambda^*(m)) \leq \left\langle \lambda^*(m), \int_Z g d\mu \right\rangle.$$

Combining the previous two inequalities and the fact that $\int_Z g d\mu \in G_m$, we deduce that

$$f\left(\int_Z g d\mu\right) + f^*(\lambda^*(m)) = \left\langle \lambda^*(m), \int_Z g d\mu \right\rangle.$$

We obtain that $\lambda^*(m) = \nabla f(\int_Z g d\mu)$ from Fenchel's relation. □

3.4.2 Stability of the dual problem

Lemma 3.4.3. *For any $\lambda_1, \lambda_2 \in \mathcal{H}$ and $x_1, x_2 \in X$, it holds that*

$$|u_{\lambda_1}(x_1) - u_{\lambda_2}(x_2)| \leq L_g \|\lambda_1\| d_X(x_1, x_2) + M\|\lambda_1 - \lambda_2\|.$$

Proof. By the triangle ineqaulity,

$$|u_{\lambda_1}(x_1) - u_{\lambda_2}(x_2)| \leq |u_{\lambda_1}(x_1) - u_{\lambda_1}(x_2)| + |u_{\lambda_1}(x_2) - u_{\lambda_2}(x_2)|.$$

By the definition of u_λ, we have

$$u_{\lambda_1}(x_1) - u_{\lambda_1}(x_2) = \inf_{y_1 \in Z_{x_1}} \langle \lambda_1 \,,\, g(x_1, y_1) \rangle - \inf_{y_2 \in Z_{x_2}} \langle \lambda_1 \,,\, g(x_2, y_2) \rangle.$$

Let $\tilde{y}_2^\epsilon$ be an ϵ-minimizer of $\inf_{y_2 \in Z_{x_2}} \langle \lambda_1 \,,\, g(x_2, y_2) \rangle$, with $\epsilon > 0$. By the Lipschitz continuity of $\mathcal{Z}$, there exists $\tilde{y}_1^\epsilon \in Z_{x_1}$ such that

$$\|g(x_1, \tilde{y}_1^\epsilon) - g(x_2, \tilde{y}_2^\epsilon)\| \leq L_g d_X(x_1, x_2).$$

By the Cauchy-Schwarz inequality, we have

$$u_{\lambda_1}(x_1) - u_{\lambda_1}(x_2) \leq \langle \lambda_1 \,,\, g(x_1, \tilde{y}_1^\epsilon) - g(x_2, \tilde{y}_2^\epsilon) \rangle + \epsilon \leq \|\lambda_1\| L_g d_X(x_1, x_2) + \epsilon.$$

By the arbitrariness of ϵ, we have $|u_{\lambda_1}(x_1) - u_{\lambda_1}(x_2)| \leq \|\lambda_1\| L_g d_X(x_1, x_2)$.

On the other hand,

$$u_{\lambda_1}(x_2) - u_{\lambda_2}(x_2) = \inf_{y \in Z_{x_2}} \langle \lambda_1 \,,\, g(x_2, y) \rangle - \inf_{y \in Z_{x_2}} \langle \lambda_2 \,,\, g(x_2, y) \rangle \leq \sup_{y \in Z_{x_2}} \langle \lambda_1 - \lambda_2 \,,\, g(x_2, y) \rangle.$$

By the Cauchy–Schwarz inequality and the definition of M, we have that

$$u_{\lambda_1}(x_2) - u_{\lambda_2}(x_2) \leq M\|\lambda_1 - \lambda_2\|.$$

The conclusion follows. □

Lemma 3.4.4 (Stability of the dual problem)**.** *For any m_0, $m_1 \in \mathcal{P}(\Omega)$, we have*

$$|\mathcal{D}_{m_0}(\lambda^*(m_0)) - \mathcal{D}_{m_1}(\lambda^*(m_1))| \leq C^* L_g \, d_1(m_0, m_1), \tag{3.4.4}$$

$$\|\lambda^*(m_0) - \lambda^*(m_1)\|^2 \leq 2C^* L_g L \, d_1(m_0, m_1), \tag{3.4.5}$$

where C^* is the a priori bound of $\|\lambda^*(\cdot)\|$ obtained in Lemma 3.4.1

Proof. According to Lemma 3.4.1, we know that $\|\lambda^*(m_0)\|$ and $\|\lambda^*(m_1)\|$ are smaller than C^*. Then, by Lemma 3.4.3, $u_{\lambda^*(m_0)}(x)$ and $u_{\lambda^*(m_1)}(x)$ are $(C^* L_g)$-Lipschitz continuous with respect to x. Hence,

$$\begin{aligned} \mathcal{D}_{m_0}(\lambda^*(m_0)) &= f^*(\lambda^*(m_0)) - \int_X u_{\lambda^*(m_0)}(x) dm_0(x) \\ &= f^*(\lambda^*(m_0)) - \int_X u_{\lambda^*(m_0)}(x) dm_1(x) + \int_X u_{\lambda^*(m_0)}(x) d(m_1 - m_0)(x) \\ &\geq \mathcal{D}_{m_1}(\lambda^*(m_0)) - C^* L_g d_1(m_0, m_1), \end{aligned} \tag{3.4.6}$$

where the third line is by the definition of the Kantorovich–Rubinstein distance. Since $\lambda^*(m_1)$ minimizes $\mathcal{D}_{m_1}$ and since $\mathcal{D}_{m_1}$ is $1/L$-strongly convex, we have

$$\mathcal{D}_{m_1}(\lambda^*(m_0)) \geq \mathcal{D}_{m_1}(\lambda^*(m_1)) + \frac{1}{2L}\|\lambda^*(m_0) - \lambda^*(m_1)\|^2. \tag{3.4.7}$$

Combining (3.4.6) and (3.4.7), we obtain that

$$\mathcal{D}_{m_0}(\lambda^*(m_0)) \geq \mathcal{D}_{m_1}(\lambda^*(m_1)) + \frac{1}{2L}\|\lambda^*(m_0) - \lambda^*(m_1)\|^2 - C^* L_g d_1(m_0, m_1).$$

In particular, we have $\mathcal{D}_{m_1}(\lambda^*(m_1)) - \mathcal{D}_{m_0}(\lambda^*(m_0)) \leq C^* L_g d_1(m_0, m_1)$. Exchanging the positions of m_0 and m_1 in (3.4.6), we obtain

$$\mathcal{D}_{m_1}(\lambda^*(m_1)) \geq \mathcal{D}_{m_0}(\lambda^*(m_0)) + \frac{1}{2L}\|\lambda^*(m_0) - \lambda^*(m_1)\|^2 - C^* L_g d_1(m_0, m_1). \tag{3.4.8}$$

Inequality (3.4.4) follows immediately and (3.4.5) is deduced by summing (3.4.6)-(3.4.8). □

3.4.3 Directional derivative of the value function

The value function of problem (P_m) is defined by

$$V\colon \mathcal{P}^1(X) \to \mathbb{R},\ m \mapsto \mathbf{val}(\mathrm{P}_m).$$

Our goal is to characterize the directional derivative of V. Define the following function:

$$v\colon \mathcal{P}^1(X) \times X \to \mathbb{R},\ (m, x) \mapsto u_{\lambda^*(m)}(x).$$

Proposition 3.4.5. *Assume that G_m is closed for any $m \in \mathcal{P}^1(X)$. Then for any $m_0, m_1 \in \mathcal{P}^1(X)$, we have*

$$\lim_{t\to 0^+} \frac{V(m_0 + t(m_1 - m_0)) - V(m_0)}{t} = \int_X v(m_0, x) d(m_1 - m_0)(x).$$

As a consequence, v is the directional derivative of V, i.e.,

$$V(m_1) - V(m_0) = \int_{t=0}^1 \int_X v(m_0 + t(m_1 - m_0), x) d(m_1 - m_0)(x) dt.$$

Proof. For any $t \in [0, 1]$, let $m_t = m_0 + t(m_1 - m_0)$. By the strong duality, we have

$$V(m_t) - V(m_0) = \mathcal{D}_{m_0}(\lambda^*(m_0)) - \mathcal{D}_{m_t}(\lambda^*(m_t)).$$

From (3.4.6), we deduce that

$$\mathcal{D}_{m_0}(\lambda^*(m_0)) - \mathcal{D}_{m_t}(\lambda^*(m_t)) \geq \int_X v(m_0, x) d(m_t - m_0)(x) = t\int_X v(m_0, x) d(m_1 - m_0)(x). \tag{3.4.9}$$

On the other hand, let μ_0 and μ_1 be solutions of (P_m) with $m = m_0$ and m_1 respectively. Let $\mu_t = \mu_0 + t(\mu_1 - \mu_0)$. It is obvious that $\pi_1\#\mu_t = m_t$. Therefore,

$$V(m_t) - V(m_0) \leq f\left(\int_Z g d\mu_t\right) - f\left(\int_Z g d\mu_0\right).$$

By Proposition 3.4.2, $\lambda^*(m_0) = \nabla f(\int_Z g d\mu_0)$. Since ∇f is L-Lipschitz, it follows that

$$f\left(\int_Z g d\mu_t\right) - f\left(\int_Z g d\mu_0\right) \leq t \int_X v(m_0, x) d(m_1 - m_0)(x) + \frac{Lt^2}{2} \left\| \int_Z g d(\mu_1 - \mu_0) \right\|^2.$$

Recall the definition of D. Combining the two inequalities above, we have

$$V(m_t) - V(m_0) \leq t \int_X v(m_0, x) d(m_1 - m_0)(x) + \frac{LDt^2}{2}. \tag{3.4.10}$$

We have $V(m_0) = -\mathcal{D}_{m_0}(\lambda^*(m_0))$ and $V(m_t) = -\mathcal{D}_{m_t}(\lambda^*(m_t))$. Using (3.4.9)-(3.4.10) and letting t go to 0^+, we obtain that

$$\lim_{t \to 0^+} \frac{V(m_t) - V(m_0)}{t} = \int_X v(m_0, x) d(m_1 - m_0)(x).$$

From Lemmas 3.4.3-3.4.4, we deduce that the function $v(m, x)$ is continuous in $\mathcal{P}^1(X) \times X$ with respect to the distance (d_1, d_X). Let us define two functions from $[0,1]$ to $\mathbb{R}$,

$$\begin{aligned} &\bar{V} \colon [0,1] \to \mathbb{R},\, t \mapsto V(m_t); \\ &\bar{v} \colon [0,1] \to \mathbb{R},\, t \mapsto \int_X v(m_t, x) d(m_1 - m_0)(x). \end{aligned}$$

For any $0 \leq t \leq T < 1$, observe that $m_T = m_t + \frac{T-t}{1-t}(m_1 - m_t)$ and $m_1 - m_t = (1-t)(m_1 - m_0)$. By using the same arguments as in (3.4.9)-(3.4.10), we have

$$(T-t)\bar{v}(t) \leq \bar{V}(T) - \bar{V}(t) \leq (T-t)\bar{v}(t) + \frac{LD(T-t)^2}{2(1-t)^2}.$$

We deduce that $\bar{v}(t)$ is the right derivative of $\bar{V}$ at t for any $t \in [0,1)$. By exchanging positions of m_0 and m_1, we can prove that $\bar{v}(t)$ is the left derivative of $\bar{V}$ at t for any $t \in (0,1]$. Therefore, $\bar{V}$ is differentiable at each point on $[0,1]$ and $\bar{v}$ is its derivative. Since $\bar{v}$ is continuous, by the fundamental theorem of calculus [Rud87, Thm. 7.21], we have that $\bar{V}(1) - \bar{V}(0) = \int_{t=0}^1 \bar{v}(t) dt$. □

3.5 Algorithms for the discretized MFO problem

We present in this section our numerical method for solving (P_m). The first step of resolution consists in discretizing m. We replace it by an empirical distribution m_N and focus next on the resolution of (P_{m_N}). By Theorem 3.3.16(1), we have

$$|\mathbf{val}(\mathrm{P}_m) - \mathbf{val}(\mathrm{P}_{m_N})| \leq L_g(C + LM) d_1(m, m_N). \tag{3.5.1}$$

We give theoretical bounds for the minimal value of $d_1(m, m_N)$ in Subsection 3.5.1. Then we discuss the resolution of (P_{m_N}) with the Frank-Wolfe algorithm in Subsection 3.5.2. Finally in Subsection 3.5.3 we propose to use a variant of the Frank-Wolfe algorithm, called Stochastic Frank-Wolfe (SFW) algorithm, introduced in [BLO+22]. This method generates a solution to (P_{m_N}) which is an empirical distribution.

3.5.1 Discretization

In view of (3.5.1), one should look for an empirical distribution $m_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$ that is as close as possible to m for the d_1-distance. This problem is commonly known as the optimal quantization problem, and for detailed information on this topic, we refer to [GG12]. Here, we present a slightly modified version of an optimal quantization result obtained in [MM16, Prop. 12]. For any subset A of X, we denote by $r_N(A)$ the minimum radius r required to cover A with N closed balls of radius r. It is defined by

$$r_N(A) := \inf_{x\in A^N} \min \left\{ r \geq 0 \,\middle|\, A \subseteq \bigcup_{i=1}^{N} B_X(x_i, r) \right\}.$$

The upper box-counting dimension (or the upper Minkowski dimension) of A [Fal04, p. 41] is defined as follows:

$$\bar{D}(A) := \inf \left\{ \bar{D} > 0 \mid \exists \bar{C} > 0 \text{ such that } r_N(A) \leq \bar{C} N^{-1/\bar{D}},\ \forall N \in \mathbb{N}_+ \right\}.$$

Lemma 3.5.1. *Let $m \in \mathcal{P}^1(X)$, and let $A \subseteq X$ be the support of m. There exists a sequence $(m_N)_{N\geq 1}$ of empirical distributions on X such that the following holds:*

1. *If $\bar{D}(A) > 1$, then there exists a constant $\tilde{C}_1$ such that for any $N \geq 1$,*
$$d_1(m, m_N) \leq \tilde{C}_1 N^{-\frac{1}{\bar{D}(A)}}.$$
2. *If $\bar{D}(A) = 1$, then there exists a constant $\tilde{C}_2$ such that for any $N \geq 1$,*
$$d_1(m, m_N) \leq \tilde{C}_2 N^{-1} \log N.$$
3. *If $\bar{D}(A) < 1$, then there exists a constant $\tilde{C}_3$ such that for any $N \geq 1$,*
$$d_1(m, m_N) \leq \tilde{C}_3 N^{-1}.$$

Proof. This follows from the proof presented in [MM16, Prop. 12], with the only difference being that in the final inequality, we employ the triangle inequality for the d_1-distance instead of the Minkowski inequality for the Wasserstein-2 distance. □

Remark 3.5.2. If A is a subset of a smooth d-dimensional submanifold of a Euclidean space, then $\bar{D}(A) \leq d$. This estimate is deduced from [Fal04, p. 48 (i)-(ii)].

3.5.2 Frank-Wolfe algorithm

For general convex optimization problems, the Frank-Wolfe algorithm relies on the resolution of a sequence of linearized problems, obtained by replacing the cost function of the problem by a first-order Taylor approximation of it. In the context of problem (P_{m_N}), the linearized problem is of the general form:

$$\inf_{\mu\in\mathcal{P}_{m_N}(Z)} \left\langle \lambda, \int_Z g d\mu \right\rangle, \tag{3.5.2}$$

for some $\lambda \in \nabla f(\mathcal{H})$.

A key observation from Lemma 3.3.3 is that a solution of the linearized problem, denoted by μ_λ, can be solved as in the proof of Lemma 3.3.3, in the simple case where m is a finitely-supported probability measure: for all $i = 1, \dots, N$, find $y_i \in \mathbf{BR}_\lambda(x_i)$ and set $\mu_\lambda = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i,y_i)}$. Therefore, one can consider applying the Frank-Wolfe algorithm to solve (P_{m_N}), in which the main task is to solve (3.5.2).

Algorithm 3.2: Frank-Wolfe Algorithm

Initialization: $\mu^0 \in \mathcal{P}_{m_N}(Z)$. Set $K \geq 1$.
for $k = 0, 1, 2, \dots, K-1$ **do**
  Compute $\lambda^k = \nabla f\left(\int_Z g d\mu^k\right)$.
  Solve (3.5.2) for $\lambda = \lambda^k$, the solution is denoted by μ_{λ^k}.
  Choose $\omega_k \in [0, 1]$.
  Set $\mu^{k+1} = (1-\omega_k)\mu^k + \omega_k \mu_{\lambda^k}$.
end for

Remark 3.5.3. If we take $\omega_k = 1/(k+1)$ for all k, then it is easy to see that $\mu^K = \frac{1}{K}\sum_{k=0}^{K-1}\mu_{\lambda^k}$. We recover the fictitious play algorithm in [CH17] applied to the Lagrangian discretization of first-order MFGs.

Lemma 3.5.4. *Let Assumption A hold true. In Algorithm 3.2, we set $\omega_k = 2/(k+2)$ for all k. Then for any $K \geq 1$,*

$$f\left(\int_Z g d\mu^K\right) - \mathbf{val}(\mathrm{P}_{m_N}) \leq \frac{2LD}{K}.$$

Proof. This is a consequence of [BLO+22, Prop. 3.4]. □

3.5.3 Stochastic Frank-Wolfe algorithm

In Algorithm 3.2, at each iteration, we generate the output by taking a convex combination of the previous iteration's result and the solution of (3.5.2). This process requires us to add N points from $\mathbf{BR}_{\lambda^k}(x_i)$, for $i = 1, \dots, N$, to stock the support of solution at each iteration. As a consequence, this approach can lead to a memory overflow issue, as K going to infinity. The large support of μ^K will also raise the difficulty of Step 2 in Algorithm 3.1, in which we will take $\bar{\mu}_0 = \mu^K$. To address this issue, we will use the stochastic Frank-Wolfe algorithm [BLO+22] to (P_{m_N}). This approach will enable us to obtain an approximate empirical solution (P_{m_N}), and can effectively handle the large support of μ^K.

Lemma 3.5.5. *Let $\mu \in \mathcal{P}(Z)$. Then μ lies in $\mathcal{P}_{m_N}(Z)$ if and only if there exists $\mu_i \in \mathcal{P}(Z_{x_i})$ for any $i = 1, \dots, N$ such that $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \otimes \mu_i$.*

Proof. If $\mu_i \in \mathcal{P}(Z_{x_i})$, then $\pi_1\#(\delta_{x_i} \otimes \mu_i) = \delta_{x_i}$. Since the push-forward operator # is linear, we have that $\pi_1\#(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} \otimes \mu_i) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} = m_N$.

Conversely, let us assume that $\mu \in \mathcal{P}_{m_N}(Z)$. By Theorem 3.2.7 and its remark, we can conclude that there exists $\mu_{x_i} \in \mathcal{P}(Z_{x_i})$ for $i = 1, \dots, N$ such that for any bounded and continuous function

$h\colon Z \to \mathbb{R}$, we have

$$\int_Z h d\mu = \frac{1}{N} \sum_{i=1}^{N} \int_{Z_{x_i}} h(x_i, y_i) d\mu_{x_i}(y_i).$$

Applying Fubini's theorem to the equality above, we have

$$\int_Z h\, d\mu = \frac{1}{N} \sum_{i=1}^{N} \int_Z h\, d\left(\delta_{x_i} \otimes \mu_{x_i}\right) = \int_Z h\, d\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \otimes \mu_{x_i}\right).$$

This implies that $\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \otimes \mu_{x_i}$. □

According to Lemma 3.5.5 and Fubini's theorem, the discretized problem (P_{m_N}) is equivalent to

$$\inf_{\mu_i \in \mathcal{P}(Z_{x_i})} f\left(\frac{1}{N} \sum_{i=1}^{N} \int_{Z_{x_i}} g(x_i, y_i) d\mu_i(y_i)\right). \tag{3.5.3}$$

Problem (3.5.3) is the randomized relaxation of an N-agent optimization problem as investigated in [BLO+22],

$$\inf_{y \in \prod_{i=1}^{N} Z_{x_i}} f\left(\frac{1}{N} \sum_{i=1}^{N} g(x_i, y_i)\right). \tag{3.5.4}$$

Problem (3.5.4) is equivalent to a version of problem (3.5.3) in which the probability measures μ_i are restricted to be Dirac measures. In particular, we can associate with each feasible element $y = (y_i)_{i=1,\ldots,N} \in \prod_{i=1}^{N} Z_{x_i}$ (for problem (3.5.4)) the tuple $(\delta_{x_i})_{i=1,\ldots,N}$, which is feasible for (3.5.3), and the probability distribution $\frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i, y_i)}$, which is feasible for (P_{m_N}).

We apply the following Stochastic Frank-Wolfe algorithm, investigated in [BLO+22], to solve problems (3.5.3) and (3.5.4). Let $\mathrm{Bern}(\omega)$ be the Bernoulli distribution with a parameter $\omega \in [0, 1]$.

Algorithm 3.3: Stochastic Frank-Wolfe Algorithm

Initialization: $y^0 \in \prod_{i=1}^{N} Z_{x_i}$. Set $K \geq 1$.
for $k = 0, 1, 2, \ldots, K-1$ **do**
  Compute $\lambda^k = \nabla f(\frac{1}{N} \sum_{i=1}^{N} g(x_i, y_i^k))$.
  for $i = 1, 2, \ldots, N$ **do**
    Find $\bar{y}_i^k \in \mathbf{BR}_{\lambda^k}(x_i)$.
  end for
  Choose $n_k \in \mathbb{N}^*$. Set $\omega_k = 2/(k+2)$.
  for $j = 1, 2, \ldots, n_k$ **do**
    for $i = 1, 2, \ldots, N$ **do**
      Simulate $P_i^{k,j} \sim \mathrm{Bern}(\omega_k)$, independently of all previously defined random variables.
      Set $\hat{y}_i^{k,j} = (1 - P_i^{k,j}) y_i^k + P_i^{k,j} \bar{y}_i^k$.
    end for
    Define $\hat{y}^{k,j} = (\hat{y}_i^{k,j})_{i=1,\ldots,N}$.
  end for
  Find $y^{k+1} \in \operatorname{argmin} \left\{ f(\frac{1}{N} \sum_{i=1}^{N} g(x_i, y_i)) \,\middle|\, y \in \{\hat{y}^{k,j},\ j = 1, 2, \ldots, n_k\} \right\}$.
end for

The interest of Algorithm 3.3 is that it provides an approximate solution to (3.5.4), and the associated empirical distribution serves as a reliable approximate solution of the problem (3.5.3), as demonstrated in the following lemma. Additionally, this empirical distribution has a fixed support size N, which does not increase with the iteration number, making the algorithm memory-efficient.

Lemma 3.5.6. *in Algorithm 3.3, whatever the numbers $(n_k)_{k\in\mathbb{N}}$, we have for any $K = 1, 2, \ldots, 2N$ that*

$$\mathbb{E}\left[f\left(\frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i^K)\right)\right] - \mathbf{val}(3.5.3) \leq \frac{4LD}{K}.$$

Proof. This is from [BLO+22, Thm. 3.7]. □

Remark 3.5.7. Lemma 3.5.6 provides a convergence result for Algorithm 3.3 in terms of expectation. An estimate of the following quantity can be found in [BLO+22, Thm. 3.7]:

$$\mathbb{P}\left[f\left(\frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i^K)\right) \geq \mathbf{val}(3.5.3) + \epsilon + \frac{4LD}{K}\right],$$

for a given $\epsilon > 0$. In particular, this probability can be made arbitrarily small, provided that the numbers n_k are large enough.

In order to obtain an approximate solution of (P_m), we combine Algorithm 3.3 with Algorithm 3.1. Let us consider the outcome y^K of Algorithm 3.3 after K iterations, for $1 \leq K \leq 2N$ and for arbitrary numbers $n_k \geq 1$ of simulations. Let $\mu_N^K = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_i^K)}$. Moving on to Algorithm 3.1, we utilize the following inputs: $m_0 = m_N$, $m_1 = m$, and $\bar{\mu}_0 = \mu_N^K$. The output of Algorithm 3.1 is denoted as $\tilde{\mu}^K$, which is an element of the set $\mathcal{P}_m(Z)$. We have the following convergence result for the combination of Algorithm 3.1 and 3.3.

Theorem 3.5.8. *Let Assumptions A-B hold true, and let $m \in \mathcal{P}^1(Z)$. Then,*

$$\mathbb{E}\left[f\left(\int_Z g d\tilde{\mu}^K\right)\right] - \mathbf{val}(\mathrm{P}_m) \leq \frac{4LD}{K} + 2L_g(C + LM)d_1(m_N, m).$$

Proof. Since $f\left(\int_Z g d\mu_N^K\right) = f\left(\frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i^K)\right)$, by Lemma 3.3.14, we have

$$f\left(\int_Z g\tilde{\mu}^K\right) - f\left(\frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i^K)\right) \leq L_g(C + LM)d_1(m_N, m), \qquad \text{almost surely.}$$

Taking expectation on both sides of the previous inequality, and applying Lemma 3.5.6 and the relation $\mathbf{val}(3.5.3) = \mathbf{val}(\mathrm{P}_{m_N})$, we have

$$\mathbb{E}\left[f\left(\int_Z g d\tilde{\mu}^K\right)\right] - \mathbf{val}(\mathrm{P}_{m_N}) \leq \frac{2LD}{K} + L_g(C + LM)d_1(m_N, m).$$

Combining with Theorem 3.3.16(1), the proof is complete. □

Remark 3.5.9. The realization of Algorithm 3.1 can be simplified in Theorem 3.5.8 thanks to the empirical structure of m_0 and $\bar{\mu}_0$, as noted in Remark 3.3.15.

3.6 Numerical simulation

We consider a Lagrangian MFG in which the agents exploit their own stock of an exhaustible resource. The model is taken from [GHS22]. We fix a time horizon $[0, T]$ where $T \in [0, +\infty)$ (the case $T = \infty$ investigated in [GHS22] is not considered here). The state variable of a representative agent is the level of the stock of resource at any time, denoted $(X_t^q)_{t\in[0,T]}$ and the control is the speed of extraction at any time, denoted q. The dynamic of a given producer with an initial position $x_0 \geq 0$ is described as follows:

$$X_t^q := x_0 - \int_0^t q_\tau d\tau, \qquad t \in [0, T],$$

where $q_t \geq 0$, for any $t \in [0, T]$. We impose that $X_T^q \geq 0$, which implies that $X_t^q \geq 0$ at any time.

We define the set of aggregate production, denoted as $\mathcal{G}$, by

$$\mathcal{G} := \{Q \in \mathbb{L}^2([0, T], \mathbb{R}) \mid 0 \leq Q(t) \leq \frac{1}{2}, \forall t \in [0, T]\}.$$

The price of the resource for this representative producer depends on its extracting speed and an aggregate production $Q \in \mathcal{G}$,

$$p_t := 1 - q_t - \epsilon Q_t, \qquad t \in [0, T],$$

where $\epsilon \in (0, 1)$ is a constant. The gain of this representative producer writes,

$$\int_0^T e^{-rt} q_t (1 - q_t - \epsilon Q_t) dt.$$

where $r \geq 0$ is a discount rate. Therefore, given an aggregate production $Q \in \mathcal{G}$ and an initial position $x_0 \geq 0$, we can formulate an optimal control problem associated with this representative producer,

$$\begin{cases} \inf_{q\in\mathcal{G}} \quad J^Q(q) := \int_0^T e^{-rt} q_t (q_t - 1 + \epsilon Q_t) dt; \\ \text{s.t.} \quad \int_0^T q_t dt \leq x_0. \end{cases} \tag{3.6.1}$$

Lemma 3.6.1. *Problem* (3.6.1) *has a unique solution* $q^Q(x_0)$. *Moreover,* $0 \leq q^Q(x_0)(t) \leq \frac{1}{2}$, *for a.e.* $t \in (0, T)$.

Proof. It is easy to see that $\mathcal{G}$ is a non-empty and convex subset of $\mathbb{L}^2([0, T], \mathbb{R})$. Following [Rud87, Thm. 3.12], if $(f_n \in \mathcal{G})_{n\geq 1}$ converges to f in $\mathbb{L}^2$ sense, then there exists a subsequence of $(f_n)_{n\geq 1}$ converges to f a.e. As a consequence, f lies in $\mathcal{G}$. Therefore, $\mathcal{G}$ is closed. Furthermore, by Hölder's inequality, we obtain that $\{q \in \mathbb{L}^2([0, T], \mathbb{R}) \mid \int_0^T q_t dt \leq x_0\}$ is non-empty, convex and closed in $\mathbb{L}^2([0, T], \mathbb{R})$. It follows that the admissible set of problem (3.6.1) is non-empty, closed and convex in Hilbert space $\mathbb{L}^2([0, T], \mathbb{R})$. On the other hand, the cost function $J^Q(\cdot)$ is strongly convex. Then the existence of the solution of (3.6.1) comes from [Bré11, Cor. 3.23] and the uniqueness is by the strong convexity of J^Q.

Let q be the solution to (3.6.1). Define $q'(t) = \min\{q(t), \frac{1}{2}\}$, for a.e. $t \in (0,T)$. Since $q' \leq q$, q' is also feasible for problem (3.6.1). Moreover, the running cost $q_0 \mapsto q_0(q_0 - 1 + \varepsilon Q_t)$ is increasing for $q_0 \geq \frac{1}{2}$. As a consequence, $J^Q(q') \leq J^Q(q)$. Therefore, q' is optimal, and since the solution is unique, we have $q = q'$, which proves that $q \leq \frac{1}{2}$. □

Let $m \in \mathcal{P}([0,+\infty))$ denote the distribution of the initial conditions of the producers. The aggregate production rate corresponding to q^Q is given by

$$Q_t^Q := \int_0^\infty q_t^Q(x_0) dm(x_0), \qquad \forall t \in [0,T].$$

Following [GHS22], we call Nash equilibrium a solution Q^* to the fix-point problem:

$$Q^* = Q^{Q^*}, \qquad Q^* \in \mathcal{G}. \tag{3.6.2}$$

3.6.1 Potential problem

In this subsection, we find an optimization problem associated with the fixed point problem (3.6.2), which is a particular case of problem (P_m). Let us specific metric spaces and admissible sets in (P_m) associated to (3.6.2):

$$X = [0,\infty), \qquad Y = \mathcal{G}, \qquad F(x) = \left\{ q \in \mathcal{G} \mid \int_0^T q_t dt \leq x \right\}, \qquad Z = \mathrm{Graph}(F), \qquad Z_x = F(x).$$

Let us define the separable Hilbert space $\mathbb{L}^2_{e^{-rt}}([0,T])$ [Rud87, Example. 4.5(b)]:

$$\mathbb{L}^2_{e^{-rt}}([0,T]) := \left\{ \zeta \colon [0,T] \to \mathbb{R} \text{ is Lebesgue measurable} \,\middle|\, \int_0^T e^{-rt} |\zeta(t)|^2 dt < +\infty \right\},$$

with a scalar product,

$$\langle f_1, f_2 \rangle_{\mathbb{L}^2_{e^{-rt}}([0,T])} = \int_0^T e^{-rt} f_1(t) f_2(t) dt.$$

It is easy to check that $Y = \mathcal{G} \subseteq \mathbb{L}^2_{e^{-rt}}([0,T])$. Then, in ($\mathrm{P}_m$), we set $\mathcal{H} = \mathbb{R} \times \mathbb{L}^2_{e^{-rt}}([0,T])$,

$$g \colon Z \to \mathcal{H}, (x,q) \mapsto \left(\int_0^T e^{-rt}(q_t^2 - q_t) dt,\, q \right),$$
$$f \colon \mathcal{H} \to \mathbb{R}, (y_1, y_2) \mapsto y_1 + \frac{\epsilon}{2} \|y_2\|^2_{\mathbb{L}^2_{e^{-rt}}([0,T])}.$$

Therefore, problem (P_m) associated with (3.6.2) writes:

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z \int_0^T e^{-rt}(q_t^2 - q_t) dt d\mu(x,q) + \frac{\epsilon}{2} \int_0^T e^{-rt} \left(\int_Z q_t d\mu(x,q) \right)^2 dt. \tag{3.6.3}$$

Remark 3.6.2. Problem (3.6.3) lies in the framework of the Lagrangian MFG (3.1.3).

Proposition 3.6.3. *If $\bar{\mu}$ is a solution of problem* (3.6.3)*, then $Q^* = \int_Z q d\bar{\mu}(x,q)$ is a Nash equilibrium of the optimal exploitation of exhaustible resources problem, i.e., Q^* is a solution of* (3.6.2).

Proof. Let us first check that Assumption A holds true for problem (3.6.3). It is easy to see that Assumption A(1) and the first and the third points in Assumption A(2) are true by the continuity of g and Lemma 3.6.1. Let us prove that G_λ is lower semi-continuous for any $\lambda \in \mathcal{H}_f$. This is a consequence of the claim that the set-valued function $\mathcal{Z}\colon X \rightsquigarrow \mathcal{H}$, $x \mapsto \{g(x,y) \mid y \in Z_x\}$ is locally Lipschitz, i.e. Lipschitz in any compact set of X. To see the local Lipschitz continuity, we fix any $x_1 < x_2$ in X. If $q \in Z_{x_1}$, then we have immediately that $q \in Z_{x_2}$. This implies that $Z_{x_1} \subseteq Z_{x_2}$. On the other hand, let $q \in Z_{x_2}$. We construct $q' \in Z_{x_1}$ by the following method:

$$q'_t = \begin{cases} q_t, & \text{if } \int_0^t q_\tau d\tau \leq x_1; \\ 0, & \text{otherwise.} \end{cases}$$

As a consequence, we have that $\|q' - q\|_{\mathbb{L}^1([0,T])} \leq x_2 - x_1$. Therefore, by Hölder's inequality,

$$\|q' - q\|^2_{\mathbb{L}^2_{e^{-rt}}([0,T])} \leq \|e^{-rt}(q' - q)\|_{\mathbb{L}^\infty([0,T])} \|q' - q\|_{\mathbb{L}^1([0,T])} \leq x_2(x_2 - x_1).$$

This implies that $Z_{x_2} \subseteq Z_{x_1} + \mathcal{B}_Y(0, \sqrt{x_2(x_2 - x_1)})$. Therefore, Assumption A follows.

Let $\bar{\mu}$ be a solution of problem (3.6.3), $\bar{\lambda} = \nabla f(\int_Z g d\bar{\mu})$ and $Q^* = \int_Z q d\bar{\mu}(x,q)$. By the definitions of f and g, we obtain that $\bar{\lambda} = (1, \epsilon Q^*)$, moreover,

$$g_{\bar{\lambda}}(x,q) = \int_0^T e^{-rt} q_t (q_t - 1 + \epsilon Q^*_t) dt.$$

By Lemma 3.6.1, $\mathbf{BR}_{\bar{\lambda}}(x_0) = \{q^{Q^*}(x_0)\}$ for any $x_0 \in X$. By Corollary 3.3.5, we have that $(\bar{\lambda}, \bar{\mu})$ satisfies the following equilibrium equation:

$$\begin{cases} \bar{\lambda} = \left(1\,,\, \epsilon \int_Z q d\bar{\mu}\right) \\ \bar{\mu}_x = \delta_{q^{Q^*}(x)}, \quad m\text{-a.e.} \end{cases}$$

Combining with Theorem 3.2.7, we obtain that $\int_Z q d\bar{\mu} = \int_X q^{Q^*}(x) dm(x)$. Recall that $Q^* = \int_Z q d\bar{\mu}$, then (3.6.2) follows. □

3.6.2 Numerical simulations

Let the initial measure m be an exponential distribution with parameter $a \geq 0$, i.e., $dm(x) = a e^{-ax} dx$ for all $x \geq 0$. Let us independently sample the distribution m for N times, denoting the samples by $x_1, x_2, ..., x_N$, and $m_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$. The time space $[0,T]$ is discretized with a step size $\Delta t = T/M$ for some $M \geq 1$. Then, a totally discretized problem associated with (3.6.3) writes:

$$\begin{cases} \inf_{q \in \mathbb{R}^{N \otimes M}} & J_N(q) := \frac{\Delta t}{N} \sum_{i=1}^N \sum_{t=0}^{M-1} e^{-rt\Delta t}(q_{i,t}^2 - q_{i,t}) + \frac{\epsilon \Delta t}{2} \sum_{t=0}^{M-1} e^{-rt} \left(\frac{1}{N}\sum_{i=1}^N q_{i,t}\right)^2, \\ \text{such that} & q_i \in S^M(x_i) := \{q \in [0, 1/2]^M \mid \Delta t \sum_{t=0}^{M-1} q_t \leq x_i\}, \qquad i = 1, 2, \ldots, N. \end{cases} \tag{3.6.4}$$

We apply Algorithm 3.3 to solve (3.6.4). At each iteration, the evaluation of a best-response, for each producer i amounts to solve a problem of the following form:

$$\begin{cases} \inf_{q_i \in \mathbb{R}^M} & \Delta t \sum_{t=0}^{M-1} e^{-rt\Delta t} q_{i,t}(q_{i,t} - 1 + \epsilon Q_t), \\ \text{such that} & q_i \in S^M(x_i), \end{cases} \tag{3.6.5}$$

for a given $Q \in [0, 1/2]^M$. This problem is a convex quadratic programming problem in $\mathbb{R}^M$ that can be dealt with by some solvers, such as GUROBI [GO18].

For the resolution of the problem, we chose the following parameters: $T = 10$, $\epsilon = r = a = 1$, $N = 100$, $M = 100$, $K = 100$, $n_k = 10$, for all k. Figure 3.1 shows the extracting speeds and the stocks of three producers with initial stocks: 0.9, 1.2, and 3.1. From Figure 3.1, we see that the producers with the higher initial stock have the same extracting speed as those with a lower initial stock, at the beginning. However, as the smaller agents exhaust their resource, the larger ones progressively raise their extraction speed. Once the extraction speed reaches its maximum value, it rapidly decreases to zero. These observations are consistent with the numerical findings of [GHS22, Sec. 3.3].

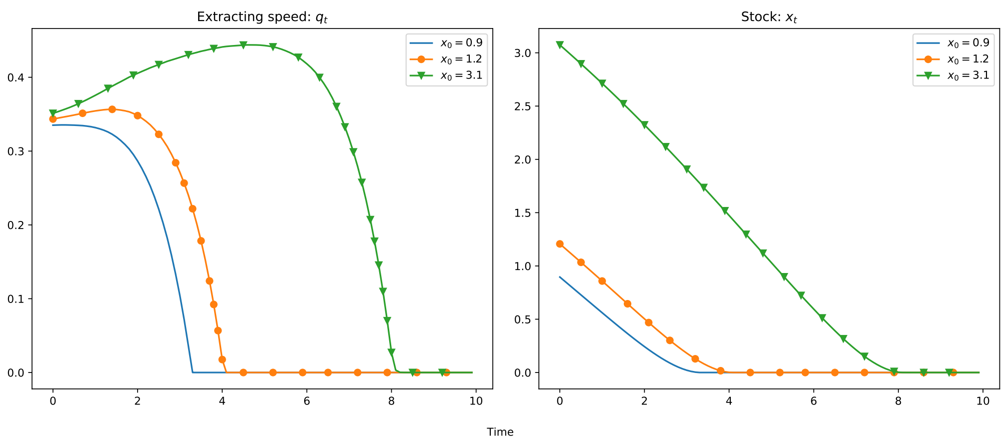


Figure 3.1: Extracting speeds and stocks of three producers with initial stocks: 0.9, 1.2, and 3.1.

To study the error caused by sampling, we independently sample the exponential distribution m for $100 * N$ times, and group them into batches of N. The empirical distribution corresponding to each batch is set as the initial distribution. Then we apply Algorithm 3.3 to compute Q^* corresponding to each initial distribution. In Figure 3.2, we show the mean and standard deviation of the results of the 100 simulations.

3.7 Appendix

3.7.1 Proof and Lemma 3.3.3

Before proving Lemma 3.3.3, let us recall the definitions of the restriction of a measure and the completion of a probability space, taken from [Rud87, Thm. 1.36].

Definition 3.7.1 (Restriction)**.** Let X_1 be a Polish space, let $\mathcal{X}$ and $\mathcal{X}'$ be two σ-algebras on X_1 such that $\mathcal{X}' \subseteq \mathcal{X}$, and let ν be a measure on $\mathcal{X}$. The restriction measure of ν on $\mathcal{X}'$ is defined as follows:

$$\nu|_{\mathcal{X}'}(A) := \nu(A), \qquad \text{for any } A \in \mathcal{X}'.$$

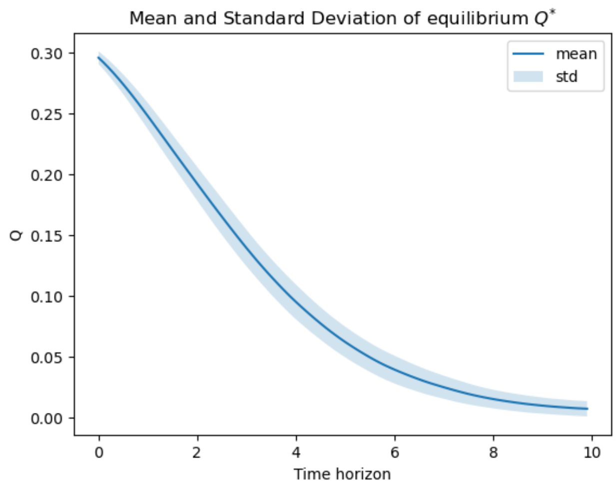


Figure 3.2: Mean and standard deviation of the equilibria of 100 batches

Definition 3.7.2 (Completion)**.** Let $(X_1, \mathcal{B}^{X_1}, \nu)$ be a probability space. Let $\mathcal{B}_\nu$ be the collection of all $E \subseteq X_1$ such that there exists A and B in $\mathcal{B}^{X_1}$, $A \subseteq E \subseteq B$, and $\nu(B - A) = 0$. For such an E, we define a function $\hat{\nu}(E)$ as

$$\hat{\nu}(E) = \nu(A).$$

Then $(X_1, \mathcal{B}_\nu, \hat{\nu})$ is a complete measure space. We say that $(X_1, \mathcal{B}_\nu, \hat{\nu})$ is the completion of $(X_1, \mathcal{B}^{X_1}, \nu)$.

Sketch of the proof of Lemma 3.3.3. The proof of the direction that the left-hand-side of (3.3.2) is greater than the right-hand-side is the same as the proof for the case that $m \in \mathcal{P}_\delta(X)$.

Let us prove the converse inequality. Let $(X, \mathcal{B}_m, \hat{m})$ be the completion of the probability space $(X, \mathcal{B}^X, m)$. Fix any $\lambda \in \mathcal{H}_f$. By Assumption A, the set-valued function $\mathbf{BR}_\lambda \colon X \rightsquigarrow Y$ has non-empty closed images. By Lemma 3.3.1, Graph($\mathbf{BR}_\lambda$) is closed in $X \times Y$, thus is a $\mathcal{B}_m \otimes \mathcal{B}^Y$-measurable set. By Lemma 3.2.5 and Theorem 3.2.3, the set-valued function $\mathbf{BR}_\lambda \colon X \rightsquigarrow Y$ is $(\mathcal{B}_m, \mathcal{B}^Y)$-measurable, and there exists a $(\mathcal{B}_m, \mathcal{B}^Y)$-measurable function $\mathbf{br}_\lambda \colon X \to Y$ such that for any $x \in X$,

$$\mathbf{br}_\lambda(x) \in \mathbf{BR}_\lambda(x).$$

We define $\mathcal{A} \colon X \to Z$, $x \mapsto (x, \mathbf{br}_\lambda(x))$. Since $\mathbf{br}_\lambda$ is $(\mathcal{B}_m, \mathcal{B}^Y)$-measurable, we have that $\mathcal{A}$ is $(\mathcal{B}_m, \mathcal{B}_m \otimes \mathcal{B}^Y)$-measurable, see [Kal97, Lem. 1.8]. Let $\mathcal{B}^Z$ be the Borel σ-algebra on Z. It is obvious that $\mathcal{B}^Z \subseteq \mathcal{B}^X \otimes \mathcal{B}^Y \subseteq \mathcal{B}_m \otimes \mathcal{B}^Y$. Let us take

$$\tilde{\mu} = \mathcal{A}\#\hat{m}|_{\mathcal{B}^Z}.$$

Then $\tilde{\mu}$ is a positive Borel measure on Z. Moreover, we deduce from Definitions 3.7.1-3.7.2 that

$$\tilde{\mu}(Z) = \mathcal{A}\#\hat{m}(Z) = \hat{m}(X) = m(X) = 1.$$

Therefore, $\tilde{\mu} \in \mathcal{P}(Z)$. Assume that the following two equalities hold true:

$$\pi_1 \# \tilde{\mu} = m, \tag{3.7.1}$$

$$\int_Z g_\lambda d\tilde{\mu} = \int_X g_\lambda \circ \mathcal{A}\, dm. \tag{3.7.2}$$

By the definitions of u_λ and $\mathcal{A}$,

$$g_\lambda \circ \mathcal{A}(x) = g_\lambda(x, \mathbf{br}_\lambda(x)) = \inf_{y \in Z_x} g_\lambda(x, y) = u_\lambda(x), \quad \forall x \in X. \tag{3.7.3}$$

Combining (3.7.1)-(3.7.3), we obtain that

$$\inf_{\mu \in \mathcal{P}_m(Z)} \int_Z g_\lambda d\mu \leq \int_Z g_\lambda d\tilde{\mu} = \int_X g_\lambda \circ \mathcal{A}\, dm = \int_X u_\lambda dm.$$

The conclusion follows. □

For completing the proof of Lemma 3.3.3, it remains to prove equalities (3.7.1)-(3.7.2). They are deduced from Lemmas 3.7.3-3.7.4:

- To prove (3.7.1), we take $\tilde{X} = X$ and $h = \pi_1$ in Lemma 3.7.3;
- To prove (3.7.2), we take $\tilde{X} = [-M\|\lambda\|, +\infty)$ and $h = g_\lambda$ in Lemma 3.7.3, and Lemma 3.7.4 implies that $g_\lambda \circ \mathcal{A} = u_\lambda$ is Borel measurable.

Recall the definition of $\mathcal{A}\colon X \to Z$ in the previous proof and recall that $\tilde{\mu} = \mathcal{A}\#\hat{m}|_{\mathcal{B}^Z}$.

Lemma 3.7.3. *Let $\tilde{X}$ be a Polish space. Let $h\colon Z \to \tilde{X}$ be a Borel measurable function. Assume that $h \circ \mathcal{A}\colon X \to \tilde{X}$ is Borel measurable. Then it holds:*

$$h\#\tilde{\mu} = (h \circ \mathcal{A})\#m.$$

As a consequence, if $\tilde{X} = [c, +\infty)$ for some $c \in \mathbb{R}$, then

$$\int_Z h\, d\tilde{\mu} = \int_X h \circ \mathcal{A}\, dm.$$

Proof. Let B be any Borel set in $\tilde{X}$. By the property of push-forward measure, $h\#\tilde{\mu}(B) = \tilde{\mu}(h^{-1}(B))$. Since h is Borel measurable, we have that $h^{-1}(B) \in \mathcal{B}^Z$. This yields that

$$h\#\tilde{\mu}(B) = \mathcal{A}\#\hat{m}(h^{-1}(B)).$$

Next, by the property of the push-forward measure,

$$\mathcal{A}\#\hat{m}(h^{-1}(B)) = \hat{m}(\mathcal{A}^{-1}h^{-1}(B)) = \hat{m}((h \circ \mathcal{A})^{-1}(B)).$$

Since $h \circ \mathcal{A}$ is Borel measurable, we have that $(h \circ \mathcal{A})^{-1}(B) \in \mathcal{B}^X$. As a consequence,

$$\hat{m}((h \circ \mathcal{A})^{-1}(B)) = m((h \circ \mathcal{A})^{-1}(B)) = (h \circ \mathcal{A})\#m(B).$$

Therefore, $h\#\tilde{\mu}(B) = (h \circ \mathcal{A})\#m(B)$ for any Borel set $B \subseteq \tilde{X}$. We conclude the first part of the proof.

In the case that $X_3 = [c, +\infty)$ for some $c \in \mathbb{R}$, since $c = \int_{X_2} c\, d\tilde{\mu} = \int_{X_1} c \circ \mathcal{A}\, dm$, it suffices to prove the conclusion for $h - c$ in instead of h. Therefore, we can assume that $X_3 = \mathbb{R}_+$. By the change-of-variable formula for push-forward measures,

$$\int_Z h d\tilde{\mu} = \int_{\mathbb{R}_+} x\, d(h\#\tilde{\mu}(x)).$$

Next, it follows from the equality $h\#\tilde{\mu} = (h \circ \mathcal{A})\#m$ that

$$\int_{\mathbb{R}_+} x\, d(h\#\tilde{\mu}(x)) = \int_{\mathbb{R}_+} x\, d((h \circ \mathcal{A})\#m(x)).$$

Again, by the change-of-variable formula, we obtain that

$$\int_{\mathbb{R}_+} x\, d((h \circ \mathcal{A})\#m(x)) = \int_X h \circ \mathcal{A}\, dm.$$

The conclusion follows. □

Lemma 3.7.4. *Under Assumption A, the function u_λ is upper semi-continuous for any $\lambda \in \mathcal{H}_f$, thus Borel measurable.*

Proof. Let $\lambda \in \mathcal{H}_f$. Since g is bounded over Z, we have that $u_\lambda(x) > -\infty$ for any $x \in X$. Fix any $x \in X$. Let $y \in \mathbf{BR}_\lambda(x)$. Let $(x_n \in X)_{n \geq 1}$ be a sequence converging to x. By the lower semi-continuity of G_λ, there exists $y_n \in Z_{x_n}$ such that $g_\lambda(x, y) = \lim_{n \to \infty} g_\lambda(x_n, y_n)$. Therefore,

$$u_\lambda(x) = g_\lambda(x, y) = \lim_{n \to \infty} g_\lambda(x_n, y_n) \geq \limsup_{n \to \infty} u_\lambda(x_n).$$

We obtain the upper semi-continuuity of u_λ for any $\lambda \in \mathcal{H}_f$. Since any upper semi-continuous function defined on a metric space is the limit of a monotonically decreasing sequence of continuous functions [Ton52, Thm. 3], we deduce that u_λ is Borel measurable. □

Chapter 4

Error estimates of a theta-scheme for second-order mean field games

4.1 Introduction

Mean field games (MFGs), introduced in 2006 independently by J.-M. Lasry and P.-L. Lions in [LL07] and M. Huang et al. in [HMC06], describe the asymptotic behavior of Nash equilibria in stochastic differential games, as the number of players goes to infinity. In this type of games, players have symmetric dynamics and payoff function. The latter function depends on the own strategy of a given player and on an interaction cost depending on the distribution of all players. Mean field games have important applications in various domains, like crowd motion [LST10], sociology, biology, macroeconomics [ABL$^+$14], trade crowding [CL18a], and finance.

Second-order MFGs (see [LL07, ACD10, Car10]) are coupled systems, including a backward Hamilton-Jacobi-Bellman (HJB) equation and a forward Fokker-Planck (FP) equation. The source term of the HJB equation depends on the solution m of the FP equation while the velocity (the optimal control) v in the transport term of the FP equation depends on the solution u of the HJB equation. Under appropriate hypotheses, we can express v as a function of ∇u at each time. Let $\mathbb{T}^d$ be the d-dimensional torus and let $Q = [0,1] \times \mathbb{T}^d$. We consider the following second-order MFG:

$$\begin{cases} \text{(i)} & -\partial_t u - \sigma \Delta u + H^c(t,x,\nabla u(t,x)) = f^c(t,x,m(t)) & (t,x) \in Q, \\ \text{(ii)} & v(t,x) = -H^c_p(t,x,\nabla u(t,x)) & (t,x) \in Q, \\ \text{(iii)} & \partial_t m - \sigma \Delta m + \operatorname{div}(vm) = 0 & (t,x) \in Q, \\ \text{(iv)} & m(0,x) = m^c_0(x), \quad u(1,x) = g^c(x) & x \in \mathbb{T}^d. \end{cases} \tag{MFG}$$

The Hamiltonian H^c is related to the Fenchel conjugate of a running cost ℓ^c:

$$H^c(t,x,p) = \sup_{v \in \mathbb{R}^d} \langle -p, v \rangle - \ell^c(t,x,v). \tag{4.1.1}$$

We introduce in this article a theta-scheme for the discretization of (MFG); our main result states that, under suitable assumptions, the solution of the theta-scheme converges to the unique solution of (MFG). To the best of the authors' knowledge, this article is the first one, in the context of MFGs, to give a precise convergence order for a fully discrete numerical scheme, namely $\mathcal{O}(h^r)$, where h is the step size of the space variable and $r \in (0,1)$ is related to regularity properties of the solution of (MFG).

Let us describe more in detail the theta-scheme which we propose. Let us denote by ∇_h, div_h and Δ_h the discrete gradient, divergence and Laplace operators of the centered finite-difference scheme (precise definitions are in Section 2). Let $\theta \in [0,1]$. At any time t, the theta-scheme of the FP equation consists of two steps:

1. An explicit scheme for an intermediate FP equation, with a weight $(1-\theta)$ for the Laplacian term:
$$\frac{m(t+1/2)-m(t)}{\Delta t} - (1-\theta)\sigma\Delta_h m(t) + \text{div}_h(mv(t)) = 0. \tag{S1}$$
2. An implicit scheme for an intermediate heat equation (without divergence term):
$$\frac{m(t+1)-m(t+1/2)}{\Delta t} - \theta\sigma\Delta_h(m(t+1)) = 0. \tag{S2}$$

Notice that when there is no divergence term ($v=0$), the above scheme (S1)-(S2) coincides with the classical theta-scheme for the heat equation [All07]. For the HJB equation, we propose an adjoint scheme; at each time t, two steps are performed: (1) an implicit scheme for an intermediate heat equation (without the Hamiltonian term) and (2) an explicit scheme for an intermediate HJB equation. The adjoint structure of the coupled system (MFG) is preserved in the resulting discretized system, which is an important property for the analysis.

Motivations of the theta-scheme. Let us describe the main properties of the theta-scheme, which justify our interest for it. If $\theta = 0$, our scheme is an explicit scheme which has a natural interpretation as a discrete mean field game. However, it is not clear whether the explicit scheme for the FP equation, when $\theta = 0$, enjoys stability properties for some ℓ^2-norm. To ensure stability, a natural idea consists in taking an implicit scheme for the second-order term, i.e. $\theta = 1$. This yields a mixed scheme (implicit for the Laplacian term and explicit for the divergence term). We emphasize that the divergence term should remain explicit, in order to guarantee that the discrete system has a structure of a discrete MFG. When $\theta = 1$, we see that (S1) is an explicit scheme of a continuity equation (without diffusion term). To ensure the monotonicity of (S1), an upwind discretization for the divergence term should be employed, instead of centered scheme. In comparison with a centered discretization, the upwind discretization has the following disadvantages: (1) the consistency error is of a lower order, (2) we need then to construct a numerical Hamiltonian (see [ACD10, ACCD13]) to preserve the adjoint structure. Finally, we propose to take $\theta \in (1/2, 1)$ in (S1)-(S2) and to keep the centered scheme for the first-order term. The ℓ^2-stability is proved in Proposition 4.4.5 for the case when $\theta > 1/2$. The monotonicity property is obtained under a CFL condition (CFL), for all $\theta < 1$, see Theorem 4.4.4. We end up with a discrete system which has a structure of a discrete MFG, has a higher order for the consistency error, and which does not require the construction of a numerical Hamiltonian.

Under suitable assumptions, MFGs have a potential structure (see [CH17, Def. 1.1]), i.e. the system (MFG) can be interpreted as the first order optimality condition of an optimal control problem of the FP equation, see [LL07, LST10, LP22]. Then some optimization algorithms can be applied to solve this optimal control problem, such as the fictitious play [CH17], the generalized conditional gradient algorithm [LP22], ADMM and Chambolle-Pock's algorithm [AL20, BLP23], etc. The last important feature of our theta-scheme is that it preserves the potential structure (when it exists), which allows the application of the previously mentioned methods directly on the discrete system. These methods avoid solving a large discrete nonlinear forward-backward system. For instance, the fictitious play [CH17] and the generalized conditional gradient algorithm [LP22] require to solve the discrete HJB and FP equations iteratively. One significant difference between the theta-scheme and the implicit scheme proposed in [ACCD13] is that the first-order terms in the discrete HJB and FP equations of the former are explicit. Thanks to this, at each time step of the discrete HJB equation, the difficulty of our method lies in solving a linear equation associated with the implicit part of the theta-scheme, which is much cheaper than solving a nonlinear algebraic equation in the totally implicit scheme [ACCD13]. We mention that the aforementioned linear equation to be solved is an implicit scheme of a heat equation. Consequently, in high-dimensional cases, we can consider splitting methods [Tho95, Sec. 4.4] to decompose the discrete Laplace operator and reduce computational complexity.

Related works. In 2010, a first result concerning the convergence of a finite-difference scheme for stationary MFGs was obtained in [ACD10]. In this paper, the authors also proposed an implicit scheme for time-dependent MFGs and proved the existence and uniqueness of the solution of this scheme. In 2013, a convergence result was obtained for the same implicit scheme in [ACCD13] when the Hamiltonian has a monomial form, i.e. $H^c(x,p) = \mathcal{H}(x) + |p|^\beta$, with $\beta \in (1,+\infty)$. The two cited works assume the existence of a classical solution for (MFG). In 2016, in the absence of this existence assumption, [AP16] proved that the solution of the implicit scheme converges to a weak solution of (MFG) when the grid steps tend to zero. No assumption on the Hamiltonian is made in [AP16], but a technical assumption, Assumption (g5), is required for the numerical Hamiltonian (the discrete counterpart of the Hamiltonian). An example of a numerical Hamiltonian satisfying (g5) is only presented for a Hamiltonian with a monomial form (as above), with $\beta \in (1,2]$.

Other discretization techniques have been considered in the literature. We mention the articles [CS14, CS15] in which a semi-Lagrangian discretization is proposed for first-order and second-order MFGs, respectiveley. The well-posedness of the resulting discrete system is established for both cases. In [CS15], the scheme's convergence is proven for non-degenerate second-order MFGs in any dimension and for degenerate second-order and first-order cases in dimension one. A sort of semi-Lagrangian discretization is proposed in [HS19] for first-order MFGs and convergence is established in general dimension. In [BC22] a semi-discretization in space, with finite differences, is investigated. It is shown that the solution of the semi-discrete master equation converges to the solution of the continuous master equation, with an explicit rate of convergence. Finally, we cite the article [AL20], which gives a good summary of the numerical methods for MFGs.

Numerical analysis. In this paper, we assume that the running cost ℓ^c is strongly convex with respect to the control variable. This is equivalent to the Lipschitz continuity of ∇H^c with respect to its third variable. This assumption plays a key role in the stability analysis. We assume that the coupling function f^c is Lipschitz continuous w.r.t. x and with respect to m, for the $\mathbb{L}^2$-norm. Note that our regularity assumptions on f^c are stronger than those of [ACCD13]. We also make a monotonicity assumption for f^c, in Lasry and Lions' sense, see [LL07, Thm. 2.4]. This assumption ensures the uniqueness of the solution of (MFG). For the consistency analysis, we assume that the exact solution of (MFG) lies in the Hölder space $\mathcal{C}^{1+r/2,2+r}(Q)$ (see [Kry96, Ch. 8.5] for the definition). In Appendix 4.6.2, we provide sufficient conditions on the data for this regularity assumption to hold, for an exponent r which is explicit. We also make use of assumptions dealing with the regularity of ℓ^c, m_0^c and g^c. Our convergence analysis relies on a consistency analysis and a stability analysis, the latter relies on a fundamental inequality and an energy estimate for the discrete FP equation.

Consistency analysis. We prove that the discrete HJB equation has a consistency error of order $\mathcal{O}(\Delta t h^r)$ at each time step. For the discrete FP equation, the consistency error is the sum of two terms: one is in the form of the discrete divergence of a term of order $\mathcal{O}(\Delta t h^{2r+d})$ (which can be dealt with by a discrete integration by parts formula in the convergence proof), the other one is of order $\mathcal{O}(\Delta t h^{r+d})$. In comparison with [ACD10, ACCD13], there is no numerical Hamiltonian in our scheme. This simplifies the consistency analysis and avoids the treatment of an additional error term.

Fundamental inequality. The fundamental inequality (Proposition 4.3.7) is established for a general class of discrete MFGs, for which the existence and uniqueness of a solution is easily obtained with a standard fixpoint approach. The fundamental inequality allows us to quantify the variation of the control variable v when a discrete MFG is subject to perturbations. It is deduced from equality (4.3.22), which is similar to the fundamental equality proved in [ACCD13, Eq. 3.20] for an implicit scheme. Our proof of the fundamental inequality also relies on the following technical lemma, given in [Nes18, Thm. 2.1.5]: If F is a convex function with L-Lipschitz gradient, then for any p, q, it follows that

$$\frac{1}{2L}\|\nabla F(p) - \nabla F(q)\|^2 \leq F(p) - F(q) - \langle \nabla F(q), p - q\rangle. \tag{4.1.2}$$

We give a second proof of the fundamental inequality, which does not rely on the fundamental equality (4.3.22). Instead we define a "relative" potential function, and deduce the fundamental inequality from upper and lower bounds of this relative potential function.

Energy estimate. We provide in Proposition 4.4.5 an upper bound of the ℓ^2-norm of the solution of the discrete FP equation under some perturbations. The proof of the energy inequality is inspired by the one for parabolic PDEs, see [Lio71, LSU88], and the one for the implicit scheme, see [ACCD13].

Numerical Hamiltonian. As we mentioned earlier, it is assumed in [AP16] that the numerical Hamiltonian satisfies a specific assumption, Assumption (g5). It turns out that when the numerical Hamiltonian is convex and has a Lipschitz gradient, then (g5) can be easily deduced from inequality (4.1.2), as we show in Lemma 4.6.3. Using this technical result, we provide an example of a numerical Hamiltonian which satisfies all the assumptions of [AP16], for the case of a running cost which is strongly convex with respect to the control variable, uniformly in time and space. See Theorem 4.6.4. This result is of independent interest since our theta-scheme does not require the construction of a numerical Hamiltonian.

Organization of the paper. In Section 2, we present the theta-scheme and state our main result.

Section 3 is dedicated to a general class of discrete MFGs (covering the theta-scheme). We prove in this section the fundamental inequality. In Section 4, some properties of the theta-scheme are demonstrated, in particular, we prove the announced energy estimate for the FP equation. The consistency analysis and the proof of the main result are given in Section 5.

4.2 The theta-scheme and the convergence result

4.2.1 Preliminaries

The set of functions from some finite set A to $\mathbb{R}$ (resp. $\mathbb{R}^d$) is denoted by $\mathbb{R}(A)$ (resp. $\mathbb{R}^d(A)$):

$$\mathbb{R}(A) = \{m \colon A \to \mathbb{R}\}, \qquad \mathbb{R}^d(A) = \{m \colon A \to \mathbb{R}^d\}.$$

Let us introduce the set of probability measures on A, defined by

$$\mathcal{P}(A) = \Big\{ m \in \mathbb{R}(A) \,\Big|\, \forall x \in A,\, m(x) \geq 0,\, \sum_{y \in A} m(y) = 1 \Big\}.$$

We denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and the scalar product in $\mathbb{R}^n$. We define below a scalar product and a norm for functions defined on a finite set.

Definition 4.2.1. Let $n \in \mathbb{N}_+$ and let A_1 and A_2 be two finite sets. For any $\mu, \nu \in \mathbb{R}^n(A_1)$ and $p \in [1, \infty)$, we define

$$\langle \mu, \nu \rangle = \sum_{x \in A_1} \langle \mu(x), \nu(x) \rangle; \qquad \|\mu\|_p = \Big(\sum_{x \in A_1} \|\mu(x)\|^p \Big)^{1/p}; \qquad \|\mu\|_\infty = \max_{x \in A_1} \|\mu(x)\|.$$

For any $\mu \in \mathbb{R}^n(A_1 \times A_2)$ and $p_1, p_2 \in [1, \infty]$, we define

$$\|\mu\|_{p_1,p_2} = \Big\| \big(\|\mu(x,\cdot)\|_{p_2} \big)_{x \in A_1} \Big\|_{p_1} = \begin{cases} \big(\sum_{x \in A_1} \|\mu(x,\cdot)\|_{p_2}^{p_1} \big)^{1/p_1}, & \text{if } p_1 \in [1,\infty), \\ \max_{x \in A_1} \|\mu(x,\cdot)\|_{p_2}, & \text{if } p_1 = \infty. \end{cases}$$

Lemma 4.2.2 (Hölder's inequality)**.** *Let $\mu, \nu \in \mathbb{R}^n(A_1 \times A_2)$. Then,*

$$\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \Big| \langle \mu(x_1,x_2), \nu(x_1,x_2) \rangle \Big| \leq \|\mu\|_{p_1,p_2} \|\nu\|_{p_1^*,p_2^*},$$

where $p_i \in [1,\infty]$ and $1/p_i + 1/p_i^* = 1$, for $i = 1, 2$.

We make use of Nemytskii operators, in order to alleviate some notations.

Definition 4.2.3 (Nemytskii operators)**.** Let $\zeta \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and let $u \colon \mathcal{X} \to \mathcal{Y}$. Then, the associated Nemytskii operator is the mapping $\zeta[u]$, defined from $\mathcal{X}$ to $\mathcal{Z}$ by

$$\zeta[u](x) = \zeta(x, u(x)).$$

4.2.2 Notations for the finite-difference scheme

The time step is $\Delta t = 1/T$, for $T \in \mathbb{N}_+$. We assume that $T > 1$. The set of time indices is denoted by $\mathcal{T}$ ($\tilde{\mathcal{T}}$ when the final time T is included):

$$\mathcal{T} = \{0, 1, \dots, T-1\}; \qquad \tilde{\mathcal{T}} = \{0, 1, \dots, T\}. \tag{4.2.1}$$

Let S be the uniform discretization of the torus $\mathbb{T}^d$ with step size $h = 1/N$, for $N \in \mathbb{N}_+$, defined by

$$S = \big\{(i_1, i_2, \dots, i_d)h \;\mid\; i_1, \dots, i_d \in \mathbb{Z}/N\mathbb{Z}\big\}. \tag{4.2.2}$$

Let $(e_i)_{i=1,\dots,d}$ be the natural canonical basis of $\mathbb{R}^d$. The discrete Laplace, gradient, and divergence operators for the centered finite-difference scheme are defined as follows:

$$\begin{aligned}
\Delta_h \mu(x) &= \sum_{i=1}^{d} \frac{\mu(x+he_i) + \mu(x-he_i) - 2\mu(x)}{h^2}, && \forall \mu \in \mathbb{R}(S),\ \forall\, x \in S,\\
\nabla_h \mu(x) &= \Big(\frac{\mu(x+he_i) - \mu(x-he_i)}{2h}\Big)_{i=1}^{d}, && \forall \mu \in \mathbb{R}(S),\ \forall\, x \in S,\\
\mathrm{div}_h \omega(x) &= \sum_{i=1}^{d} \frac{\omega_i(x+he_i) - \omega_i(x-he_i)}{2h}, && \forall\, \omega \in \mathbb{R}^d(S),\ \forall\, x \in S,
\end{aligned}$$

where ω_i is the i^{th} coordinate of ω. The forward discrete gradient is defined by

$$\nabla_h^+ \mu(x) = \Big(\frac{\mu(x+he_i) - \mu(x)}{h}\Big)_{i=1}^{d}, \qquad \forall \mu \in \mathbb{R}(S),\ \forall x \in S. \tag{4.2.3}$$

Lemma 4.2.4 (Integration by parts formula)**.** *For any $\omega \in \mathbb{R}^d(S)$ and for any $\mu, \nu \in \mathbb{R}(S)$, it holds that*

$$-\sum_{x \in S} \mu(x) \mathrm{div}_h \omega(x) = \sum_{x \in S} \langle \nabla_h \mu(x), \omega(x) \rangle\,; \tag{4.2.4}$$

$$-\sum_{x \in S} \nu(x) \Delta_h \mu(x) = \sum_{x \in S} \big\langle \nabla_h^+ \nu(x), \nabla_h^+ \mu(x) \big\rangle\,. \tag{4.2.5}$$

The proof is given in the Appendix 4.6.1.

Lemma 4.2.5. *For any $\mu \in \mathbb{R}(S)$, the following inequality holds:*

$$\|\nabla_h \mu\|_2^2 \leq \|\nabla_h^+ \mu\|_2^2. \tag{4.2.6}$$

The proof is given in the Appendix 4.6.1. The following lemma shows some general properties of the implicit scheme associated with the heat equation $\frac{\partial m}{\partial t} - c\Delta m = 0$, used in our theta-scheme.

Lemma 4.2.6. *Let $X \in \mathbb{R}^{|S|}$. Consider the scheme*

$$\frac{Y(x) - X(x)}{\Delta t} - c\Delta_h Y(x) = 0, \qquad \forall\, x \in S, \tag{4.2.7}$$

with unknown $Y \in \mathbb{R}^{|S|}$. The following holds true.

1. *(Existence and uniqueness) The scheme* (4.2.7) *has a unique solution Y.*
2. *(Monotonicity) If $X \geq 0$, then $Y \geq 0$. Moreover, if $X \in \mathcal{P}(S)$, then $Y \in \mathcal{P}(S)$.*
3. *(Lipschitz constant) If X is L-Lipschitz, then Y has the same Lipschitz constant L.*
4. *(Continuity of the discrete gradient and Laplacian) Suppose that $\Delta_h X$ is α-Hölder continuous with constant L', where $0 < \alpha \leq 1$. Then there exists a constant C, independent of Δt and h, such that*
$$\|\nabla_h X - \nabla_h Y\|_\infty \leq C\Delta t h^{\alpha-1}, \qquad \|\Delta_h X - \Delta_h Y\|_\infty \leq C\Delta t h^{\alpha-2}.$$

The proof is given in the Appendix 4.6.1.

4.2.3 The theta-scheme and the main result

We describe the MFG system of interest. Let us fix a running cost ℓ^c, a coupling cost f^c, an initial condition m_0^c and a terminal cost g, where

$$\ell^c\colon Q \times \mathbb{R}^d \to \mathbb{R}, \qquad f^c\colon Q \times \mathcal{D} \to \mathbb{R}, \qquad m_0^c \in \mathcal{D}, \qquad g^c\colon \mathbb{T}^d \to \mathbb{R},$$

and where the set $\mathcal{D}$ is defined by $\mathcal{D} = \left\{\mu \in \mathbb{L}^2(\mathbb{T}^d) \mid \mu \geq 0, \int_{\mathbb{T}^d} \mu(x)dx = 1\right\}$. Recall the formulation of the continuous mean field game:

$$\begin{cases} \text{(i)} & -\partial_t u - \sigma\Delta u + H^c\left(t,x,\nabla u(x,t)\right) = f^c(t,x,m(t)) & (t,x) \in Q, \\ \text{(ii)} & v(t,x) = -H_p^c\left(t,x,\nabla u(x,t)\right) & (t,x) \in Q, \\ \text{(iii)} & \partial_t m - \sigma\Delta m + \operatorname{div}(vm) = 0 & (t,x) \in Q, \\ \text{(iv)} & m(0,x) = m_0^c(x), \quad u(1,x) = g^c(x) & x \in \mathbb{T}^d, \end{cases}$$

where $H^c(t,x,p) = \sup_{v\in\mathbb{R}^d}\langle -p, v\rangle - \ell^c(t,x,v)$. We make the following assumptions on the data functions.

Assumption A. The following holds:

1. *Regularity.* The running cost ℓ^c is continuously differentiable with respect to v. There exist positive constants L_ℓ^c, L_g^c, and L_f^c such that for any $(t,x) \in Q$, for any $v \in \mathbb{R}^d$, and for any $m \in \mathcal{D}$,
   - $\ell^c(\cdot,x,v)$, $\ell^c(t,\cdot,v)$, and $\ell_v^c(\cdot,x,v)$ are L_ℓ^c-Lipschitz continuous
   - g^c is L_g^c-Lipschitz continuous
   - $f^c(\cdot,x,m)$, $f^c(t,\cdot,m)$, and $f^c(t,x,\cdot)$ are L_f^c-Lipschitz continuous (with respect to the $\|\cdot\|_{\mathbb{L}^2}$-norm for the third variable).
2. *Strong convexity.* There exists $\alpha > 0$ such that for any $(t,x) \in Q$, $\ell^c(t,x,\cdot)$ is strongly convex with modulus α^c, i.e.
$$\ell^c(t,x,v_2) \geq \ell^c(t,x,v_1) + \langle \ell_v^c(t,x,v_1), v_2 - v_1\rangle + \frac{\alpha^c}{2}\|v_2 - v_1\|^2, \quad \forall v_1, v_2 \in \mathbb{R}^d.$$

3. *Monotonicity.* The global cost f^c is monotone, i.e., for any $t \in [0,T]$, for any m_1 and $m_2 \in \mathcal{D}$,

$$\int_{\mathbb{T}^d} \Big(f^c(t,x',m_1) - f^c(t,x',m_2) \Big) \big(m_1(x') - m_2(x')\big) dx' \geq 0.$$

Lemma 4.2.7. *Let Assumption A hold true. Then H^c is continuously differentiable with respect to p and H^c_p is $(1/\alpha)$-Lipschitz continuous with respect to p. Moreover, H^c and H^c_p are respectively L^c_ℓ- and (L^c_ℓ/α)-Lipschitz continuous with respect to t.*

The proof is given in Appendix 4.6.1. Following [Kry96, page 117], we introduce the following spaces. Given $r \in (0,1)$, $\mathcal{C}^{r/2,r}(Q)$ denotes the set of real-valued functions over Q which are Hölder continuous with exponent r (resp. $r/2$) with respect to x (resp. t). We denote by $\mathcal{C}^{1+r/2,2+r}(Q)$ the set of real-values functions Q which are such that m, $\partial_t m$, $\partial_{x_i} m$, $\partial_{x_i x_j} m$ lie in $\mathcal{C}^{r/2,r}(Q)$, for any $i, j = 1, \dots d$.

We make the following assumption on the solution of (MFG).

Assumption B. The continuous mean field game (MFG) has a unique solution (u^*, v^*, m^*), with $u^*, m^* \in \mathcal{C}^{1+r/2,2+r}(Q)$ and $v^* \in \mathcal{C}^r(Q) \cap \mathbb{L}^\infty([0,1]; \mathcal{C}^{1+r}(\mathbb{T}^d))$, where $r \in (0,1)$.

In Appendix 4.6.2, we propose a set of regularity assumptions on ℓ^c, f^c, m^c_0 and g^c (Assumption C). We show in Theorem 4.6.2 that Assumptions A and C together imply the Assumption B, for an explicit value of r.

Assumptions A and B are supposed to be satisfied throughout the article.

Let us now discretize the data functions. Let us define $B_h(x) = \prod_{i=1}^d [x - he_i/2,\, x + he_i/2)$. We introduce two operators $\mathcal{I}_h \colon \mathbb{R}(\mathbb{T}^d) \to \mathbb{R}(S)$ and $\mathcal{R}_h \colon \mathbb{R}(S) \to \mathbb{R}(\mathbb{T}^d)$, defined as follows: For any $m^c \in \mathbb{R}(\mathbb{T}^d)$ and for any $m \in \mathbb{R}(S)$,

$$\begin{aligned} \mathcal{I}_h(m^c)(x) &= \int_{B_h(x)} m^c(y) dy, \quad \forall x \in S; \\ \mathcal{R}_h(m)(y) &= \frac{m(x)}{h^d}, \qquad\qquad \forall x \in S,\ y \in B_h(x). \end{aligned} \tag{4.2.8}$$

The discrete counterparts of the data functions ℓ^c, H^c, m^c_0, and g^c are the functions defined as follows: For any $t \in \tilde{\mathcal{T}}$, $x \in S$ and $p \in \mathbb{R}^d$,

$$\begin{aligned} &\ell(t,x,p) = \ell^c(t\Delta t, x, p), \quad H(t,x,p) = H^c(t\Delta t, x, p), \\ &m_0(x) = \mathcal{I}_h(m^c_0)(x), \qquad g(x) = g^c(x). \end{aligned} \tag{4.2.9}$$

The discrete counterpart of f^c is the function $f \colon \mathcal{T} \times S \times \mathbb{R}(S)$ to $\mathbb{R}$ defined by

$$f(t,x,m) = \frac{1}{h^d} \int_{y \in B_h(x)} f^c\Big(t\Delta t, y, \mathcal{R}_h(m)\Big) dy. \tag{4.2.10}$$

Taking any $\theta \in [0,1]$, we introduce the theta-scheme of (MFG): find $(u,v,m) \in \mathbb{R}(\bar{\mathcal{T}} \times S) \times \mathbb{R}^d(\mathcal{T} \times$

$S) \times \mathbb{R}(\bar{\mathcal{T}} \times S)$ such that $\forall (t,x) \in \mathcal{T} \times S$,

$$\begin{cases} \text{(i)} \quad \begin{cases} -\frac{u(t+1,x)-u(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h u(t+1/2,x) = 0, \\ -\frac{u(t+1/2,x)-u(t,x)}{\Delta t} - (1-\theta)\sigma\Delta_h u(t+1/2,x) + H[\nabla_h u(\cdot+1/2,\cdot)](t,x) = f(t,x,m(t)); \end{cases} \\ \\ \text{(ii)} \quad v(t,x) = -H_p[\nabla_h u(\cdot+1/2,\cdot)](t,x); \\ \\ \text{(iii)} \quad \begin{cases} \frac{m(t+1/2,x)-m(t,x)}{\Delta t} - (1-\theta)\Delta_h m(t,x) + \mathrm{div}_h(vm)(t,x) = 0, \\ \frac{m(t+1,x)-m(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h m(t+1,x) = 0; \end{cases} \\ \\ \text{(iv)} \quad m(0,x) = m_0(x), \quad u(T,x) = g(x). \end{cases} \tag{θ-MFG}$$

Denoting $B_1 = \mathrm{Id} - \theta\sigma\Delta t\Delta_h$, the first equation in the dynamic programming equation can be rewritten as follows: $B_1 u(t+1/2,\cdot) = u(t+1,\cdot)$. By Lemma 4.2.6, B_1 is invertible. This allows us to consider $u(t+1/2,\cdot)$ as an auxiliary variable, uniquely determined by $u(t+1,\cdot)$, and thus to regard the unknown value function u of the theta-scheme as an element of $\mathbb{R}(\bar{\mathcal{T}} \times S)$. The same argument also holds for the other auxiliary variable $m(t+1/2,\cdot)$.

We fix now a constant M, defined as follows:

$$M = \frac{1}{\alpha^c}\Big(2 \max_{(t,x)\in Q} \|\ell_v^c(t,x,0)\| + \sqrt{d}(L_\ell^c + L_f^c + L_g^c)\Big). \tag{4.2.11}$$

The constant M is an upper bound of $\|v\|_{\infty,\infty}$, as will be seen in Theorem 4.4.4. We consider the following condition on $(\Delta t, h)$:

$$\Delta t \leq \frac{h^2}{2d(1-\theta)\sigma}, \qquad h \leq \frac{2(1-\theta)\sigma}{M}. \tag{CFL}$$

Remark 4.2.8. Let us reformulate the explicit part of (θ-MFG)(iii) by isolating $m(t+1/2,x)$:

$$\begin{aligned} m(t+1/2,x) = \Big(1 - \frac{2d(1-\theta)\sigma\Delta t}{h^2}\Big)m(t,x) + \Delta t \sum_{i=1}^{d} \Big(\frac{(1-\theta)\sigma}{h^2} - \frac{v_i(t,x+he_i)}{2h}\Big)m(t,x+he_i) \\ + \Delta t \sum_{i=1}^{d} \Big(\frac{(1-\theta)\sigma}{h^2} + \frac{v_i(t,x-he_i)}{2h}\Big)m(t,x-he_i). \end{aligned} \tag{4.2.12}$$

The coefficients preceding $m(t,x)$ and $m(t,x\pm he_i)$ in (4.2.12) are affine functions with respect to $v(t,x)$ and $v(t,x\pm he_i)$ respectively, and these coefficients are positive under the condition (CFL) since M is an upper bound of $\|v\|_{\infty,\infty}$. Moreover, summing (4.2.12) over x yields that $\sum_{x\in S} m(t+1/2,x) = \sum_{x\in S} m(t,x)$. Therefore, under the condition (CFL), if $m(t) \in \mathcal{P}(S)$, then $m(t+1/2) \in \mathcal{P}(S)$. Since $m(t+1)$ is the solution of an implicit scheme for the heat equation (with source term $m(t+1/2,x)$), we have that $m(t+1) \in \mathcal{P}(S)$ if $m(t+1/2) \in \mathcal{P}(S)$, by Lemma 4.2.6. In other words, probability distributions on S are preserved by the discrete Fokker-Planck equation under the condition (CFL).

Remark 4.2.9. Let us discuss the choice of θ in the theta-scheme (θ-MFG). If we set $\theta = 1$, we cannot guarantee the positivity of the coefficients preceding $m(t, x \pm he_i)$ in (4.2.12). As a result, we cannot use the same argument presented in Remark 4.2.8 to ensure the preservation of probability distributions of the discrete FP equation. On the other hand, to obtain an energy estimate (ℓ^2-stability) of the discrete FP equation, we require that $\theta > 1/2$, as demonstrated in Proposition 4.4.5.

Theorem 4.2.10. *Let Assumptions A and B hold true. Let $\theta \in (1/2, 1)$ and let $(\Delta t, h)$ satisfy the condition* (CFL)*. Then, the theta-scheme* (θ-MFG) *has a unique solution (u_h, v_h, m_h). Moreover, there exists a constant $C > 0$, independent of Δt and h, such that*

$$\|u_h - u_h^*\|_{\infty,\infty} + \|m_h - m_h^*\|_{\infty,1} \leq Ch^r,$$

where $u_h^*, m_h^* \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ are defined by $u_h^*(t,x) = u^*(t\Delta t, x)$ and $m_h^*(t) = \mathcal{I}_h(m^*(t\Delta t))$.

The proof of Theorem 4.2.10 is given in Section 4.5.2.

4.3 General properties of discrete mean field games

We consider in this section a general class of discrete time and finite state space mean field games, for which we establish the existence and uniqueness of a solution as well as a fundamental inequality. We will show in Section 4.4 that the theta-scheme falls into this class of problems.

4.3.1 Notations and assumptions

In this section, the state space S is an arbitrary discrete set in $\mathbb{R}^d$, not necessarily a discretization of $\mathbb{T}^d$. Let us introduce the set of discrete curves of probability measures and the set of transition processes, defined by

$$\begin{aligned}
\mathcal{P}(\tilde{\mathcal{T}}, S) &= \left\{ m \in \mathbb{R}(\tilde{\mathcal{T}} \times S) \,\middle|\, \forall t \in \tilde{\mathcal{T}},\, m(t,\cdot) \in \mathcal{P}(S) \right\}, \\
\Pi(\mathcal{T}, S) &= \left\{ \pi \in \mathbb{R}(\mathcal{T} \times S \times S) \,\middle|\, \forall (t,x) \in \mathcal{T} \times S,\, \pi(t,x,\cdot) \in \mathcal{P}(S) \right\}.
\end{aligned}$$

Remark 4.3.1. Any $\pi \in \mathbb{R}(\mathcal{T} \times S \times S)$ is a transition process if and only if for any $m \in \mathcal{P}(S)$ and for any $t \in \mathcal{T}$, we have $m' \in \mathcal{P}(S)$, for $m'(y) = \sum_{x \in S} \pi(t,x,y) m(x)$, for all $y \in S$.

We introduce now a running cost ℓ, a coupling cost f, an initial condition m_0 and a terminal cost g, where

$$\ell \colon \mathcal{T} \times S \times \mathbb{R}^d \to \mathbb{R}, \qquad f \colon \mathcal{T} \times S \times \mathbb{R}(S) \to \mathbb{R}, \qquad m_0 \in \mathcal{P}(S), \qquad g \in \mathbb{R}(S).$$

In this section, ℓ, f, m_0, and g are considered independently of the definition (4.2.9). We will consider again definition (4.2.9) in the next section when we interpret the theta-scheme as a discrete MFG.

To formulate the discrete MFG system, we need a control bound $\bar{D} > 0$. The admissible control space, denoted by $\mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$, is the set of all elements $v \in \mathbb{R}^d(\mathcal{T} \times S)$ such that $\|v\|_{\infty,\infty} \leq \bar{D}$. The probability of the motion from one state $x \in S$ to another state $y \in S$ at a time $t \in \mathcal{T}$ under some control $v \in \mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$ is given by

$$\pi[v](t, x, y) \coloneqq \pi(t, x, y, v(t, x)),$$

where π is a function from $\mathcal{T} \times S \times S \times \mathbb{R}^d$ to $\mathbb{R}$. We assume that $\pi[v]$ is a transition process for any admissible control v, i.e.,

$$\pi[v] \in \Pi(\mathcal{T}, S), \qquad \forall v \in \mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S). \tag{4.3.1}$$

For any $D \in (0, \infty]$, we denote by $\ell^D \colon \mathcal{T} \times S \times \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ the function defined by

$$\ell^D(t, x, v) = \begin{cases} \ell(t, x, v), & \text{if } \|v\| \leq D, \\ \infty, & \text{otherwise.} \end{cases} \tag{4.3.2}$$

When $D = \infty$, $\ell^D = \ell$. The Hamiltonian H^D is defined as follows:

$$H^D(t, x, p) = \sup_{v \in \mathbb{R}^d} \langle -p, v \rangle - \ell^D(t, x, v) = \sup_{v \in \mathbb{R}^d,\, \|v\| \leq D} \langle -p, v \rangle - \ell(t, x, v). \tag{4.3.3}$$

We consider the following assumptions on the previous data.

Assumption 4.1. The following holds:

1. *Regularity.* There exist positive constants L_ℓ, L_g, L_f, and $L_{f'}$ such that for any $t \in \mathcal{T}$, for any $v \in R^d$, and for any $m \in \mathcal{P}(S)$, the functions $\ell(t, \cdot, v)$, $g(\cdot)$, and $f(t, \cdot, m)$ are resp. L_ℓ, L_g, and L_f-Lipschitz continuous, i.e.

   $$\begin{aligned} |\ell(t, x_1, v) - \ell(t, x_2, v)| &\leq L_\ell \|x_1 - x_2\|, \\ |g(x_1) - g(x_2)| &\leq L_g \|x_1 - x_2\|, \\ |f(t, x_1, m) - f(t, x_2, m)| &\leq L_f \|x_1 - x_2\|, \end{aligned}$$

   for all x_1 and x_2 in S. Moreover, the function $f(t, x, \cdot)$ is L'_f-Lipschitz w.r.t. m for the $\| \cdot \|_2$ norm , i.e., for all m_1 and m_2 in $\mathcal{P}(S)$,

   $$|f(t, x, m_1) - f(t, x, m_2)| \leq L'_f \|m_1 - m_2\|_2.$$

2. *Strong convexity.* There exist $\alpha > 0$ such that for any $t \in \mathcal{T}$ and for any $x \in S$, the function $\ell(t, x, \cdot)$ is α-strongly convex, i.e.,

   $$\ell(t, x, v_2) \geq \ell(t, x, v_1) + \langle p, v_2 - v_1 \rangle + \frac{\alpha}{2} \|v_2 - v_1\|^2,$$

   for all v_1 and v_2 in $\mathbb{R}^n$ and for all $p \in \partial_p \ell(t, x, v_1)$.

3. *Monotonicity.* For any $t \in \mathcal{T}$, for any m_1 and m_2 in $\mathcal{P}(S)$,

   $$\sum_{x \in S} \Big(f(t, x, m_1) - f(t, x, m_2) \Big) (m_1(x) - m_2(x)) \geq 0.$$

Lemma 4.3.2. *Let $D \in (0, \infty]$. The following holds true.*

1. *The Hamiltonian H^D is continuously differentiable with respect to p.*
2. *For any $t \in \mathcal{T}$, for any $x \in S$, and for any $v \in \mathbb{R}^d$, we have $H^D(t,x,p) = -\langle p, v\rangle - \ell^D(t,x,v)$ if and only if $v = -H_p^D(t,x,p)$.*
3. *The partial derivative H_p^D is $\frac{1}{\alpha}$-Lipschitz continuous with respect to p.*
4. *For any $t \in \mathcal{T}$, for any $x \in S$, for any $v \in \mathbb{R}^d$, and for any $p_0 \in \partial_v \ell(t,x,0)$,*
$$\left\| H_p^D(t,x,p) \right\| \leq \frac{1}{\alpha}\Big(2\|p_0\| + \|p\|\Big). \tag{4.3.4}$$

The proof is given in the Appendix 4.6.1. A direct consequence of Lemma 4.3.2 is the following.

Corollary 4.3.3. *Let $(t,x,p) \in \mathcal{T} \times S \times \mathbb{R}^d$. Let $p_0 \in \partial_v \ell(t,x,0)$. Let D_1 and $D_2 \in (0,\infty]$ be such that $D_i \geq \frac{1}{\alpha}\big(2\|p_0\| + \|p\|\big)$, for $i = 1, 2$. Then*
$$H^{D_1}(t,x,p) = H^{D_2}(t,x,p) \quad \text{and} \quad H_p^{D_1}(t,x,p) = H_p^{D_2}(t,x,p).$$

Lemma 4.3.4. *Let $D \in (0,\infty]$, let $t \in \mathcal{T}$ and let $x \in S$. For any v, for any $\bar{p} \in \mathbb{R}^d$, for any $m \geq 0$ and for any $\bar{m} \in \mathbb{R}$, it holds that*
$$\ell^D(t,x,v)m - \ell^D(t,x,\bar{v})\bar{m} \geq -H^D(t,x,\bar{p})(m - \bar{m}) - \langle \bar{p}, mv - \bar{m}\bar{v}\rangle + \frac{\alpha}{2}\|v - \bar{v}\|^2 m, \tag{4.3.5}$$
where $\bar{v} = -H_p^D(t,x,\bar{p})$.

The proof is given in the Appendix 4.6.1.

4.3.2 The discrete MFG model

The discrete MFG model of interest in this section is a coupled system of three variables: a value function $u \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$, a policy $v \in \mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$, and a curve of probability distributions $m \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$. It consists of a Kolmogorov equation, a dynamic programming equation, and a feedback relation.

- Given $v \in \mathbb{R}^d_{\bar{D}}(\mathcal{T} \times S)$, denote by $\mathbf{FP}(v) \in \mathbb{R}(\mathcal{T} \times S)$ the solution m to the Kolmogorov equation
$$\begin{cases} m(t+1,y) = \sum_{x \in S} \pi[v](t,x,y)m(t,x), & \forall (t,y) \in \mathcal{T} \times S, \\ m(0,x) = m_0(x), & \forall x \in S. \end{cases} \tag{4.3.6}$$
- Given $\mu \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, denote by $\mathbf{HJB}(\mu) \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ the solution u to the dynamic programming equation
$$\begin{cases} u(t,x) = \inf\limits_{\omega \in \mathbb{R}^d} \Big(\tilde{\ell}^{\bar{D}}_\mu(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega)u(t+1,y)\Big), & \forall (t,x) \in \mathcal{T} \times S; \\ u(T,x) = g(x), & \forall x \in S, \end{cases} \tag{4.3.7}$$
where $\tilde{\ell}^{\bar{D}}_\mu(t,x,\omega) = \ell^{\bar{D}}(t,x,\omega) + f(t,x,\mu(t))$.

- Given $u \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$, denote by $\mathbf{V}(u)$ the policy v defined by

$$v(t,x) = \underset{\omega \in \mathbb{R}^d}{\operatorname{argmin}} \Big(\ell^{\bar{D}}(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega) u(t+1,y) \Big), \quad \forall (t,x) \in \mathcal{T} \times S. \tag{4.3.8}$$

The uniqueness of the minimizer in the above definition is a consequence of Lemma 4.3.2.

The discrete MFG consists in finding a triplet (u,v,m) such that $u = \mathbf{HJB}(m)$, $v = \mathbf{V}(u)$, and $m = \mathbf{FP}(v)$. This is equivalent to find a fixpoint to the map ϕ, defined by

$$\phi \colon m \in \mathcal{P}(\tilde{\mathcal{T}}, S) \mapsto \mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}(m) \in \mathcal{P}(\tilde{\mathcal{T}}, S).$$

It is easy to verify that ϕ is indeed valued in $\mathcal{P}(\tilde{\mathcal{T}} \times S)$. Let $m \in \mathcal{P}(\tilde{\mathcal{T}} \times S)$ and let $v = \mathbf{V} \circ \mathbf{HJB}(m)$. By definition, $\|v\|_{\infty,\infty} \leq \bar{D}$. Therefore, by assumption (4.3.1), $\pi[v]$ is a transition process. Then $\mathbf{FP}(v) \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, by Remark 4.3.1.

The discrete MFG can be formulated as the following coupled system: for all $(t,x) \in \mathcal{T} \times S$,

$$\begin{cases} \text{(i)} \quad u(t,x) = \inf_{\omega \in \mathbb{R}^d} \tilde{\ell}_m^{\bar{D}}(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega) u(t+1,y); \\ \text{(ii)} \quad v(t,x) = \operatorname{argmin}_{\omega \in \mathbb{R}^d} \ell^{\bar{D}}(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega) u(t+1,y); \\ \text{(iii)} \quad m(t+1,x) = \sum_{y \in S} \pi[v](t,y,x) m(t,y); \\ \text{(iv)} \quad m(0,x) = m_0(x), \quad u(T,x) = g(x). \end{cases} \tag{4.3.9}$$

As mentioned in Remark 4.2.8, the coefficients preceding $m(t, x \pm he_i)$ in (4.2.12) are affine functions with respect to $v(t, x \pm he_i)$. Furthermore, $m(t+1)$ can be seen as a linear function of $m(t+1/2)$ independent of v from the implicit part of (θ-MFG)(iii). Therefore, in the theta-scheme (θ-MFG), we can express $m(t+1,x)$ as a linear combination of $m(t,y)$ for $y \in S$, where the coefficients preceding $m(t,y)$ are affine functions with respect to $v(t,y)$. Comparing this with the coefficients $\pi[v](t,y,x) = \pi(t,y,x,v(t,y))$ in (4.3.9)(iii), in order to study (θ-MFG) as a particular case of (4.3.9), we find it convenient to consider $\pi(t,x,y,\omega)$ in an affine form of ω, i.e.,

$$\pi(t,x,y,\omega) = \pi_0(t,x,y) + \Delta t \langle \pi_1(t,x,y), \omega \rangle, \qquad \forall (t,x,y,\omega) \in \mathcal{T} \times S^2 \times \mathbb{R}^d, \tag{4.3.10}$$

where $\pi_0 \in \mathbb{R}(\mathcal{T} \times S \times S)$ and $\pi_1 \in \mathbb{R}^d(\mathcal{T} \times S \times S)$. The exact formulas for π_0 and π_1 associated with (θ-MFG) are given in (4.4.3)-(4.4.4).

In the sequel of this section, we consider π given by (4.3.10). We make the following assumption on π_0 and π_1.

Assumption 4.2. The elements π_0 and π_1 satisfy the following condition:

$$\begin{cases} \pi_0(t,x,\cdot) \in \mathcal{P}(S), & \forall (t,x) \in \mathcal{T} \times S, \\ \sum_{y \in S} \pi_1(t,x,y) = 0, & \forall (t,x) \in \mathcal{T} \times S, \\ \pi_0(t,x,y) \geq \Delta t \bar{D} \|\pi_1(t,x,y)\|, & \forall (t,x,y) \in \mathcal{T} \times S \times S. \end{cases}$$

Lemma 4.3.5. *For π given by* (4.3.10)*, Assumption 4.2 is equivalent to* (4.3.1).

The proof of the previous lemma is left to the reader.

Thanks to (4.3.10), we can simplify (4.3.9) (i)-(ii) with the help of $H^{\bar{D}}$ (defined by (4.3.3)). Let us define p_0, p_1, q_0, and q_1 as follows: for all $(t,x) \in \mathcal{T} \times S$,

$$p_0(t,x) = \sum_{s \in S} \pi_0(t,x,s) u(t+1,s), \qquad p_1(t,x) = \sum_{s \in S} \pi_1(t,x,s) u(t+1,s); \tag{4.3.11}$$

$$q_0(t,x) = \sum_{s \in S} \pi_0(t,s,x) m(t,s), \qquad q_1[v](t,x) = \sum_{s \in S} \langle \pi_1(t,s,x), v(t,s) m(t,s) \rangle. \tag{4.3.12}$$

Observe that the dependence of p_0 and p_1 with respect to u is not explicitly mentioned, similarly, the dependence of q_0 and q_1 with respect to m and v is not explicitly mentioned and will be clear from the context. Then system (4.3.9) equivalently writes: for all $(t,x) \in \mathcal{T} \times S$,

$$\begin{cases} \text{(i)} \quad u(t,x) = \big(- H^{\bar{D}}[p_1](t,x) + f(t,x,m(t)) \big) \Delta t + p_0(t,x); \\ \text{(ii)} \quad v(t,x) = -H_p^{\bar{D}}[p_1](t,x); \\ \text{(iii)} \quad m(t+1,x) = q_0(t,x) + \Delta t q_1[v](t,x); \\ \text{(iv)} \quad m(0,x) = m_0(x), \quad u(T,x) = g(x). \end{cases} \tag{DMFG}$$

Theorem 4.3.6 (Existence)**.** *Under Assumptions 4.1 and 4.2,* (DMFG) *has at least one solution. Furthermore, if* $(\bar{u}, \bar{v}, \bar{m})$ *is a solution of* (DMFG)*, then* $\bar{m} \in \mathcal{P}(\tilde{\mathcal{T}}, S)$*.*

Proof (first part). We equip the finite-dimensional space $\mathbb{R}(\tilde{\mathcal{T}} \times S)$ with the norm $\| \cdot \|_{\infty,1}$. The set $\mathcal{P}(\tilde{\mathcal{T}}, S)$ is non-empty, convex, and compact. In order to prove the existence of a solution, we need to show the existence of fixpoint for the map ϕ, defined in (4.3.2). By the Brouwer fixed-point theorem, it suffices to show that ϕ is a continuous mapping, which we do in appendix 4.6.1. □

4.3.3 A fundamental inequality

Let us define a perturbed version of (DMFG) with additional terms $(\eta, \delta) \in \mathbb{R}^2(\mathcal{T} \times S)$ in the right-hand side: for all $(t,x) \in \mathcal{T} \times S$,

$$\begin{cases} \text{(i)} \quad u(t,x) = \Big(- H^{\bar{D}}[p_1](t,x) + f(t,x,m(t)) \Big) \Delta t + p_0(t,x) + \eta(t,x); \\ \text{(ii)} \quad v(t,x) = -H_p^{\bar{D}}[p_1](t,x); \\ \text{(iii)} \quad m(t+1,x) = q_0(t,x) + \Delta t q_1[v](t,x) + \delta(t,x); \\ \text{(iv)} \quad m(0,x) = m_0(x), \quad u(T,x) = g(x). \end{cases} \tag{PDMFG}$$

The fundamental inequality proved in the next proposition is an essential tool in the stability analysis for the system (DMFG).

Proposition 4.3.7 (Fundamental inequality)**.** *Let Assumptions 4.1 and 4.2 hold true. Let* $(\bar{u}, \bar{v}, \bar{m})$ *be a solution of* (DMFG) *and let* (u, v, m) *satisfy* (PDMFG) *with* $m \geq 0$*. Then, the following* *inequality holds:*

$$\frac{\Delta t\alpha}{2}\sum_{t\in\mathcal{T}}\sum_{x\in S}\|(v-\bar{v})(t,x)\|^2(m+\bar{m})(t,x)\leq \sum_{t\in\mathcal{T}}\sum_{x\in S}(u-\bar{u})(t+1,x)\delta(t,x)+(\bar{m}-m)(t,x)\eta(t,x). \tag{4.3.13}$$

This fundamental inequality is of the same nature as the one established in [ACCD13, Sec. 3.3]. We provide two different proofs of Proposition 4.3.7 in the next subsection. The fundamental inequality allows us to show the uniqueness of the solution to (DMFG).

Lemma 4.3.8 (Uniqueness)**.** *Under Assumptions 4.1-4.2,* (DMFG) *has a unique solution.*

Proof. The existence result was already established in Theorem 4.3.6. Let (u_1, v_1, m_1) and (u_2, v_2, m_2) be two solutions of (DMFG). By Theorem 4.3.6, $m_1 \geq$ and $m_2 \geq 0$. Viewing (u_2, v_2, m_2) as a solution to (PDMFG) with $(\eta, \delta) = (0, 0)$, we deduce from the fundamental inequality that

$$\|v_1(t,x)-v_2(t,x)\|(m_1(t,x)+m_2(t,x))=0.$$

Thus for any $(t,x)\in\mathcal{T}\times S$, either $v_1(t,x)=v_2(t,x)$, or $m_1(t,x)=m_2(t,x)=0$. Let $\mu=m_1-m_2$, then μ satisfies the following equation: for any $(t,x)\in\mathcal{T}\times S$,

$$\begin{cases} \mu(t+1,x) = \sum_{s\in S}\pi[v_1](t,s,x)\mu(t,s)+\Delta t\sum_{s\in S}\langle\pi_1(t,s,x),(v_1-v_2)m_2(t,s)\rangle, \\ \quad\mu(0,x) = 0. \end{cases}$$

It follows by induction that $\mu = 0$, i.e. $m_1 = m_2$. Then $u_1 = \mathbf{HJB}(m_1) = \mathbf{HJB}(m_2) = u_2$ and $v_1 = \mathbf{V}(u_1) = \mathbf{V}(u_2) = v_2$, which concludes the proof. □

4.3.4 Two proofs of the fundamental inequality

In this subsection, (u, v, m) is a solution of (PDMFG) and (p_0, p_1) is defined by (4.3.11). Let $(\bar{u}, \bar{v}, \bar{m})$ be a solution to (DMFG). Let $(\bar{p}_0, \bar{p}_1)$ be defined by (4.3.11), for the triplet $(\bar{u}, \bar{v}, \bar{m})$. The following sum-by-parts formulas will be used in both two methods of proof. For all $t \in \mathcal{T}$,

$$\sum_{x\in S}\bar{p}_0\bar{m}(t,x)+\Delta t\langle\bar{p}_1,\bar{m}\bar{v}\rangle(t,x)=\sum_{y\in S}\bar{u}(t+1,y)\bar{m}(t+1,y); \tag{4.3.14}$$

$$\sum_{x\in S}\bar{p}_0 m(t,x)+\Delta t\langle\bar{p}_1,mv\rangle(t,x)=\sum_{y\in S}\bar{u}(t+1,y)m(t+1,y)-\sum_{y\in S}\bar{u}(t+1,y)\delta(t,y); \tag{4.3.15}$$

$$\sum_{x\in S}p_0\bar{m}(t,x)+\Delta t\langle p_1,\bar{m}\bar{v}\rangle(t,x)=\sum_{y\in S}u(t+1,y)\bar{m}(t+1,y); \tag{4.3.16}$$

$$\sum_{x\in S}p_0 m(t,x)+\Delta t\langle p_1,mv\rangle(t,x)=\sum_{y\in S}u(t+1,y)m(t+1,y)-\sum_{y\in S}u(t+1,y)\delta(t,y). \tag{4.3.17}$$

For proving (4.3.14), one simply needs to multiply the first equation in (4.3.11) by $\bar{m}(t,x)$, to multiply the second equation in (4.3.11) by $\bar{m}\bar{v}(t,x)$ and to sum the results over x. This yields

$$\sum_{x\in S}\bar{p}_0\bar{m}(t,x)+\Delta t\langle\bar{p}_1,\bar{m}\bar{v}\rangle(t,x)=\sum_{x\in S}\sum_{s\in S}\bar{u}(t+1,s)\Big(\pi_0(t,x,s)\bar{m}(t,x)+\Delta t\langle\pi_1(t,x,s),\bar{v}\bar{m}(t,s)\rangle\Big).$$

Then (4.3.14) follows from (DMFG)-(iii). The proofs of the other three equations can be obtained similarly. We provide now two different proofs of Proposition 4.3.7.

1. Direct method We follow [ACCD13]. Summing the difference of (4.3.15) and (4.3.14) over $t \in \mathcal{T}$, we get

$$\sum_{t=1}^{T}\sum_{x\in S}\bar{u}(m-\bar{m})(t,x) = \sum_{t\in\mathcal{T}}\sum_{x\in S}\bar{p}_0(m-\bar{m})(t,x)+\Delta t\langle \bar{p}_1, mv-\bar{m}\bar{v}\rangle(t,x)+\bar{u}(t+1,x)\delta(t,x). \quad (4.3.18)$$

In addition, summing the difference of (4.3.17) and (4.3.16) over $t \in \mathcal{T}$, we get

$$\sum_{t=1}^{T}\sum_{x\in S}u(m-\bar{m})(t,x) = \sum_{t\in\mathcal{T}}\sum_{x\in S}p_0(m-\bar{m})(t,x)+\Delta t\langle p_1, mv-\bar{m}\bar{v}\rangle(t,x)+u(t+1,x)\delta(t,x). \quad (4.3.19)$$

Taking the difference of (4.3.19) and (4.3.18), we have

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{x\in S}(u-\bar{u})(m-\bar{m})(t,x)\\ &\quad= \sum_{t\in\mathcal{T}}\sum_{x\in S}(p_0-\bar{p}_0)(m-\bar{m})(t,x)+\Delta t\langle p_1-\bar{p}_1, mv-\bar{m}\bar{v}\rangle(t,x)+(u-\bar{u})(t+1,x)\delta(t,x).\end{aligned} \quad (4.3.20)$$

Moreover, taking the difference of (PDMFG) (i) and (DMFG) (i), multiplying the result by $m-\bar{m}$, summing over $(t,x) \in \mathcal{T}\times S$, we obtain that

$$\begin{aligned}&\sum_{t\in\mathcal{T}}\sum_{x\in S}(u-\bar{u})(m-\bar{m})(t,x)\\ &\quad= \sum_{t\in\mathcal{T}}\sum_{x\in S}(p_0-\bar{p}_0)(m-\bar{m})(t,x)+\Delta t\big(H^{\bar{D}}[\bar{p}_1]-H^{\bar{D}}[p_1]\big)(m-\bar{m})(t,x)\\ &\qquad+\sum_{t\in\mathcal{T}}\sum_{x\in S}\Delta t\big(f(t,x,m(t))-f(t,x,\bar{m}(t))\big)(m-\bar{m})(t,x)+\eta(m-\bar{m})(t,x).\end{aligned} \quad (4.3.21)$$

Comparing (4.3.20) and (4.3.21) and using the relations $v=-H_p^{\bar{D}}[p_1]$, $\bar{v}=-H_p^{\bar{D}}[\bar{p}_1]$, we obtain the following equality:

$$\begin{aligned}&\Delta t\sum_{t\in\mathcal{T}}\sum_{x\in S}m\Big(H^{\bar{D}}[\bar{p}_1]-H^{\bar{D}}[p_1]-\langle H_p^{\bar{D}}[p_1],\bar{p}_1-p_1\rangle\Big)(t,x)\\ &\quad+\Delta t\sum_{t\in\mathcal{T}}\sum_{x\in S}\bar{m}\Big(H^{\bar{D}}[p_1]-H^{\bar{D}}[\bar{p}_1]-\langle H_p^{\bar{D}}[\bar{p}_1],p_1-\bar{p}_1\rangle\Big)(t,x)\\ &\quad+\Delta t\sum_{t\in\mathcal{T}}\sum_{x\in S}\big(f(t,x,m(t))-f(t,x,\bar{m}(t))\big)(m-\bar{m})(t,x)\\ &\qquad=\sum_{t\in\mathcal{T}}\sum_{x\in S}(u-\bar{u})(t+1,x)\delta(t,x)+(\bar{m}-m)(t,x)\eta(t,x).\end{aligned} \quad (4.3.22)$$

Since $H^{\bar{D}}$ is convex and $H_p^{\bar{D}}$ is $1/\alpha$-Lipschitz, we obtain with inequality (4.1.2) (see [Nes18, Thm. 2.1.5]) that

$$H^{\bar{D}}[\bar{p}_1]-H^{\bar{D}}[p_1]-\langle H_p^{\bar{D}}[p_1],\bar{p}_1-p_1\rangle \geq \frac{\alpha}{2}\big\|H_p^{\bar{D}}[p_1]-H_p^{\bar{D}}[\bar{p}_1]\big\|^2=\frac{\alpha}{2}\|v-\bar{v}\|^2;$$
$$H^{\bar{D}}[p_1]-H^{\bar{D}}[\bar{p}_1]-\langle H_p^{\bar{D}}[\bar{p}_1],p_1-\bar{p}_1\rangle \geq \frac{\alpha}{2}\big\|H_p^{\bar{D}}[p_1]-H_p^{\bar{D}}[\bar{p}_1]\big\|^2=\frac{\alpha}{2}\|v-\bar{v}\|^2.$$

We substitute the last two inequalities into (4.3.22). Then, inequality (4.3.13) follows from the non-negativity of m and $\bar{m}$ and the monotonicity of f in Assumption 4.1.

2. Variational method Let us define a "relative" potential function $\tilde{J}_{\bar{m}} \colon \mathbb{R}^d(\mathcal{T}\times S)\times\mathcal{P}(\tilde{\mathcal{T}}, S) \to \mathbb{R}$,

$$\tilde{J}_{\bar{m}}(v,m) = \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} m(t,x)\Big(\ell^{\bar{D}}(t,x,v(t,x)) + f(t,x,\bar{m}(t))\Big) + \sum_{x\in S} g(x)m(T,x).$$

Note that in the above function $\tilde{J}_{\bar{m}}$, the third variable of f is fixed to $\bar{m}$. The second proof of Proposition 4.3.7 consists in proving a lower bound and an upper bound of $\tilde{J}_{\bar{m}}(v,m) - \tilde{J}_{\bar{m}}(\bar{v},\bar{m})$, from which the fundamental inequality directly follows.

Step 1. Let us prove that

$$\tilde{J}_{\bar{m}}(v,m) - \tilde{J}_{\bar{m}}(\bar{v},\bar{m}) \geq \sum_{t\in\mathcal{T}} \sum_{x\in S} \bar{u}(t+1,x)\delta(t,x) + \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \frac{\alpha}{2}\|v-\bar{v}\|^2 m(t,x).$$

By Lemma 4.3.4, we have

$$\begin{aligned}
\Big(\ell^{\bar{D}}[v]m - \ell^{\bar{D}}[\bar{v}]\bar{m}\Big)\Delta t \geq{}& \Big(-H^{\bar{D}}[\bar{p}_1](m-\bar{m}) - \langle \bar{p}_1, mv - \bar{m}\bar{v}\rangle + \frac{\alpha}{2}\|v-\bar{v}\|^2 m\Big)\Delta t \\
={}& \left(\bar{u} - \bar{p}_0 - \Delta t f(t,x,\bar{m}(t))\right)(m-\bar{m}) - \Delta t\langle \bar{p}_1, mv - \bar{m}\bar{v}\rangle + \Delta t\frac{\alpha}{2}\|v-\bar{v}\|^2 m \\
={}& \bar{u}(m-\bar{m}) + (\bar{p}_0\bar{m} + \Delta t\langle \bar{p}_1, \bar{m}\bar{v}\rangle) - (\bar{p}_0 m + \Delta t\langle \bar{p}_1, mv\rangle) \\
&- \Delta t f(t,x,\bar{m}(t))(m-\bar{m}) + \Delta t\frac{\alpha}{2}\|v-\bar{v}\|^2 m.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\tilde{J}_{\bar{m}}(v,m) - \tilde{J}_{\bar{m}}(\bar{v},\bar{m}) \\
&\quad= \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S}\Big(\ell^{\bar{D}}[v]m - \ell^{\bar{D}}[\bar{v}]\bar{m} + f(t,x,\bar{m}(t))(m-\bar{m})\Big)(t,x) + \sum_{x\in S} g(x)(m-\bar{m})(T,x) \\
&\quad\geq \sum_{t\in\mathcal{T}} \sum_{x\in S} \bar{u}(m-\bar{m}) + (\bar{p}_0\bar{m} + \Delta t\langle \bar{p}_1, \bar{m}\bar{v}\rangle) - (\bar{p}_0 m + \Delta t\langle \bar{p}_1, mv\rangle) + \Delta t\frac{\alpha}{2}\|v-\bar{v}\|^2 m \\
&\qquad\quad + \sum_{x\in S} g(x)(m-\bar{m})(T,x) \\
&\quad= \sum_{t\in\mathcal{T}} \sum_{x\in S} \bar{u}(t+1,x)\delta(t,x) + \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \frac{\alpha}{2}\|v-\bar{v}\|^2 m(t,x),
\end{aligned}$$

where the last equality was obtained with (4.3.14) and (4.3.15).

Step 2. Let us prove that

$$\tilde{J}_{\bar{m}}(v,m) - \tilde{J}_{\bar{m}}(\bar{v},\bar{m}) \leq \sum_{t\in\mathcal{T}} \sum_{x\in S} u(t+1,x)\delta(t,x) - \eta(m-\bar{m})(t,x) - \Delta t\frac{\alpha}{2}\|v-\bar{v}\|^2 \bar{m}(t,x).$$

Since v satisfies (PDMFG)-(ii), by Fenchel's relation [HUL93, Cor. 1.4.4], we have

$$\ell^{\bar{D}}[v] = -\langle p_1, v\rangle - H^{\bar{D}}[p_1], \qquad -p_1 \in \partial\ell^{\bar{D}}[v].$$

Then, by the α-strong convexity of ℓ and the last equality, we have

$$\ell^{\bar{D}}[\bar{v}] \geq \ell^{\bar{D}}[v] + \langle -p_1, \bar{v} - v \rangle + \frac{\alpha}{2}\|\bar{v} - v\|^2 = -H^{\bar{D}}[p_1] - \langle p_1, \bar{v} \rangle + \frac{\alpha}{2}\|\bar{v} - v\|^2.$$

Using the nonnegativity of m and $\bar{m}$, we obtain that

$$\begin{aligned}
\left(\ell^{\bar{D}}[v]m - \ell^{\bar{D}}[\bar{v}]\bar{m}\right)\Delta t &\leq \left(-H^{\bar{D}}[p_1](m - \bar{m}) - \langle p_1, mv - \bar{m}\bar{v} \rangle - \frac{\alpha}{2}\|\bar{v} - v\|^2 \bar{m}\right)\Delta t \\
&= (u - p_0 - \Delta t f(t,x,m(t)) - \eta)(m - \bar{m}) - \Delta t \langle p_1, mv - \bar{m}\bar{v} \rangle - \Delta t \frac{\alpha}{2}\|\bar{v} - v\|^2 \bar{m} \\
&= u(m - \bar{m}) + (p_0 \bar{m} + \Delta t \langle p_1, \bar{m}\bar{v} \rangle) - (p_0 m + \Delta t \langle p_1, mv \rangle) \\
&\quad - (\Delta t f(t,x,m(t)) + \eta)(m - \bar{m}) - \Delta t \frac{\alpha}{2}\|\bar{v} - v\|^2 \bar{m}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\tilde{J}_{\bar{m}}(v,m) - \tilde{J}_{\bar{m}}(\bar{v},\bar{m}) \\
&\quad = \Delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \left(\ell^{\bar{D}}[v]m - \ell^{\bar{D}}[\bar{v}]\bar{m} + f(t,x,\bar{m}(t))(m - \bar{m})\right)(t,x) + \sum_{x \in S} g(x)(m - \bar{m})(T,x) \\
&\quad \leq \sum_{t \in \mathcal{T}} \sum_{x \in S} u(m - \bar{m}) + (p_0 \bar{m} + \Delta t \langle p_1, \bar{m}\bar{v} \rangle) - (p_0 m + \Delta t \langle p_1, mv \rangle) + \sum_{x \in S} g(x)(m - \bar{m})(T,x) \\
&\qquad - \sum_{t \in \mathcal{T}} \sum_{x \in S} \Delta t \left(f(t,x,m(t)) - f(t,x,\bar{m}(t))\right)(m - \bar{m})(t,x) \\
&\qquad - \sum_{t \in \mathcal{T}} \sum_{x \in S} \eta(m - \bar{m})(t,x) + \Delta t \frac{\alpha}{2}\|v - \bar{v}\|^2 \bar{m}(t,x) \\
&\quad \leq \sum_{t \in \mathcal{T}} \sum_{x \in S} u(t+1,x)\delta(t,x) - \eta(t,x)(m - \bar{m})(t,x) - \Delta t \frac{\alpha}{2}\|v - \bar{v}\|^2 \bar{m}(t,x),
\end{aligned}$$

where the last inequality is a consequence of (4.3.16), (4.3.17), and the monotonicity of f.

4.4 Stability analysis for the theta-scheme

We turn back to the stability analysis of the theta-scheme. It consists of two steps: the fundamental inequality, which is obtained by formulating (DMFG) as a discrete MFG, and an energy estimate for the Kolmogorov equation.

From now on ℓ, H, g, m_0, and f are again to be understood according to the definitions given in (4.2.9) and (4.2.10).

4.4.1 Reformulation of the theta-scheme as a discrete MFG

The goal of this subsection is to show the equivalence between the scheme (θ-MFG) and a discrete MFG of the form (DMFG). Given $D \in (0,\infty]$, define ℓ^D as in (4.3.2) and H^D as in (4.3.3). Note that for $D = \infty$, $H^D = H$. Consider the following system, with unknown variables $u \in \mathbb{R}(\bar{\mathcal{T}} \times S)$,

$v \in \mathbb{R}^d(\mathcal{T} \times S)$, and $m \in \mathbb{R}(\bar{\mathcal{T}} \times S)$:

$$\begin{cases} \text{(i)} \quad \begin{cases} (\mathrm{Id} - \theta\sigma\Delta t\Delta_h)\, u(t+1/2) = u(t+1), \\ u(t,x) = \big[- H^D[\nabla_h u(\cdot + 1/2, \cdot)](t,x) + f(t,x,m(t))\big]\Delta t \\ \qquad\qquad +\big(\mathrm{Id} + (1-\theta)\sigma\Delta t\Delta_h\big)u(t+1/2)(x), \end{cases} & \forall (t,x) \in \mathcal{T} \times S; \\ \text{(ii)} \quad v(t,x) = -H_p^D[\nabla_h u(\cdot + 1/2, \cdot)](t,x), & \forall (t,x) \in \mathcal{T} \times S; \\ \text{(iii)} \quad \begin{cases} m(t+1/2) = \big(\mathrm{Id} + (1-\theta)\sigma\Delta t\Delta_h\big)m(t) - \Delta t \mathrm{div}_h\big(v(t)m(t)\big), \\ (\mathrm{Id} - \theta\sigma\Delta t\Delta_h)\, m(t+1) = m(t+1/2), \end{cases} & \forall t \in \mathcal{T}; \\ \text{(iv)} \quad m(0,x) = m_0(x), \quad u(T,x) = g(x), & \forall x \in S. \end{cases} \tag{θ-MFG(D)}$$

Multiplying the dynamic programming equation (i) and the Kolmogorov equation (iii) of (θ-MFG) by Δt, we easily see that (θ-MFG) is equivalent to (θ-MFG(D)) with $D = \infty$.

We recall here the definition of the matrix B_1 and introduce a new matrix B_2:

$$B_1 = \mathrm{Id} - \theta\sigma\Delta t\Delta_h, \qquad B_2 = (1-\theta)\sigma\Delta_h. \tag{4.4.1}$$

By Lemma 4.2.6, the matrix B_1 is invertible.

We regard the variables u and m of the system (θ-MFG(D)) as elements of $\mathbb{R}(\bar{\mathcal{T}} \times S)$, since the auxiliary variables $u(t+1/2,\cdot)$ and $m(t+1/2,\cdot)$ are uniquely determined by $u(t+1,\cdot)$ and $m(t,\cdot)$. In the sequel, we will make use of the following convention: Given $u \in \mathbb{R}(\bar{\mathcal{T}} \times S)$, we denote

$$u(t+1/2,\cdot) = B_1^{-1}u(t+1,\cdot), \quad \forall t \in \mathcal{T}. \tag{4.4.2}$$

Lemma 4.4.1. *For any $D > 0$, the system* (θ-MFG(D)) *is equivalent to the system* (DMFG) *with running cost ℓ, control bound $\bar{D} = D$, coupling function f, final cost g, initial distribution m_0 and with π_0 and π_1 defined by:*

$$\pi_0(t,x,y) = B_1^{-1}(y,x) + \Delta t(B_1^{-1}B_2)(y,x), \tag{4.4.3}$$

$$\pi_1(t,x,y) = \Big(\frac{B_1^{-1}(y,x+he_i) - B_1^{-1}(y,x-he_i)}{2h}\Big)_{i=1}^d. \tag{4.4.4}$$

Proof. We make use of the notations p_0, p_1, q_0 and q_1, defined as in (4.3.11)-(4.3.12). By the definition of B_1 and B_2, the implicit steps in equations (i) and (iii) are equivalent to

$$u(t+1/2) = B_1^{-1}u(t+1) \quad \text{and} \quad m(t+1) = B_1^{-1}m(t+1/2).$$

Next we verify the equivalences between each of the three equations of the two systems.

Step 1. Using the definition of B_1 and π_1, we have that

$$\begin{aligned}\nabla_h u(t+1/2,x) &= \left(\frac{u(t+1/2,x+he_i)-u(t+1/2,x-he_i)}{2h}\right)_{i=1}^d\\ &= \left(\frac{\sum_{y\in S}\Big(B_1^{-1}(x+he_i,y)-B_1^{-1}(x-he_i,y)\Big)u(t+1,y)}{2h}\right)_{i=1}^d\\ &= \sum_{y\in S}\pi_1(t,x,y)u(t+1,y) = p_1(t,x).\end{aligned}$$

The equivalence of the feedback relations follows.

Step 2. The dynamic programming equation is equivalent to

$$u(t,x) = \Big[-H^D[p_1](t,x)+f(t,x,m(t))\Big]\Delta t + \big(\mathrm{Id}+\Delta t B_2\big)B_1^{-1}u(t+1)(x).$$

Observe that

$$\begin{aligned}&\big(\mathrm{Id}+\Delta t B_2\big)B_1^{-1}u(t+1)(x)\\ &\quad= \sum_{y\in S}B_1^{-1}(x,y)u(t+1,y)+\Delta t\sum_{z\in S}\sum_{y\in S}B_2(x,z)B_1^{-1}(z,y)u(t+1,y)\\ &\quad= \sum_{y\in S}B_1^{-1}(y,x)u(t+1,y)+\Delta t\sum_{y\in S}\Big(\sum_{z\in S}B_1^{-1}(y,z)B_2(z,x)\Big)u(t+1,y)\\ &\quad= \sum_{y\in S}\pi_0(t,x,y)u(t+1,y) = p_0(x),\end{aligned}$$

The equivalence with the dynamic programming equation of (DMFG) follows.

Step 3. The Kolmogorov equation in (θ-MFG(D)) is equivalent to

$$\begin{aligned}m(t+1,y) &= \sum_{x\in S}\Big(B_1^{-1}(y,x)+\Delta t(B_1^{-1}B_2)(y,x)\Big)m(t,x)\\ &\qquad -\Delta t\sum_{i=1}^d\sum_{x\in S}B_1^{-1}(y,x)\frac{v_i(t,x+he_i)m(t,x+he_i)-v_i(t,x-he_i)m(t,x-he_i)}{2h}\\ &= \sum_{x\in S}\Big(B^{-1}(y,x)+\Delta t(B_1^{-1}B_2)(y,x)\Big)m(t,x)\\ &\qquad +\Delta t\sum_{x\in S}\sum_{i=1}^d\frac{B_1^{-1}(y,x+he_i)-B_1^{-1}(y,x-he_i)}{2h}v_i(t,x)m(t,x)\\ &= \sum_{x\in S}\pi_0(t,x,y)m(t,x)+\Delta t\langle\pi_1(t,x,y),v(t,x)\rangle m(t,x)\\ &= q_0(t,y)+\Delta t q_1[v](t,y).\end{aligned}$$

The lemma is proved. □

Lemma 4.4.2. *The maps ℓ, f, and g satisfy Assumption 4.1 with the following constants:*

$$\alpha=\alpha^c,\quad L_\ell = L_\ell^c,\quad L_f=L_f^c,\quad L_f' = L_f^c h^{-d/2},\quad \textit{and}\quad L_g=L_g^c.$$

Proof. The Lipschitz-continuity of ℓ, f, and g, and the strong convexity of ℓ are straightforward. For all $(t,x) \in \mathcal{T} \times S$ and μ_1, $\mu_2 \in \mathcal{P}(S)$, we have

$$\begin{aligned} f(t,x,\mu_1) - f(t,x,\mu_2) &= \frac{1}{h^d} \int_{B_h(x)} f^c\Big(t\Delta t, x, \mathcal{R}_h(\mu_1)\Big) - f^c\Big(t\Delta t, x, \mathcal{R}_h(\mu_2)\Big) dy \\ &\leq L_f^c \big\| \mathcal{R}_h(\mu_1) - \mathcal{R}_h(\mu_2) \big\|_{\mathbb{L}^2} \\ &= L_f^c \Big(\sum_{x \in S} h^d \Big(\frac{\mu_1(x)}{h^d} - \frac{\mu_2(x)}{h^d} \Big)^2 \Big)^{1/2} \\ &= L_f^c h^{-d/2} \|\mu_1 - \mu_2\|_2. \end{aligned}$$

This proves the Lipschitz continuity of f with respect to its third variable. Let us consider again μ_1, $\mu_2 \in \mathcal{P}(S)$ and $t \in \mathcal{T}$. We have

$$\begin{aligned} &\sum_{x \in S} (f(t,x,\mu_1) - f(t,x,\mu_2))(\mu_1(x) - \mu_2(x)) \\ &\quad = \frac{1}{h^d} \sum_{x \in S} \int_{y \in B_h(x)} f^c(t\Delta t, y, \mathcal{R}_h(\mu_1)) - f^c(t\Delta t, y, \mathcal{R}_h(\mu_2)) dy (\mu_1(x) - \mu_2(x)) \\ &\quad = \sum_{x \in S} \int_{y \in B_h(x)} \Big(f^c(t\Delta t, y, \mathcal{R}_h(\mu_1)) - f^c(t\Delta t, y, \mathcal{R}_h(\mu_2)) \Big) \Big(\frac{\mu_1(x)}{h^d} - \frac{\mu_2(x)}{h^d} \Big) dy \\ &\quad = \sum_{x \in S} \int_{y \in B_h(x)} \Big(f^c(t\Delta t, y, \mathcal{R}_h(\mu_1)) - f^c(t\Delta t, y, \mathcal{R}_h(\mu_2)) \Big) \Big(\mathcal{R}_h(\mu_1)(y) - \mathcal{R}_h(\mu_2)(y) \Big) dy \\ &\quad = \int_{\mathbb{T}^d} \Big(f^c(t\Delta t, y, \mathcal{R}_h(\mu_1)) - f^c(t\Delta t, y, \mathcal{R}_h(\mu_2)) \Big) \Big(\mathcal{R}_h(\mu_1)(y) - \mathcal{R}_h(\mu_2)(y) \Big) dy \geq 0. \end{aligned}$$

This proves the monotonicity assumption. The lemma is proved. □

Lemma 4.4.3 (Lipschitz continuity)**.** *Let $D \in (0, \infty]$. Let (m,u,v) be a solution to (θ-MFG(D)). Suppose that $(\Delta t, h)$ satisfies the condition* (CFL)*. Then for all $t \in \bar{\mathcal{T}}$, $u(t,\cdot)$ and $u(t+1/2,\cdot)$ are $(L_g^c + L_f^c + L_\ell^c)$-Lipschitz continuous. Moreover, $\|v\|_{\infty,\infty} \leq M$, where M was defined in* (4.2.11).

Proof. We define, for any $t \in \bar{\mathcal{T}}$, $L_t = L_g^c + \Delta t(T-t)(L_f^c + L_\ell^c)$. We prove by induction that for any $t \in \bar{\mathcal{T}}$, $u(t,\cdot)$ is L_t-Lipschitz continuous. The claim is obvious for $t = T$, by the terminal condition. Suppose that $u(t+1,\cdot)$ is L_{t+1}-Lipschitz for some $t \in \mathcal{T}$. The first equation in (θ-MFG(D)) is equivalent to the dynamic programming equation:

$$\begin{cases} (\mathrm{Id} - \theta\sigma\Delta t\Delta_h)\, u(t+1/2) = u(t+1); \\ u(t,x) = \Delta t \inf_\omega \Big\{ f(t,x,m(t)) + \ell^D(t,x,\omega) + \Big\langle \omega, \nabla_h u(t+1/2,x) \Big\rangle \Big\} \\ \qquad\qquad + \big(\mathrm{Id} + (1-\theta)\sigma\Delta t\Delta_h\big) u(t+1/2)(x), \qquad \forall\, x \in S. \end{cases} \tag{4.4.5}$$

By the third statement of Lemma 4.2.6, we have that $u(t+1/2,\cdot)$ is L_{t+1}-Lipschitz. Therefore, $\|\nabla_h u(t+1/2,\cdot)\| \leq \sqrt{d}(L_g^c + L_f^c + L_\ell^c)$. Next let us take $y \in S$ and let us set $\omega_y = v(t,y)$. We have

$$\omega_y = \underset{\|\omega\| \leq D}{\operatorname{argmin}} \Big(\ell(t,y,\omega) + \langle \omega, \nabla_h u(t+1/2,y) \rangle \Big).$$

By inequality (4.3.4) of Lemma 4.3.2, we have $\|\omega_y\| \leq M$. Let $r' = (1-\theta)\sigma\Delta t/h^2$, then, for any x, we have

$$\begin{aligned} u(t,x) - u(t,y) \leq & \Big(f(t,x,m(t)) - f(t,z,m(t)) + \ell(t,x,\omega_y) - \ell(t,y,\omega_y) \Big) \Delta t \\ & + (1 - 2dr') \Big(u(t+1/2,x) - u(t+1/2,y) \Big) \\ & + \sum_{i=1}^{d} \Big(r' + \frac{(\omega_z)_i}{2h} \Big) \Big(u(t+1/2, x+he_i) - u(t+1/2, y+he_i) \Big) \\ & + \sum_{i=1}^{d} \Big(r' - \frac{(\omega_z)_i}{2h} \Big) \Big(u(t+1/2, x-he_i) - u(t+1/2, y-he_i) \Big). \end{aligned}$$

The coefficients $\big(1-2dr'\big)$, $\big(r' + \frac{(\omega_z)_i}{2h}\big)$, and $\big(r' - \frac{(\omega_z)_i}{2h}\big)$ are positive by the condition (CFL). Thus,

$$u(t,x) - u(t,z) \leq \|x-z\| \left(\Delta t (L_f^c + L_\ell^c) + L_{t+1} \right) \leq L_t \|x-z\|,$$

where $L_t = L_g + \Delta t(T-t)(L_f + L_\ell) \leq L_g + L_f + L_\ell$. Since x and y are arbitrary, we deduce that $u(t,\cdot)$ is L_t-Lipschitz. The claim is proved. We have meanwhile established that $\|v\|_{\infty,\infty} \leq M$. The lemma is proved, since for any $t \in \bar{\mathcal{T}}$, $L_t \leq L_g^c + L_f^c + L_\ell^c$. □

Theorem 4.4.4. *Let the condition* (CFL) *hold true. Then the discretized MFG system* (θ-MFG) *is equivalent to the system* (θ-MFG(D)) *with $D = M$, which is itself equivalent to a discrete MFG of the form* (DMFG)*, satisfying Assumptions 4.1 and 4.2, with π_0 and π_1 defined by* (4.4.3) *and* (4.4.4) *and with control bound $\bar{D} = M$. As a consequence,* (θ-MFG) *has a unique solution.*

Proof. By construction, (4.4.6) is equivalent to (θ-MFG(D)) with $D = \infty$. As a direct consequence of Corollary 4.3.3 and Lemma 4.4.3, the system (θ-MFG(D)) with $D = \infty$ is equivalent to (θ-MFG(D)) with $D = M$. We already know that (θ-MFG(D)) is equivalent to (DMFG), by Lemma 4.4.1 and that Assumption 4.1 is satisfied, by Lemma 4.4.2. It remains to verify that (θ-MFG(D)) satisfy Assumption 4.2, for $\bar{D} = M$. To do this, we need to verify that for any $v \in \mathbb{R}^d(\mathcal{T} \times S)$ with $\|v\|_{\infty,\infty} \leq D$, $\pi[v]$ is a transition process, by Lemma 4.3.5. By Remark 4.3.1, this is equivalent to show that for any $t \in \mathcal{T}$, for any $m(t) \in \mathcal{P}(S)$, for any $v \in \mathbb{R}^d(\mathcal{T} \times S)$ such that $\|v\|_{\infty,\infty} \leq M$, we have $m(t+1) \in \mathcal{P}(S)$, where $m(t+1)$ is defined by equation (θ-MFG(D))-(iii). We conclude that $m(t+1) \in \mathcal{P}(S)$ from Remark 4.2.8. □

4.4.2 Energy estimate for the discrete FP equation

We investigate here the ℓ^2-stability of the discrete Fokker-Planck equation of the theta-scheme. To this aim we consider the following perturbed equation:

$$\begin{cases} \left(\mathrm{Id} - \theta\sigma\Delta t \Delta_h\right) \mu(t+1) = & \big(\mathrm{Id} + (1-\theta)\sigma\Delta t\Delta_h\big)\mu(t) - \Delta t \mathrm{div}_h\big(v(t)\mu(t)\big) \\ & -\Delta t \mathrm{div}_h\big(\delta_v(t)\big) + \Delta t \delta(t), \\ \mu(0) = & \mu_0, \end{cases} \tag{4.4.6}$$

where $\delta_v \in \mathbb{R}^d(S), \delta \in \mathbb{R}(S)$. Note that we have no sign condition on μ. The first error term $-\Delta t \operatorname{div}_h\big(\delta_v(t)\big)$ represents a perturbation in the form of a discrete divergence and $\Delta t \delta(t)$ is another general perturbation term.

Proposition 4.4.5 (Energy inequality)**.** *Let $\theta > 1/2$ and μ be a solution of (4.4.6). Let $v \in \mathbb{R}^d(\mathcal{T} \times S)$ be such that $\|v\|_{\infty,\infty} \leq M$. Then, there exists some constant c independent of h and Δt such that*

$$\max_{t \in \tilde{\mathcal{T}}} \big\|\mu(t)\big\|_2^2 \leq c \left(\big\|\mu_0\big\|_2^2 + (1-\theta)\sigma \big\|\nabla_h^+ \mu_0\big\|_2^2 + \sum_{\tau \in \mathcal{T}} \Delta t \left(\big\|\delta_v(\tau)\big\|_2^2 + \big\|\delta(\tau)\big\|_2^2 \right) \right). \tag{4.4.7}$$

Proof. Recall that the forward discrete gradient was defined in (4.2.3). Computing the scalar product with $\mu(t+1)$ of both sides of (4.4.6) and applying the integration by parts formulas (4.2.4)-(4.2.5), we obtain that

$$\big\langle \mu(t+1) - \mu(t), \mu(t+1) \big\rangle + \theta \sigma \Delta t \beta_1 = (1-\theta)\sigma \Delta t \beta_2 + \Delta t \big(\gamma_1 + \gamma_2 + \gamma_3\big), \tag{4.4.8}$$

where

$$\begin{aligned}
\beta_1 &= -\big\langle \mu(t+1), \Delta_h \mu(t+1) \big\rangle = \big\|\nabla_h^+ \mu(t+1)\big\|_2^2, \\
\beta_2 &= \big\langle \mu(t+1), \Delta_h \mu(t) \big\rangle = -\big\langle \nabla_h^+ \mu(t+1), \nabla_h^+ \mu(t) \big\rangle, \\
\gamma_1 &= -\big\langle \operatorname{div}_h\big(\mu(t) v(t)\big), \mu(t+1) \big\rangle = \sum_{x \in S} \big\langle \nabla_h \mu(t+1,x), \mu v(t,x) \big\rangle, \\
\gamma_2 &= -\big\langle \operatorname{div}_h\big(\delta_v(t)\big), \mu(t+1) \big\rangle = \sum_{x \in S} \big\langle \nabla_h \mu(t+1,x), \delta_v(t,x) \big\rangle, \\
\gamma_3 &= \big\langle \delta(t), \mu(t+1) \big\rangle.
\end{aligned}$$

Using Young's inequality, it is easy to prove that

$$\Big\langle \mu(t+1) - \mu(t), \mu(t+1) \Big\rangle \geq \frac{1}{2} \Big(\|\mu(t+1)\|_2^2 - \|\mu(t)\|_2^2 \Big).$$

Combining the above inequality with (4.4.8), we obtain that

$$\begin{aligned}
&\frac{1}{2}\Big(\|\mu(t+1)\|_2^2 - \|\mu(t)\|_2^2 \Big) + \theta \sigma \Delta t \big\|\nabla_h^+ \mu(t+1)\big\|_2^2 \\
&\qquad \leq -(1-\theta)\sigma \Delta t \left\langle \nabla_h^+ \mu(t+1), \nabla_h^+ \mu(t) \right\rangle + \Delta t \big(\gamma_1 + \gamma_2 + \gamma_3\big).
\end{aligned} \tag{4.4.9}$$

Applying Young's inequality to the right-hand side of (4.4.9) and using inequality (4.2.6), we obtain that for all positive numbers $\alpha_0, \alpha_1, \alpha_2$ and α_3, we have

$$\begin{aligned}
-\big\langle \nabla_h^+ \mu(t+1), \nabla_h^+ \mu(t) \big\rangle &\leq \frac{\alpha_0}{2} \big\|\nabla_h^+ \mu(t+1)\big\|_2^2 + \frac{1}{2\alpha_0} \big\|\nabla_h^+ \mu(t)\big\|_2^2; \\
\gamma_1 &\leq \frac{\alpha_1}{2} \big\|\nabla_h^+ \mu(t+1)\big\|_2^2 + \frac{M^2}{2\alpha_1} \big\|\mu(t)\big\|_2^2; \\
\gamma_2 &\leq \frac{\alpha_2}{2} \big\|\nabla_h^+ \mu(t+1)\big\|_2^2 + \frac{1}{2\alpha_2} \big\|\delta_v(t)\big\|_2^2; \\
\gamma_3 &\leq \frac{\alpha_3}{2} \big\|\delta(t)\big\|_2^2 + \frac{1}{2\alpha_3} \big\|\mu(t+1)\big\|_2^2.
\end{aligned}$$

Taking $\alpha_0 = 1$, $\alpha_1 = \alpha_2 = \sigma(2\theta - 1) > 0$, and $\alpha_3 = 1$, we have

$$(1-\Delta t)\big\|\mu(t+1)\big\|_2^2 - (1-\Delta t)\big\|\mu(t)\big\|_2^2 + (1-\theta)\sigma\Delta t\Big(\big\|\nabla_h^+\mu(t+1)\big\|_2^2 - \big\|\nabla_h^+\mu(t)\big\|_2^2\Big)$$
$$\leq \Delta t\Big(c_1\big\|\mu(t)\big\|_2^2 + c_2\big\|\delta_v(t)\big\|_2^2 + \big\|\delta(t)\big\|_2^2\Big),$$

where $c_1 = 1 + \frac{M^2}{\sigma(2\theta-1)}$ and $c_2 = \frac{1}{\sigma(2\theta-1)}$. Summing the above equation over t, it follows that

$$(1-\Delta t)\big\|\mu(t+1)\big\|_2^2 \leq \Delta t \sum_{\tau=0}^{t} \Big(c_1\big\|\mu(\tau)\big\|_2^2 + c_2\big\|\delta_v(\tau)\big\|_2^2 + \big\|\delta(\tau)\big\|_2^2\Big) + c_3,$$

where $c_3 = (1-\Delta t)\|\mu_0\|_2^2 + (1-\theta)\sigma\Delta t\|\nabla_h^+\mu_0\|_2^2$. Since $1 - \Delta t \geq 1/2$, by the discrete Gronwall inequality [Cla87], there exists some constant c independent of $(\Delta t, h)$ such that

$$\max_{t\in\tilde{\mathcal{T}}} \big\|\mu(t)\big\|_2^2 \leq c\left(\big\|\mu_0\big\|_2^2 + (1-\theta)\sigma\big\|\nabla_h^+\mu_0\big\|_2^2 + \sum_{\tau\in\mathcal{T}}\Delta t\Big(\big\|\delta_v(\tau)\big\|_2^2 + \big\|\delta(\tau)\big\|_2^2\Big)\right).$$

The proposition is proved. □

Remark 4.4.6. By taking $\alpha_1 = \alpha_2 < \sigma(2\theta - 1)$ in the proof, we can get a refined energy estimate with an additional term $\sum_{t\in\tilde{\mathcal{T}}}\Delta t\big\|\nabla_h^+\mu(t)\big\|_2^2$ on the left-hand side of (4.4.7). This refined energy estimate is consistent with the continuous case [LSU88, Thm. 2.1].

4.5 Consistency analysis of the theta-scheme

This section is dedicated to the consistency analysis of the theta-scheme and to the proof of Theorem 4.2.10. To alleviate the proofs, we will make use of the big $\mathcal{O}$ notation: Given $f_1, f_2 \in \mathbb{R}^n(\mathcal{T}\times S)$ and $\gamma > 0$, the notation $f_1 - f_2 = \mathcal{O}(h^\gamma)$ (or $f_1 = f_2 + \mathcal{O}(h^\gamma)$) means that there exists some constant C independent of h and Δt such that $\|f_1 - f_2\|_{\infty,\infty} \leq Ch^\gamma$. In particular, $f_1 = \mathcal{O}(h^\gamma)$ means that $\|f_1\|_{\infty,\infty} \leq Ch^\gamma$.

All along the section, (CFL) is supposed to be satisfied. Therefore, we have $\Delta t = \mathcal{O}(h^2)$.

4.5.1 Consistency error

Let us recall that (u^*, v^*, m^*) is the unique solution to the continuous system (MFG). The restriction of (u^*, m^*) on the grid, denoted by (u_h^*, v_h^*), is defined as in Theorem 4.2.10. Making use of the convention (4.4.2), we define $v_h^* \in \mathbb{R}^d(\mathcal{T}\times S)$ by

$$v_h^*(t,x) = -H_p^M[\nabla_h u_h^*(\cdot + 1/2, \cdot)](t,x). \tag{4.5.1}$$

Then, (u_h^*, v_h^*, m_h^*) can be considered as a solution of the perturbed discrete mean field game (PDMFG) with perturbation terms η and δ specified later in Lemma 4.5.5.

Lemma 4.5.1. *For any $t \in \bar{\mathcal{T}}$, $u_h^*(t,\cdot)$ is $(L_\ell^c + L_f^c + L_g^c)$-Lipschitz continuous. For any $t \in \mathcal{T}$, $u_h^*(t+1/2,.)$ is also $(L_\ell^c + L_f^c + L_g^c)$-Lipschitz continuous. Moreover, $\|v_h^*\|_\infty \leq M$ and*

$$H^M[\nabla_h u_h^*(\cdot + 1/2, \cdot)](t,x) = H^c[\nabla_h u_h^*(\cdot + 1/2, \cdot)](t,x)$$
$$H_p^M[\nabla_h u_h^*(\cdot + 1/2, \cdot)](t,x) = H_p^c[\nabla_h u_h^*(\cdot + 1/2, \cdot)](t,x).$$

Proof. It can be proved, with similar ideas to those of the proof of Lemma 4.4.3, that $u^*(t,\cdot)$ is $(L_\ell^c + L_f^c + L_g^c)$-Lipschitz continuous, for any $t \in [0,1]$. The first claim of the lemma follows immediately. The other claims can be shown with the same arguments as those of the proof of Lemma 4.4.3. □

Below we state (without proof) elementary consistency estimates, all directly deduced from Assumption B:

$$\begin{aligned}
\tfrac{u^*(t+\Delta t,x)-u^*(t,x)}{\Delta t} - \tfrac{\partial u^*(t,x)}{\partial t} &= \mathcal{O}(\Delta t^{r/2}), & \tfrac{m^*(t+\Delta t,x)-m^*(t,x)}{\Delta t} - \tfrac{\partial m^*(t,x)}{\partial t} &= \mathcal{O}(\Delta t^{r/2}),\\
\nabla_h u^*(t,x) - \nabla u^*(t,x) &= \mathcal{O}(h^{1+r}), & \mathrm{div}_h(m^*v^*)(t,x) - \mathrm{div}(m^*v^*)(t,x) &= \mathcal{O}(h^r),\\
\Delta_h u^*(t,x) - \Delta u^*(t,x) &= \mathcal{O}(h^r), & \Delta_h m^*(t,x) - \Delta m^*(t,x) &= \mathcal{O}(h^r).
\end{aligned} \tag{4.5.2}$$

We also observe that the discrete differential operators commute with integrals. For example,

$$\begin{aligned}
\Delta_h(\mathcal{I}_h(m^*))(t,x) &= \frac{1}{h^2}\sum_{i=1}^d \int_{B_h(x)} \Big(m^*(t,y+he_i) + m^*(t,y-he_i) - 2m^*(t,y)\Big)dy\\
&= \int_{B_h(x)} \Delta_h m^*(t,y)dy = \mathcal{I}_h(\Delta_h m^*)(t,x).
\end{aligned} \tag{4.5.3}$$

In the following three lemmas, we investigate the consistency errors associated with the coupling cost, the Hamiltonian, and the divergence term of the Fokker-Planck equation.

Lemma 4.5.2. *For the global cost term, there holds: for all $(t,x) \in \mathcal{T} \times S$,*

$$f\big(t,x,m_h^*(t)\big) - f^c\big((t+1)\Delta t, x, m^*((t+1)\Delta t)\big) = \mathcal{O}(h). \tag{4.5.4}$$

Proof. Since m^* is Lipschitz continuous in time, uniformly in x, we have that

$$\|m^*((t+1)\Delta t) - m^*(t\Delta t)\|_{\mathbb{L}^2} = \mathcal{O}(\Delta t).$$

Then the Lipschitz continuity of f^c with respect to t and m implies that

$$f^c\big((t+1)\Delta t, x, m^*((t+1)\Delta t)\big) - f^c\big(t\Delta t, x, m^*(t\Delta t)\big) = \mathcal{O}(\Delta t).$$

Using the definition of f (provided in (4.2.10)) and the Lipschitz continuity of f^c, we have

$$\begin{aligned}
&\big|f(t,x,m_h^*(t)) - f^c(t\Delta t, x, m^*(t\Delta t))\big|\\
&\quad= \Big|\frac{1}{h^d}\int_{B_h(x)} \Big(f^c(t\Delta t, y, \mathcal{R}_h\mathcal{I}_h(m^*(t\Delta t))) - f^c(t\Delta t, x, m^*(t\Delta t))\Big)dy\Big|\\
&\quad\leq L_f^c\Big(\sqrt{d}h + \big\|\mathcal{R}_h\mathcal{I}_h(m^*(t\Delta t)) - m^*(t\Delta t)\big\|_{\mathbb{L}^2}\Big).
\end{aligned}$$

Then we estimate $\left\|\mathcal{R}_h\mathcal{I}_h(m^*(t\Delta t)) - m^*(t\Delta t)\right\|_{\mathbb{L}^2}$ as follows:

$$\begin{aligned}\left\|\mathcal{R}_h\mathcal{I}_h(m^*(t\Delta t)) - m^*(t\Delta t)\right\|_{\mathbb{L}^2} &= \Big(\sum_{x\in S}\int_{y\in B_h(x)}\Big|\frac{\mathcal{I}_h(m^*(t\Delta t))(x)}{h^d} - m^*(t\Delta t, y)\Big|^2 dy\Big)^{1/2} \\ &\leq \Big(\sum_{x\in S}\int_{y\in B_h(x)}\int_{z\in B_h(x)}\frac{|m^*(t\Delta t, z) - m^*(t\Delta t, y)|^2}{h^d}dzdy\Big)^{1/2} \\ &= \Big(\sum_{x\in S}\int_{y\in B_h(x)}\int_{z\in B_h(x)}\frac{\mathcal{O}(h^2)}{h^d}dzdy\Big)^{1/2} = \mathcal{O}(h),\end{aligned}$$

where the second line is a consequence of Jensen's inequality. The lemma is proved. □

Lemma 4.5.3. *It holds: for all $(t, x) \in \mathcal{T} \times S$,*

$$H^M\Big(t, x, \nabla_h u_h^*(t + 1/2, x)\Big) - H^c\Big((t + 1)\Delta t, x, \nabla u^*((t + 1)\Delta t, x)\Big) = \mathcal{O}(h^{1+r}). \tag{4.5.5}$$

Moreover,

$$\begin{aligned}&v^*((t + 1)\Delta t, x) - v_h^*(t, x) \\ &\quad= H_p^M\Big(t, x, \nabla_h u_h^*(t + 1/2, x)\Big) - H_p^c\Big((t + 1)\Delta t, x, \nabla u^*((t + 1)\Delta t, x)\Big) = \mathcal{O}(h^{1+r}).\end{aligned} \tag{4.5.6}$$

Proof. By Lemma 4.5.1, we have $\|\nabla_h u_h^*(t + 1/2, x)\| \leq C$ and $\|\nabla u^*((t + 1)\Delta, x)\| \leq C$, where $C = \sqrt{d}(L_\ell^c + L_f^c + L_g^c)$. Let Ω denote the closed ball of radius Ω. Since H^c is uniformly Lipschitz with respect to t and continuously differentiable with respect to p (see Lemma 4.2.7), we deduce that $H^c(\cdot, x, \cdot)$ is Lipschitz continuous on $[0, T] \times \Omega$, uniformly in x. Let L_H denote the corresponding modulus. Then,

$$\begin{aligned}&\Big|H^M\Big(t, x, \nabla_h u_h^*(t + 1/2, x)\Big) - H^c\Big((t + 1)\Delta t, x, \nabla u^*((t + 1)\Delta t, x)\Big)\Big| \\ &\qquad\leq L_H\|\nabla_h u_h^*(t + 1/2, x) - \nabla u^*((t + 1)\Delta t, x)\| + \mathcal{O}(\Delta t).\end{aligned}$$

It is easy to deduce from the regularity of u^* (Assumption B) that $\Delta_h u_h^*(t + 1, \cdot)$ is Hölderian with exponent r. Then, using the fourth statement of Lemma 4.2.6 and the consistency estimate (4.5.2), we obtain that

$$\nabla_h u_h^*(t + 1/2, x) = \nabla_h u_h^*(t + 1, x) + \mathcal{O}(\Delta t h^{r-1}) = \nabla u^*((t + 1)\Delta t, x) + \mathcal{O}(\Delta t h^{r-1} + h^{1+r}).$$

The estimate (4.5.5) follows and estimate (4.5.6) can be proved similarly. □

Lemma 4.5.4. *For the divergence term, there holds: for all $(t, x) \in \mathcal{T} \times S$,*

$$\mathrm{div}_h(v_h^* m_h^*(t, x)) - \int_{B_h(x)} \mathrm{div}(v^* m^*)((t + 1)\Delta t, y)dy = \mathcal{O}(h^{r+d}) + \mathrm{div}_h(\epsilon_1); \tag{4.5.7}$$

$$\mathrm{div}_h(v_h^* m_h^*(t, x)) - \int_{B_h(x)} \mathrm{div}(v^* m^*)(t\Delta t, y)dy = \mathcal{O}(h^{r+d}) + \mathrm{div}_h(\epsilon_2), \tag{4.5.8}$$

where $\epsilon_1 = \mathcal{O}(h^{1+r+d})$, and $\epsilon_2 = \mathcal{O}(h^{2r+d})$.

Proof. In order to prove (4.5.7), let us decompose $v_h^* m_h^*$ as the sum of three terms, γ_1, γ_2, and γ_3, defined by

$$\begin{aligned}
\gamma_1(t,x) &= \Big(v_h^*(t,x) - v^*((t+1)\Delta t, x)\Big) m_h^*(t,x);\\
\gamma_2(t,x) &= v^*((t+1)\Delta t, x)\Big(m_h^*(t,x) - m_h^*(t+1,x)\Big);\\
\gamma_3(t,x) &= v^*((t+1)\Delta t, x) m_h^*(t+1,x).
\end{aligned}$$

Step 1: Estimation of γ_1. Since $m_h^*(t,x) = \mathcal{O}(h^d)$, we directly obtain with Lemma 4.5.3 that $\gamma_1(t,x) = \mathcal{O}(h^{1+r+d})$.

Step 2: Estimation of γ_2. By the definition of m_h^*, we have

$$m_h^*(t,x) - m_h^*(t+1,x) = \int_{B_h(x)} \Big(m^*(t\Delta t, y) - m^*((t+1)\Delta t, y)\Big) dy = \mathcal{O}(\Delta t h^d).$$

Then $\gamma_2 = \mathcal{O}(h^{2+d})$, since v^* is uniformly bounded.

Step 3: Estimation of $\mathrm{div}_h \gamma_3$. Using the definitions of γ_3 and m_h^*, we obtain that

$$\begin{aligned}
\mathrm{div}_h(\gamma_3)(t,x) &= \int_{B_h(0)} \mathrm{div}_h\Big(v^*((t+1)\Delta t, \cdot) m^*((t+1)\Delta t, \cdot + y)\Big)(x)\, dy\\
&= \int_{B_h(0)} \mathrm{div}\Big(v^*((t+1)\Delta t, \cdot) m^*((t+1)\Delta t, \cdot + y)\Big)(x)\, dy + \mathcal{O}(h^{r+d})\\
&= \int_{B_h(0)} \mathrm{div}\Big(v^*((t+1)\Delta t, \cdot + y) m^*((t+1)\Delta t, \cdot + y)\Big)(x)\, dy + \mathcal{O}(h^{r+d})\\
&= \int_{B_h(x)} \mathrm{div}(v^* m^*)((t+1)\Delta t)(y)\, dy + \mathcal{O}(h^{r+d}).
\end{aligned}$$

The second equality follows from the fact that $(v^* m^*)((t+1)\Delta t, \cdot + y) \in \mathcal{C}^{1+r}(\mathbb{T}^d)$. For the third one, we use that v^* and $D_x v$ are Hölderian with exponent r. Then, the estimate (4.5.7) holds true.

Step 4: Proof of (4.5.8). Since $v^* m^*(t\Delta t, \cdot)$ and $v^* m^*((t+1)\Delta t, \cdot)$ lie in $\mathcal{C}^{1+r}(\mathbb{T}^d)$, we first have that

$$\begin{aligned}
&\mathrm{div}(v^* m^*)(t\Delta t, y) - \mathrm{div}_h(v^* m^*)(t\Delta t, y) = \mathcal{O}(h^r);\\
&\mathrm{div}(v^* m^*((t+1)\Delta t, y)) - \mathrm{div}_h(v^* m^*)((t+1)\Delta t, y) = \mathcal{O}(h^r).
\end{aligned}$$

Since $v^* m^*(\cdot, y) \in \mathcal{C}^r([0,1])$, we have

$$v^* m^*((t+1)\Delta t, y) - v^* m^*(t\Delta t, y) = \mathcal{O}(\Delta t^r) = \mathcal{O}(h^{2r}).$$

Then we have

$$\begin{aligned}
&\int_{B_h(x)} \mathrm{div}_h\Big(v^* m^*((t+1)\Delta t, \cdot) - v^* m^*(t\Delta t, \cdot)(y)\Big) dy\\
&\qquad = \mathrm{div}_h\Big(\int_{B_h(0)} \big(v^* m^*((t+1)\Delta t, \cdot + y) - v^* m^*(t\Delta t, \cdot + y)\big) dy\Big)(x).
\end{aligned}$$

The right-hand side is a discrete divergence of a term of order $\mathcal{O}(h^{2r+d})$. The estimate (4.5.8) follows. □

We are ready to derive a complete consistency estimate for the triplet (u_h^*, v_h^*, m_h^*) defined at the beginning of the section.

Lemma 4.5.5 (Consistency error)**.** *The triplet* (u_h^*, v_h^*, m_h^*) *is a solution to the perturbed discrete mean field game* (PDMFG) *with perturbation terms* η *and* δ *satisfying*

$$\eta = \mathcal{O}(\Delta t h^r), \qquad B_1 \delta = \mathcal{O}(\Delta t h^{r+d}) + \Delta t \operatorname{div}_h(\epsilon_3), \qquad \textit{where } \epsilon_3 = \mathcal{O}(h^{2r+d}).$$

Proof. **Step 1.** The perturbation term η of the dynamic programming equation is defined by

$$-\frac{u_h^*(t+1,x) - u_h^*(t,x)}{\Delta t} - \sigma \Delta_h u_h^*(t+1/2,x) + H^M[\nabla_h u_h^*(\cdot+1/2,\cdot)](t,x) = f(t,x,m_h^*(t)) + \frac{\eta(t,x)}{\Delta t}.$$

The continuous HJB equation, satisfied by u^*, reads at time $(t+1)\Delta t$ as follows:

$$-\frac{\partial u^*((t+1)\Delta t, x)}{\partial t} - \sigma \Delta u^*((t+1)\Delta t, x) + H^c[\nabla u^*]((t+1)\Delta t, x) = f^c\Big((t+1)\Delta t, x, m^*((t+1)\Delta t)\Big).$$

Then η can be put in the form $\eta = \Delta t(r_1 + r_2 + r_3 + r_4)$, where

$$\begin{aligned}
r_1(t,x) &= \frac{\partial u^*((t+1)\Delta t, x)}{\partial t} - \frac{u_h^*(t+1,x) - u_h^*(t,x)}{\Delta t};\\
r_2(t,x) &= \sigma\Big(\Delta u^*((t+1)\Delta t, x) - \Delta_h u_h^*(t+1/2,x)\Big);\\
r_3(t,x) &= H^M[\nabla_h u_h^*(\cdot+1/2,\cdot)](t,x) - H^c[\nabla u^*]((t+1)\Delta t, x);\\
r_4(t,x) &= f^c\Big((t+1)\Delta t, x, m^*((t+1)\Delta t)\Big) - f(t,x,m_h^*(t)).
\end{aligned}$$

By (4.5.2), we have $r_1 = \mathcal{O}(\Delta t^{r/2}) = \mathcal{O}(h^r)$. Since $u^*((t+1)\Delta t, \cdot) \in \mathcal{C}^{2+r}(\mathbb{T}^d)$, it follows that $\Delta_h u_h^*(t+1,\cdot)$ is r-Hölder continuous. Using the fourth statement of Lemma 4.2.6 and (4.5.2), we obtain that

$$\Delta_h u_h^*(t+1/2,x) = \Delta_h u_h^*(t+1,x) + \mathcal{O}(\Delta t h^{r-2}) = \Delta u^*((t+1)\Delta t, x) + \mathcal{O}(\Delta t h^{r-2} + h^r).$$

Thus $r_2 = \mathcal{O}(h^r)$. Lemmas 4.5.2 and Lemma 4.5.3 yield $r_3 = \mathcal{O}(h^{1+r})$ and $r_4 = \mathcal{O}(h)$. It follows that $\eta(t,x) = \mathcal{O}(\Delta t h^r)$.

Step 2. For the estimation of the perturbation term of the discrete Fokker-Planck equation, we first show that m_h^* is the solution to

$$\frac{m_h^*(t+1,x) - m_h^*(t,x)}{\Delta t} - \sigma\theta \Delta_h m_h^*(t+1,x) - (1-\theta)\sigma\Delta_h m_h^*(t,x) + \operatorname{div}_h(v_h^* m_h^*(t,x)) = \frac{\bar{\delta}(t,x)}{\Delta t}, \tag{4.5.9}$$

for some error term $\bar{\delta}$. It directly follows from (4.5.9) that m_h^* is the solution to the perturbed Fokker-Planck equation in (PDMFG) with $\delta = B_1^{-1}\bar{\delta}$, i.e. $B_1\delta = \bar{\delta}$. Thus it only remains to calculate and to estimate $\bar{\delta}$. The Fokker-Planck equation, satisfied by m^*, writes as follows at times $t\Delta t$ and $(t+1)\Delta t$:

$$\begin{aligned}
&\frac{\partial m^*(t\Delta t, x)}{\partial t} - \sigma \Delta m^*(t\Delta t, x) + \operatorname{div}(v^* m^*(t\Delta t, x)) = 0;\\
&\frac{\partial m^*((t+1)\Delta t, x)}{\partial t} - \sigma \Delta m^*((t+1)\Delta t, x) + \operatorname{div}(v^* m^*((t+1)\Delta t, x)) = 0.
\end{aligned}$$

Let us integrate over $B_h(x)$ the convex combination of the last two equations:

$$(1-\theta)\int_{y\in B_h(x)} \frac{\partial m^*(t\Delta t,y)}{\partial t} - \sigma\Delta m^*(t\Delta t,y) + \operatorname{div}(v^* m^*(t\Delta t,y))dy$$
$$+\,\theta\int_{y\in B_h(x)} \frac{\partial m^*((t+1)\Delta t,y)}{\partial t} - \sigma\Delta m^*((t+1)\Delta t,y) + \operatorname{div}(v^* m^*((t+1)\Delta t,y))dy = 0.$$

Then $\bar{\bar{\delta}} = \Delta t(\bar{r}_1 + \bar{r}_2 + \bar{r}_3 + \tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3)$, where

$$\bar{r}_1(t,x) = \theta\Big(\frac{m_h^*(t+1,x) - m_h^*(t,x)}{\Delta t} - \int_{y\in B_h(x)} \frac{\partial m^*((t+1)\Delta t,y)}{\partial t}dy\Big);$$
$$\bar{r}_2(t,x) = \sigma\theta\Big(\int_{y\in B_h(x)} \Delta m^*((t+1)\Delta t,y)dy - \Delta_h m_h^*(t+1,x)\Big);$$
$$\bar{r}_3(t,x) = \theta\Big(\operatorname{div}_h(v_h^* m_h^*(t,x)) - \int_{y\in B_h(x)} \operatorname{div}(v^* m^*((t+1)\Delta t,y))dy\Big);$$
$$\tilde{r}_1(t,x) = (1-\theta)\Big(\frac{m_h^*(t+1,x) - m_h^*(t,x)}{\Delta t} - \int_{y\in B_h(x)} \frac{\partial m^*(t\Delta t,y)}{\partial t}dy\Big);$$
$$\tilde{r}_2(t,x) = \sigma(1-\theta)\Big(\int_{y\in B_h(x)} \Delta m^*(t\Delta t,y)dy - \Delta_h m_h^*(t,x)\Big);$$
$$\tilde{r}_3(t,x) = (1-\theta)\Big(\operatorname{div}_h(v_h^* m_h^*(t,x)) - \int_{y\in B_h(x)} \operatorname{div}(v^* m^*(t\Delta t,y))dy\Big).$$

Using the basic consistency estimates in (4.5.2) and the commutation property shown in (4.5.3), we have $\bar{r}_1 = \mathcal{O}(\Delta t^{r/2}h^d) = \mathcal{O}(h^{r+d})$, $\bar{r}_2 = \mathcal{O}(h^{r+d})$, $\tilde{r}_1 = \mathcal{O}(\Delta t^{r/2}h^d) = \mathcal{O}(h^{r+d})$, and $\tilde{r}_2 = \mathcal{O}(h^{r+d})$. Lemma 4.5.4 shows that $\bar{r}_3 = \mathcal{O}(h^{r+d}) + \theta\operatorname{div}_h\epsilon_1$ and $\tilde{r}_3 = \mathcal{O}(h^{r+d}) + (1-\theta)\operatorname{div}_h\epsilon_2$. Taking $\epsilon_3 = \theta\epsilon_1 + (1-\theta)\epsilon_2$, the conclusion follows. □

4.5.2 Proof of Theorem 4.2.10

All constants in the proof are independent of Δt and h. The existence and uniqueness of the solution (u_h, v_h, m_h) to the theta-scheme was established in Theorem 4.4.4. The triplet (u_h, v_h, m_h) is also the unique solution to (DMFG) with control bound M. We proved in Lemma 4.5.5 that (u_h^*, v_h^*, m_h^*) is a solution to (PDMFG) with perturbation terms η and δ estimated as follows:

$$\eta = \mathcal{O}(\Delta t h^r), \qquad B_1\delta = \mathcal{O}(\Delta t h^{r+d}) + \Delta t\operatorname{div}_h(\epsilon_3), \qquad \text{where } \epsilon_3 = \mathcal{O}(h^{2r+d}).$$

Step 1. Using similar arguments to the ones of the proof of Theorem 4.3.6 (see in particular estimate (4.6.6)), we easily obtain that

$$\|u_h^* - u_h\|_{\infty,\infty} \leq L_f^c \frac{\|m_h^* - m_h\|_{\infty,2}}{h^{d/2}} + \|\eta\|_{1,\infty}. \tag{4.5.10}$$

Step 2. Next we apply the fundamental inequality (Proposition 4.3.7) to (u_h, v_h, m_h) and (u_h^*, v_h^*, m_h^*). We obtain

$$\frac{\Delta t\alpha}{2}\big\|\|v_h^* - v_h\|^2(m_h^* + m_h)\big\|_{1,1} \leq \sum_{t\in\mathcal{T}}\sum_{x\in S}\Big((u_h^* - u_h)(t+1,x)\delta(t,x) + (m_h - m_h^*)(t,x)\eta(t,x)\Big). \tag{4.5.11}$$

Let us bound the right-hand side of the obtained inequality. Using the symmetry of B_1 and the convention (4.4.2), we first obtain that

$$\begin{aligned}\sum_{t\in\mathcal{T}}\sum_{x\in S}(u_h^*-u_h)(t+1,x)\delta(t,x) &= \sum_{t\in\mathcal{T}}\Delta t\left\langle B_1^{-1}(u_h^*-u_h)(t+1,\cdot),B_1\delta(t,\cdot)\right\rangle\\ &= \sum_{t\in\mathcal{T}}\Delta t\left\langle (u_h^*-u_h)(t+1/2,\cdot),B_1\delta(t,\cdot)\right\rangle\end{aligned}$$

It follows that there exist two constants C_0 and C_1 such that

$$\begin{aligned}&\sum_{t\in\mathcal{T}}\sum_{x\in S}(u_h^*-u_h)(t+1,x)\delta(t,x)\\ &\quad\leq \sum_{t\in\mathcal{T}}\sum_{x\in S}\Big(\Delta t(u_h^*-u_h)(t+1/2,x)\mathrm{div}_h(\epsilon_3(t,x))\Big)+\|u_h^*-u_h\|_{\infty,\infty}\|B_1\delta-\Delta t\mathrm{div}_h(\epsilon_3)\|_{1,1}\\ &\quad\leq \sum_{t\in\mathcal{T}}\sum_{x\in S}\Big(-\Delta t\Big\langle\nabla_h(u_h^*-u_h)(t+1/2,x),\epsilon_3(t,x)\Big\rangle\Big)+C_0\Big(L_f^c\frac{\|m_h^*-m_h\|_{\infty,2}}{h^{d/2}}+\|\eta\|_{1,\infty}\Big)h^r\\ &\quad\leq C_1\big(h^{2r}+\|m_h^*-m_h\|_{\infty,2}h^{r-d/2}\big).\end{aligned}$$

The first inequality is a consequence of Hölder's inequality and the second one derives from the discrete integration by parts formula combined with inequality (4.5.10). The third one follows from the Lipschitz continuity of $u_h^*(t+1/2,\cdot)$ and $u_h(t+1/2,\cdot)$, which was proved in Lemmas 4.4.3 and 4.5.1.

By Hölder's inequality, there also exists a constant C_2 such that

$$\sum_{t\in\mathcal{T}}\sum_{x\in S}(m_h-m_h^*)(t,x)\eta(t,x)\leq\|m_h^*-m_h\|_{\infty,2}\|\eta\|_{1,2}\leq C_2\|m_h^*-m_h\|_{\infty,2}h^{r-d/2}.$$

Then, there exists a constant C_3 such that

$$\epsilon\leq C_3\big(h^{2r}+\|m_h^*-m_h\|_{\infty,2}h^{r-d/2}\big),\quad \text{where: } \epsilon=\Delta t\left\|\|v_h^*-v_h\|^2m_h^*\right\|_{1,1}. \tag{4.5.12}$$

Step 3. We next find an upper bound of $\|m_h^*-m_h\|_{\infty,2}$ involving ϵ, using the energy estimate established in Proposition 4.4.5. Let $\mu=m_h^*-m_h$. Then μ satisfies the perturbed discrete Fokker-Planck equation defined in (4.4.6):

$$\left\{\begin{aligned}(\mathrm{Id}-\theta\sigma\Delta t\Delta_h)\,\mu(t+1)=&\ \big(\mathrm{Id}+(1-\theta)\sigma\Delta t\,\Delta_h\big)\mu(t)-\Delta t\,\mathrm{div}_h\big(v_h(t)\mu(t)\big)\\ &-\Delta t\,\mathrm{div}_h\big(\delta_v(t)\big)+\Delta t\,\delta'(t),\\ \mu(0)=&\ 0,\end{aligned}\right.$$

where

$$\delta_v(t,x)=(v_h^*-v_h)m_h^*(t,x)-\epsilon_3(t,x)\qquad\text{and}\qquad\delta'=\mathcal{O}(h^{r+d}).$$

From Theorem 4.4.4 we know that $\|v_h\|_{\infty,\infty}\leq M$. Thus, the energy inequality (4.4.7) implies that there exists a constant C_4 such that

$$\max_{t\in\tilde{\mathcal{T}}}\left\|\mu(t)\right\|_2^2\leq C_4\sum_{\tau\in\mathcal{T}}\Delta t\left(\left\|\delta_v(\tau)\right\|_2^2+\left\|\delta'(\tau)\right\|_2^2\right).$$

Applying inequality $(a+b)^2 \leq 2a^2 + 2b^2$ to $\|\delta_v(\tau)\|_2^2$, there exists a constant C_5 such that

$$\big\|\delta_v(\tau)\big\|_2^2 \leq 2\big\|(v_h^* - v_h)m_h^*(\tau)\big\|_2^2 + 2\big\|\epsilon_3(\tau)\big\|_2^2 \leq 2\big\|(v_h^* - v_h)m_h^*(\tau)\big\|_2^2 + C_5 h^{4r+d},$$

$$\big\|\delta'(\tau)\big\|_2^2 \leq C_5 h^{2r+d}.$$

Since $\|m_h^*\|_{\infty,\infty} = \mathcal{O}(h^d)$, there exists a constant C_6 such that

$$\sum_{\tau\in\mathcal{T}} \Delta t \left(\big\|\delta_v(\tau)\big\|_2^2 + \big\|\delta'(\tau)\big\|_2^2\right) \leq C_6 h^d\Big(\epsilon + h^{2r}\Big).$$

Therefore, for some constant C_7,

$$\|m_h^* - m_h\|_{\infty,2}^2 = \|\mu\|_{\infty,2}^2 \leq C_7 h^d\big(\epsilon + h^{2r}\big). \tag{4.5.13}$$

Step 4. Let us combine inequality (4.5.12) with (4.5.13). We obtain that

$$\begin{aligned}\|m_h^* - m_h\|_{\infty,2}^2 &\leq C_7(C_3+1)h^{2r+d} + C_7C_3\|m_h^* - m_h\|_{\infty,2}h^{r+d/2} \\ &\leq C_7(C_3+1)h^{2r+d} + \frac{\|m_h^* - m_h\|_{\infty,2}^2}{2} + \frac{C_7^2C_3^2}{2}h^{2r+d}.\end{aligned}$$

Therefore, for some constant C_8,

$$\|m_h^* - m_h\|_{\infty,2} \leq C_8 h^{r+d/2}. \tag{4.5.14}$$

Applying Hölder's inequality to (4.5.14) and using (4.5.10), we obtain the existence of a constant C_9 such that

$$\|u_h - u_h^*\|_{\infty,\infty} + \|m_h - m_h^*\|_{\infty,1} \leq C_9 h^r.$$

The conclusion follows.

4.6 Appendix

4.6.1 Technical lemmas and proofs

Proof of Lemma 4.2.4. We prove (4.2.4):

$$\begin{aligned}-\sum_{x\in S}\mu(x)\mathrm{div}_h\omega(x) &= -\sum_{x\in S}\sum_{i=1}^{d}\mu(x)\frac{\omega_i(x+he_i) - \omega_i(x-he_i)}{2h} \\ &= -\sum_{x\in S}\sum_{i=1}^{d}\omega_i(x)\frac{\mu(x-he_i) - \mu(x+he_i)}{2h} = \sum_{x\in S}\left\langle \nabla_h\mu(x), \omega(x)\right\rangle.\end{aligned}$$

We prove (4.2.5):

$$
\begin{aligned}
-\sum_{x\in S}\nu(x)\Delta_h\mu(x) &= -\sum_{x\in S}\sum_{y\in S}\nu(x)\Delta_h(x,y)\mu(y)\\
&= \frac{1}{h^2}\sum_{x\in S}\nu(x)\sum_{i=1}^{d}\big(2\mu(x)-(\mu(x+he_i)+\mu(x-he_i))\big)\\
&= \frac{1}{h^2}\sum_{i=1}^{d}\Big(\sum_{x\in S}\mu(x+he_i)\big(\nu(x+he_i)-\nu(x)\big)-\sum_{x\in S}\mu(x-he_i)\big(\nu(x)-\nu(x-he_i)\big)\Big)\\
&= \sum_{x\in S}\left\langle \nabla_h^+\nu(x),\nabla_h^+\mu(x)\right\rangle.
\end{aligned}
$$

The lemma is proved. □

Proof of Lemma 4.2.5. Applying the inequality $(a+b)^2 \le 2a^2+2b^2$, we obtain

$$
\begin{aligned}
\|\nabla_h\mu\|_2^2 &= \frac{1}{4h^2}\sum_{i=1}^{d}\sum_{x\in S}\big(\mu(x+he_i)-\mu(x-he_i)\big)^2\\
&\le \frac{1}{2h^2}\sum_{i=1}^{d}\sum_{x\in S}\big(\mu(x+he_i)-\mu(x)\big)^2+\big(\mu(x)-\mu(x-he_i)\big)^2 = \|\nabla_h^+\mu\|_2^2.
\end{aligned}
$$

Inequality (4.2.6) follows. □

Proof of Lemma 4.2.6. Let $r = c\Delta t/h^2$. Consider the mapping $\mathbb{S}_X(\mu)\colon \mathbb{R}(S)\to\mathbb{R}(S)$, defined by

$$
\mathbb{S}_X(\mu)(x) = \frac{1}{1+2dr}\left(r\sum_{j=1}^{d}\mu(x+he_j)+r\sum_{j=1}^{d}\mu(x-he_j)+X(x)\right). \tag{4.6.1}
$$

Then Y is a solution to (4.2.7) if and only if it is a fixed point of $\mathbb{S}_X$. For any μ_1 and μ_2 in $\mathbb{R}(S)$ and for any $x\in S$,

$$
\begin{aligned}
\Big|\mathbb{S}_X(\mu_1)(x)-\mathbb{S}_X(\mu_2)(x)\Big| &= \frac{1}{1+2dr}\left|r\sum_{j=1}^{d}(\mu_1-\mu_2)(x+he_j)+r\sum_{j=1}^{d}(\mu_1-\mu_2)(x-he_j)\right|\\
&\le \frac{2dr}{1+2dr}\,\|\mu_1-\mu_2\|_\infty.
\end{aligned} \tag{4.6.2}
$$

Therefore, $\mathbb{S}_X$ is a contraction for the $\|\cdot\|_\infty$ norm. As a consequence, it has a unique fixed point Y, which is then the unique solution to (4.2.7). Point (1) is proved.

Let us prove point (2). Assume that $X\ge 0$. Since $\mathbb{S}_X$ is a contraction, we have that $Y = \lim_{n\to\infty}\mathbb{S}_X^n(\mu)$ for any $\mu\in\mathbb{R}(S)$. In particular, taking $\mu = X$,

$$
Y = \lim_{n\to\infty}\mathbb{S}_X^n(X).
$$

It is easy to verify that for any $\mu\in\mathbb{R}(S)$, if $\mu\ge 0$, then $\mathbb{S}_X(\mu)\ge 0$. Therefore, we deduce that $\mathbb{S}_X^n(X)\ge 0$ for any n by induction, and therefore $Y\ge 0$. If we assume that $X\in\mathcal{P}(S)$, then for any $\mu\in\mathcal{P}(S)$, we can deduce that $\mathbb{S}_X(\mu)\in\mathcal{P}(S)$. This yields that $Y\in\mathcal{P}(S)$.

Point (3) is proved similarly, assuming that X is L-Lipschitz, observing that if μ is L-Lipschitz continuous, then $\mathbb{S}_X(\mu)$ is L-Lipschitz continuous.

Let us prove the last statement. Taking any $i \in \{1, 2, \dots, d\}$, we define $\bar{\omega}$, $\omega \in \mathbb{R}(S)$ as follows:

$$\bar{\omega}(x) = (\nabla_h Y)_i(x), \qquad \omega(x) = (\nabla_h X)_i(x), \qquad \forall x \in S.$$

Then $\bar{\omega}$ is the fixed point of $\mathbb{S}_\omega$ (replace X by ω in (4.6.1)). Let $\gamma = 2dr/(1+2dr)$. Using $\bar{\omega} = \lim_{n\to\infty} \mathbb{S}_\omega^n(\omega)$, we deduce from (4.6.2) that

$$\|\bar{\omega} - \omega\|_\infty \leq \sum_{k=0}^{\infty} \|\mathbb{S}_\omega^{k+1}(\omega) - \mathbb{S}_\omega^k(\omega)\|_\infty \leq \sum_{k=0}^{\infty} \gamma^k \|\mathbb{S}_\omega(\omega) - \omega\|_\infty = \frac{1}{1-\gamma}\|\mathbb{S}_\omega(\omega) - \omega\|_\infty.$$

It further follows that

$$\begin{aligned}\|\bar{\omega} - \omega\|_\infty &\leq \frac{\Delta t}{1-\gamma}\frac{c}{1+2dr}\big\|\Delta_h \omega\big\|_\infty \\ &\leq \frac{\Delta t}{1-\gamma}\frac{c}{1+2dr}\max_{x\in S}\left|\frac{\Delta_h X(x+he_i) - \Delta_h X(x-he_i)}{2h}\right| \leq 2^{\alpha-1} c L' \Delta t h^{\alpha-1},\end{aligned}$$

where the last inequality is a consequence of the α-Hölder continuity of $\Delta_h X$. Finally, let $\bar{\omega}_i^+(x) = (\nabla_h^+ Y)_i(x)$ and let $\omega_i^+(x) = (\nabla_h^+ X)_i(x)$. By the same argument, we have that $\|\bar{\omega}_i^+ - \omega_i^+\|_\infty \leq 2^{\alpha-1} c L'(\Delta t h^{\alpha-1})$. Then, for any $x \in S$, it follows from the triangle inequality that

$$\Big|\Delta_h Y(x) - \Delta_h X(x)\Big| = \Big|\sum_{i=1}^{d} \frac{\bar{\omega}_i^+(x) - \bar{\omega}_i^+(x-he_i)}{h} - \sum_{i=1}^{d} \frac{\omega_i^+(x) - \omega_i^+(x-he_i)}{h}\Big| = 2^\alpha d c L'(\Delta t h^{\alpha-2}).$$

The lemma is proved. □

Proof of Lemma 4.2.7. The differentiability of H^c with respect to p and the Lipschitz continuity of H_p^c are proved in [HUL93, Thm. 4.2.1, page 82]. For any $t_1, t_2 \in [0,1]$, we have

$$\begin{aligned} H^c(t_1,x,p) - H^c(t_2,x,p) &= \sup_{v_1\in\mathbb{R}^d}\Big(\langle -p, v_1\rangle - \ell^c(t_1,x,v_1)\Big) - \sup_{v_2\in\mathbb{R}^d}\Big(\langle -p, v_2\rangle - \ell^c(t_1,x,v_2)\Big) \\ &\leq \sup_{v\in\mathbb{R}^d}\Big(\ell^c(t_2,x,v) - \ell^c(t_1,x,v)\Big) \\ &\leq L_\ell^c |t_1 - t_2|. \end{aligned}$$

Using the relation of Fenchel, $-H_p^c(t,x,p) = \mathrm{argmax}_v \langle -p, v\rangle - \ell^c(t,x,v)$, and the continuous differentiability of ℓ^c , we have the first order optimality condition

$$p + \ell_v^c\Big(t, x, -H_p^c(t,x,p)\Big) = 0.$$

Fix $x \in \mathbb{T}^d$ and $p \in \mathbb{R}^d$. Take any t_1 and t_2 in $[0,1]$. By the above equation,

$$\ell_v^c\Big(t_1, x, -H_p^c(t_1,x,p)\Big) = \ell_v^c\Big(t_2, x, -H_p^c(t_2,x,p)\Big).$$

The strong convexity of ℓ^c implies that

$$\begin{aligned}
\alpha\big\|H_p^c(t_1,x,p) - H_p^c(t_2,x,p)\big\|^2 \\
&\leq \big\langle \ell_v^c\big(t_1,x,-H_p^c(t_1,x,p)\big) - \ell_v^c\big(t_1,x,-H_p^c(t_2,x,p)\big), H_p^c(t_2,x,p) - H_p^c(t_1,x,p)\big\rangle \\
&= \big\langle \ell_v^c\big(t_2,x,-H_p^c(t_2,x,p)\big) - \ell_v^c\big(t_1,x,-H_p^c(t_2,x,p)\big), H_p^c(t_2,x,p) - H_p^c(t_1,x,p)\big\rangle \\
&\leq L_\ell^c\big|t_1 - t_2\big|\big\|H_p^c(t_1,x,p) - H_p^c(t_2,x,p)\big\|,
\end{aligned}$$

where the last inequality is a consequence of the Lipschitz continuity of ℓ_v^c with respect to t. The lemma is proved. □

Proof of Lemma 4.3.2. The first three claims can be shown with the same arguments as those of the proof of Lemma 4.2.7. Since H_p^D is $\frac{1}{\alpha}$-Lipschitz continuous with respect to p, it is enough to prove (4.3.4) for $p = 0$. Let $v(t,x) = -H_p^D(t,x,0)$. Since $v(t,x)$ is optimal in (4.3.3), with $p = 0$, we deduce that $v(t,x)$ minimizes $\ell(t,x,\cdot)$ over the closed ball of radius D. Using the strong convexity of ℓ, it follows that

$$\ell(t,x,0) + \langle p_0, v(t,x)\rangle + \frac{\alpha}{2}\|v(t,x,0)\|^2 \leq \ell(t,x,v(t,x)) \leq \ell(t,x,0),$$

from which we deduce that $\|v(t,x)\| \leq \frac{2}{\alpha}\|p_0\|$, by Cauchy-Schwarz inequality. □

Proof of Lemma 4.3.4. By Fenchel's relation [HUL93, Cor. 1.4.4], we know that

$$H^D(t,x,\bar{p}) = -\langle \bar{p}, \bar{v}\rangle - \ell^D(t,x,\bar{v}) \quad \text{and} \quad -\bar{p} \in \partial_v \ell^D(t,x,\bar{v}). \tag{4.6.3}$$

Using the strong convexity of ℓ^D, we obtain that

$$\ell^D(t,x,v) \geq \ell^D(t,x,\bar{v}) - \langle \bar{p}, v - \bar{v}\rangle + \frac{\alpha}{2}\|v - \bar{v}\|^2. \tag{4.6.4}$$

Summing up (4.6.3) and (4.6.4), we obtain the following inequality:

$$H^D(t,x,\bar{p}) + \ell^D(t,x,v) + \langle \bar{p}, v\rangle \geq \frac{\alpha}{2}\left\|v - \bar{v}\right\|^2. \tag{4.6.5}$$

Multiplying (4.6.5) by $\bar{m}$, multiplying (4.6.3) by m, and taking the difference, we obtain the desired inequality. □

Proof of Theorem 4.3.6, second part. We prove here the continuity of the mapping ϕ. Since ϕ is the composition of (4.3.7), (4.3.8) and (4.3.6), it suffices to show that these three mappings are continuous.

Step 1: Continuity of **HJB**. Take any μ_1 and μ_2 in $\mathcal{P}_{m_0}(\tilde{\mathcal{T}}, S)$. Let $u_1 = \mathbf{HJB}(\mu_1)$ and $u_2 = \mathbf{HJB}(\mu_2)$. By Assumption 4.2, we have that for any $x \in S$,

$$\begin{aligned}
|(u_1 - u_2)(t,x)| &\leq \sup_{\|\omega\| \leq D} \big|\tilde{\ell}_{\mu_1}(t,x,\omega) - \tilde{\ell}_{\mu_2}(t,x,\omega)\big|\Delta t + \Big|\sum_{y\in S}\pi(t,x,y,\omega)\big(u_1(t+1,y) - u_2(t+1,y)\big)\Big| \\
&\leq L_f'\|(\mu_1 - \mu_2)(t,\cdot)\|_2\Delta t + \|(u_1 - u_2)(t+1,\cdot)\|_\infty,
\end{aligned}$$

where the last inequality follows from the Lipschitz continuity of f and Assumption 4.2. Since $\mu_1(t,\cdot), \mu_2(t,\cdot) \in \mathcal{P}(S)$ for any $t \in \tilde{\mathcal{T}}$, we have that $\mu_1(t,s), \mu_2(t,s) \in [0,1]$ for any $(t,s) \in \tilde{\mathcal{T}} \times S$, which implies that $\|\mu_1 - \mu_2\|_{\infty,\infty} \leq 1$. Combining this with the fact that $\|(u_2 - u_1)(T, \cdot)\|_\infty = 0$, it follows that

$$\|u_1 - u_2\|_{\infty,\infty} \leq L'_f \Delta t \sum_{t \in \mathcal{T}} \|(\mu_1 - \mu_2)(t, \cdot)\|_2 \leq L'_f \|\mu_1 - \mu_2\|_{\infty,2} \leq L'_f \|\mu_1 - \mu_2\|_{\infty,1}^{1/2}. \tag{4.6.6}$$

Step 2: Continuity of $\mathbf{V}$. Let $v_1 = \mathbf{V}(u_1)$ and $v_2 = \mathbf{V}(u_2)$. By the equivalent form of (4.3.8), we have $v_1(t,x) = -H_p[p_{1,1}](t,x)$ and $v_2(t,x) = -H_p[p_{1,2}](t,x)$, where

$$p_{1,1}(t,x) = \sum_{s \in S} \pi_1(t,x,s) u_1(t+1,s) \quad \text{and} \quad p_{1,2}(t,x) = \sum_{s \in S} \pi_1(t,x,s) u_2(t+1,s).$$

By the $(1/\alpha)$-Lipschitz continuity of $H_p(t,x,p)$ on p, we have for any $(t,x) \in \mathcal{T} \times S$

$$\|v_1(t,x) - v_2(t,x)\| \leq \frac{1}{\alpha} \|p_{1,1}(t,x) - p_{1,2}(t,x)\| \leq \frac{1}{\alpha} \|\pi_1\|_{\infty,\infty,1} \|u_1 - u_2\|_{\infty,\infty},$$

where $\|\pi_1\|_{\infty,\infty,1} = \max_{t,x} \sum_s \|\pi_1(t,x,s)\|$.

Step 3: Continuity of $\mathbf{FP}$. Let $m_1 = \mathbf{FP}(v_1)$ and $m_2 = \mathbf{FP}(v_2)$. Then,

$$\begin{cases} (m_1 - m_2)(t+1, y) = \sum_{x \in S} \pi[v_1](t,x,y)(m_1 - m_2)(t,x) + \delta_{v_1,v_2,m_2}(t,y), \\ (m_1 - m_2)(0,y) = 0, \end{cases}$$

where $\delta_{v_1,v_2,m_2}(t,y) = \Delta t \sum_{x \in S} \pi_1(t,x,y)(v_1 - v_2)(t,x) m_2(t,x)$. Since $\pi[v_1]$ is a transition process, we have

$$\|m_1(t+1,\cdot) - m_2(t+1,\cdot)\|_1 \leq \|m_1(t,\cdot) - m_2(t,\cdot)\|_1 + \|\delta_{v_1,v_2,m_2}(t,\cdot)\|_1.$$

The second term $\|\delta_{v_1,v_2,m_2}(t,\cdot)\|_1$ is estimated with Hölder's inequality:

$$\begin{aligned} \|\delta_{v_1,v_2,m_2}(t,\cdot)\|_1 &= \Delta t \Big| \sum_{y \in S, x \in S} \pi_1(t,x,y)(v_1 - v_2)(t,x) m_2(t,x) \Big| \\ &\leq \Delta t \|m_2(t,\cdot)\|_1 \|\pi_1\|_{\infty,\infty,1} \|v_1 - v_2\|_{\infty,\infty} = \Delta t \|\pi_1\|_{\infty,\infty,1} \|v_1 - v_2\|_{\infty,\infty}, \end{aligned}$$

where the last equality is a consequence of $m_2 \in \mathcal{P}(\tilde{\mathcal{T}}, S)$. Therefore, we have

$$\|m_1 - m_2\|_{\infty,1} \leq \|\pi_1\|_{\infty,\infty,1} \|v_1 - v_2\|_{\infty,\infty}.$$

The continuity of ϕ follows. □

4.6.2 On the regularity of the continuous MFG system

Recall that $Q = [0,1] \times \mathbb{T}^d$. For any $R > 0$, let $\mathbf{B}_R \coloneqq Q \times B(0,R)$, where $B(0,R)$ is the closed ball in $\mathbb{R}^d$ with center 0 and radius R. Let us refer the reader fto [Kry08, pages 8 and 51] for the definitions of the Sobolev space $W_p^k(Q)$ and the anisotropic Sobolev space $W_p^{1,2}(Q)$. For any $\delta \in (0,1)$, we define the local Hölder space

$$\mathcal{C}_{\text{loc}}^{\delta/2,\delta,\delta}(Q \times \mathbb{R}^d) = \Big\{ w \in \mathcal{C}(Q \times \mathbb{R}^d) \;\Big|\; w\,|_{\mathbf{B}_R} \in \mathcal{C}^{\delta/2,\delta,\delta}(\mathbf{B}_R), \text{ for any } R > 0 \Big\},$$

where $w\,|_{\mathbf{B}_R}$ is the restriction of w in $\mathbf{B}_R$ and where $\mathcal{C}^{\delta/2,\delta,\delta}(\mathbf{B}_R)$ denotes the sets of functions from $\mathbf{B}_R$ to $\mathbb{R}$ which are Hölder continuous with respect to their first (resp. second and third) variable with exponent $\delta/2$ (resp. δ).

Assumption C. There exist $C > 0$ and $0 < \bar{r} < 1$ such that for all $(t, x) \in Q$, for all $v \in \mathbb{R}^d$ and for all $m \in \mathbb{L}^\infty(\mathbb{T}^d)$ satisfying $m \geq 0$ and $\int_{\mathbb{T}^d} m(x)dx = 1$, it holds:

- $\ell^c(t, x, v) \leq C\|v\|^2 + C$ and $|f^c(t, x, m)| \leq C$;
- ℓ^c and ℓ^c_v are continuously differentiable, and $\ell^c_{vx}, \ell^c_{vv} \in \mathcal{C}^{\bar{r}/2,\bar{r},\bar{r}}_{\text{loc}}(Q \times \mathbb{R}^d)$;
- $m^c_0 \in \mathcal{C}^{2+\bar{r}}(\mathbb{T}^d)$, and $g^c \in \mathcal{C}^3(\mathbb{T}^d)$.

Lemma 4.6.1. *Let Assumptions A and C hold true. Then the Hamiltonian H^c is continuously differentiable and H^c_p is also continuously differentiable. Moreover, $H^c_{px} \in \mathcal{C}^{\bar{r}/2,\bar{r},\bar{r}}_{\text{loc}}(Q \times \mathbb{R}^d)$ and $H^c_{pp} \in \mathcal{C}^{\bar{r}/2,\bar{r},\bar{r}}_{\text{loc}}(Q \times \mathbb{R}^d)$.*

Proof. Fix any $(t_0, x_0, p_0) \in (0, 1) \times \mathbb{T}^d \times \mathbb{R}^d$. By the strong convexity of ℓ^c w.r.t. v, there exists a unique $v_0 \in \mathbb{R}^d$ such that $H^c(t_0, x_0, p_0) = -\langle p_0, v_0\rangle - \ell^c(t_0, x_0, v_0)$. The first order optimality condition writes

$$-p_0 - \ell^c_v(t_0, x_0, v_0) = 0.$$

Since ℓ^c is strongly convex, we have that $\ell^c_{vv}(t_0, x_0, v_0)$ is invertible. By the implicit function theorem, there exist a neighborhood $\mathcal{A}$ of (t_0, x_0, p_0) and a function $v(t, x, p)$ from $\mathcal{A}$ to $\mathbb{R}^d$ such that for all $(t, x, p) \in \mathcal{A}$,

$$-p - \ell^c_v(t, x, v(t, x, p)) = 0. \tag{4.6.7}$$

Since ℓ^c_v is continuously differentiable, $v(t, x, p)$ is continuously differentiable. Moreover,

$$\begin{aligned} v_x(t, x, p) &= \Big(\ell^c_{vv}(t, x, v(t, x, p))\Big)^{-1} \ell^c_{vx}(t, x, v(t, x, p)), \\ v_p(t, x, p) &= \Big(\ell^c_{vv}(t, x, v(t, x, p))\Big)^{-1}. \end{aligned}$$

By the regularity of ℓ^c_{vv} and ℓ^c_{vx}, we deduce that $v_x, v_p \in \mathcal{C}^{\bar{r}/2,\bar{r},\bar{r}}(\mathcal{A})$. The convexity of ℓ^c and the first order optimality condition (4.6.7) imply that

$$H^c(t, x, p) = -\langle p, v(t, x, p)\rangle - \ell^c(t, x, v(t, x, p)), \qquad \forall (t, x, p) \in \mathcal{A}.$$

We deduce that $H^c\mid_{\mathcal{A}} \in \mathcal{C}^1(\mathcal{A})$ by the regularity of v and ℓ^c. Differentiating the above equation with respect to p and using (4.6.7), we obtain that $H^c_p(t, x, p) = -v(t, x, p)$, for all $(t, x, p) \in \mathcal{A}$. Then, deriving H^c_p with respect to x and p, we obtain

$$H^c_{px}(t, x, p) = -v_x(t, x, p), \qquad H^c_{pp}(t, x, p) = -v_p(t, x, p), \qquad \forall (t, x, p) \in \mathcal{A}.$$

The conclusion follows from the regularity of v, v_x and v_p. □

Theorem 4.6.2. *Under Assumptions A and C, the continuous system* (MFG) *has a unique solution (u^*, v^*, m^*) satisfying Assumption B for any $r < \bar{r}$.*

Proof. Fixing any $0 < r < \bar{r}$, we will prove that Assumption B is satisfied for r. Under Assumptions A and C, according to [BHP21, Thm. 1], there exists $r' \in (0, \bar{r}]$ such that the continuous system (MFG) has a unique classical solution (u^*, v^*, m^*) with

$$u^*,\ m^* \in \mathcal{C}^{1+r'/2,2+r'}(Q), \quad v^* \in \mathcal{C}^{r'}(Q), \quad \text{and} \quad \nabla v^* \in \mathcal{C}^{r'}(Q, \mathbb{R}^{d\times d}). \tag{4.6.8}$$

Step 1: Regularity of ∇u^*. By (4.6.8), we have that $u^* \in \mathcal{C}^1(Q)$. This implies that $\nabla u^* \in \mathbb{L}^\infty(Q)$. Let $u^*_{x_i}$ be the partial derivative of u^* w.r.t. x_i. Then, $u^*_{x_i}$ is a weak solution of the following linear equation:

$$\begin{cases} -\partial_t w(t,x) - \sigma\Delta w(t,x) + H^c_p[\nabla u^*]\nabla w(t,x) = \check{f}_0(t,x), & (t,x) \in Q\,, \\ w(1,x) = g_{x_i}(x), & x \in \mathbb{T}^d, \end{cases}$$

where

$$\check{f}_0(t,x) = D_{x_i} f^c(t,x,m^*(t)) - H^c_{x_i}[\nabla u^*](t,x),$$

where D_{x_i} denotes the weak derivative w.r.t. x_i. By the regularity of H^c_p, H^c, f^c, and u^*, we deduce that $H^c_p[\nabla u^*] \in \mathbb{L}^\infty(Q)$ and $\check{f}_0 \in \mathbb{L}^\infty(Q)$. Moreover, the regularity of g implies that $g_{x_i} \in W^2_\infty(\mathbb{T}^d)$. Then [BHP21, Thm. 4] shows that $u^*_{x_i}$ is the unique weak solution and $u^*_{x_i} \in W^{1,2}_p(Q) \subset W^1_p(Q)$ for any $p > d+1$. Morrey's inequality [AF03, Lem. 4.28] implies that

$$u^*_{x_i} \in \mathcal{C}^\gamma(Q), \qquad \text{with} \;\; \gamma = 1 - \frac{d+1}{p}.$$

Taking $p = \frac{d+1}{1-r/\bar{r}}$, we have that $u^*_{x_i} \in \mathcal{C}^{r/\bar{r}}(Q)$. The same result follows for ∇u^*.

Step 2: Regularity of u^*. Let $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ such that $\varphi(x) = 1$ for $x \in B(0, 2\sqrt{d})$ and $\varphi(x) = 0$ for $x \notin \Omega \coloneqq B(0, 3\sqrt{d})$. It is straightforward that $B(0, 2\sqrt{d})$ contains a neighborhood of $\mathbb{T}^d$. Let us set $Q' = (0,1) \times \Omega$.

Since u^* can be identified to a periodic function over $\mathbb{R}^d$, we define $\check{u} = u^*\varphi$. Then, $\check{u}$ is the solution of the following equation:

$$\begin{cases} -\partial_t \check{u}(t,x) - \sigma\Delta\check{u}(t,x) = \check{f}_1(t,x), & (t,x) \in Q'\,, \\ \check{u}(t,x) = 0, & (t,x) \in (0,1) \times \partial\Omega, \\ \check{u}(1,x) = g(x)\varphi(x), & x \in \Omega, \end{cases}$$

where

$$\check{f}_1(t,x) = \varphi(x)\Big(f^c(t,x,m^*(t)) - H^c(t,x,\nabla u^*(t,x))\Big) - 2\sigma\langle\nabla\varphi(x), \nabla u^*(t,x)\rangle - \sigma u^*(t,x)\Delta\varphi(x).$$

By the regularity of f^c and m^*, we deduce the following: For any $(t_1,x_1),(t_2,x_2) \in Q'$,

$$\begin{aligned} f^c(t_1,x_1,m^*(t_1)) - f^c(t_2,x_2,m^*(t_2)) &\leq L^c_f(|t_1-t_2| + \|x_1-x_2\|) + L^c_f\|m^*(t_1) - m^*(t_2)\|_{\mathbb{L}^2} \\ &\leq L^c_f(|t_1-t_2| + \|x_1-x_2\|) + L^c_f\|m^*(t_1) - m^*(t_2)\|_{\mathbb{L}^\infty} \\ &\leq C(|t_1-t_2|^{r/2} + \|x_1-x_2\|^r), \end{aligned}$$

for some constant C. Using the regularity properties of u^*, ∇u^* and H^c, we obtain that $\check{f}_1 \in \mathcal{C}^{r/2,r}(Q')$. The final condition lies in $\mathcal{C}^r(\Omega)$ by Assumption C. The boundary conditions satisfying the requirements in [LSU88, Thm. 5.2], we deduce that $\check{u} \in \mathcal{C}^{1+r/2,2+r}(\bar{Q}')$, where $\bar{Q}'$ is the closure of Q'. By the definition of φ, we have that $u^*(t,x) = \check{u}(t,x)$ for all $(t,x) \in Q$. The regularity of u^* follows.

Step 3: Regularity of v^*. By (MFG) (ii) and the regularity of H^c, we have

$$\begin{aligned} v^*(t,x) &= -H^c_p(t,x,\nabla u^*(t,x)); \\ \nabla v^*(t,x) &= -H^c_{px}(t,x,\nabla u^*(t,x)) - H^c_{pp}(t,x,\nabla u^*(t,x)) D_{xx} u^*(t,x). \end{aligned}$$

Since H^c_p is continuously differentiable and $\nabla u^* \in \mathcal{C}^{r/\bar{r}}(Q)$, we deduce that $v^* \in \mathcal{C}^{r/\bar{r}}(Q) \subset \mathcal{C}^r(Q)$. By a similar argument, from the regularity of H^c_{px}, and H^c_{pp} in Lemma 4.6.1 and the regularity of ∇u^* and $D_{xx}u^*$, we have $\nabla v^* \in \mathcal{C}^{r/2,r}(Q)$.

Step 4: Regularity of m^*. Since $m^* \in \mathcal{C}^{1+r'/2,2+r'}(Q)$ and Q is bounded, we have $m^* \in W^{1,2}_p(Q)$ for any $p > d+2$. By [LSU88, Lem. 3.3], it holds that

$$m^* \in \mathcal{C}^{\gamma/2,\gamma}(Q) \text{ and } \nabla m^* \in \mathcal{C}^{\gamma/2,\gamma}(Q), \text{ with } \gamma = 1 - \frac{d+2}{p}.$$

Taking $p = \frac{d+2}{1-r}$, it follows that $m^* \in \mathcal{C}^{r/2,r}(Q)$ and $\nabla m^* \in \mathcal{C}^{r/2,r}(Q)$.

Let us define $\check{m} = m^*\varphi$. Then $\check{m}$ satisfies the following equation:

$$\begin{cases} \partial_t \check{m}(t,x) - \sigma\Delta\check{m}(t,x) + \langle v^*, \nabla\check{m}\rangle(t,x) + \operatorname{div}(v^*)\check{m}(t,x) = \check{f}_2(t,x), & (t,x) \in Q', \\ \check{m}(t,x) = 0, & (t,x) \in (0,1) \times \partial\Omega, \\ \check{m}(0,x) = m_0(x)\varphi(x), & x \in \Omega, \end{cases}$$

where

$$\check{f}_2(t,x) = -2\sigma\langle\nabla\varphi(x), \nabla m^*(t,x)\rangle - \sigma m^*(t,x)\Delta\varphi(x) + \langle v^*(t,x), \nabla\varphi(x)\rangle m^*(t,x).$$

From the regularity of v^*, m^* and ∇m^*, we deduce that $\check{f}_2 \in \mathcal{C}^{r/2,r}(Q')$. Combining with the regularity of v^*, ∇v^* and $m_0\varphi$, we deduce that $\check{m} \in \mathcal{C}^{1+r/2,2+r}(\bar{Q'})$ by [LSU88, Thm. 5.2]. Therefore, $m^* \in \mathcal{C}^{1+r/2,2+r}(Q)$. □

4.6.3 Construction of a numerical Hamiltonian

This section, as a complementary material to the rest of the article, is dedicated to the construction of a numerical Hamiltonian satisfying the assumptions of [AP16], in a general framework (see equation (4.6.11)). Our main assumption is the strong convexity of the running cost with respect to the control variable.

Given a vector $q \in \mathbb{R}^{2d}$, we denote

$${}^\dagger q = (q_1, q_3, \ldots, q_{2d-1}), \quad q^\dagger = (q_2, q_4, \ldots, q_{2d}).$$

Following the terminology of [AP16], we call *numerical Hamiltonian* a function $\mathcal{H}\colon [0,1]\times\mathbb{T}^d\times\mathbb{R}^{2d} \to \mathbb{R}$ satisfying the following conditions: For any $(t,x) \in [0,1] \times \mathbb{T}^d$,

(g1) [Monotonicity] $\mathcal{H}(t,x,\cdot)$ is nonincreasing w.r.t. ${}^\dagger q_i$ and nondecreasing w.r.t. $q^\dagger_i$ for all $i = 1, 2, \ldots, d$;

(g2) [Consistency] For any q such that ${}^\dagger q = q^\dagger$, it holds $\mathcal{H}(t,x,q) = H^c(t,x,q^\dagger)$;

(g3) [Regularity] $\mathcal{H}(t,x,\cdot)$ is continuously differentiable;

(g4) [Convexity] $\mathcal{H}(t,x,\cdot)$ is convex;

(g5) There exists positive constants c_1, c_2, c_3 and c_4, independent of (t,x), such that for any $q \in \mathbb{R}^{2d}$,

$$\langle \mathcal{H}_q(t,x,q), q\rangle - \mathcal{H}(t,x,q) \geq c_1\|\mathcal{H}_q(t,x,q)\|^2 - c_2; \tag{4.6.9}$$

$$\|\mathcal{H}_q(t,x,q)\| \leq c_3\|q\| + c_4. \tag{4.6.10}$$

Lemma 4.6.3. *Consider a function $\mathcal{H}\colon [0,1]\times\mathbb{T}^d\times\mathbb{R}^{2d}\to\mathbb{R}$ satisfying* (g3)-(g4)*. Assume that $\mathcal{H}_q$ is uniformly Lipschitz continuous w.r.t. q, $\mathcal{H}(t,x,0)$ is bounded from above, and $\mathcal{H}_q(t,x,0)$ is uniformly bounded. Then $\mathcal{H}$ satisfies* (g5)*.*

Proof. Let L be the Lipschitz constant of $\mathcal{H}_q(t,x,\cdot)$. Applying inequality (4.1.2), we obtain that

$$\frac{1}{2L}\|\mathcal{H}_q(t,x,0)-\mathcal{H}_q(t,x,q)\|^2 \leq \mathcal{H}(t,x,0)-\mathcal{H}(t,x,q)+\langle\mathcal{H}_q(t,x,q),q\rangle.$$

Applying inequality $\|a-b\|^2 \geq \|b\|^2/2 - \|a\|^2$, we deduce that

$$\langle\mathcal{H}_q(t,x,q),q\rangle - \mathcal{H}(t,x,q) \geq \frac{1}{4L}\|\mathcal{H}_q(t,x,q)\|^2 - \frac{1}{2L}\|\mathcal{H}_q(t,x,0)\|^2 - \mathcal{H}(t,x,0).$$

Since $\mathcal{H}(t,x,0)$ is bounded from above and since $\mathcal{H}_q(t,x,0)$ is uniformly bounded, (4.6.9) is satisfied. Inequality (4.6.10) is obvious by the uniform Lipschitz continuity of $\mathcal{H}_q$. □

Assume that the running cost $\ell^c(t,x,v)\colon [0,1]\times\mathbb{T}^d\times\mathbb{R}^d\to\mathbb{R}$ is uniformly α^c-convex w.r.t. v with some $\alpha^c>0$. Then, ℓ^c can be decomposed as

$$\ell^c(t,x,v) = \ell_0^c(t,x,v) + \frac{\alpha^c}{2}\|v\|^2,$$

where ℓ_0^c is convex w.r.t. v. We propose the following definition for a numerical Hamiltonian:

$$\mathcal{H}(t,x,q) = \sup_{\substack{v\in\mathbb{R}^d,\, v\geq 0\\ u\in\mathbb{R}^d,\, u\leq 0}} \Big(-\langle v, {}^\dagger q\rangle - \langle u, q^\dagger\rangle - \ell_0^c(t,x,v+u) - \frac{\alpha^c}{2}\big(\|v\|^2+\|u\|^2\big)\Big). \tag{4.6.11}$$

Theorem 4.6.4. *Assume that $\ell^c\colon [0,1]\times\mathbb{T}^d\times\mathbb{R}^d\to\mathbb{R}$ is α^c-convex with respect to its third variable, ℓ^c is bounded from below by some constant c, and for some $v_0\in\mathbb{R}^d$, there exists a constant $C(v_0)<+\infty$ such that for all $(t,x)\in[0,1]\times\mathbb{T}^d$, $\ell^c(t,x,v_0)\leq C(v_0)$. Then the function $\mathcal{H}$ defined by* (4.6.11) *is a numerical Hamiltonian, for the Hamiltonian H^c defined by* (4.1.1)*.*

Proof. The condition (g1) can be easily deduced from the nonnegativity and nonpositivity constraints for v and u in (4.6.11).

Step 1: Proof of (g2). Let us take any $q\in\mathbb{R}^{2d}$, such that ${}^\dagger q = q^\dagger$. Then, we deduce that

$$\mathcal{H}(t,x,q) = \sup_{v\geq 0,\, u\leq 0} -\langle v+u, q^\dagger\rangle - \ell_0^c(t,x,v+u) - \frac{\alpha^c}{2}\big(\|v\|^2+\|u\|^2\big).$$

Since $v \geq 0$ and $u \leq 0$, we have that $\|v\|^2 + \|u\|^2 \geq \|u+v\|^2$. Then,

$$\mathcal{H}(t,x,q) \leq \sup_{v\geq 0,\, u\leq 0} -\langle v+u, q^\dagger\rangle - \ell_0^c(t,x,v+u) - \frac{\alpha^c}{2}(\|v+u\|^2) = H^c(t,x,q^\dagger).$$

Conversely, take $v^* = -H_p^c(t,x,q^\dagger)$, $v_+^* = \{\max\{0, v_i^*\}\}_{i=1,\dots d}$ and $v_-^* = \{\min\{0, v_i^*\}\}_{i=1,\dots d}$. We have that $v_+^* \geq 0$, $v_-^* \leq 0$, $v^* = v_+^* + v_-^*$ and $\|v^*\|^2 = \|v_+^*\|^2 + \|v_-^*\|^2$. Thus, by Fenchel's relation, it follows that

$$H^c(t,x,q^\dagger) = -\langle v_+^* + v_-^*, q^\dagger\rangle - \ell_0^c(t,x,v_+^*+v_-^*) - \frac{\alpha^c}{2}\big(\|v_+^*\|^2 + \|v_-^*\|^2\big) \leq \mathcal{H}(t,x,q).$$

Step 2: Proof of (g3)-(g4). Consider the function $\bar{\ell}^c\colon [0,1]\times\mathbb{T}^d\times\mathbb{R}^{2d}\to\mathbb{R}$ defined by

$$\bar{\ell}^c(t,x,w) = \ell_0^c(t,x,{}^\dagger w + w^\dagger) + \frac{\alpha^c}{2}(\|{}^\dagger w\|^2 + \|w^\dagger\|^2) + \chi^+({}^\dagger w) + \chi^-(w^\dagger),$$

where $\chi^+(x) = 0$ (resp. $\chi^-(x) = 0$) if $x \geq 0$ (resp. $x \leq 0$) and infinity otherwise. It is obvious that $\bar{\ell}^c$ is uniformly α^c-convex w.r.t. w. The definition (4.6.11) implies that

$$\mathcal{H}(t,x,q) = (\bar{\ell}^c)^*(t,x,-q).$$

By [HUL93, Thm. 4.2.1], $\mathcal{H}$ is convex and continuously differentiable w.r.t. q and $\mathcal{H}_q$ is uniformly $1/\alpha^c$-Lipschitz w.r.t. q.

Step 3: Proof of (g5). We apply Lemma 4.6.3 for the proof. Taking $q=0$, by the consistency of $\mathcal{H}$, we have for any $(t,x)\in[0,1]\times\mathbb{T}^d$ that

$$\mathcal{H}(t,x,0) = H^c(t,x,0) = -\inf_{v\in\mathbb{R}^d}\Big(\ell_0^c(t,x,v) + \frac{\alpha^c}{2}\|v\|^2\Big) \leq -c,$$

By Fenchel's relation, it follows that

$$-\mathcal{H}_q(t,x,0) = \operatorname*{argmin}_{v\geq 0, u\leq 0}\Big(\ell_0^c(t,x,v+u) + \frac{\alpha^c}{2}(\|v\|^2 + \|u\|^2)\Big).$$

Let us set $v^*(t,x) = \operatorname{argmin}_{v\in\mathbb{R}^d}\ell^c(t,x,v)$. By a similar argument to the one of Step 1, we have that ${}^\dagger\mathcal{H}_q(t,x,0) = -v^*(t,x)_+$ and $\mathcal{H}_q(t,x,0)^\dagger = -v^*(t,x)_-$. In order to prove that $\mathcal{H}_q(t,x,0)$ is uniformly bounded, it suffices to show the boundedness of $v^*(t,x)$. By the strong convexity and boundedness assumptions of ℓ^c, we deduce that for any $(t,x)\in[0,1]\times\mathbb{T}^d$,

$$C(v_0) \geq \ell^c(t,x,v_0) \geq \ell^c(t,x,v^*(t,x)) + \frac{\alpha^c}{2}\|v^*(t,x)-v_0\|^2 \geq c + \frac{\alpha^c}{2}\|v^*(t,x)-v_0\|^2.$$

This implies that $\|v^*\|_\infty \leq \|v_0\| + \sqrt{2(C(v_0)-c)/\alpha^c}$. The conclusion follows. □

Chapter 5

A mesh-independent method for second-order potential mean field games

5.1 Introduction

5.1.1 Context and main contributions

This article is concerned with the numerical resolution of second-order mean field games (MFGs). These models describe the asymptotic behavior of Nash equilibria in stochastic differential games, as the number of players goes to infinity. They were introduced independently in [LL07] and [HMC06] and have applications in various domains, such as economics, biology, finance and social networks, see [DTT17]. In this article, we consider the following standard second-order MFG on the space $Q \coloneqq [0,1] \times \mathbb{T}^d$,

$$\begin{cases} \text{(i)} & -\partial_t u - \sigma \Delta u + H^c\left(t,x,\nabla u(t,x)\right) = f^c(t,x,m(t)) & (t,x) \in Q, \\ \text{(ii)} & v(t,x) = -H_p^c\left(t,x,\nabla u(t,x)\right) & (t,x) \in Q, \\ \text{(iii)} & \partial_t m - \sigma \Delta m + \operatorname{div}(vm) = 0 & (t,x) \in Q, \\ \text{(iv)} & m(0,x) = m_0^c(x), \quad u(1,x) = g^c(x) & x \in \mathbb{T}^d, \end{cases} \tag{MFG}$$

where the Hamiltonian H^c is related to the Fenchel conjugate of a running cost $\ell^c \colon Q \times \mathbb{R}^d \to \mathbb{R}$:

$$H^c(t,x,p) \coloneqq \sup_{v \in \mathbb{R}^d} \langle -p, v \rangle - \ell^c(t,x,v).$$

The existence and uniqueness of the classical solution of (MFG) is proved in [LL07] under assumptions on the coupling function f^c.

Several discretization schemes have been proposed and analyzed for the resolution of (MFG). They consist of a backward discrete Hamilton-Jacobi-Bellman (HJB) equation and a forward discrete Fokker-Planck (FP) equation: They preserve the nature of the problem as a coupled system of two equations. An implicit finite difference scheme has been introduced in [ACD10] and convergence results for this scheme have been obtained in [ACCD13] and [AP16]. Other schemes, based on semi-Lagrangian discretizations have been investigated in [CS14, CS15, HS19] for first-order and (possibly) degenerate second-order MFGs. This work will focus on a scheme called theta-scheme, recently introduced by the authors in [BLP22]. In short, this scheme involves a Crank-Nicolson discretization of the diffusion term and an explicit discretization of the first-order non-linear term.

The resolution of the discretized coupled system is in general a difficult task. We restrict our attention to the case of potential MFGs, for which optimization methods can be leveraged. The system (MFG) is said to be potential (or variational) if there exists a function $F^c\colon [0,1] \times \mathcal{D} \to \mathbb{R}$ such that for any $t \in [0,1]$ and $m_1, m_2 \in \mathcal{D}$ (see the definition of $\mathcal{D}$ in (5.4.1)),

$$F^c(t,m_1) - F^c(t,m_2) = \int_0^1 \int_{x\in\mathbb{T}^d} f^c(t,x,m_1 + s(m_2-m_1))(m_2(x)-m_1(x))dxds. \tag{5.1.1}$$

We further restrict our attention to the case of a convex potential function F^c. In the presence of such F^c, the system (MFG) can be interpreted as the first-order optimality condition of an optimal control problem driven by the FP equation,

$$\begin{cases} \inf_{(m,v)} J^c(m,v) \coloneqq \int_Q \ell^c(t,x,v)m(t,x)dtdx + \int_0^1 F^c(t,m(t))dt + \int_{\mathbb{T}^d} g^c(x)m(T,x)dx, \\ \text{such that} \begin{cases} \partial_t m - \sigma\Delta m + \operatorname{div}(vm) = 0, & \forall (t,x) \in Q, \\ m(0,x) = m_0^c(x), & \forall x \in \mathbb{T}^d. \end{cases} \end{cases} \tag{OC}$$

Problem (OC) is equivalent to a convex optimal control problem, obtained through the classical Benamou-Brenier transform [BCS17]. Then, some numerical algorithms can be applied to find a solution of (OC), such as ADMM [BC15, And17], the Chambolle-Pock algorithm [AL20], the fictitious play [HS19] and the generalized Frank-Wolfe (GFW) algorithm [LP22]. Some articles propose to discretize the optimal control problem (OC), see for example [LST10, And17]. In this context, it is very desirable that the potential structure of the continuous MFG is preserved at the level of the discretized coupled system, so that one can apply in a direct fashion suitable optimization methods to the discrete system. This is in particular the case for the implicit scheme proposed in [ACCD13] and solved in [AL20] with the Chambolle-Pock algorithm. As we establish in this article, the theta-scheme of [BLP22] also preserves the potential structure of the MFG system.

We focus in this article on the resolution of the discrete MFG system with the Generalized Frank-Wolfe (GFW) algorithm see [BLM09]. This algorithm is an iterative method, consisting in solving at each iteration a partially linearized version of the potential problem (OC). The linearized problem to be solved is equivalent to a stochastic optimal control problem that can be solved by dynamic programming. As we will explain more in detail, this allows to interpret the GFW method as a best-response procedure. For a specific choice of stepsize, it coincides with the fictitious play method of [CH17]. Others works have investigated the fictitious play method for MFGs: [PPL+20] proves the convergence of the continuous method, in a discrete setting with common noise; [HS19] proves a general result for fully discrete MFGs, which can be applied to discretized first-order MFGs. The article [GPL+22] shows the connexion between fictitious play and the Frank-Wolfe algorithm for potential discrete MFGs.

The general objective of the article is to show that the performance of the GFW algorithm is not impacted by a refinement of the discretization grid. The main results of our article are two mesh-independence properties for the resolution of (MFG) with the theta-scheme and the GFW algorithm. The terminology mesh-independence was coined in the article [ABPR86]. It is said that an algorithm satisfy a mesh-independence property when approximately the same number of iterations is required to satisfy a stopping criterion, when comparing an infinite-dimensional problem and its discrete counterpart. In a more precise fashion, we will say that the GFW algorithm has a mesh-independent sublinear rate of convergence if there exists a constant $C > 0$, independent of the discretization parameters, such that

$$J_h(m_h^k, v_h^k) - J_h^* \leq \frac{C}{k}, \quad \forall k \geq 1.$$

In the above estimate, J_h denotes the discretized counterpart of J, J_h^* denotes the value of the discretized optimal control problem, and (m_h^k, v_h^k) denotes the candidate to optimality obtained at iteration k. Similarly, we will say that the GFW algorithm has a mesh-independent linear rate of convergence if there exist two constants $C > 0$ and $\delta \in (0,1)$ such that

$$J_h(m_h^k, v_h^k) - J_h^* \leq C\delta^k, \quad \forall k \geq 1.$$

We establish that the GFW algorithm has a mesh-independent sublinear (resp. a linear) rate for two different choices of stepsize. Our analysis is close to the one performed in [LP22], in which the sublinear and the linear convergence of the GFW algorithm is demonstrated for the continuous model and for the same choices of stepsizes as in the present study. While the sublinear convergence of the GFW method (in a general setting) is classical, the linear rate of convergence relies on recent techniques from [KW22]. To the best of our knowledge, in the context of mean field games, the mesh-independence property has never been established so far for any other method. Though it seems a natural property, it may not hold in general. In particular, it might not hold for primal-dual methods, whose application relies on a saddle-point formulation of the convex counterpart of (OC) of the form, in which the Fokker-Planck equation is “dualized”. This saddle-point formulation involves a linear operator, encoding the (discrete) Fokker-Planck equation (see for example [AL20, Sec. 3.2]). As the discretization parameters decrease, the operator norm of these operators (for the Euclidean norm) increases, which has an impact on the convergence properties of methods such as the Chambolle-Pock algorithm. In contrast, the discrete Fokker-Planck equation remains satisfied at each iteration of the GFW equation.

This article is organized as follows. In Section 5.2, we introduce some preliminary results and notations. In Section 5.3, we introduce a general class of potential discrete MFGs, containing the theta-scheme. We establish the sublinear and the linear convergence of the GFW method in this discrete setting. We give explicit formulas for the convergence constants. These constants essentially depend on the Lipschitz-modulus of the coupling function of the MFG and on two bounds, for different norms, of the solution of the discretized Fokker-Planck equation, denoted C_1 and C_2. We recall the theta-scheme in Section 5.4, we show that it preserves the potential structure of the MFG, and we prove that the GFW algorithm has a mesh-independent sublinear and linear rates of convergence. The technical analysis relies on precise estimates of the constants C_1 and C_2, obtained thanks to a general energy estimate and an L^∞-estimate for the discrete Fokker-Planck equation.

5.1.2 Notation

We discretize the interval $[0,1]$ with a time step $\Delta t = 1/T$, where $T \in \mathbb{N}_+$. The time set is denoted by $\mathcal{T}$ ($\tilde{\mathcal{T}}$ if the final time step T is included). Given a finite subset S of $\mathbb{R}^d$, we denote by $\mathbb{R}(S)$ (resp. $\mathbb{R}^d(S)$) the set of functions from S to $\mathbb{R}$ (resp. $\mathbb{R}^d$). We also denote by $\mathcal{P}(S)$ the of probability measures over S. We call curve of probability measures any function $m\colon \tilde{\mathcal{T}} \times S \to \mathbb{R}$ such that $m(t,\cdot) \in \mathcal{P}(S)$, for any $t \in \mathcal{T}$. The set of probability curves is denoted by $\mathcal{P}(\tilde{\mathcal{T}}, S)$. In mathematical terms,

$$
\begin{aligned}
&\mathcal{T} = \{0,1,\ldots,T-1\}, \qquad \tilde{\mathcal{T}} = \{0,1,\ldots,T\};\\
&\mathbb{R}(S) = \{m\colon S \to \mathbb{R}\}, \qquad \mathbb{R}^d(S) = \{m\colon S \to \mathbb{R}^d\};\\
&\mathcal{P}(S) = \Big\{ m \in \mathbb{R}(S) \;|\; \forall x \in S,\, m(x) \geq 0,\, \sum_{y \in S} m(y) = 1 \Big\};\\
&\mathcal{P}(\tilde{\mathcal{T}}, S) = \Big\{ m \in \mathbb{R}(\tilde{\mathcal{T}} \times S) \;|\; \forall t \in \tilde{\mathcal{T}},\, m(t,\cdot) \in \mathcal{P}(S) \Big\}.
\end{aligned}
$$

We denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and the scalar product in $\mathbb{R}^d$. Let S_1 and S_2 be two finite sets. Let $\mu \in \mathbb{R}^d(S_1 \times S_2)$. For any $x \in S_1$ and for any p_1 and $p_2 \in [1,\infty]$, we denote by $\|\mu(x,\cdot)\|_{p_2}$ the L^{p_2}-norm of the function $y \mapsto \mu(x,y)$, defined as follows:

$$
\|\mu(x,\cdot)\|_{p_2} = \begin{cases} \Big(\sum_{y \in S_2} \|\mu(x,y)\|^{p_2} \Big)^{1/p_2}, & \text{if } p_2 \in [1,\infty),\\ \max_{y \in S_2} \|\mu(x,y)\|, & \text{if } p_2 = \infty. \end{cases}
$$

We next define

$$
\|\mu\|_{p_1,p_2} = \begin{cases} \Big(\sum_{x \in S_1} \|\mu(x,\cdot)\|_{p_2}^{p_1} \Big)^{1/p_1}, & \text{if } p_1 \in [1,\infty),\\ \max_{x \in S_1} \|\mu(x,\cdot)\|_{p_2}, & \text{if } p_1 = \infty. \end{cases}
$$

Lemma 5.1.1 (Hölder's inequality)**.** *Let S_1 and S_2 be two finite sets. Let μ and $\nu \in \mathbb{R}^n(S_1 \times S_2)$ and let p_1 and $p_2 \in [1,\infty]$. It holds that*

$$
\sum_{x_1 \in S_1} \sum_{x_2 \in S_2} \Big| \langle \mu(x_1,x_2), \nu(x_1,x_2) \rangle \Big| \leq \|\mu\|_{p_1,p_2} \|\nu\|_{p_1^*,p_2^*},
$$

where $1/p_i + 1/p_i^* = 1$, for $i = 1,2$.

Definition 5.1.2 (Nemytskii operators)**.** Let ζ be a function from $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{Z}$ and let u be a function from $\mathcal{X}$ to $\mathcal{Y}$. Then the Nemytskii operator is the mapping $\zeta[u]$ from $\mathcal{X}$ to $\mathcal{Z}$ defined by

$$
\zeta[u](x) = \zeta(x, u(x)), \quad \forall x \in \mathcal{X}.
$$

5.2 Potential discrete mean field games

In this section we introduce a general class of discrete MFGs containing the θ-scheme of [BLP22]. We provide a first potential formulation of the discrete MFGs and show their equivalence with a convex optimization problem, using the classical Benamou-Brenier transformation. The analysis being rather standard (see for example [BLP23]), we mostly give succinct proofs.

5.2.1 Problem formulation

We fix $T \in \mathbb{N}_+$ and a finite subset S of $\mathbb{R}^d$. For the description of the MFG model, we fix a running cost ℓ, a coupling cost f, an initial condition m_0, and a terminal cost g, where

$$\ell \colon \mathcal{T} \times S \times \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}, \qquad f \colon \mathcal{T} \times S \times \mathbb{R}(S) \to \mathbb{R}, \qquad m_0 \in \mathcal{P}(S), \qquad g \in \mathbb{R}(S).$$

For a given $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, we denote by $\bar{\ell}_m$ the map defined by

$$\bar{\ell}_m \colon (t, x, \omega) \in \mathcal{T} \times S \mapsto \ell(t, x, \omega) + f(t, x, m(t)),$$

where $m(t) = (m(t, x))_{x \in S}$. We require the following assumptions for ℓ, m, f, and m_0.

Assumption 5.1. The following holds:

1. *Bounded domain.* For all $(t, x) \in \mathcal{T} \times S$, $\ell(t, x, \cdot)$ is lower semi-continuous with a non-empty domain. There exists $D > 0$ such that for any $t \in \mathcal{T}$, for any $x \in S$, and for any $\|v\| > D$, we have $\ell(t, x, v) = +\infty$.

2. *Regularity.* There exists L_f such that for any (t, x) and for any m_1, and m_2 in $\mathcal{P}(S)$, we have

$$|f(t, x, m_1) - f(t, x, m_2)| \leq L_f \|m_1 - m_2\|_2.$$

3. *Strong convexity.* There exists $\alpha > 0$ such that for any $t \in \mathcal{T}$ and for any $x \in S$, the function $\ell(t, x, \cdot)$ is α-strongly convex, i.e.,

$$\ell(t, x, v_2) \geq \ell(t, x, v_1) + \langle p, v_2 - v_1 \rangle + \frac{\alpha}{2} \|v_2 - v_1\|^2,$$

for all v_1 and v_2 in $\mathbb{R}^n$ and for all $p \in \partial_p \ell(t, x, v_1)$.

4. *Monotonicity.* For any $t \in \mathcal{T}$, for any m_1 and m_2 in $\mathcal{P}(S)$,

$$\sum_{x \in S} \big(f(t, x, m_1) - f(t, x, m_2)\big)(m_1(x) - m_2(x)) \geq 0.$$

We now fix two elements $\pi_0 \in \mathbb{R}(\mathcal{T} \times S^2)$ and $\pi_1 \in \mathbb{R}^d(\mathcal{T} \times S^2)$ and define the map π by

$$\pi \colon (t, x, y, \omega) \in \mathcal{T} \times S \times S \times \mathbb{R}^d \mapsto \pi_0(t, x, y) + \Delta t \langle \pi_1(t, x, y), \omega \rangle.$$

The map π describes the probability of an agent located at time t in state x, using the control ω, to reach state y at time $t + 1$. Our analysis will exploit the fact that π is affine with respect to ω. Recalling the constant D introduced in Assumption 5.1, we consider the following assumption.

Assumption 5.2. The elements π_0 and π_1 satisfy the following:

$$\begin{cases} \pi_0(t, x, \cdot) \in \mathcal{P}(S), & \forall (t, x) \in \mathcal{T} \times S, \\ \sum_{y \in S} \pi_1(t, x, y) = 0, & \forall (t, x) \in \mathcal{T} \times S, \\ \pi_0(t, x, y) \geq \Delta t D \|\pi_1(t, x, y)\|, & \forall (t, x, y) \in \mathcal{T} \times S \times S. \end{cases} \tag{5.2.1}$$

An immediate consequence of Assumption 5.2 is the following: For all $(t,x) \in \mathcal{T} \times S$, for all $\omega \in \mathbb{R}^d$, if $\|\omega\| \leq D$, then $\pi(t,x,\cdot,\omega) \in \mathcal{P}(S)$.

Assumption 5.3. There exists a function $F \colon \mathcal{T} \times \mathcal{P}(S) \to \mathbb{R}^d$ such that for any $t \in \mathcal{T}$ and for any m_1 and m_2 in $\mathcal{P}(S)$, it holds

$$F(t,m_1) - F(t,m_2) = \int_0^1 \sum_{x \in S} f(t,x,m_1 + s(m_2 - m_1))(m_2(x) - m_1(x))ds.$$

We have the following convexity property for the potential function F.

Lemma 5.2.1. *For any $t \in \mathcal{T}$ and for any m_1 and m_2 in $\mathcal{P}(S)$, it holds that*

$$F(t,m_2) \geq F(t,m_1) + \sum_{x \in S} f(t,x,m_1)(m_2(x) - m_1(x)). \tag{5.2.2}$$

Proof. Inequality (5.2.2) follows from the definition of F and the monotonicity in Assumption 5.1. □

Until the end of Section 5.3, we assume that Assumptions 5.1, 5.2, and 5.3 are satisfied. Following [BLP22, Sec. 2.3], we consider the following discrete mean field game system, involving the variables $u \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$, $v \in \mathbb{R}^d(\mathcal{T} \times S)$, and $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$:

$$\begin{cases} \text{(i)} & u = \mathbf{HJB}(m), \\ \text{(ii)} & v = \mathbf{V}(u), \\ \text{(iii)} & m = \mathbf{FP}(v), \end{cases} \tag{DMFG}$$

where the Hamilton-Jacobi-Bellman mapping **HJB**, the optimal control mapping **V**, and the Fokker-Planck mapping **FP** are defined as follows:

- Given $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, $u = \mathbf{HJB}(m) \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ is the solution to

$$\begin{cases} u(t,x) = \inf\limits_{\omega \in \mathbb{R}^d} \bar{\ell}_m(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega) u(t+1,y), & \forall (t,x) \in \mathcal{T} \times S, \\ u(T,x) = g(x), & \forall x \in S. \end{cases} \tag{5.2.3}$$

- Given $u \in \mathbb{R}(\mathcal{T} \times S)$, $v = \mathbf{V}(u) \in \mathbb{R}^d(\mathcal{T} \times S)$ is defined by

$$v(t,x) = \operatorname*{argmin}_{\omega \in \mathbb{R}^d} \; \ell(t,x,\omega)\Delta t + \sum_{y \in S} \pi(t,x,y,\omega) u(t+1,y), \quad \forall (t,x) \in \mathcal{T} \times S. \tag{5.2.4}$$

- Given $v \in \mathbb{R}^d(\mathcal{T} \times S)$, $m = \mathbf{FP}(v) \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ is defined as the solution to

$$\begin{cases} m(t+1,y) = \sum_{x \in S} \pi(t,x,y,v(t,x)) m(t,x), & \forall (t,y) \in \mathcal{T} \times S, \\ m(0,x) = m_0(x), & \forall x \in S. \end{cases} \tag{5.2.5}$$

Lemma 5.2.2 (Continuity of **HJB**). *For any m_1 and m_2 in $\mathcal{P}(\tilde{\mathcal{T}}, S)$, we have*

$$\|\mathbf{HJB}(m_1) - \mathbf{HJB}(m_2)\|_{\infty,\infty} \leq L_f \|m_1 - m_2\|_{\infty,2}. \tag{5.2.6}$$

Proof. See [BLP22, Eq. A.6]. □

Lemma 5.2.3 (Continuity of **FP**). *Let v_1 and v_2 in $\mathbb{R}^d(\mathcal{T} \times S)$ be such that $\|v_1\|_{\infty,\infty} \leq D$ and $\|v_2\|_{\infty,\infty} \leq D$. There exists a constant C, independent of v_1 and v_2, such that*

$$\left\|\mathbf{FP}(v_1) - \mathbf{FP}(v_2)\right\|_{\infty,2}^2 \leq C\Delta t \sum_{t\in\mathcal{T}} \sum_{x \in S} \|(v_1 - v_2) m_1(t,x)\|^2. \tag{5.2.7}$$

As an immediate consequence, the mapping **FP** *is continuous.*

Proof. Let $m_1 = \mathbf{FP}(v_1)$ and $m_2 = \mathbf{FP}(v_2)$. Let $\mu = m_1 - m_2$. It is easy to verify that

$$\begin{cases} \mu(t+1,y) = \sum_{x\in S} \pi(t,x,y,v_1(t,x))\mu(t,x) + \Delta t \sum_{x \in S} \langle \pi_1(t,x,y), (v_2 - v_1) m_2(t,x)\rangle, \\ \mu(0,y) = 0. \end{cases} \tag{5.2.8}$$

Inequality (5.2.7) immediately follows from Gronwall's inequality, keeping in mind that all norms are equivalent on the finite-dimensional vector space $\mathbb{R}^d(\mathcal{T} \times S)$. □

Theorem 5.2.4 (Existence). *Under Assumptions 5.1 and 5.2, system* (DMFG) *has a solution.*

Proof. We follow the proof of [BLP22, Thm. 3.6]. We note first that by Assumption 5.2, the composed mapping $\mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}$ is valued in $\mathcal{P}(\tilde{\mathcal{T}}, S)$. By Brouwer's fixed point theorem, it suffices to show that $\mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}$ is continuous. The continuity of **HJB** and **FP** was established in Lemmas 5.2.2-5.2.3. The continuity of **V** is deduced from the strong convexity of ℓ, see step 2 of the proof of [BLP22, Thm. 3.6]. □

5.2.2 Potential formulation

Similarly to the case of continuous MFGs (see [LL07, BCS17] for example), the system (DMFG) has a potential formulation. Consider the following optimal control problem:

$$\inf_{\substack{m \in \mathcal{P}(\tilde{\mathcal{T}}, S) \\ v \in \mathbb{R}^d(\mathcal{T}\times S)}} J(m,v), \quad \text{subject to: } (m,v) \in \mathcal{A}, \tag{P}$$

where the cost function J and the set $\mathcal{A}$ are defined by

$$J(m,v) = \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \ell[v](t,x) m(t,x) + \Delta t \sum_{t \in \mathcal{T}} F(t, m(t)) + \sum_{x\in S} g(x) m(T,x);$$

$$\mathcal{A} = \left\{ (m,v) \in \mathcal{P}(\tilde{\mathcal{T}}, S) \times \mathbb{R}^d(\mathcal{T} \times S) \,\middle|\, m = \mathbf{FP}(v),\ \|v\|_{\infty,\infty} \leq D \right\}.$$

Problem (P) is a non-convex problem which can be made convex with the classical Benamou-Brenier transform (see [BCS17] for example) defined by

$$\chi \colon (m,v) \in \mathcal{A} \mapsto (m, mv) \in \mathcal{P}(\tilde{\mathcal{T}}, S) \times \mathbb{R}^d(\mathcal{T} \times S). \tag{5.2.9}$$

Here mv is the pointwise product of m and v: $mv(t,x) := m(t,x)v(t,x)$, for all $t \in \mathcal{T}$ and for all $x \in S$. We set $\tilde{\mathcal{A}} = \chi(\mathcal{A})$ and consider the cost function $\tilde{J}$, defined by

$$\tilde{J}(m,w) = \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \tilde{\ell}[m,w](t,x) + \Delta t \sum_{t\in\mathcal{T}} F(t,m(t)) + \sum_{x\in S} g(x)m(T,x),$$

where the function $\tilde{\ell}[m,w]\colon \mathcal{T} \times S \to \bar{\mathbb{R}}$ is defined by

$$\tilde{\ell}[m,w](t,x) = \begin{cases} \ell\Big(t,x,\frac{w(t,x)}{m(t,x)}\Big)m(t,x), & \text{if } m(t,x) \neq 0, \\ 0, & \text{if } m(t,x) = w(t,x) = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

The new problem of interest is

$$\inf_{\substack{m\in\mathcal{P}(\tilde{\mathcal{T}},S)\\ w\in\mathbb{R}^d(\mathcal{T}\times S)}} \tilde{J}(m,w), \quad \text{subject to: } (m,w) \in \tilde{\mathcal{A}}. \tag{$\tilde{P}$}$$

In the rest of the section, we investigate some properties of Problems (P) and $(\tilde{P})$, as well as their relationship with (DMFG)

Lemma 5.2.5. *It holds that* $\mathrm{val}(P) = \mathrm{val}(\tilde{P})$.

Proof. It follows from the definitions of $\mathcal{A}$, J, $\tilde{J}$, and χ that for any $(m,v) \in \mathcal{A}$, we have $\tilde{J}(\chi(m,v)) = J(m,v)$. The lemma follows immediately. □

We next discuss the convexity of Problem $(\tilde{P})$.

Lemma 5.2.6 (Convexity)**.** *The set $\tilde{\mathcal{A}}$ is convex. The cost function $\tilde{J}(m,w)$ is convex.*

Proof. Take any $(m_1,w_1),(m_2,w_2)$ in $\tilde{\mathcal{A}}$ and any $\lambda \in (0,1)$. By the definition of $\tilde{\mathcal{A}}$, there exist $v_1, v_2 \in \mathbb{R}^d(\mathcal{T} \times S)$, such that $(m_i,v_i) \in \mathcal{A}$ and $w_i = m_i v_i$ for $i = 1,2$. Let

$$\begin{aligned} m &= \lambda m_1 + (1-\lambda)m_2, \\ w &= \lambda w_1 + (1-\lambda)w_2, \\ v(t,x) &= \begin{cases} 0, & \text{if } m_1(t,x) = m_2(t,x) = 0, \\ \frac{\lambda m_1 v_1 + (1-\lambda)m_2 v_2}{\lambda m_1 + (1-\lambda)m_2}(t,x), & \text{otherwise.} \end{cases} \end{aligned}$$

We can check that $(m,v) \in \mathcal{A}$ and that $(m,w) = \chi(m,v)$. The convexity of $\tilde{\mathcal{A}}$ follows.

The function $\tilde{J}$ is defined as the sum of three terms. The last one is linear, thus convex. The second one is also convex, by Lemma 5.2.1. Finally, for any (t,x), the map $(m,w) \mapsto \tilde{\ell}[m,w](t,x)$ is convex (see [Com18, Proposition 2.3]). The convexity of $\tilde{J}$ follow. □

Given $m' \in \mathcal{P}(\tilde{\mathcal{T}},S)$, we consider the cost function $\tilde{J}_{m'}$, defined by

$$\tilde{J}_{m'}(m,w) = \Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \tilde{\ell}[m,w](t,x) + f(t,x,m'(t))m(t,x) + \sum_{x\in S} g(x)m(T,x),$$

for $(m, w) \in \mathcal{P}(\tilde{T}, S) \times \mathbb{R}^d(\mathcal{T} \times S)$. We will regard $\tilde{J}_{m'}$ as a partial linearization of J around m'. We define the corresponding optimal control problem:

$$\inf_{(m,w)\in\tilde{\mathcal{A}}} \tilde{J}_{m'}(m, w). \tag{$P_{m'}$}$$

Lemma 5.2.7. *Let (m_1, w_1) and (m_2, w_2) be in $\tilde{\mathcal{A}}$. Then*

$$\tilde{J}(m_2, w_2) - \tilde{J}(m_1, w_1) \geq \tilde{J}_{m_1}(m_2, w_2) - \tilde{J}_{m_1}(m_1, w_1).$$

Proof. This is an immediate consequence of the definitions of $\tilde{J}$, $\tilde{J}_{m'}$, and Lemma 5.2.1. □

The next lemma provides us with a solution to Problem $(P_{m'})$.

Lemma 5.2.8. *Let $m' \in \mathcal{P}(\tilde{\mathcal{T}}, S)$. Let us set $\tilde{v} = \mathbf{V} \circ \mathbf{HJB}(m')$, $\tilde{m} = \mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}(m')$, and $\tilde{w} = \tilde{m}\tilde{v}$. Then $(\tilde{m}, \tilde{w})$ is the unique solution to $(P_{m'})$. Moreover, for any $(m, w) \in \tilde{\mathcal{A}}$ and for any v such that $w = mv$, it holds that*

$$\tilde{J}_{m'}(m, w) - \tilde{J}_{m'}(\tilde{m}, \tilde{w}) \geq \frac{\alpha}{2}\Delta t \sum_{t\in\mathcal{T}} \sum_{x\in S} \|(v - \tilde{v})(t, x)\|^2 m(t, x). \tag{5.2.10}$$

Proof. The first inequality is proved in [BLP22, Sec. 3.5, page 14]. It is of similar nature to the fundamental equality of [ACCD13]. As a consequence of (5.2.10), the pair $(\tilde{m}, \tilde{w})$ is a solution to $(P_{m'})$. It remains to prove uniqueness. Let $(\hat{m}, \hat{w}) \in \tilde{\mathcal{A}}$ be a solution to $(P_{m'})$. Let $\hat{v}$ be such that $\hat{w} = \hat{m}\hat{v}$. Then, by inequality (5.2.10), we have $\sum_{t\in\mathcal{T}} \sum_{x\in S} \|\hat{v}(t, x) - \tilde{v}(t, x)\|^2 \hat{m}(t, x) = 0$. It follows next that $(\hat{v} - \tilde{v})(t, x)\hat{m}(t, x) = 0$ for all (t, x). Applying Lemma 5.2.3, we immediately obtain that $\hat{m} = \tilde{m}$. Finally, we have $\hat{w} - \tilde{w} = \hat{m}\hat{v} - \tilde{m}\tilde{v} = \hat{m}(\hat{v} - \tilde{v}) = 0$, which concludes the proof of uniqueness. □

Lemma 5.2.9. *System* (DMFG) *has a unique solution $(\bar{m}, \bar{u}, \bar{v})$. Moreover, $(\bar{m}, \bar{w}) \coloneqq \chi(\bar{m}, \bar{v})$ is the unique solution to $(\tilde{P})$.*

Proof. Let (m', u', v') be a solution to (DMFG). Let $w' = m'v'$. Combining Lemma 5.2.7 and Lemma 5.2.8, we deduce that for any $(m, w) \in \tilde{\mathcal{A}}$,

$$\tilde{J}(m, w) - \tilde{J}(m', w') \geq \tilde{J}_{m'}(m, w) - \tilde{J}_{m'}(m', w') \geq 0.$$

Thus (m', w') is a solution to $(\tilde{P})$.

We next prove that (m', w') is the unique solution to $(\tilde{P})$. Let (m, w) be a solution to $(\tilde{P})$. The above inequality shows that (m, w) is also a solution to $(P_{m'})$. Thus by Lemma 5.2.8, $(m, w) = (m', w')$.

It remains to prove the uniqueness of the solution to (DMFG). Let (m, u, v) be a solution to (DMFG). As was proved above, (m, mv) is a solution to $(\tilde{P})$ and therefore $m = m'$. It follows that $u = \mathbf{HJB}(m) = \mathbf{HJB}(m') = u'$ and that $v = \mathbf{V}(u) = \mathbf{V}(u') = v'$, which concludes the proof. □

5.3 Generalized Frank-Wolfe algorithm: the discrete case

We investigate in this section the convergence of the GFW algorithm, applied to the (convex) potential problem $(\tilde{P})$. In this section, Assumptions 5.1-5.3 are supposed to be satisfied. We recall that (DMFG) has a unique solution $(\bar{u}, \bar{v}, \bar{m})$ and that by Lemma 5.2.9, $(\bar{m}, \bar{w}) = \chi(\bar{m}, \bar{v})$ is the unique solution of problem $(\tilde{P})$.

5.3.1 Algorithm and convergence results

We first define the mapping $\mathbf{BR}\colon \mathcal{P}(\tilde{\mathcal{T}}, S) \to \tilde{\mathcal{A}}$. Given $m' \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, we obtain $(\tilde{m}, \tilde{w}) = \mathbf{BR}(m')$ by successively computing

$$\tilde{v} = \mathbf{V} \circ \mathbf{HJB}(m'), \quad \tilde{m} = \mathbf{FP}(\tilde{v}), \quad \text{and} \quad \tilde{w} = \tilde{m}\tilde{v}.$$

We refer to $\mathbf{BR}$ as the best-response mapping: given a prediction m' of the equilibrium distribution of the agents, $\tilde{v}$ (as defined above) is the optimal feedback for the underlying optimal control problem and $\tilde{m}$ the resulting distribution. As was demonstrated in Lemma 5.2.8, $\mathbf{BR}(m')$ is also the unique solution to the linearized problem $(P_{m'})$. This allows us to write the GFW algorithm for the resolution of $(\tilde{P})$ in the form of a best-response algorithm.

Algorithm 5.1: Generalized Frank-Wolfe Algorithm

Initialization: $(m^0, w^0) \in \tilde{\mathcal{A}}$;
First iteration: $(m^1, w^1) = (\bar{m}^0, \bar{w}^0) = \mathbf{BR}(m^0)$;
for $k = 1, 2, \ldots$ **do**

> **Step 1: Resolution of the partial linearized problem.**
> Set $(\bar{m}^k, \bar{w}^k) = \mathbf{BR}(m^k)$;
>
> **Step 2: Update.**
> Choose $\lambda_k \in [0, 1]$;
> Set $(m^{k+1}, w^{k+1}) = (1 - \lambda_k)(m^k, w^k) + \lambda_k(\bar{m}^k, \bar{w}^k)$;

end

Note that the choice of the stepsize λ_k will be discussed in Proposition 5.3.1. We introduce now three constants, C_1, C_2, and C_3, that will be used for the convergence analysis of Algorithm 5.1. The constants C_1 and C_2 are defined by

$$C_1 = \sup_{\|v\|_{\infty,\infty} \leq D} \|\mathbf{FP}(v)\|_{\infty,2}^2 \quad \text{and} \quad C_2 = \sup_{m \in \mathcal{P}(\tilde{\mathcal{T}}, S)} \|\mathbf{FP} \circ \mathbf{V} \circ \mathbf{HJB}(m)\|_{\infty,\infty}. \tag{5.3.1}$$

The finiteness of C_1 and C_2 follows from the compactness of $\mathcal{P}(\tilde{\mathcal{T}}, S)$ and from the continuity of the three mappings $\mathbf{HJB}$, $\mathbf{V}$, and $\mathbf{FP}$. Note that

$$C_1 = \sup_{(m,v) \in \mathcal{A}} \|m\|_{\infty,2}^2 = \sup_{(m,w) \in \tilde{\mathcal{A}}} \|m\|_{\infty,2}^2. \tag{5.3.2}$$

By Lemma 5.2.3, there exists a constant $C_3 > 0$ such that for any (m_1, v_1) and (m_2, v_2) in $\mathcal{A}$,

$$\|m_1 - m_2\|_{\infty,2}^2 \leq C_3 \Delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \|(v_1 - v_2) m_1(t,x)\|^2. \tag{5.3.3}$$

We next introduce three other constants, D_1, D_2, and c, defined by:

$$D_1 = C_1 L_f |S|^{1/2}, \quad D_2 = (L_f + |S|^{1/2}) \sqrt{\frac{2C_2C_3}{\alpha}}, \quad c = \max\left\{1 - \frac{\alpha}{4C_2C_3L_f|S|^{1/2}}, \frac{1}{2}\right\}. \tag{5.3.4}$$

Proposition 5.3.1. *We consider the sequence $(m_k, w_k)_{k\geq 1}$ generated by Algorithm (5.1).*

1. Sublinear rate. *Assume that $\lambda_k = 2/(k+2)$, for all $k \geq 1$. Then,*

$$\tilde{J}(m^k, w^k) - \tilde{J}(\bar{m}, \bar{w}) \leq \frac{8D_1}{k}, \qquad \forall k \geq 1. \tag{5.3.5}$$

2. Linear rate. *Assume that*

$$\lambda_k = \min\left\{ \frac{\tilde{J}_{m^k}(m^k, w^k) - \tilde{J}_{m^k}(\bar{m}^k, \bar{w}^k)}{L_f |S|^{1/2} \|m^k - \bar{m}^k\|_{\infty,2}^2}, 1 \right\}, \tag{5.3.6}$$

for all $k \geq 1$. Then,

$$\tilde{J}(m^k, w^k) - \tilde{J}(\bar{m}, \bar{w}) \leq 4D_1 c^k, \qquad \forall k \geq 1. \tag{5.3.7}$$

The following subsection is dedicated to the proof of Proposition 5.3.1. The used stepsize rule (5.3.6) is motivated in the proof of convergence (see (5.3.13)).

5.3.2 Convergence analysis

Lemma 5.3.2 (A priori bounds)**.** *For any $k \geq 1$, we have $(m^k, w^k) \in \tilde{\mathcal{A}}$. As a consequence,*

$$\|m^k\|_{\infty,2}^2 \leq C_1 \quad \textit{and} \quad \|m^k\|_{\infty,\infty} \leq C_2. \tag{5.3.8}$$

Proof. Using the definitions of C_1 and C_2 (given in (5.3.1)) and using (5.3.2), we have that

$$\|\bar{m}^k\|_{\infty,2}^2 \leq C_1, \|\bar{m}^k\|_{\infty,\infty} \leq C_2, \quad \text{and} \quad (\bar{m}^k, \bar{w}^k) \in \tilde{\mathcal{A}}, \tag{5.3.9}$$

for any $k \geq 1$. To conclude the proof of the lemma, it suffices to observe that for any $k \geq 1$, (m^k, w^k) is a convex combination of $(\bar{m}^\kappa, \bar{w}^\kappa) \in \tilde{\mathcal{A}}$ for $0 \leq \kappa \leq k-1$. Then we have $(m^k, w^k) \in \tilde{\mathcal{A}}$, since $\tilde{\mathcal{A}}$ is convex, by Lemma 5.2.6. Inequality (5.3.8) follows from (5.3.9) and from the triangle inequality. □

Lemma 5.3.3. *For any $t \in \mathcal{T}$ and for any m_1 and m_2 in $\mathcal{P}(S)$, we have*

$$F(t, m_2) \leq F(t, m_1) + \sum_{x \in S} f(t, x, m_1)(m_2(x) - m_1(x)) + \frac{L_f |S|^{1/2}}{2} \|m_2 - m_1\|_2^2. \tag{5.3.10}$$

Proof. Combining Assumption 5.3 and the Lipschitz-continuity of f (Assumption 5.1), we deduce that

$$F(t,m_2) \leq F(t,m_1) + \sum_{x\in S} f(t,x,m_1)(m_2(x)-m_1(x)) + \frac{L_f}{2}\|m_2-m_1\|_2\|m_2-m_1\|_1.$$

By Hölder's inequality, we have that $\|m_2-m_1\|_1 \leq |S|^{1/2}\|m_2-m_1\|_2$. Inequality (5.3.10) follows. $\square$

Algorithm 5.1 generates two sequences $(m^k,w^k)_{k\geq 0}$ and $(\bar{m}^k,\bar{w}^k)_{k\geq 0}$ in $\tilde{\mathcal{A}}$. For the analysis, we need to fix two sequences $(v^k)_{k\geq 0}$ and $(\bar{v}^k)_{k\geq 0}$ such that $w^k = m^k v^k$ and $\bar{w}^k = \bar{m}^k\bar{v}^k$. We introduce the following six sequences of positive numbers:

$$\delta_k = \|m^k - \bar{m}\|^2_{\infty,2}, \qquad \bar{\delta}_k = \|m^k - \bar{m}^k\|^2_{\infty,2};$$

$$\gamma_k = \tilde{J}(m^k,w^k) - \tilde{J}(\bar{m},\bar{w}), \qquad \bar{\gamma}_k = \tilde{J}_{m^k}(m^k,w^k) - \tilde{J}_{m^k}(\bar{m}^k,\bar{w}^k);$$

$$\epsilon_k = \Delta t \sum_{t\in\mathcal{T}}\sum_{x\in S}\|v^k-\bar{v}\|^2 m^k(t,x), \qquad \bar{\epsilon}_k = \Delta t \sum_{t\in\mathcal{T}}\sum_{x\in S}\|v^k-\bar{v}^k\|^2 m^k(t,x).$$

The following lemma establishes various relationships between these sequences, independent of the choice of stepsize $(\lambda_k)_{k\geq 1}$. For convenience, we fix $\lambda_0 = 1$ and make use of the fact that $(m^1,w^1) = (1-\lambda_0)(m^0,w^0) + \lambda_0(\bar{m}^0,\bar{w}^0)$.

Lemma 5.3.4. *For any choice of stepsizes $(\lambda_k)_{k\geq 1}$, we have $\bar{\delta}_k \leq 4C_1$, for any $k\geq 1$. Moreover we have that*

$$\gamma_k \leq \bar{\gamma}_k \qquad \text{and} \qquad \gamma_{k+1} \leq \gamma_k - \lambda_k\bar{\gamma}_k + \lambda_k^2\frac{L_f|S|^{1/2}}{2}\bar{\delta}_k. \tag{5.3.11}$$

We also have the following estimates:

$$\frac{\alpha}{2C_2C_3}\delta_k \leq \frac{\alpha}{2}\epsilon_k \leq \gamma_k \qquad \text{and} \qquad \frac{\alpha}{2C_2C_3}\bar{\delta}_k \leq \frac{\alpha}{2}\bar{\epsilon}_k \leq \bar{\gamma}_k. \tag{5.3.12}$$

Proof. **Step 1.** The inequality $\bar{\delta}_k = \|m^k-\bar{m}^k\|^2_{\infty,2} \leq 4C_1$ follows from the bounds $\|m^k\|^2_{\infty,2}\leq C_1$ and $\|\bar{m}^k\|^2_{\infty,2}\leq C_1$ obtained in Lemma 5.3.2.

Step 2. We next prove that $\gamma_k \leq \bar{\gamma}_k$. Recalling that $(\bar{m}^k,\bar{w}^k)$ minimizes $\tilde{J}_{m^k}(\cdot)$ over $\tilde{\mathcal{A}}$, we obtain that

$$\bar{\gamma}_k = \tilde{J}_{m^k}(m^k,w^k) - \tilde{J}_{m^k}(\bar{m}^k,\bar{w}^k) \geq \tilde{J}_{m^k}(m^k,w^k) - \tilde{J}_{m^k}(\bar{m},\bar{w}).$$

Using next Lemma 5.2.7, we deduce that $\bar{\gamma}_k \geq J(m^k,w^k) - J(\bar{m},\bar{w}) = \gamma_k$. Let us prove the upper bound of γ_{k+1}. Since $\tilde{\ell}$ is convex (see the proof of Lemma 5.2.6), we have

$$\tilde{\ell}[m^{k+1},w^{k+1}](t,x) \leq (1-\lambda_k)\tilde{\ell}[m^k,w^k](t,x) + \lambda_k\tilde{\ell}[\bar{m}^k,\bar{w}^k](t,x).$$

Moreover, by Lemma 5.3.3, we have

$$F(t,m^{k+1}(t)) \leq F(t,m^k(t)) + \lambda_k\sum_{x\in S} f(t,x,m^k(t))(\bar{m}^k(t,x)-m^k(t,x)) + \frac{\lambda_k^2 L_f|S|^{1/2}}{2}\bar{\delta}_k.$$

Then (5.3.11) follows from the above two inequalities and the definitions of $\tilde{J}$, $\tilde{J}_{m^k}$, γ_k, and $\bar{\gamma}_k$.

Step 3. Let us prove (5.3.12). From the definition of C_3 and Lemma 5.3.2 we have that

$$\begin{aligned}\delta_k &\leq C_3 \delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \|(v_k - \bar{v})(t,x) m_k(t,x)\|^2 \\ &\leq C_2 C_3 \delta t \sum_{t \in \mathcal{T}} \sum_{x \in S} \|(v_k - \bar{v})(t,x)\|^2 m_k(t,x) = C_2 C_3 \epsilon_k.\end{aligned}$$

Moreover, using Lemmas 5.2.7 and 5.2.8, we obtain that

$$\gamma_k \geq \tilde{J}_{\bar{m}}(m^k) - \tilde{J}_{\bar{m}}(\bar{m}) \geq \frac{\alpha}{2} \epsilon_k.$$

We prove in a similar fashion that $\bar{\delta}_k \leq C_2 C_3 \bar{\epsilon}_k$ and that $\bar{\gamma}_k \geq \frac{\alpha}{2} \bar{\epsilon}_k$. □

Lemma 5.3.5. *In Algorithm 5.1, let* $u^k = \mathbf{HJB}(m^k)$*. Then, whatever the choice of stepsizes* $(\lambda_k)_{k \geq 1}$*, we have*

$$\|u^k - \bar{u}\|_{\infty,\infty} + \|m^k - \bar{m}\|_{\infty,1} \leq D_2 \sqrt{\gamma_k}.$$

Proof. By Cauchy-Schwarz inequality and Lemma 5.3.4, we have

$$\|m^k - \bar{m}\|_{\infty,1} \leq \sqrt{|S|} \sqrt{\delta_k} \leq \sqrt{|S|} \sqrt{\frac{2 C_2 C_3}{\alpha}} \sqrt{\gamma_k}.$$

By Lemma 5.2.2, we have $\|u^k - \bar{u}\|_{\infty,\infty} \leq L_f \sqrt{\delta_k}$. Recalling the definition of D_2, we obtain the announced result. □

Proof of Proposition 5.3.1. **Sublinear case**. Combining the two inequalities in (5.3.11) and using the bound $\bar{\delta}_k \leq 4C_1$, we obtain that

$$\gamma_{k+1} \leq (1 - \lambda_k)\gamma_k + 2\lambda_k^2 C_1 L_f |S|^{1/2} = (1 - \lambda_k)\gamma_k + 2\lambda_k^2 D_1.$$

In particular, for $k = 0$, we have $\lambda_0 = 1$, therefore $\gamma_1 \leq 2D_1$. It is then easy to prove by induction the inequality $\gamma_k \leq 8D_1/k$ (see [Jag13, Thm. 1] for example).

Linear case. Note first that the stepsize rule (5.3.6) writes

$$\lambda_k = \min \left\{ \frac{\bar{\gamma}_k}{L_f |S|^{1/2} \bar{\delta}_k} \, , \, 1 \right\}.$$

It is easy to verify that λ_k minimizes the upper-bound (5.3.11), that is to say:

$$\lambda_k = \underset{\lambda \in [0,1]}{\operatorname{argmin}} - \lambda \bar{\gamma}_k + \lambda^2 \frac{L_f |S|^{1/2}}{2 \bar{\delta}_k}. \tag{5.3.13}$$

Let $k \geq 1$. We consider the following two cases:

1. If $\bar{\gamma}_k \geq L_f |S|^{1/2} \bar{\delta}_k$, then $\lambda_k = 1$. We deduce from (5.3.11) that

$$\gamma_{k+1} \leq \gamma_k - \bar{\gamma}_k + \frac{L_f |S|^{1/2}}{2} \bar{\delta}_k \leq \gamma_k - \frac{\bar{\gamma}_k}{2} \leq \frac{\gamma_k}{2} \leq c \gamma_k,$$

   where the second inequality follows from the assumption $\bar{\gamma}_k \geq L_f |S|^{1/2} \bar{\delta}_k$, the third one from the inequality $\gamma_k \leq \bar{\gamma}_k$ and the last one from the fact that $c \geq 1/2$.

2. If $\bar{\gamma}_k \leq L_f|S|^{1/2}\bar{\delta}_k$, then $\lambda_k = \frac{\bar{\gamma}_k}{L_f|S|^{1/2}\bar{\delta}_k}$. We deduce from (5.3.11) and from the inequality $\gamma_k \leq \bar{\gamma}_k$ that

$$\gamma_{k+1} \leq \gamma_k - \frac{\bar{\gamma}_k}{2L_f|S|^{1/2}\bar{\delta}_k}\bar{\gamma}_k \leq \left(1 - \frac{\bar{\gamma}_k}{2L_f|S|^{1/2}\bar{\delta}_k}\right)\gamma_k.$$

By (5.3.12), we know that $\bar{\gamma}_k/\bar{\delta}_k \geq \frac{\alpha}{2C_2C_3}$. It follows that

$$\gamma_{k+1} \leq \left(1 - \frac{\alpha}{4C_2C_3L_f|S|^{1/2}}\right)\gamma_k \leq c\gamma_k.$$

It follows that $\gamma_k \leq \gamma_1 c^{k-1}$ for all $k \geq 1$. Since $c^{-1} \leq 2$ and $\gamma_1 \leq 2D_1$, estimate (5.3.7) holds true. □

5.4 Mesh-independent convergence of the GFW algorithm for second-order MFGs

We establish in this section two mesh-independence principles for the GFW algorithm, applied to a discretization of (MFG) with the θ-scheme proposed in [BLP22]. First, we recall the theta-scheme and the main convergence result of [BLP22]. Then, we show that the potential structure of the continuous (MFG) is preserved by the theta-scheme at the discrete level. Therefore, Algorithm 5.1 can be applied to the theta-scheme. The mesh-independence principles are stated in Theorems 5.4.5 and 5.4.9 and demonstrated in Subsections 5.4.3 and 5.4.4.

5.4.1 The theta-scheme and error estimates

In this subsection, we recall the theta-scheme of the continuous system (MFG) investigated in [BLP22] and its main result. Let us define

$$\mathcal{D} = \left\{\mu \in \mathbb{L}^2(\mathbb{T}^d) \,|\, \mu \geq 0, \int_{\mathbb{T}^d} \mu(x)dx = 1\right\}. \tag{5.4.1}$$

We make the following assumptions on the data function $\ell^c \colon Q \times \mathbb{R}^d \to \mathbb{R}$, $g^c \colon \mathbb{T}^d \to \mathbb{R}$, $f^c \colon Q \times \mathcal{D} \to \mathbb{R}$, and $m^c \colon \mathbb{T}^d \to \mathbb{R}$ of the continuous model (MFG).

Assumption A. The following holds:

1. *Regularity.* The function ℓ^c is continuously differentiable with respect to v. There exist positive constants L^c_ℓ, L^c_g, and L^c_f such that for any $(t,x) \in Q$, for any $v \in \mathbb{R}^d$, and for any $m \in \mathcal{D}$,
   - $\ell^c(\cdot,x,v)$, $\ell^c(t,\cdot,v)$, and $\ell^c_v(\cdot,x,v)$ are L^c_ℓ-Lipschitz continuous
   - g^c is L^c_g-Lipschitz continuous
   - $f^c(\cdot,x,m)$, $f^c(t,\cdot,m)$, and $f^c(t,x,\cdot)$ are L^c_f-Lipschitz continuous (with respect to the $\|\cdot\|_{\mathbb{L}^2}$-norm for the third variable).

2. *Strong convexity.* There exists $\alpha^c > 0$ such that for any $(t,x) \in Q$, $\ell^c(t,x,\cdot)$ is strongly convex with modulus α^c, i.e.

$$\ell^c(t,x,v_2) \geq \ell^c(t,x,v_1) + \langle \ell_v^c(t,x,v_1), v_2 - v_1 \rangle + \frac{\alpha^c}{2}\|v_2 - v_1\|^2, \quad \forall v_1, v_2 \in \mathbb{R}^d.$$

3. *Monotonicity.* The function f^c is monotone, i.e., for any $t \in [0,T]$, for any m_1 and $m_2 \in \mathcal{D}$,

$$\int_{\mathbb{T}^d} \Big(f^c(t,x',m_1) - f^c(t,x',m_2) \Big) \big(m_1(x') - m_2(x') \big) dx' \geq 0.$$

Assumption B. The continuous mean field game (MFG) admits a unique solution (u^*, v^*, m^*), with $u^*, m^* \in \mathcal{C}^{1+r/2, 2+r}(Q)$ and $v^* \in \mathcal{C}^r(Q) \cap \mathbb{L}^\infty([0,1]; \mathcal{C}^{1+r}(\mathbb{T}^d))$, where $r \in (0,1)$.

Note that more explicit assumptions on the problem data allows to verify Assumption B, see [BLP22, Appendix B].

The time and space discretization parameters are denoted $\Delta t > 0$ and $h > 0$. As in the discrete model, we suppose that $\Delta t = 1/T$ and that $h = 1/N$, for two integers T and N. The discrete state space is defined as

$$S = \mathbb{T}^d / h\mathbb{Z}^d = \big\{ (i_1, i_2, \ldots, i_d)h \ \mid \ i_1, \ldots, i_d \in \mathbb{Z}/N\mathbb{Z} \big\}.$$

Given $x \in \mathbb{T}^d$, let us set $B_h(x) = \prod_{i=1}^n [x - he_i/2, \, x + he_i/2)$. Given $y \in \mathbb{T}^d$, we denote by $x_h[y]$ the unique point x in S such that $y \in B_h(x)$. We consider the mappings $\mathcal{I}_h \colon \mathbb{R}(\mathbb{T}^d) \to \mathbb{R}(S)$ and $\mathcal{R}_h \colon \mathbb{R}(S) \to \mathbb{R}(\mathbb{T}^d)$, defined as follows: For any $m^c \in \mathbb{R}(\mathbb{T}^d)$ and for any $m \in \mathbb{R}(S)$,

$$\begin{aligned}
\mathcal{I}_h(m^c)(x) &= \int_{B_h(x)} m^c(y) dy, && \forall x \in S; \\
\mathcal{R}_h(m)(y) &= \frac{m(x_h[y])}{h^d}, && \forall y \in \mathbb{T}^d.
\end{aligned}$$

We consider the constant $M > 0$, defined by

$$M = \frac{1}{\alpha^c} \Big(2 \max_{(t,x) \in Q} \|\ell_v^c(t,x,0)\| + \sqrt{d}(L_\ell^c + L_f^c + L_g^c) \Big). \tag{5.4.2}$$

Note that M is independent of Δt and h. We define the truncated running cost $\hat{\ell}^c$ as follows:

$$\hat{\ell}^c(t,x,v) = \begin{cases} \ell^c(t,x,v), & \text{if } \|v\| \leq M, \\ +\infty, & \text{otherwise.} \end{cases}$$

The discrete counterparts of ℓ^c, g^c, f^c, and m_0^c are defined as follows: For any $t \in \mathcal{T}$, for any $x \in S$, for any $v \in \mathbb{R}^d$, for any $m \in \mathcal{P}(S)$,

$$\begin{aligned}
\ell(t,x,p) &= \hat{\ell}^c(t\Delta t, x, v) \\
g(x) &= g^c(x), \\
f(t,x,m) &= \frac{1}{h^d} \int_{B_h(x)} f^c\big(t\Delta t, y, \mathcal{R}_h(m)\big) dy, \\
m_0(x) &= \mathcal{I}_h(m_0^c)(x).
\end{aligned} \tag{5.4.3}$$

We denote by $H\colon \mathcal{T}\times S\times \mathbb{R}^d$ the associated Hamiltonian, defined by

$$H(t,x,p) = \sup_{v\in\mathbb{R}^d} -\langle p,v\rangle - \ell(t,x,v). \tag{5.4.4}$$

Let us now discetize the differential operators appearing in the continuous MFG system. Let $(e_i)_{i=1,\dots,d}$ denote the canonical basis of $\mathbb{R}^d$. The discrete Laplace, gradient and divergence operators for the centered finite-difference scheme are defined as follows:

$$\begin{aligned}
\Delta_h\mu(x) &= \sum_{i=1}^d \frac{\mu(x+he_i)+\mu(x-he_i)-2\mu(x)}{h^2}, && \forall\, \mu\in\mathbb{R}(S),\ \forall\, x\in S,\\
\nabla_h\mu(x) &= \Big(\frac{\mu(x+he_i)-\mu(x-he_i)}{2h}\Big)_{i=1}^d, && \forall\, \mu\in\mathbb{R}(S),\ \forall\, x\in S,\\
\mathrm{div}_h\mu(x) &= \sum_{i=1}^d \frac{\mu_i(x+he_i)-\mu_i(x-he_i)}{2h}, && \forall\, \mu\in\mathbb{R}^d(S),\ \forall\, x\in S,
\end{aligned}$$

where μ_i is the i-th coordinate of μ.

Finally, we fix $\theta\in(1/2,1)$. The theta-scheme for (MFG), introduced in [BLP22, Sec. 2.3], writes:

$$\begin{cases}
\text{(i)} & u=\mathbf{HJB}_\theta(m),\\
\text{(ii)} & v=\mathbf{V}_\theta(u),\\
\text{(iii)} & m=\mathbf{FP}_\theta(v),
\end{cases} \tag{Theta-mfg}$$

where the Hamilton-Jacobi-Bellman mapping $\mathbf{HJB}_\theta\colon \mathcal{P}(\tilde{\mathcal{T}},S)\to\mathbb{R}(\tilde{\mathcal{T}}\times S)$, $m\mapsto u$, is defined by

$$\begin{cases}
-\frac{u(t+1,x)-u(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h u(t+1/2,x) = 0,\\
-\frac{u(t+1/2,x)-u(t,x)}{\Delta t} - (1-\theta)\sigma\Delta_h u(t+1/2,x) + H[\nabla_h u(\cdot+1/2,\cdot)](t,x) = f(t,x,m(t)),\\
u(T,x)=g(x),
\end{cases}$$

the optimal control mapping $\mathbf{V}_\theta\colon \mathbb{R}(\tilde{\mathcal{T}}\times S)\to\mathbb{R}^d(\mathcal{T}\times S)$, $u\mapsto v$, is defined by

$$\begin{cases}
-\frac{u(t+1,x)-u(t+1/2,x)}{\Delta t} - \theta\sigma\Delta_h u(t+1/2,x) = 0,\\
v(t,x) = -H_p(t,x,\nabla_h u(t+1/2,x)),
\end{cases}$$

and the Fokker-Planck mapping $\mathbf{FP}_\theta\colon \mathbb{R}^d(\mathcal{T}\times S)\to\mathbb{R}(\tilde{\mathcal{T}}\times S)$, $v\mapsto m$, is defined by

$$\begin{cases}
\frac{m(t+1,x)-m(t,x)}{\Delta t} - \theta\sigma\Delta_h m(t+1,x) - (1-\theta)\sigma\Delta_h m(t,x) + \mathrm{div}_h\big(vm(t,x)\big) = 0;\\
m(0,x) = m_0(x).
\end{cases}$$

Let us briefly motivate the theta-scheme. For the Fokker-Planck equation, a Crank-Nicolson discretization of the Laplace operator and an explicit discretization of the first-order term are utilized. An adjoint scheme is used for the HJB equation.

Given $u(t+1,\cdot)$, one first needs to solve an implicit scheme for the heat equation, corresponding to a diffusion equal to $\theta\sigma$. This yields the intermediate function $u(t+1/2,\cdot)$, which is not "saved" in the output $\mathbf{HJB}_\theta$. Then $u(t,\cdot)$ is obtained by solving an explicit scheme, containing the first order term and a diffusion equal to $(1-\theta)\sigma$.

Remark 5.4.1. The evaluation of the mapping $\mathbf{HJB}_\theta$, which is a crucial step in the GFW algorithm, requires to solve successively linear equations (for the implicit part) and explicit equations. It is therefore easier to implement than fully implicit schemes, which would require to solve general non-linear equations (with a policy iteration algorithm, for example).

Let us consider the following CFL condition:

$$\Delta t \leq \frac{h^2}{2d(1-\theta)\sigma}, \qquad h \leq \frac{2(1-\theta)\sigma}{M}. \tag{CFL}$$

Lemma 5.4.2. *Let Assumption A and condition* (CFL) *hold true. Then the system* (Theta-mfg) *is a particular case of* (DMFG)*, with ℓ, f, m_0 and g defined by* (5.4.3)*. Furthermore, Assumptions 5.1-5.2 hold with the constants*

$$D = M, \quad L_f = L_f^c h^{-d/2}, \quad \textit{and} \quad \alpha = \alpha^c.$$

Proof. See [BLP22, Lem. 4.1, Lem. 4.2, Thm. 4.4]. □

Let us emphasize that the well-posedness of the mappings $\mathbf{HJB}_\theta$, $\mathbf{V}_\theta$, and $\mathbf{FP}_\theta$ is a corollary of the fact that (Theta-mfg) is a discrete MFG. The explicit formulas for π_0 and π_1 are provided in [BLP22, Lemma 4.1]. We define the traces $u_h^* \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ and $m_h^* \in \mathcal{P}(\tilde{\mathcal{T}}, S)$ of the continuous solution (u^*, m^*) as follows:

$$u_h^*(t,x) = u^*(t\Delta, x) \quad \text{and} \quad m_h^*(t,\cdot) = \mathcal{I}_h(m^*(t\Delta t,\cdot)), \quad \forall t \in \tilde{\mathcal{T}}, \forall x \in S. \tag{5.4.5}$$

The main result of [BLP22] is the following error estimate.

Theorem 5.4.3. *Let $1/2 < \theta < 1$. Let Assumptions A-B and condition* (CFL) *hold true. Then* (Theta-mfg) *admits a unique solution (u_h, v_h, m_h). There exists C^* independent of Δt and h such that*

$$\|u_h - u_h^*\|_{\infty,\infty} + \|m_h - m_h^*\|_{\infty,1} \leq C^* h^r.$$

Proof. See [BLP22, Thm. 2.10]. □

5.4.2 GFW algorithm for the theta-scheme and main results

Let us first give the assumption on the potential structure of f.

Assumption C. There exists a function $F^c \colon [0,1] \times \mathcal{D} \to \mathbb{R}^d$ such that (5.1.1) holds.

The following lemma shows that the discretization of f^c proposed in (5.4.3) preserves the potential structure of the MFG system at the discrete level.

Lemma 5.4.4. *Let Assumption C hold true. Then, f defined by* (5.4.3) *satisfies Assumption 5.3, with the primitive function F defined by*

$$F(t,m) = F^c(t\Delta t, \mathcal{R}_h(m)).$$

Proof. Taking any $(t, m_1, m_2) \in \mathcal{T} \times \mathcal{P}(S)^2$, we have that

$$\begin{aligned} F(t,m_1) - F(t,m_2) &= F^c(t\Delta t, \mathcal{R}_h(m_1)) - F^c(t\Delta t, \mathcal{R}_h(m_2)) \\ &= \int_0^1 \int_{x\in\mathbb{T}^d} f^c(t\Delta t, x, \mathcal{R}_h(m_1 + s(m_2 - m_1)))(\mathcal{R}_h(m_2)(x) - \mathcal{R}_h(m_1)(x))dxds \\ &= \int_0^1 \sum_{x\in S} \frac{1}{h^d} \int_{y\in B_h(x)} f^c(t\Delta t, y, \mathcal{R}_h(m_1 + s(m_2 - m_1)))dy(m_2(x) - m_1(x))ds \\ &= \int_0^1 \sum_{x\in S} f(t,x,m_1 + s(m_2 - m_1))(m_2(x) - m_1(x))ds, \end{aligned}$$

where the second equality follows from Assumption C and the linearity of operator $\mathcal{R}_h$, the third equality comes from the fact that $\mathcal{R}_h(m)$ is constant in $B_h(x)$, and the last equality derives from the definition of f. □

Following the definition of **BR** in Sec. 5.3, we define the best response mapping for (Theta-mfg): for any $m' \in \mathcal{P}(\tilde{\mathcal{T}}, S)$,

$$\mathbf{BR}_\theta(m') = \chi(\tilde{m}, \tilde{v}), \qquad \text{where } \tilde{v} = \mathbf{V}_\theta \circ \mathbf{HJB}_\theta(m') \quad \text{and} \quad \tilde{m} = \mathbf{FP}_\theta(\tilde{v}).$$

The GFW algorithm writes as follows.

Algorithm 5.2: Generalized Frank-Wolfe Algorithm for (Theta-mfg)

Initialization: $m_h^0 \in \mathcal{P}(\tilde{\mathcal{T}}, S)$;
First iteration: $(m_h^1, w_h^1) = \mathbf{BR}_\theta(m_h^0)$;
for $k = 1, 2, \ldots$ **do**

    Step 1: Resolution of the partial linearized problem.
    Set $(\bar{m}_h^k, \bar{w}_h^k) = \mathbf{BR}_\theta(m_h^k)$;

    Step 2: Update.
    Choose $\lambda_k \in [0,1]$;
    Set $(m_h^{k+1}, w_h^{k+1}) = (1-\lambda_k)(m_h^k, w_h^k) + \lambda_k(\bar{m}_h^k, \bar{w}_h^k)$;

end

From now on, all notations introduced in Sections 5.2 and 5.3 for general discrete MFGs will be restricted to (Theta-mfg), without the adjunction of the subscript θ. For example, we will denote by γ_k the optimality gap in Algorithm 5.2. The following result is our first mesh-independence principle. Recall that (u_h, m_h) and (u_h^*, m_h^*) have been introduced in Theorem 5.4.3.

Theorem 5.4.5 (Sublinear rate)**.** *Let Assumptions A-C and condition* (CFL) *hold true. In Algorithm 5.2, take $\lambda_k = 2/(k+2)$, for any $k \geq 1$. Then there exist two constants C_θ and $\bar{C}_\theta$, both independent of Δt and h, such that*

$$\gamma_k \leq \frac{C_\theta}{k}, \quad \forall k \geq 1, \tag{5.4.6}$$

and such that

$$\begin{aligned} \|u_h^k - u_h\|_{\infty,\infty} + \|m_h^k - m_h\|_{\infty,1} &\leq \frac{\bar{C}_\theta}{\sqrt{kh^{d/2}}}, \\ \|u_h^k - u_h^*\|_{\infty,\infty} + \|m_h^k - m_h^*\|_{\infty,1} &\leq \bar{C}_\theta \left(h^r + \frac{1}{\sqrt{kh^{d/2}}} \right), \end{aligned} \tag{5.4.7}$$

for any $k \geq 1$. In particular, for $k \geq h^{-(2r+d/2)}$, we have

$$\|u_h^k - u_h^*\|_{\infty,\infty} + \|m_h^k - m_h^*\|_{\infty,1} \leq 2\bar{C}_\theta h^r. \tag{5.4.8}$$

Remark 5.4.6. Let us emphasize that only the estimate (5.4.6) is mesh-independent. The estimates provided in (5.4.7) get worse as $h \to 0$.

The proof of Theorem 5.4.5 is given in Subsec. 5.4.3. It is based on Proposition 5.3.1, more precisely on estimates of the three fundamental constants C_1, C_2, and C_3. We will derive from these estimates some estimates of the constants D_1, D_2, and c (as defined in (5.3.4)). A key point is that the estimate of the constant D_1 is mesh-independent. Our analysis mainly relies on a stability result for the discrete Fokker-Planck equation, for the ℓ^2-norm.

In order to establish mesh-independent estimates of D_2 and c, so as to have a mesh-independent linear rate of convergence, we need to establish an ℓ^∞-estimate for the discrete Fokker-Planck, assuming that the involved vector field v derives from a value function u. In the continuous case, such an estimate is relatively easy to deduce from the semi-concavity of u. The transposition of this analysis to a discrete setting is quite delicate; our analysis is restricted to the case of a separable running cost (with respect to the control variables).

A function $l\colon \mathbb{R}^d \to \mathbb{R}$ is said to be semi-concave if $l(x) - L\|x\|^2/2$ is concave, for some $L \geq 0$. This definition makes no sense when l is a function defined on torus. In fact, one can check that a periodical function is concave if and only if it is constant. Besides the previous definition, a second one based on a quadratic inequality is introduced in [CS04, Def. 1.1.1] for functions on an open set. We use the latter one for our definition of semi-concavity on a torus and its discretization.

Definition 5.4.7. [Semi-concave functions on the torus] Let L be a positive constant. A function $l^c\colon \mathbb{T}^d \to \mathbb{R}$ is said to be L-semi-concave if

$$l^c(x) \geq \frac{l^c(x+y) + l^c(x-y)}{2} - L\|y\|^2, \qquad \forall x,\, y \in \mathbb{T}^d. \tag{5.4.9}$$

Remark 5.4.8. If a function $l^c\colon \mathbb{T}^d \to \mathbb{R}$ is $\mathcal{C}^2$, then it is semi-concave, as can be easily verified.

We consider the following assumption.

Assumption D. The following holds:

1. *Semi-concavity.* There exists $L^c > 0$, such that for any $(t, v, m) \in [0,1] \times \mathbb{R}^d \times \mathcal{D}$ with $\|v\| \leq M$, the functions $\ell^c(t, x, v)$, $f^c(t, x, m)$, and $g^c(x)$ are L^c-semi-concave with respect to x.
2. *Separability.* There exist functions $\ell_i^c \colon Q \times \mathbb{R} \to \mathbb{R}$ for $i = 1, 2, \ldots, d$, such that for any $(t, x, v) \in Q \times \mathbb{R}^d$,
$$\ell^c(t, x, v) = \sum_{i=1}^d \ell_i^c(t, x, v_i),$$
where v_i is the i-th coordinate of v.

In the next theorem we establish a linear and mesh-independent rate of convergence for the GFW algorithm with the line-search rule (5.3.6).

Theorem 5.4.9 (Linear rate)**.** *Let Assumptions A-D and condition* (CFL) *hold true. In Algorithm 5.2, set λ_k with the rule* (5.3.6)*, for all $k \geq 1$. Then there exist three constants $c_\theta \in (0,1)$, $C_\theta > 0$ and $\bar{C}_\theta$, independent of Δt and h, such that*

$$\gamma_k \leq C_\theta c_\theta^k, \quad \forall k \geq 1, \tag{5.4.10}$$

and such that

$$\begin{aligned} \|u_h^k - u_h\|_{\infty,\infty} + \|m_h^k - m_h\|_{\infty,1} &\leq \bar{C}_\theta c_\theta^{k/2}, \\ \|u_h^k - u_h^*\|_{\infty,\infty} + \|m_h^k - m_h^*\|_{\infty,1} &\leq \bar{C}_\theta \left(h^r + c_\theta^{k/2} \right), \end{aligned} \tag{5.4.11}$$

for any $k \geq 1$. In particular, taking $k \geq 2r \log(h)/\log(c_\theta)$, we have

$$\|u_h^k - u_h^*\|_{\infty,\infty} + \|m_h^k - m_h^*\|_{\infty,1} \leq 2\bar{C}_\theta h^r. \tag{5.4.12}$$

Note that both estimates (5.4.10) and (5.4.11) are now mesh-independent. The proof of Theorem 5.4.9 is given in Subsec. 5.4.4.

5.4.3 Proof of the sublinear rate of convergence

We prove in this section Theorem 5.4.5. Assumptions A-C are supposed to be satisfied all along the subsection, as well as the condition (CFL). Our analysis relies on an energy estimate, obtained in Lemma 5.4.10, which allows us to find first estimates of the convergence constants C_1, C_2, C_3, D_1, and D_2 (in Lemma 5.4.11).

We define the forward discrete gradient as follows:

$$\nabla_h^+ \nu(x) = \Big(\frac{\nu(x + he_i) - \nu(x)}{h} \Big)_{i=1}^d, \qquad \forall \nu \in \mathbb{R}(S), \ \forall x \in S.$$

Lemma 5.4.10 (Energy estimate)**.** *Let $\mu \in \mathbb{R}(\tilde{\mathcal{T}} \times S)$ satisfy the following equation:*

$$\begin{cases} (Id - \theta\sigma\Delta t\Delta_h)\, \mu(t+1) = \big(Id + (1-\theta)\sigma\Delta t\Delta_h\big)\mu(t) - \Delta t \operatorname{div}_h\big(v(t)\mu(t)\big) - \Delta t \operatorname{div}_h\big(\delta_v(t)\big), \\ \mu(0) = \mu_0, \end{cases}$$

where $\delta_v(t) \in \mathbb{R}^d(S)$ *and* $\|v\|_{\infty,\infty} \le M$*. Then,*

$$\max_{t\in\tilde{\mathcal{T}}} \big\|\mu(t)\big\|_2^2 \le c(\sigma,\theta,M)\left(\big\|\mu_0\big\|_2^2 + (1-\theta)\sigma\Delta t\big\|\nabla_h^+\mu_0\big\|_2^2 + \frac{1}{\sigma(2\theta-1)}\sum_{\tau\in\mathcal{T}}\Delta t\big\|\delta_v(\tau)\big\|_2^2\right), \quad (5.4.13)$$

where

$$c(\sigma,\theta,M) := 1 + \frac{M^2}{\sigma(2\theta-1)}\exp\left(\frac{M^2}{\sigma(2\theta-1)}\right). \quad (5.4.14)$$

Proof. We deduce from the proof of [BLP22, Prop. 4.5] that

$$\begin{aligned}\frac{1}{2}\Big(\|\mu(t+1)\|_2^2 - \|\mu(t)\|_2^2\Big) + \theta\sigma\Delta t\big\|\nabla_h^+\mu(t+1)\big\|_2^2 \\ \le -(1-\theta)\sigma\Delta t\big\langle\nabla_h^+\mu(t+1),\nabla_h^+\mu(t)\big\rangle + \Delta t\big(\gamma_1+\gamma_2\big),\end{aligned} \quad (5.4.15)$$

where

$$\gamma_1 = \sum_{x\in S}\big\langle\nabla_h\mu(t+1,x),\mu v(t,x)\big\rangle, \qquad \gamma_2 = \sum_{x\in S}\big\langle\nabla_h\mu(t+1,x),\delta_v(t,x)\big\rangle.$$

Let $\alpha_1 = \sigma(2\theta-1) > 0$. Applying Young's inequality to the right-hand side of (5.4.15), we have the following inequalities:

$$\begin{aligned}-\big\langle\nabla_h^+\mu(t+1),\nabla_h^+\mu(t)\big\rangle &\le \frac{1}{2}\big\|\nabla_h^+\mu(t+1)\big\|_2^2 + \frac{1}{2}\big\|\nabla_h^+\mu(t)\big\|_2^2;\\ \gamma_1 &\le \frac{\alpha_1}{2}\big\|\nabla_h\mu(t+1)\big\|_2^2 + \frac{1}{2\alpha_1}\big\|\mu v(t)\big\|_2^2 \le \frac{\alpha_1}{2}\big\|\nabla_h^+\mu(t+1)\big\|_2^2 + \frac{M^2}{2\alpha_1}\big\|\mu(t)\big\|_2^2;\\ \gamma_2 &\le \frac{\alpha_1}{2}\big\|\nabla_h^+\mu(t+1)\big\|_2^2 + \frac{1}{2\alpha_1}\big\|\delta_v(t)\big\|_2^2.\end{aligned}$$

Combining the above inequalities and (5.4.15), it follows that

$$\|\mu(t+1)\|_2^2 - \|\mu(t)\|_2^2 + (1-\theta)\sigma\Delta t\Big(\big\|\nabla_h^+\mu(t+1)\big\|_2^2 - \big\|\nabla_h^+\mu(t)\big\|_2^2\Big) \le \Delta t\Big(\frac{M^2}{\alpha_1}\big\|\mu(t)\big\|_2^2 + \frac{1}{\alpha_1}\big\|\delta_v(t)\big\|_2^2\Big).$$

Summing the previous inequality over t, it follows that for any $t \in \mathcal{T}$,

$$\|\mu(t+1)\|_2^2 \le \gamma + \frac{M^2\Delta t}{\alpha_1}\sum_{\tau=0}^{t}\big\|\mu(\tau)\big\|_2^2,$$

where

$$\gamma = \|\mu_0\|_2^2 + (1-\theta)\sigma\Delta t\|\nabla_h^+\mu_0\|_2^2 + \frac{1}{\alpha_1}\sum_{\tau\in\mathcal{T}}\Delta t\big\|\delta_v(\tau)\big\|_2^2.$$

We deduce from the discrete Gronwall inequality [Cla87] that

$$\max_{t\in\tilde{\mathcal{T}}}\big\|\mu(t)\big\|_2^2 \le \gamma + T\gamma\frac{M^2\Delta t}{\alpha_1}\exp\left(\sum_{\tau\in\mathcal{T}}\frac{M^2\Delta t}{\alpha_1}\right) = \gamma\left(1+\frac{M^2}{\alpha_1}\exp\left(\frac{M^2}{\alpha_1}\right)\right).$$

The conclusion follows. □

We define next two constants E_1 and E_3, both independent of Δt and h:

$$E_1 = c(\sigma,\theta,M)\Big(\|m_0^c\|_{\mathbb{L}^\infty}^2 + \frac{1}{2}\|\nabla m_0^c\|_{\mathbb{L}^\infty}^2\Big) \quad \text{and} \quad E_3 = \frac{c(\sigma,\theta,M)}{\sigma(2\theta-1)}.$$

A constant E_2 will be introduced later on.

Lemma 5.4.11. *The constants C_1, C_2, and C_3 (as defined in* (5.3.1)-(5.3.3)*) satisfy the following inequalities:*

$$C_1 \leq E_1 h^d, \qquad C_2 \leq \sqrt{C_1} \leq E_1^{1/2} h^{d/2}, \qquad C_3 \leq E_3. \tag{5.4.16}$$

As a consequence, for the constants D_1 and D_2 defined in (5.3.4)*, we have*

$$D_1 \leq E_1 L_f^c, \quad D_2 \leq (1+L_f^c)\sqrt{\frac{2E_1^{3/2}E_3 L_f^c}{\alpha^c}}\, h^{-d/4}. \tag{5.4.17}$$

Proof. The condition (CFL) implies that

$$d(1-\theta)\sigma\Delta t \leq \frac{h^2}{2} \leq \frac{1}{2}. \tag{5.4.18}$$

We let the reader verify that

$$\|m_0\|_2^2 + (1-\theta)\sigma\Delta t\|\nabla_h^+ m_0\|_2^2 \leq \left(\|m_0^c\|_{\mathbb{L}^\infty}^2 + d(1-\theta)\sigma\Delta t\|\nabla m_0^c\|_{\mathbb{L}^\infty}^2\right) h^d.$$

Combining the above two estimates, we deduce that

$$\|m_0\|_2^2 + (1-\theta)\sigma\Delta t\|\nabla_h^+ m_0\|_2^2 \leq \Big(\|m_0^c\|_{\mathbb{L}^\infty}^2 + \frac{1}{2}\|\nabla m_0^c\|_{\mathbb{L}^\infty}^2\Big) h^d.$$

Using this inequality in Lemma 5.4.10, applied with $\delta_v(t) = 0$ and $\mu_0 = m_0$, we obtain the estimate of C_1. Then, the estimate of C_2 is deduced from Hölder's inequality. The estimate of C_3 follows from equality (5.2.8) and Lemma 5.4.10 by taking $\mu_0 = 0$ and $\delta_v = (v_1 - v_2)m_1$. Finally, (5.4.17) follows from (5.3.4) and the previous estimates. □

Proof of Theorem 5.4.5. Inequality (5.4.6) directly follows from Proposition 5.3.1 and from the estimate $D_1 \leq E_1 L_f^c$, with $C_\theta = 8E_1 L_f^c$. Then, using Lemma 5.3.5 and the estimate of D_2 obtained in (5.4.17), we deduce that

$$\|u^k - u_h\|_{\infty,\infty} + \|m^k - m_h\|_{\infty,2} \leq D_2\sqrt{\gamma_k} \leq 4E_1 L_f^c(1+L_f^c)\sqrt{\frac{E_1E_3}{\alpha^c}}\frac{1}{\sqrt{k h^{d/2}}}.$$

Recall that C^* was introduced in Theorem 5.4.3. The inequalities in (5.4.7) hold true with

$$\bar{C}_\theta = \max\left\{4E_1 L_f^c(1+L_f^c)\sqrt{\frac{E_1E_3}{\alpha^c}}\,,\, C^*\right\}.$$

The theorem is proved. □

5.4.4 Proof of the linear rate of convergence

We prove in this subsection Theorem 5.4.9. Assumptions A-D are supposed to be satisfied all along the subsection, as well as the condition (CFL). At a technical level, we look for a more precise estimate of the constant C_2 (see Lemma 5.4.22), using an ℓ^∞-stability result for the map $\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta$ (see Lemma 5.4.21), whose proof is inspired from the continuous case (see for example [CL18a, Lem. 5.3]). Our analysis begins with a series of technical lemmas which will allow us to establish the semi-concavity of the value function.

Definition 5.4.12. [Semi-concave functions on S] A function $l\colon S \to \mathbb{R}$ is said to be L-semi-concave if

$$l\,(x) \geq \frac{l(x+y)+l(x-y)}{2} - L\|y\|^2, \qquad \forall x,\, y \in S. \tag{5.4.19}$$

Lemma 5.4.13. *Let L, L_1, and L_2 be positive constants. The following statements hold:*

1. *Let $l_1\colon S \to \mathbb{R}$ be L_1-semi-concave and $l_2\colon S \to \mathbb{R}$ be L_2-semi-concave. For any $\lambda_1, \lambda_2 \geq 0$, the function $\lambda_1 l_1 + \lambda_2 l_2$ is $(\lambda_1 L_1 + \lambda_2 L_2)$-semi-concave.*
2. *Let $\{l_\omega\colon S \to \mathbb{R}\}_{\omega\in\Omega}$ be a family of L-semi-concave functions. Let $l\colon S \to \mathbb{R}$, $l(x) = \inf_{\omega\in\Omega} l_\omega(x)$. Suppose that for all $x \in S$, it holds that $l(x_0) > -\infty$. Then l is L-semi-concave.*

Proof. The first point is obtained by the definition (5.4.19) and the non-negativity of λ_1 and λ_2. Let us prove the second point. Let $x_0 \in S$. We have $l(x_0) > -\infty$. Then for any $\epsilon > 0$, there exists $\omega_\epsilon \in \Omega$ such that

$$l(x_0) \geq l_{\omega_\epsilon}(x_0) - \epsilon \geq \frac{l_{\omega_\epsilon}(x_0+y) + l_{\omega_\epsilon}(x_0-y)}{2} - L\|y\|^2 - \epsilon \geq \frac{l(x_0+y)+l(x_0-y)}{2} - L\|y\|^2 - \epsilon,$$

where the second inequality follows from the semi-concavity of l_{ω_ϵ}. We deduce the L-semi-concavity at the point x_0 by the arbitrariness of ϵ. ◻

Lemma 5.4.14. *The functions ℓ, f, and g, defined in* (5.4.3)*, are L^c-semi-concave with respect to x.*

Proof. The semi-concavity of ℓ and g is a direct consequence of their definitions in (5.4.3). Let us prove the semi-concavity of f. By the definition of f in (5.4.3), taking any $t \in \mathcal{T}$, $m \in \mathcal{P}(\mathcal{T}, S)$, and $x, y \in S$, we have

$$\begin{aligned}
&f(t,x+y,m) + f(t,x-y,m)\\
&\quad= \frac{1}{h^d}\int_{z\in B_h(x+y)} f^c(t,z,\mathcal{R}_h(m))dz + \frac{1}{h^d}\int_{z\in B_h(x-y)} f^c(t,z,\mathcal{R}_h(m))dz\\
&\quad= \frac{1}{h^d}\int_{z\in B_h(0)} f^c(t,x+y+z,\mathcal{R}_h(m)) + f^c(t,x-y+z,\mathcal{R}_h(m))dz\\
&\quad\leq \frac{1}{h^d}\int_{z\in B_h(0)} 2f^c(t,x+z,\mathcal{R}_h(m)) + 2L^c\|y\|^2 dz\\
&\quad= \frac{1}{h^d}\int_{z\in B_h(x)} 2f^c(t,z,\mathcal{R}_h(m))dz + 2L^c\|y\|^2 = 2f(t,x,m) + 2L^c\|y\|^2.
\end{aligned}$$

The conclusion follows. ◻

Definition 5.4.15. Let A be a function from $\mathbb{R}(S)$ to $\mathbb{R}$. We say that A is *translation invariant* if for any $X \in \mathbb{R}(S)$ and any $y \in S$,

$$A(X(\cdot)) = A(X(\cdot - y)).$$

Lemma 5.4.16. *Suppose that $X, Y \in \mathbb{R}(S)$ satisfy the following equation for some $c > 0$:*

$$(Id - c\Delta t \Delta_h)Y = X. \tag{5.4.20}$$

Let $A\colon \mathbb{R}(S) \to \mathbb{R}$ be a translation invariant function. We have the following statements:

- *If A is convex and l.s.c, then $A(Y) \leq A(X)$.*
- *If A is concave and u.s.c, then $A(Y) \geq A(X)$.*

Proof. For any $Z \in \mathbb{R}(S)$ and $y \in S$, we define $\tau^y Z \in \mathbb{R}(S)$ by

$$\tau^y Z(\cdot) = Z(\cdot - y).$$

Let $\gamma = c\Delta t/h^2$. We define $\mathbb{S}_X\colon \mathbb{R}(S) \to \mathbb{R}(S)$,

$$\mathbb{S}_X(\mu) = \frac{1}{1+2d\gamma}\left\{X + \gamma\sum_{j=1}^{d}\left(\tau^{he_j}\mu + \tau^{-he_j}\mu\right)\right\}.$$

By the proof of [BLP22, Lemma 2.6], $\mathbb{S}_X$ is a contraction mapping and Y is the fixed point of $\mathbb{S}_X$.

Suppose that A is l.s.c. and convex. Suppose that $\mu \in \mathbb{R}(S)$ satisfy $A(\mu) \leq A(X)$. By the convexity of A, we have

$$\begin{aligned} A(\mathbb{S}_X(\mu)) &\leq \frac{1}{1+2d\gamma}\left\{A(X) + \gamma\sum_{j=1}^{d}\left(A(\tau^{he_j}\mu) + A(\tau^{-he_j}\mu)\right)\right\} \\ &= \frac{1}{1+2d\gamma}\left\{A(X) + \gamma\sum_{j=1}^{d}\left(A(\mu) + A(\mu)\right)\right\} \leq A(X), \end{aligned}$$

where the second line follows from the translation invariance of A. Therefore, $A(\mathbb{S}_X^k(X)) \leq A(X)$ for any $k \geq 1$. Since $Y = \lim_{k\to\infty}\mathbb{S}_X^k(X)$, by the lower-semi-continuity of A, we have

$$A(Y) \leq \liminf_{k\to\infty} A(\mathbb{S}_X^k(X)) \leq A(X).$$

For the case where A is u.s.c and concave, it suffices to apply the previous result to $-A$. ☐

Lemma 5.4.17. *Let $X, Y \in \mathbb{R}(S)$ satisfy* (5.4.20)*. Then, the following statements hold.*

1. *Maximum/minimum principle:*

$$\min_{x\in S} X(x) \leq \min_{x\in S} Y(x) \leq \max_{x\in S} Y(x) \leq \max_{x\in S} X(s).$$

2. *Conservation of the mass:*
$$\sum_{x\in S} Y(s) = \sum_{x\in S} X(s).$$

3. *Conservation of the Lipschitz constant: If $X(x)$ is L-Lipschitz, then $Y(x)$ is L-Lipschitz.*

4. *Conservation of the semi-concavity constant: If $X(x)$ is L-semi-concave, then $Y(x)$ is L-semi-concave.*

Proof. We use Lemma 5.4.16 for the proof. The key point is the choice of the translation invariant function A in Lemma 5.4.16. Keep in mind that the maximum (resp. minimum) of a family of linear functions is l.s.c. and convex (resp. u.s.c. and concave) in finite dimensions.

For point (1), it suffices to take $A(X) = \min_{x\in S}\{X(x)\}$ and $A(X) = \max_{x\in S}\{X(x)\}$. For point (2), we take $A(X) = \sum_{x\in S} X(x)$. For point (3), we take
$$A(X) = \max_{x\in S} \max_{y\in S, y\neq 0} \frac{X(x+y) - X(x)}{\|y\|}.$$
Finally, for point (4), we take
$$A(X) = \max_{x\in S} \max_{y\in S, y\neq 0} \frac{X(x+y) + X(x-y) - 2X(x)}{\|y\|^2}.$$
The conclusion follows. □

Lemma 5.4.18 (Semi-concavity of the value function). *Let Assumptions A, D(1) and condition* (CFL) *hold true. Then for any $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, the unction $u = \mathbf{HJB}_\theta(m)$ is $3L^c$-semi-concave with respect to x.*

Proof. Observe that $\mathbf{HJB}_\theta(m)$ is equivalent to the formulation below:
$$\begin{cases} (Id - \theta\sigma\Delta t\Delta_h)\, u(t+1/2) = u(t+1); \\ u(t,x) = \Delta t \inf_{\|\omega\|\le M} \Big\{ f(t,x,m(t)) + \ell(t,x,\omega) + \big\langle \omega, \nabla_h u(t+1/2,x)\big\rangle \Big\} \\ \qquad\qquad + \big(Id + (1-\theta)\sigma\Delta t\Delta_h\big)u(t+1/2)(x), \qquad \forall x\in S; \\ u(T,x) = g(x), \qquad \forall x \in S. \end{cases} \tag{5.4.21}$$
We prove the lemma by induction. For $t = T$, by the terminal condition, it is obvious that
$$u\,(T,x) \ge \frac{u(T,x+y) + u(T,x-y)}{2} - L^c\|y\|^2, \qquad \forall x,\, y\in S.$$
Suppose that for some $t \in \mathcal{T}$, we have
$$u\,(t+1,x) \ge \frac{u(t+1,x+y) + u(t+1,x-y)}{2} - (2(T-1-t)\Delta t + 1)L^c\|y\|^2, \qquad \forall x,\, y\in S.$$
Since $u(t+1)$ and $u(t+1/2)$ satisfy the implicit scheme (5.4.20), by Lemma 5.4.17(4), we have
$$u\,(t+1/2,x) \ge \frac{u(t+1/2,x+y) + u(t+1/2,x-y)}{2} - (2(T-1-t)\Delta t + 1)L^c\|y\|^2, \qquad \forall x,\, y\in S.$$

Let $r' = (1-\theta)\sigma\Delta t/h^2$. The second equation in (5.4.21) can be written as follows:

$$u(t,x) = \inf_{\|\omega\|\le M} l_\omega(t,x),$$

where

$$\begin{aligned} l_\omega(t,x) :=& (1-2dr')u(t+1/2,x) + \sum_{i=1}^d \left(r' + \frac{\omega_i}{2h}\right)u(t+1/2,x+he_i) \\ &+ \sum_{i=1}^d \left(r' - \frac{\omega_i}{2h}\right)u(t+1/2,x-he_i) + \Delta t(f(t,x,m(t)) + \ell(t,x,\omega)). \end{aligned}$$

By condition (CFL), the coefficients of the above equation are positive for any $\|w\| \le M$. Then by Lemma 5.4.13(1), the semi-concavity of $u(t+1/2,\cdot)$, f, and ℓ, we have that l_ω is $(2(T-t)\Delta t+1)L^c$-semi-concave. Since $u(t,x) > -\infty$ for any $x \in S$, we deduce from Lemma 5.4.13(2) that

$$u(t,x) \ge \frac{u(t,x+y)+u(t,x-y)}{2} - (2(T-t)\Delta t+1)L^c\|y\|^2, \qquad \forall x,\, y \in S.$$

The conclusion follows by induction. □

We have the following regularity result for the discrete Hamiltonian H (defined in (5.4.4)).

Lemma 5.4.19. *Let Assumptions A and D(2) hold true. Then, for any $(t,x) \in \mathcal{T} \times S$ and $\|p\| \le \sqrt{d}(L_\ell^c + L_f^c + L_g^c)$,*

$$H(t,x,p) = \sum_{i=1}^d H_i^c(t\Delta t, x, p_i), \qquad H_p(t,x,p) = \left(\frac{\partial H_i^c}{\partial p_i}(t\Delta t,x,p_i)\right)_{i=1}^d, \tag{5.4.22}$$

where p_i is the i-th coordinate of p and

$$H_i^c(t,x,p_i) = \sup_{v_i} -v_i p_i - \ell_i^c(t,x,v_i). \tag{5.4.23}$$

Moreover, $\frac{\partial H_i^c}{\partial p_i}(t\Delta t,x,p_i)$ is L_ℓ^c/α^c-Lipschitz with respect to x.

Proof. Equality (5.4.22) is from [BLP22, Lemma 5.1] and the separable form of ℓ^c. The Lipschitz continuity of $\frac{\partial H_i^c}{\partial p_i}$ is proved with the same argument as the one the proof of [BLP22, Lemma 2.7]. □

Lemma 5.4.20 (Lipschitz continuity of the value function)**.** *Let Assumption A and condition* (CFL) *hold true. For any $m \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, let $u = \mathbf{HJB}_\theta(m)$ and $v = \mathbf{V}_\theta(u)$. Then u is $(L_\ell^c+L_f^c+L_g^c)$-Lipschitz with respect to x and $\|v\|_{\infty,\infty} \le M$.*

Proof. See [BLP22, Lemma 4.3]. □

Lemma 5.4.21 (ℓ^∞-stability)**.** *Let Assumptions A, D and condition* (CFL) *hold true. Then,*

$$\sup_{\mu\in\mathcal{P}(\tilde{\mathcal{T}},S)} \|\mathbf{FP}_\theta \circ \mathbf{V}_\theta \circ \mathbf{HJB}_\theta(\mu)\|_{\infty,\infty} \le \exp\left(\frac{d(L_\ell^c + 6L^c)}{\alpha^c}\right)\|m_0\|_\infty.$$

Proof. Let $\mu \in \mathcal{P}(\tilde{\mathcal{T}}, S)$, let $u = \mathbf{HJB}_\theta(\mu)$, let $v = \mathbf{V}_\theta(u)$, and let $m = \mathbf{FP}_\theta(v)$. Observe that $m = \mathbf{FP}_\theta(v)$ is equivalent to the formulation below:

$$\begin{cases} m(t+1/2) = \big(Id + (1-\theta)\sigma\Delta t\Delta_h\big)m(t) - \Delta t\mathrm{div}_h\big(v(t)m(t)\big); \\ (Id - \theta\sigma\Delta t\Delta_h)\, m(t+1) = m(t+1/2); \\ m(0) = m_0. \end{cases} \tag{5.4.24}$$

Let us first compare $\|m(t+1/2,\cdot)\|_\infty$ and $\|m(t+1,\cdot)\|_\infty$. Since $m(t+1)$ and $m(t+1/2)$ satisfy the implicit scheme (5.4.20), by Lemma 5.4.17(1), we have

$$\|m(t+1,\cdot)\|_\infty \leq \|m(t+1/2,\cdot)\|_\infty. \tag{5.4.25}$$

Then, we compare $\|m(t,\cdot)\|_\infty$ and $\|m(t+1/2,\cdot)\|_\infty$. Let $r' = (1-\theta)\sigma\Delta t/h^2$. The first equation in (5.4.24) shows that for any $(t,x) \in \mathcal{T} \times S$,

$$\begin{aligned} m(t+1/2,x) = (1-2d\gamma')m(t,x) + \sum_{i=1}^{d}\left(\gamma' - \Delta t\frac{v_i(t,x+he_i)}{2h}\right)m(t,x+he_i) \\ + \sum_{i=1}^{d}\left(\gamma' + \Delta t\frac{v_i(t,x-he_i)}{2h}\right)m(t,x-he_i). \end{aligned}$$

Condition (CFL) implies that all the coefficients in the above equation are positive. Therefore, for any $(t,x) \in \mathcal{T} \times S$,

$$m(t+1/2,x) \leq \left(1 - \Delta t\sum_{i=1}^{d}\frac{v_i(t,x+he_i) - v_i(t,x-he_i)}{2h}\right)\|m(t,\cdot)\|_\infty = (1-\Delta t\mathrm{div}_h v)\,\|m(t,\cdot)\|_\infty. \tag{5.4.26}$$

By Lemma 5.4.20 and Lemma 5.4.17(3), we have $\|\nabla_h u(t+1/2,x)\| \leq \sqrt{d}(L_\ell^c + L_f^c + L_g^c)$ for any $(t,x) \in \mathcal{T} \times S$. Then, formula (5.4.22) implies that

$$\begin{aligned} -\mathrm{div}_h v(t,x) &= \frac{1}{2h}\sum_{i=1}^{d}\frac{\partial H_i^c}{\partial p_i}(t\Delta t, x+he_i, (\nabla_h u(t+1/2,x+he_i))_i) \\ &\qquad - \frac{\partial H_i^c}{\partial p_i}(t\Delta t, x-he_i, (\nabla_h u(t+1/2,x-he_i))_i) \\ &\leq \frac{dL_\ell^c}{\alpha^c} + \frac{1}{2h}\sum_{i=1}^{d}\frac{\partial H_i^c}{\partial p_i}(t\Delta t, x, (\nabla_h u(t+1/2,x+he_i))_i) \\ &\qquad - \frac{\partial H_i^c}{\partial p_i}(t\Delta t, x, (\nabla_h u(t+1/2,x-he_i))_i), \end{aligned} \tag{5.4.27}$$

where the last inequality follows from the Lipschitz-continuity of $\frac{\partial H_i^c}{\partial p_i}$ with respect to x (established in Lemma 5.4.19). Since H_i^c is convex on p_i, the derivative $\frac{\partial H_i^c}{\partial p_i}$ is non-decreasing with respect to p_i. Furthermore, we know that $\frac{\partial H_i^c}{\partial p_i}$ is $1/\alpha^c$-Lipschitz on p_i by the strong convexity of ℓ^c. It follows that for any $(t,x) \in Q$ and $p_i^1, p_i^2 \in \mathbb{R}$,

$$\frac{\partial H_i^c}{\partial p_i}(t,x,p_i^1) - \frac{\partial H_i^c}{\partial p_i}(t,x,p_i^2) \leq \max\left\{0, \frac{1}{\alpha^c}(p_i^1 - p_i^2)\right\}.$$

Applying the above inequality to (5.4.27), we have

$$
\begin{aligned}
-\operatorname{div}_h v(t,x) &\leq \frac{dL_\ell^c}{\alpha^c} + \frac{1}{2\alpha^c h}\sum_{i=1}^d \max\left\{0, (\nabla_h u(t+1/2, x+he_i))_i - (\nabla_h u(t+1/2, x-he_i))_i\right\} \\
&= \frac{dL_\ell^c}{\alpha^c} + \frac{1}{\alpha^c}\sum_{i=1}^d \max\left\{0, \frac{u(t+1/2, x+2he_i) + u(t+1/2, x-2he_i) - 2u(t+1/2, x)}{4h^2}\right\}.
\end{aligned}
$$

By Lemma 5.4.18, for any $(t,x,y) \in \mathcal{T} \times S^2$ and $y \neq 0$, we have

$$
\frac{u(t+1/2, x+y) + u(t+1/2, x-y) - 2u(t+1/2, x)}{\|y\|^2} \leq 6L^c.
$$

Taking $y = 2he_i$, it follows that

$$
-\operatorname{div}_h v(t,x) \leq \frac{d(L_\ell^c + 6L^c)}{\alpha^c}. \tag{5.4.28}
$$

Combining (5.4.25), (5.4.26), and (5.4.28), we have

$$
\|m(t+1,\cdot)\|_\infty \leq \left(1 + \Delta t \frac{d(L_\ell^c + 6L^c)}{\alpha^c}\right) \|m(t,\cdot)\|_\infty.
$$

Since $\Delta t = 1/T$, the conclusion follows. □

We are now ready to derive an improved estimate of C_2 (in comparison with the one in (5.4.16)). We define the constant E_2 as follows:

$$
E_2 = \exp\left(\frac{d(L_\ell^c + 6L^c)}{\alpha^c}\right) \|m_0^c\|_{\mathbb{L}^\infty}.
$$

It is independent of Δt and h.

Lemma 5.4.22. *The constants C_1, C_2, and C_3 defined in* (5.3.1)-(5.3.3) *satisfy the following inequalities:*

$$
C_1 \leq E_1 h^d, \qquad C_2 \leq E_2 h^d, \qquad C_3 \leq E_3. \tag{5.4.29}
$$

As a consequence, for the constants defined in (5.3.4)*, we have*

$$
D_1 \leq E_1 L_f^c, \quad D_2 \leq (1 + L_f^c)\sqrt{\frac{2E_1E_2E_3L_f^c}{\alpha^c}}, \quad c \leq c_\theta := \max\left\{1 - \frac{\alpha^c}{4E_2E_3L_f^c}, \frac{1}{2}\right\}. \tag{5.4.30}
$$

Proof. The estimates of C_1 and C_3 are the same as in (5.4.16), and the estimate of C_2 is a direct consequence of Lemma 5.4.21 and the regularity of m_0^c. Then (5.4.30) is deduced from (5.4.29) and (5.3.4). □

Proof of Theorem 5.4.9. Inequality (5.4.10) holds true with $C_\theta = 4E_1L_f^c$, as a direct consequence of Proposition 5.3.1. Inequality (5.4.11) is established in similar fashion to inequality (5.4.7). It holds true with

$$
\bar{C}_\theta = \max\left\{2E_1L_f^c(1+L_f^c)\sqrt{\frac{2E_2E_3}{\alpha^c}}, C^*\right\}.
$$

The theorem is proved. □

5.4.5 Discussion on convergence constants

In this subsection, we study the dependence of the convergence constants C_θ and c_θ (appearing in Theorem 5.4.9) with respect to the viscosity parameter σ and the Lipschitz constant L_f^c of the coupling term f^c. First, let us recall the constant $c(\sigma,\theta,M)$, introduced in (5.4.14),

$$c(\sigma,\theta,M) = 1 + \frac{M^2}{\sigma(2\theta-1)} \exp\left(\frac{M^2}{\sigma(2\theta-1)}\right).$$

It is not difficult to see that $c(\sigma,\theta,M)$ decreases and converges to 1 as σ goes to $+\infty$. The constant E_2 is independent of σ and L_f^c by its definition (assuming that the change of L_f^c has no impact on the semi-concavity constant of f^c).

By the proofs in the previous subsection, we can give the following explicit formulas of C_θ and c_θ in Theorem 5.4.9 (without using E_1 and E_3):

$$C_\theta = 4c(\sigma,\theta,M)L_f^c\left(\|m_0^c\|^2_{\mathbb{L}^\infty} + \frac{1}{2}\|\nabla m_0^c\|^2_{\mathbb{L}^\infty}\right), \qquad c_\theta = \max\left\{1 - \frac{\sigma\alpha^c(2\theta-1)}{4c(\sigma,\theta,M)L_f^c E_2}, \frac{1}{2}\right\}. \tag{5.4.31}$$

Lemma 5.4.23. *For the constants in* (5.4.31)*, we have the following.*

1. *Fix L_f^c and θ. There exists $\sigma^* > 0$ and $C_1^* > 0$, such that for any $\sigma \geq \sigma^*$, we have*

$$C_\theta \leq C_1^*\left(\|m_0^c\|^2_{\mathbb{L}^\infty} + \frac{1}{2}\|\nabla m_0^c\|^2_{\mathbb{L}^\infty}\right), \qquad c_\theta = \frac{1}{2}.$$

2. *Fix σ and θ. There exists $L^* > 0$ and $C_2^* > 0$, such that for any $L_f^c \leq L^*$, we have*

$$C_\theta \leq C_2^*\left(\|m_0^c\|^2_{\mathbb{L}^\infty} + \frac{1}{2}\|\nabla m_0^c\|^2_{\mathbb{L}^\infty}\right), \qquad c_\theta = \frac{1}{2}.$$

Proof. Point (1) follows from the monotonicity of $c(\sigma,\theta,M)$ w.r.t. σ and the fact that M is independent of σ. We can take $C_1^* = 5L_f^c$ for example. To prove (2), we first notice that if $L_c^c \leq 1$, then $M \leq M^*$, where M^* is defined by (5.4.2), replacing L_f^c with 1. The monotonicity of $c(\sigma,\theta,M)$ w.r.t. M shows that $c(\sigma,\theta,M) \leq c(\sigma,\theta,M^*)$. Since $c(\sigma,\theta,M^*)L_f^c$ goes to 0 as L_f^c goes to 0, we prove the existence of L^*, and $C_2^* = 4c(\sigma,\theta,M^*)L^*$. □

From the proof of Proposition 5.3.1, we know that $c_\theta = 1/2$ implies that $\lambda_k = 1$ for any $k \geq 0$. In other words, Algorithm 5.2 is equivalent to the so-called best-response iteration, i.e., for any $k \geq 0$,

$$(m^{k+1}, w^{k+1}) = \mathbf{BR}_\theta(m^k).$$

Combined with Lemma 5.4.23, we have the following observations.

1. *High-viscosity* case: let L_f^c be fixed, if σ is large enough, then the best response iteration has a linear convergence rate with a factor 1/2.
2. *Weak-coupling* case: let σ be fixed, if L_f^c is small enough, then the best response iteration has a linear convergence rate with a factor 1/2.

5.5 Numerical tests

5.5.1 Problem formulation

In this section, we consider an example of (MFG) in dimension one. We identify the torus with the segment $[0,1]$. The initial distribution is concentrated around the point 0.5, the running cost is a quadratic function of the control, and the terminal cost $g(x)$ decreases to 0 as x goes to zero. Additionally, we consider a non-local congestion term which penalizes the density of the agents within the intervals $[0.2, 0.3]$ and $[0.7, 0.8]$. We will refer to $[0.2, 0.3] \cup [0.7, 0.8]$ as the *congestion-sensitive zone*.

To model this situation, let us introduce the functions $\varphi_{A,k} \in \mathcal{C}^\infty(\mathbb{R})$ and $\phi_{A,k,l_1,l_2} \in \mathcal{C}^\infty(\mathbb{R})$, parameterized by $A > 0$, $k > 0$, $0 < l_1 < l_2 < 1$ and defined by

$$\varphi_{A,k}(x) = \begin{cases} Ae^{-\frac{1}{1-(kx)^2}}, & \text{if } |x| < \frac{1}{k}, \\ 0 & \text{otherwise,} \end{cases} \qquad \phi_{A,k,l_1,l_2}(x) = \begin{cases} \varphi_{A,k}(x-l_1), & \text{if } x < l_1, \\ Ae^{-1} & \text{if } l_1 \leq x \leq l_2, \\ \varphi_{A,k}(x-l_2), & \text{otherwise.} \end{cases}$$

The data of our one-dimensional MFG is parameterized by five positive numbers a_1, a_2, k_0, k_1, and k_2 and defined by: For any $(t,x) \in Q$, any $v \in \mathbb{R}$, and any $m \in \mathcal{D}(\mathbb{T})$,

- $\ell^c(t,x,v) = \frac{1}{2}v^2$;
- $m_0^c(x) = \phi_{1,k_0,0.49,0.51}(x) / \|\phi_{1,k_0,0.49,0.51}\|_{\mathbb{L}^1}$;
- $g^c(x) = \phi_{a_1,k_1,1/k_1,1-1/k_1}(x)$;
- $f^c(t,x,m) = h^c(x) \int_0^1 h^c(y)m(y)dy$, where $h^c(x) = \phi_{a_2,k_2,0.24,0.25}(x) + \phi_{a_2,k_2,0.75,0.76}(x)$.

We take $a_1 = 2$, $a_2 = 20$, $k_0 = 10$, $k_1 = 3$, and $k_2 = 20$. The functions m_0^c, g^c, and h^c are shown in Figure 5.1. Moreover, we fix the viscosity coefficient $\sigma = 0.02$.

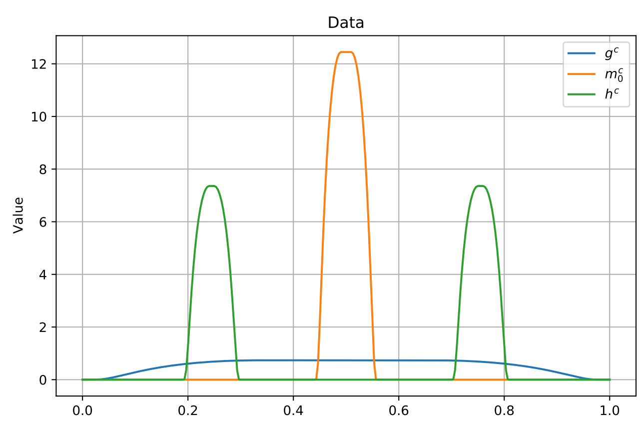


Figure 5.1: Data of the one-dimensional MFG with $a_1 = 2$, $a_2 = 20$, $k_0 = 10$, $k_1 = 3$, and $k_2 = 20$.

We can verify that this one-dimensional MFG satisfies Assumptions A, C, D, with the constants in Assumption A satisfying

$$\alpha^c = 1, \quad L_\ell^c = 0, \quad L_g^c \leq a_1 k_1, \quad L_f^c \leq \frac{a_2^2 k_2}{e}.$$

Furthermore, [BLP22, Assumption C, Appx. B] holds true for this example, which implies Assumption B for any $r < 1$ by [BLP22, Thm. B.2].

5.5.2 Results

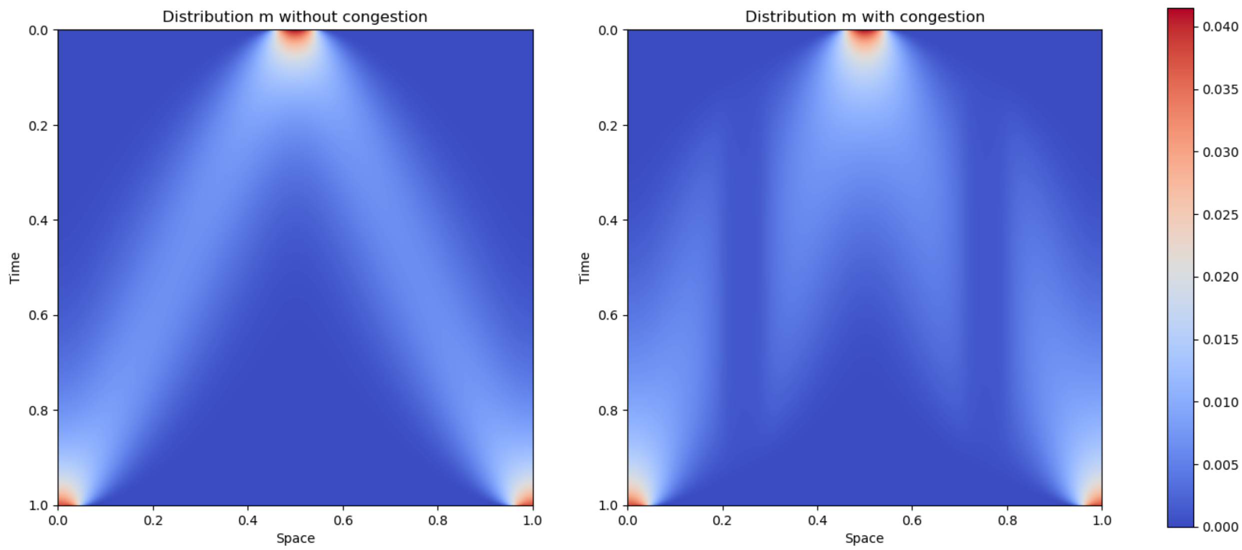
For the discretization of the system, we first choose the parameters

$$\theta = 0.8, \quad h = 1/300, \quad \text{and} \quad \Delta t = \frac{h^2}{2(1-\theta)\sigma} = 1/720,$$

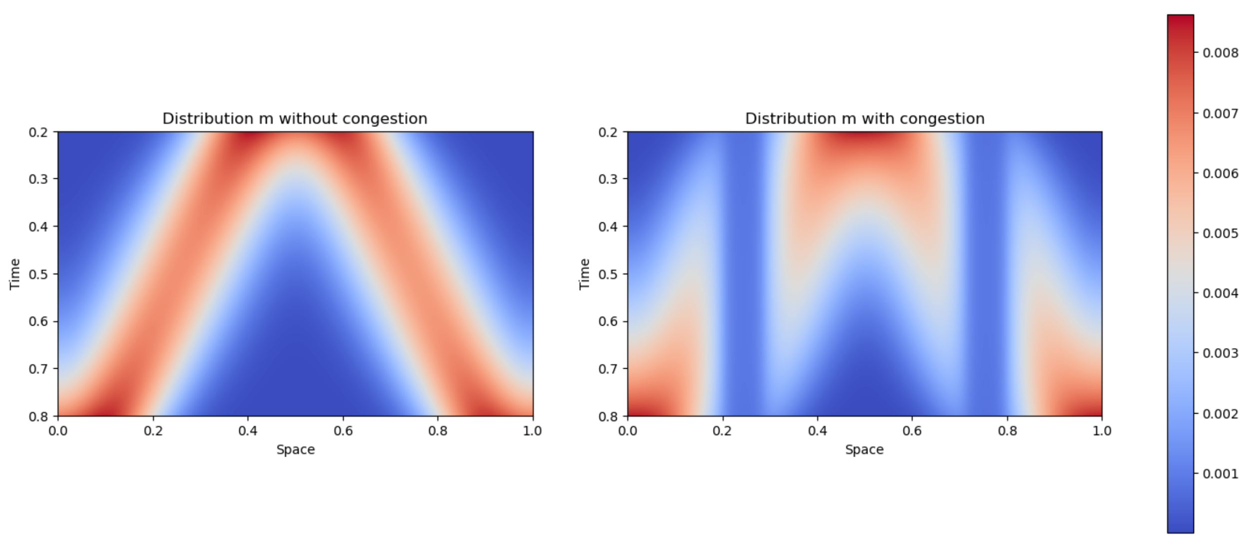
and we present the outcome of Algorithm (Theta-mfg) after 1000 iterations of the GFW algorithm, for step-sizes determined by line-search. For a better interpretation of the result, we also present the solution of the problem obtained by removing the congestion term f^c, which is a simple stochastic optimal control problem that can be solved in one iteration of the GFW method.

We present the equilibrium distribution of the agents in Figure 5.2a (without congestion term on the left, with congestion on the right). Note that the vertical axis corresponds to the time variable and is oriented downwards. We also present the restriction of the equilibrium distribution to the time interval $[0.2, 0.8]$ in Figure 5.2b, with another color scale. As the time progresses, the agents are transported towards the target points 0 and 1. The congestion term leads to a reduced density in the congestion-sensitive zone: We see two dark blue vertical areas corresponding to this zone. We also see that at time $t \approx 0.35$, a significant part of the agents is still located around 0.5 and has not crossed yet the sensitive zone, in comparison with the case without f^c. Similarly, we present the optimal control v in Figure 5.3 (without congestion term on the left, with congestion term on the right). Unsurprisingly, the agents must have a high velocity (in absolute value) in the sensitive zone. It is interesting to see that for t close to zero and for the agents not that close to 0.5, there is an incentive to "rush" to the sensitive zone. Finally, we display the value functions for the two problems in Figure 5.3. In the present setting, note that the optimal control is the discrete gradient of the value function.

We next investigate the convergence of Algorithm 5.2 (for the same discretization parameters as above). We execute Algorithm 5.2 with 1000 iterations, utilizing the open-loop choice $\lambda_k = 2/(k+2)$ and the closed-loop choice (5.3.6) (referred to as the line-search method). We present the convergence results in Figure 5.4. Evaluating γ_k, equal to $\mathcal{J}(m_k, w_k) - \mathcal{J}(\bar{m}, \bar{w})$ by definition, is difficult since the exact solution $(\bar{m}, \bar{w})$ is not known. On the other hand, the quantity $\bar{\gamma}_k$, which serves as an upper bound of γ_k by (5.3.11) can directly computed in view of its definition, based on (m_k, w_k) and $(\bar{m}_k, \bar{w}_k)$. Therefore, instead of evaluating γ_k, we display the evolution of $\bar{\gamma}_k$, see Figure 5.4. The two figures of Figure 5.4 are the same, with different scales for the horizontal axis. In the left part of Figure 5.4, we see that Algorithm 5.2 exhibits a convergence rate of order $\mathcal{O}(1/k^4)$ for the choice $\lambda_k = 2/(k+2)$, which is better than the theoretical convergence rate $\mathcal{O}(1/k)$

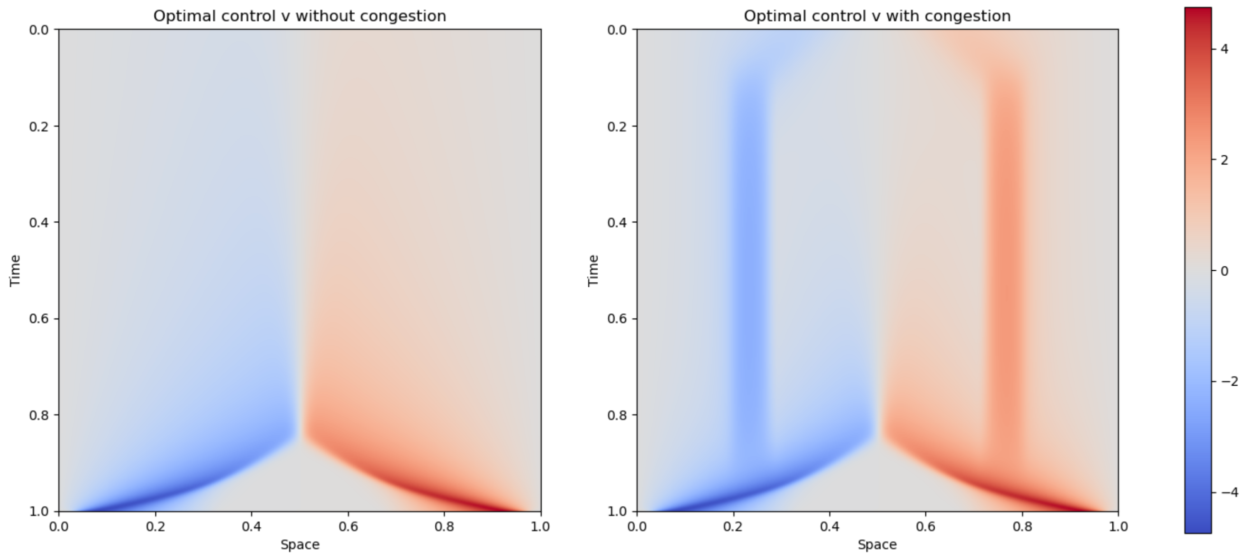


(a) Comparison of distributions in the time horizon $[0, 1]$: the case without f^c (left), the case with f^c (right).

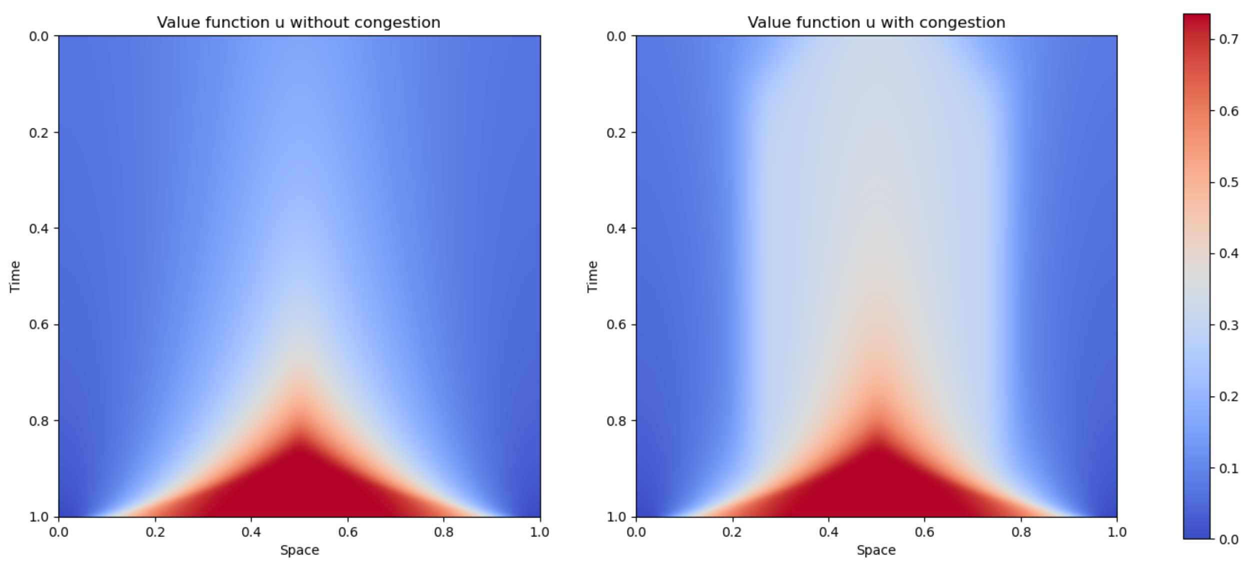


(b) Comparison of distributions in the time horizon $[0.2, 0.8]$: the case without f^c (left), the case with f^c (right).

Figure 5.2: Distributions



(a) Comparison of optimal controls: the case without f^c (left), the case with f^c (right).



(b) Comparison of value functions: the case without f^c (left), the case with f^c (right).

Figure 5.3: Optimal controls and value functions

obtained from (5.3.5). In the right part of Figure 5.4, a linear convergence rate can be observed for the line-search case, as predicted in (5.3.7).

Finally, we present numerical results concerning the mesh-independence of Algorithm 5.2 applied to (Theta-mfg). To see this, we discretize the state space with steps sizes: $h = 1/250$, $h = 1/500$, and $h = 1/1000$. The corresponding step sizes for the time space are: $\Delta t = 1/500$, $\Delta t = 1/2000$, and $\Delta t = 1/8000$. The convergence results associated with these discretization steps are displayed in Figure 5.5. From the left part of Figure 5.5, it can be observed that the convergence rate of Algorithm 5.2 remains unaffected by the choice of h when $\lambda_k = 2/(k+2)$. The right part of Figure 5.5 shows that the convergence rate of Algorithm 5.2 can even benefit from a refinement of the discretization parameters in the line-search case. These results are consistent with mesh-independence properties outlined in Theorems 5.4.5 and 5.4.9.

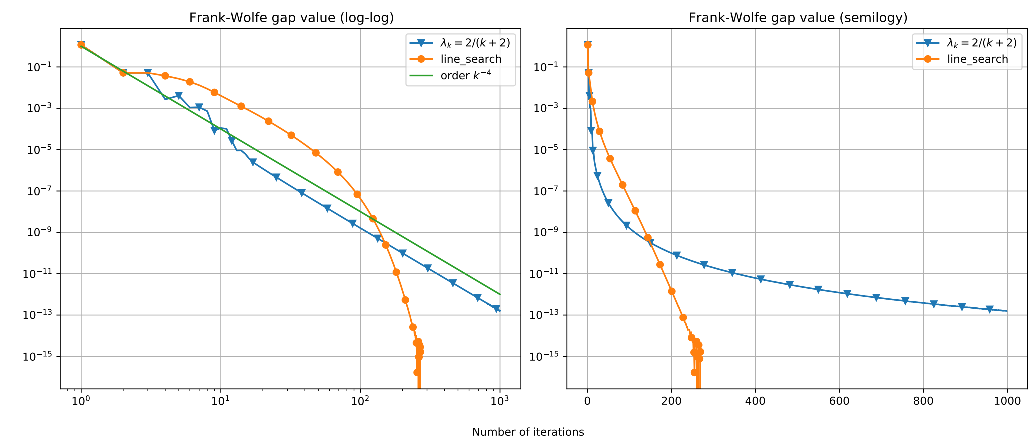


Figure 5.4: Convergence results of Algorithm 5.2.

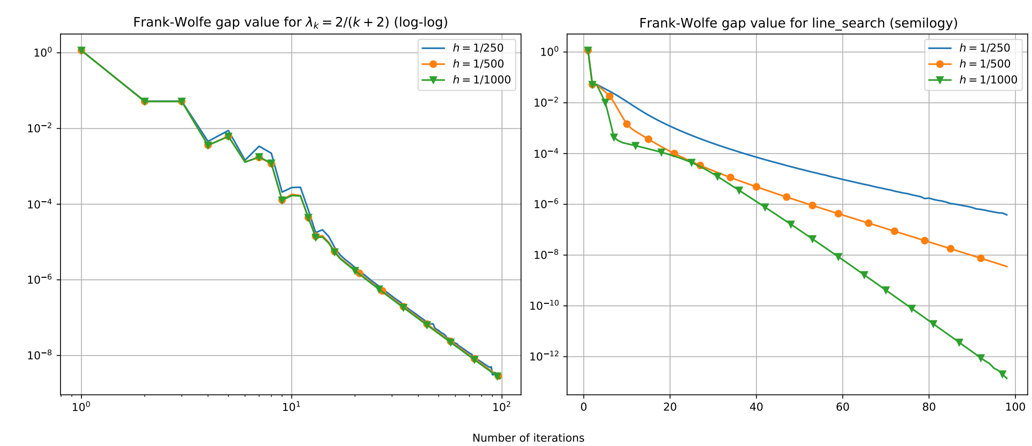


Figure 5.5: Mesh-independence property of Algorithm 5.2.

Bibliography

[ABL$^+$04] M. Atallah, M. Bykova, J. Li, K. Frikken, and M. Topkara. Private collaborative forecasting and benchmarking. In *Proceedings of the 2004 ACM workshop on Privacy in the electronic society*, pages 103–114, 2004.

[ABL$^+$14] Y. Achdou, F.J. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll. Partial differential equation models in macroeconomics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2028):20130397, 2014.

[ABPR86] E.L. Allgower, K. Böhmer, F.A. Potra, and W.C. Rheinboldt. A mesh-independence principle for operator equations and their discretizations. *SIAM Journal on Numerical Analysis*, 23(1):160–169, 1986.

[ACCD12] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM Journal on Control and Optimization*, 50(1):77–109, 2012.

[ACCD13] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: Convergence of a finite difference method. *SIAM Journal on Numerical Analysis*, 51(5):2585–2612, 2013.

[ACD10] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010.

[AE76] J.-P. Aubin and I. Ekeland. Estimates of the duality gap in nonconvex optimization. *Mathematics of Operations Research*, 1(3):225–245, 1976.

[AF03] R.A. Adams and J.F. Fournier. *Sobolev spaces*. Elsevier, 2003.

[AF09] J.-P. Aubin and Hélène Frankowska. *Set-valued analysis*. Springer Science & Business Media, 2009.

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005.

[AL20] Y. Achdou and M. Laurière. Mean field games and applications: Numerical aspects. *Mean field games*, pages 249–307, 2020.

[All07] G. Allaire. *Numerical analysis and optimization: an introduction to mathematical modelling and numerical simulation*. OUP Oxford, 2007.

[And17] R. Andreev. Preconditioning the augmented lagrangian method for instationary mean field games with diffusion. *SIAM Journal on Scientific Computing*, 39(6):A2763–A2783, 2017.

[AP16] 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 Journal on Numerical Analysis*, 54(1):161–186, 2016.

[BBC+21] K. Bestuzheva, M. Besançon, W.-K. Chen, A. Chmiela, T. Donkiewicz, J. van Doornmalen, L. Eifler, O. Gaul, G. Gamrath, A. Gleixner, et al. The scip optimization suite 8.0. *arXiv preprint arXiv:2112.08872*, 2021.

[BBG+20] O. Beaude, P. Benchimol, S. Gaubert, P. Jacquot, and N. Oudjane. A privacy-preserving method to optimize distributed resource allocation. *SIAM Journal on Optimization*, 30(3):2303–2336, 2020.

[BC11] H.H. Bauschke and P.L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*, volume 408. Springer, 2011.

[BC15] J.-D. Benamou and G. Carlier. Augmented lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, 2015.

[BC22] C. Bertucci and A. Cecchin. Mean field games master equations: from discrete to continuous state space. *arXiv preprint arXiv:2207.03191*, 2022.

[BCD97] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, volume 12. Springer, 1997.

[BCS17] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 141–171, 2017.

[Ber12] D. Bertsekas. *Dynamic programming and optimal control: Volume I*, volume 1. Athena scientific, 2012.

[BHP21] J.F. Bonnans, S. Hadikhanloo, and L. Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Applied Mathematics & Optimization*, 83(3):1431–1464, 2021.

[BKL12] P. Bonami, M. Kilinç, and J. Linderoth. Algorithms and software for convex mixed integer nonlinear programs. In *Mixed integer nonlinear programming*, pages 1–39. Springer, 2012.

[BLM09] K. Bredies, D. A Lorenz, and P. Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *Computational Optimization and applications*, 42(2):173–193, 2009.

[BLO+22] J.F. Bonnans, K. Liu, N. Oudjane, L. Pfeiffer, and C. Wan. Large-scale nonconvex optimization: randomization, gap estimation, and numerical resolution. *arXiv preprint*, 2022.

[BLP22] J.F. Bonnans, K. Liu, and L. Pfeiffer. Error estimates of a theta-scheme for second-order mean field games. *arXiv preprint arXiv:2212.08128*, 2022.

[BLP23] J.F. Bonnans, P. Lavigne, and L. Pfeiffer. Discrete potential mean field games: duality and numerical resolution. *Mathematical Programming*, pages 1–38, 2023.

[BPC11] S. Boyd, N. Parikh, and E. Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[Bré11] H. Brézis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011.

[BSR17] N. Boyd, G. Schiebinger, and B. Recht. The alternating descent conditional gradient method for sparse inverse problems. *SIAM Journal on Optimization*, 27(2):616–639, 2017.

[BT03] A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[BT09] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202, 2009.

[BT13] A. Beck and L. Tetruashvili. On the convergence of block coordinate descent type methods. *SIAM Journal on Optimization*, 23(4):2037–2060, 2013.

[Car10] P. Cardaliaguet. Notes on mean field games. Technical report, Technical report, 2010.

[Cas75] J. Cassels. Measures of the non-convexity of sets and the Shapley–Folkman–Starr theorem. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 78, pages 433–436. Cambridge University Press, 1975.

[CB18] L. Chizat and F. Bach. On the global convergence of gradient descent for over-parameterized models using optimal transport. *Advances in neural information processing systems*, 31, 2018.

[CC18] P. Cannarsa and R. Capuani. Existence and uniqueness for mean field games with state constraints. *PDE models for multi-agent phenomena*, pages 49–71, 2018.

[CCRW23] J. Claisse, G. Conforti, Z. Ren, and S. Wang. Mean field optimization problem regularized by fisher information. *arXiv preprint arXiv:2302.05938*, 2023.

[CGPT15] P. Cardaliaguet, P.J. Graber, A. Porretta, and D. Tonon. Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differential Equations and Applications NoDEA*, 22:1287–1317, 2015.

[CH17] P. Cardaliaguet and S. Hadikhanloo. Learning in mean field games: the fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017.

[CL18a] P. Cardaliaguet and C.-A. Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.

[CL18b] M.-W. Cheung and R. Lahkar. Nonatomic potential games: the continuous strategy case. *Games and Economic Behavior*, 108:341–362, 2018.

[Cla87] D.S. Clark. Short proof of a discrete Gronwall inequality. *Discrete applied mathematics*, 16(3):279–281, 1987.

[CLV20] C. Coey, M. Lubin, and J.P. Vielma. Outer approximation with conic certificates for mixed-integer convex problems. *Mathematical Programming Computation*, 12(2):249–293, 2020.

[CMS16] P. Cardaliaguet, A. R Mészáros, and F. Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016.

[Coh13] D.L. Cohn. *Measure theory*, volume 1. Springer, 2013.

[Com18] P.L. Combettes. Perspective functions: Properties, constructions, and examples. *Set-Valued and Variational Analysis*, 26(2):247–264, 2018.

[Con16] L. Condat. Fast projection onto the simplex and the $l1$ ball. *Mathematical Programming*, 158(1):575–585, 2016.

[CP11] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40:120–145, 2011.

[CP16] A. Chambolle and T. Pock. On the ergodic convergence rates of a first-order primal–dual algorithm. *Mathematical Programming*, 159(1-2):253–287, 2016.

[CS04] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, volume 58. Springer Science & Business Media, 2004.

[CS14] E. Carlini and F.J. Silva. A fully discrete semi-lagrangian scheme for a first order mean field game problem. *SIAM Journal on Numerical Analysis*, 52(1):45–67, 2014.

[CS15] E. Carlini and F.J. Silva. A semi-lagrangian scheme for a degenerate second order mean field game system. *Discrete & Continuous Dynamical Systems*, 35(9):4269, 2015.

[CV06] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*, volume 580. Springer, 2006.

[CW05] P.L. Combettes and V.R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale modeling & simulation*, 4(4):1168–1200, 2005.

[Dan63] G. Dantzig. *Linear programming and extensions*. Princeton university press, 1963.

[DDDM04] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 57(11):1413–1457, 2004.

[Del15] B. Delyon. Exponential inequalities for dependent processes. *HAL preprint hal-01072019*, 2015.

[DH78] J.C. Dunn and S. Harshbarger. Conditional gradient algorithms with open loop step size rules. *Journal of Mathematical Analysis and Applications*, 62(2):432–444, 1978.

[dP20] A. d'Aspremont and M. Pilanci. Global convergence of frank wolfe on one hidden layer networks. *arXiv preprint arXiv:2002.02208*, 2020.

[DTT17] B. Djehiche, A. Tcheukam, and H. Tembine. Mean-field-type games in engineering. *AIMS Electronics and Electrical Engineering*, 1(1):18–73, 2017.

[DU18] L. Ding and M. Udell. Frank-Wolfe style algorithms for large scale optimization. In *Large-Scale and Distributed Optimization*, pages 215–245. Springer, 2018.

[EG94] H.C. Elman and G.H. Golub. Inexact and preconditioned uzawa algorithms for saddle point problems. *SIAM Journal on Numerical Analysis*, 31(6):1645–1661, 1994.

[Fal04] K. Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.

[FF13] M. Falcone and R. Ferretti. *Semi-Lagrangian approximation schemes for linear and Hamilton—Jacobi equations*. SIAM, 2013.

[FG15] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3-4):707–738, 2015.

[FMSF21] J. Fadili, C. Molinari, and A. Silveti-Falls. Inexact and stochastic generalized conditional gradient with augmented lagrangian and proximal step. *Journal of Nonsmooth Analysis and Optimization*, 2, 2021.

[FR75] W.H. Fleming and R.W. Rishel. *Deterministic and stochastic optimal control*, volume 1 of *Appl. Math. (N. Y.)*. Springer, New York, 1975.

[FR16] O. Fercoq and P. Richtárik. Optimization in high dimensions via accelerated, parallel, and proximal coordinate descent. *Siam review*, 58(4):739–771, 2016.

[FW56] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[GG12] A. Gersho and R.M. Gray. *Vector quantization and signal compression*, volume 159. Springer Science & Business Media, 2012.

[GHS22] P. Graewe, U. Horst, and R. Sircar. A maximum principle approach to a deterministic mean field game of control with absorption. *SIAM Journal on Control and Optimization*, 60(5):3173–3190, 2022.

[GMS10] D.A. Gomes, J. Mohr, and R.R. Souza. Discrete time, finite state space mean field games. *Journal de mathématiques pures et appliquées*, 93(3):308–328, 2010.

[GO18] LLC Gurobi Optimization. Gurobi optimizer reference manual, 2018.

[GPA21] D. Ghaderyan, F.L. Pereira, and A.P. Aguiar. A fully distributed method for distributed multiagent system in a microgrid. *Energy Reports*, 7:2294–2301, 2021.

[GPL+22] M. Geist, J. Pérolat, M. Laurière, R. Elie, S. Perrin, O. Bachem, R. Munos, and O. Pietquin. Concave utility reinforcement learning: The mean-field game viewpoint. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, pages 489–497, 2022.

[HKMS20] H. Hassani, A. Karbasi, A. Mokhtari, and Z. Shen. Stochastic conditional gradient++:(non) convex minimization and continuous submodular maximization. *SIAM Journal on Optimization*, 30(4):3315–3344, 2020.

[HL16] E. Hazan and H. Luo. Variance-reduced and projection-free stochastic optimization. In *International Conference on Machine Learning*, pages 1263–1271. PMLR, 2016.

[HMC06] 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 & Systems*, 6(3):221–252, 2006.

[HS19] S. Hadikhanloo and F.J. Silva. Finite mean field games: fictitious play and convergence to a first order continuous mean field game. *Journal de Mathématiques Pures et Appliquées*, 132:369–397, 2019.

[HUL93] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms II: advanced theory and bundle methods*. Springer Verlag, 1993.

[Jag13] M. Jaggi. Revisiting frank-wolfe: Projection-free sparse convex optimization. In *International conference on machine learning*, pages 427–435. PMLR, 2013.

[JBGO18] P. Jacquot, O. Beaude, S. Gaubert, and N. Oudjane. Analysis and implementation of an hourly billing mechanism for demand response management. *IEEE Transactions on Smart Grid*, 10(4):4265–4278, 2018.

[Kal97] O. Kallenberg. *Foundations of modern probability*, volume 2. Springer, 1997.

[KBB15] W. Krichene, A. Bayen, and P.L. Bartlett. Accelerated mirror descent in continuous and discrete time. *Advances in neural information processing systems*, 28, 2015.

[KCD22] T. Kerdreux, I. Colin, and A. D'Aspremont. Stable bounds on the duality gap of separable nonconvex optimization problems. *Mathematics of Operations Research*, 2022.

[Kob22] Z. Kobeissi. On classical solutions to the mean field game system of controls. *Communications in Partial Differential Equations*, 47(3):453–488, 2022.

[Kry96] N.V. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, volume 12. American Mathematical Soc., 1996.

[Kry08] N.V. Krylov. *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces*, volume 96. American Mathematical Soc., 2008.

[KW22] K. Kunisch and D. Walter. On fast convergence rates for generalized conditional gradient methods with backtracking stepsize. *Numerical Algebra, Control and Optimization*, 2022.

[Lio71] J.-L. Lions. *Optimal control of systems governed by partial differential equations*. Springer Verlag, 1971.

[LJ16] S. Lacoste-Julien. Convergence rate of Frank-Wolfe for non-convex objectives. *arXiv preprint arXiv:1607.00345*, 2016.

[LJJ15] S. Lacoste-Julien and M. Jaggi. On the global linear convergence of frank-wolfe optimization variants. *Advances in neural information processing systems*, 28, 2015.

[LL07] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[LOP22] K. Liu, N. Oudjane, and L. Pfeiffer. Decentralized resolution of finite-state, non-convex, and aggregative optimal control problems. *arXiv preprint arXiv:2204.07080*, 2022.

[LOW22] K. Liu, N. Oudjane, and C. Wan. Approximate nash equilibria in large nonconvex aggregative games. *Mathematics of Operations Research*, 2022.

[LP22] P. Lavigne and L. Pfeiffer. Generalized conditional gradient and learning in potential mean field games. *arXiv preprint arXiv:2109.05785*, 2022.

[LST10] A. Lachapelle, J. Salomon, and G. Turinici. Computation of mean field equilibria in economics. *Mathematical Models and Methods in Applied Sciences*, 20(04):567–588, 2010.

[LSU88] O.A. Lady**V**zenskaja, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[LYFC19] F. Locatello, A. Yurtsever, O. Fercoq, and V. Cevher. Stochastic Frank-Wolfe for composite convex minimization. *Advances in Neural Information Processing Systems*, 32, 2019.

[Mal09] S. Mallat. *A wavelet tour of signal processing*. Elsevier, 2009.

[MBP14] J. Mairal, F. Bach, and J. Ponce. Sparse modeling for image and vision processing. *arXiv preprint arXiv:1411.3230*, 2014.

[McD89] C. McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

[MHK20] A. Mokhtari, H. Hassani, and A. Karbasi. Stochastic conditional gradient methods: From convex minimization to submodular maximization. *Journal of machine learning research*, 2020.

[MM16] Q. Mérigot and J.-M. Mirebeau. Minimal geodesics along volume-preserving maps, through semidiscrete optimal transport. *SIAM Journal on Numerical Analysis*, 54(6):3465–3492, 2016.

[MMM19] S. Mei, T. Misiakiewicz, and A. Montanari. Mean-field theory of two-layers neural networks: dimension-free bounds and kernel limit. In *Conference on Learning Theory*, pages 2388–2464. PMLR, 2019.

[MMN18] S. Mei, A. Montanari, and P.-M. Nguyen. A mean field view of the landscape of two-layer neural networks. *Proceedings of the National Academy of Sciences*, 115(33):E7665–E7671, 2018.

[MS19] G. Mazanti and F. Santambrogio. Minimal-time mean field games. *Mathematical Models and Methods in Applied Sciences*, 29(08):1413–1464, 2019.

[Nes18] Y. Nesterov. *Lectures on convex optimization*, volume 137. Springer, 2018.

[Pac18] F. Pacaud. *Decentralized optimization for energy efficiency under stochasticity.* PhD Thesis, Université Paris-Est, October 2018.

[PC19] G. Peyré and M. Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[PPL+20] S. Perrin, J. Pérolat, M. Laurière, M. Geist, R. Elie, and O. Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. *Advances in Neural Information Processing Systems*, 33:13199–13213, 2020.

[Roc97] R.T. Rockafellar. *Convex analysis*, volume 11. Princeton university press, 1997.

[Rud87] W. Rudin. *Real and complex analysis*, volume 3. McGraw-Hill, 1987.

[SAB+23] A. Seguret, C. Alasseur, J.F. Bonnans, A. De Paola, N. Oudjane, and V. Trovato. Decomposition of convex high dimensional aggregative stochastic control problems. *Applied Mathematics & Optimization*, 88(1):8, 2023.

[Sar22] C. Sarrazin. Lagrangian discretization of variational mean field games. *SIAM Journal on Control and Optimization*, 60(3):1365–1392, 2022.

[SFMF20] A. Silveti-Falls, C. Molinari, and J. Fadili. Generalized conditional gradient with augmented Lagrangian for composite minimization. *SIAM Journal on Optimization*, 30(4):2687–2725, 2020.

[SMZ+16] R.-H. Shi, Y. Mu, H. Zhong, J. Cui, and S. Zhang. Secure multiparty quantum computation for summation and multiplication. *Scientific reports*, 6(1):1–9, 2016.

[SS21] F. Santambrogio and W. Shim. A Cucker–Smale inspired deterministic mean field game with velocity interactions. *SIAM Journal on Control and Optimization*, 59(6):4155–4187, 2021.

[SST09] S. Shalev-Shwartz and A. Tewari. Stochastic methods for l1 regularized loss minimization. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 929–936, 2009.

[Sta69] R.M. Starr. Quasi-equilibria in markets with non-convex preferences. *Econometrica: journal of the Econometric Society*, pages 25–38, 1969.

[TBL22] T. Tang, K. Balasubramanian, and T.C.M. Lee. High-probability bounds for robust stochastic frank-wolfe algorithm. In *The 38th Conference on Uncertainty in Artificial Intelligence*, 2022.

[Tho95] J.W. Thomas. *Numerical partial differential equations: finite difference methods*, volume 22. Springer Science & Business Media, 1995.

[Tib96] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[Ton52] H. Tong. Some characterizations of normal and perfectly normal spaces. *Duke Math. J.*, pages 289–292, 1952.

[Tsy09] A.B. Tsybakov. *Introduction to Nonparametric Estimation.* Springer, 2009.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58. American Mathematical Soc., 2003.

[Vil09] C. Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

[Wan17] M. Wang. Vanishing price of decentralization in large coordinative nonconvex optimization. *SIAM Journal on Optimization*, 27(3):1977–2009, 2017.

[YFC19] A. Yurtsever, O. Fercoq, and V. Cevher. A conditional-gradient-based augmented lagrangian framework. In *International Conference on Machine Learning*, pages 7272–7281, 2019.

[YSC19] A. Yurtsever, S. Sra, and V. Cevher. Conditional gradient methods via stochastic path-integrated differential estimator. In *International Conference on Machine Learning*, pages 7282–7291. PMLR, 2019.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Analyse numérique et méthodes pour les problèmes d'optimisation de type champ moyen

Mots clés : Commande optimale ; Champ moyen ; Equation de Fokker-Planck ; Algorithme de Frank-Wolfe ; Schéma de différences finies

Résumé : Cette thèse traite de l'analyse numérique et des méthodes pour les problèmes d'optimisation et les jeux potentiels impliquant un grand nombre d'agents. Nous considérons des modèles asymptotiques obtenus par une approximation de champ moyen ; ils présentent des propriétés de convexité d'un grand intérêt. Nous nous concentrons sur les problèmes d'optimisation agrégative de grande dimension, pour lesquels la fonction coût dépend d'un terme d'agrégat, qui est la somme des contributions des agents à un bien commun. Nous nous concentrons également sur des modèles potentiels de jeux à champ moyen (MFG), qui sont des modèles asymptotiques pour les jeux différentiels. La thèse comporte quatre contributions.
1) Nous proposons une relaxation de type champ moyen pour les problèmes d'optimisation agrégative, obtenue par randomisation. Une estimation d'ordre $\mathcal{O}(1/N)$ du saut de relaxation est démontrée, où N représente le nombre d'individus. Nous développons et prouvons la convergence d'une variante stochastique de l'algorithme de Frank-Wolfe, appelée algorithme SFW, pour résoudre le problème agrégatif original.
2) Nous formulons une classe générale de problèmes d'optimisation impliquant un ensemble de distributions de probabilités avec une marginale prescrite, égale à m. Nous les appelons problèmes d'optimisation à champ moyen (MFO). Notre cadre contient les problèmes agrégatifs relaxés ainsi que certains MFGs potentiels en formulation lagrangienne. Nous démontrons un résultat de stabilité par rapport à une perturbation de m. Nous en déduisons une estimation d'erreur pour une méthode numérique reposant sur une discrétisation de m et l'algorithme SFW.
3) Nous introduisons un nouveau schéma de différences finies, appelé *thêta-schéma*, pour résoudre les MFG monotones du second ordre. Nous donnons un résultat de convergence précis pour le thêta-schéma, d'ordre $\mathcal{O}(h^r)$, où h est le pas de discrétisation en espace et $r \in (0,1)$ est lié à la continuité de Hölder de la solution du problème continu et de certaines de ses dérivées.
4) Nous considérons la résolution de MFGs potentiels du second ordre avec l'algorithme de Frank-Wolfe généralisé, combiné avec le thêta-schéma. Nous prouvons des taux de convergence sous-linéaire et linéaire pour cet algorithme. Plus important encore, ces taux possèdent la propriété d'indépendance au maillage, c'est-à-dire que les constantes de convergence sont indépendantes des paramètres de discrétisation.

Title : Numerical analysis and methods for mean-field-type optimization problems

Keywords : Optimal control ; Mean fields ; Fokker-Planck equation ; Frank-Wolfe algorithm ; Finite difference scheme

Abstract : This thesis deals with the numerical analysis and methods for optimization problems and potential games involving a large number of agents. We consider asymptotic models obtained through a mean-field approximation ; they exhibit convexity properties of great interest. We focus on large-scale aggregative optimization problems, for which the objective function depends on an aggregate term, which is the sum of the contributions of the agents to some common good. We also focus on potential Mean Field Game (MFG) models, which are limit models for differential games. The thesis consists of four contributions.
1) We propose a mean-field relaxation for aggregative optimization problems, obtained by randomization. The relaxation gap is estimated to be of order $\mathcal{O}(1/N)$, where N represents the number of individuals. We develop and prove the convergence of a stochastic variant of the Frank-Wolfe algorithm, called SFW algorithm, to address the original aggregative problem.
2) We formulate a general class of optimization problems involving a set of probability distributions with a prescribed marginal m. We call them *Mean Field Optimization* (MFO) problems. Our framework contains the relaxed aggregative problems as well as some Lagrangian potential MFGs. We demonstrate a stability result with respect to perturbations of m. It enables us to derive an error estimate for a numerical method relying on a discretization of m and the SFW algorithm.
3) We introduce a novel finite-difference scheme, called *theta-scheme*, for solving monotone second-order MFGs. We give a precise convergence result for the theta-scheme, of order $\mathcal{O}(h^r)$, where h is the step length of the space variable and $r \in (0,1)$ is related to the Hölder continuity of the solution of the continuous problem and some of its derivatives.
4) We consider the resolution of potential second-order MFGs with the generalized Frank-Wolfe algorithm, combined with the theta-scheme. We prove a sublinear and a linear rate of convergence for this algorithm. More importantly, these rates possess the mesh-independence property, i.e., the convergence constants are independent of the discretization parameters.

Institut Polytechnique de Paris
91120 Palaiseau, France