Generalized conditional gradient and learning in potential mean field games

Pierre Lavigne, Laurent Pfeiffer

Generalized conditional gradient and learning in potential mean field games*

Pierre Lavigne† Laurent Pfeiffer‡

September 26, 2022

Abstract

We investigate the resolution of second-order, potential, and monotone mean field games with the generalized conditional gradient algorithm, an extension of the Frank-Wolfe algorithm. We show that the method is equivalent to the fictitious play method. We establish rates of convergence for the optimality gap, the exploitability, and the distances of the variables to the unique solution of the mean field game, for various choices of stepsizes. In particular, we show that linear convergence can be achieved when the stepsizes are computed by linesearch.

Key-words: mean field games, generalized conditional gradient, fictitious play, learning, exploitability.

AMS classification: 90C52, 91A16, 91A26, 91B06, 49K20, 35F21, 35Q91.

1 Introduction

Framework Mean field games (MFG), introduced by J.-M. Lasry and P.-L. Lions in [52, 53, 54] and M. Huang, R. Malhamé, and P. Caines in [44], are a class of mathematical problems which allow to approximate differential games involving a very large number of agents. The general situation of interest is as follows: each agent aims at minimizing some cost function, depending on his own decision variables and some coupling terms, common to all agents. There are two fundamental assumptions in MFG theory: the coupling terms depend on the distribution of the agents and each agent has a negligible contribution to the coupling terms. Mean field games can typically be formulated as a coupled system of two equations, characterizing the decisions of the agents as functions of the coupling terms and vice versa. Mean field games have found various applications such as epidemic control [27, 29], electricity management [3, 24], finance and banking [19, 22, 23, 31, 48], social network [4], economics [1, 42], crowd motion [49].

*This work was supported by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH.

†Institut Louis Bachelier. E-mail: pierre.lavigne@institutlouisbachelier.org.

‡Inria and Laboratoire des Signaux et Systèmes, CNRS (UMR 8506), CentraleSupélec, Université Paris-Saclay, 91190 Gif-sur-Yvette, France. E-mail: laurent.pfeiffer@inria.fr.

In this article, we consider the standard situation where each agent optimizes a stochastic differential system and where the agents only differ from each other by their initial condition; therefore, for given coupling terms, the agents share the same optimal feedback function v, that can be characterized through the Hamilton-Jacobi-Bellman equation. We consider two coupling terms: a first standard one, denoted γ, deduced from the density of the states of the agents through an interaction function f. The density of the states, denoted by m, is itself deduced from v through the Fokker-Planck equation. The second coupling term, denoted P, can be interpreted as a price variable and depends on the distribution of the agents with respect to their states and controls through a price function ϕ. Our model thus falls into the class of mean field games of controls. Interactions through the density of players typically appear in epidemic or crowd motion models, while interactions through the controls v typically appear in economics, finance or energy management models.

We assume in this work that the interaction cost and the price function derive from convex potentials. Our MFG system has then a potential structure, that is, it can be interpreted as the first-order necessary and sufficient optimality conditions for some convex optimal control problem of the Fokker-Planck equation, referred to as the potential problem. This class of games has been widely investigated, we refer the reader to [7, 17, 21, 56, 61] for interactions through the density m and to [9, 37, 38, 40, 39, 41] for price interactions.

Various methods from convex optimization have been been employed to solve the potential problem, see [2] for a survey. A first approach consists in formulating the potential problem as a saddle-point problem and to solve it with primal-dual algorithms, see [10, 12, 13]. Another approach consists in applying the augmented Lagrangian algorithm to the dual problem of the potential problem, see [5, 7, 10]. Other methods have been investigated such as the Sinkhorn algorithm [6].

Generalized condition gradient algorithm and learning The *generalized condition gradient* (GCG) algorithm is an extension of the conditional gradient algorithm, also called Frank-Wolfe algorithm, first introduced in [33]. The conditional gradient method allows to minimize a convex objective function on a convex and compact set. The idea is to linearize the cost function at each iteration $k \in \mathbb{N}$, at a given point $\bar{x}_k$, and to find a minimizer x_k of the corresponding linearized problem. Then a new point $\bar{x}_{k+1} = (1-\delta_k)\bar{x}_k + \delta_k x_k$ is computed for some step size $\delta_k \in [0,1]$. A classical choice of step size is given by $\delta_k = 2/(k+2)$ (see [28, 45]) which yields the convergence of the objective function in $O(1/k)$. For a recent description of the conditional gradient algorithm and its extensions, we refer to [46, Chapter 1]. In our study we consider the generalized conditional gradient algorithm (first introduced in [11]), which is based on a partial linearization of the objective function instead of a full linearization. Most of the basic existing convergence results for the conditional gradient remain true for the GCG method, which also exhibits faster convergence rates in some cases, typically when the partially linearized cost function enjoys some coercivity property and when δ_k is obtained with a linesearch procedure. Improved rates of convergence have been recently obtained in [47], in an infinite-dimensional setting.

A fundamental issue in game theory is the formation of an equilibrium. It is often unrealistic to consider that the agents can perfectly anticipate the behavior of the others, in particular in the sophisticated situation underlying an MFG model.

On contrast, it is more realistic to assume that the game is repeated many times and that the agents update their decisions according to a more or less complex procedure called *learning* procedure. See [34] for a general reference. We consider in this article the *fictitious play*, a learning procedure introduced in [14, 62] in which the agents play at each iteration of the game an optimal decision (also called *best-responses*), corresponding to a predicted value of the coupling term (also called *belief*), which is then updated. An application of the fictitious play to potential games can be found in [57]. In the context of MFGs, the fictitious play has been investigated in [18, 30, 43, 60]. The convergence results for learning methods can be of various forms. In potential games, one can study the convergence of the potential cost along a sequence generated by the fictitious play algorithm. In general, one can consider the *exploitability* of the game at each iteration and try to show its convergence to zero. For a given value of the coupling terms, the exploitability is the highest reduction of cost that a representative agent can achieve by changing his current decision to the best-response, assuming that the coupling terms remain the same. This notion has recently received a growing attention [25, 26, 35, 58, 59, 60]. The convergence of the exploitability has been addressed in [60] in the context of continuous-time learning and discrete mean field games, and a convergence rate is provided.

A key message of this article is that, in the context of second-order potential mean field games, the GCG algorithm can be interpreted as a fictitious play method. It relies on the following fact: at each step of the method, the problem to be solved (arising from a semi-linearization of the potential problem) coincides with the individual control problem of the agents, for a given belief of the coupling terms. We will see that the *primal-dual gap* of the GCG method is equal to the exploitability. The update formula $\bar{x}_{k+1} = (1-\delta_k)\bar{x}_k + \delta_k x_k$ corresponds to the learning step in the fictitious play algorithm, where the agents update their belief by averaging the past and the new distributions of states and controls. This interpretation has already been highlighted in a very recent work [35], for a class of potential mean field games with some discrete structure. The connection between the Frank-Wolfe algorithm and fictitious play has also been been investigated in [64] for a general class of potential games.

Contributions The article [18] is the most related to ours. It considers a second-order potential MFG, similar to our model but without price interaction. It is proved that any cluster point (there exists at least one) of the sequences of value functions and probability distributions generated by the fictitious play is a solution to the MFG. In the case of a convex potential, the entire sequence converge.

The connection between the GCG algorithm and fictitious play, in the context of second-order MFGs, is the first contribution of our work. As we already mentioned, this connection was already established in [35], in a different MFG setting. Taking advantage of this connection, we prove a general convergence result for the optimality gap (associated with the potential cost), when the stepsizes are predefined. These results easily lead to convergence rates, in particular, for $\delta_k = p/(k+p)$ (with $p > 0$), we prove a convergence rate of order $\mathcal{O}(k^{-p})$. This covers the case of the fictitious play (with $p = 1$) and thus improves the convergence result of [18] in the convex potential case. Our most important contribution is the proof of the linear convergence of the optimality gap when the stepsizes are determined with classical linesearch rules, which is of major interest from a numerical perspective. Let us

emphasize that the proof of convergence follows the techniques from [47]. With the exception of [65], which is restricted to linear-quadratic MFGs, we are not aware of any other numerical method for convex potential MFGs that can achieve this speed of convergence.

Plan of the paper We formulate in Section 2 the mean field game of study. We describe the generalized conditional gradient algorithm in Section 3 and we explain its connection with fictitious play. We state our main convergence result in Theorem 10. Section 4 contains technical results related to the Fokker-Planck and the Hamilton-Jacobi-Bellman equations. We provide a stability result for optimal control problems in Section 5. Finally, we proof our main result in Section 6 and we give an other convergence result for the variables of the system, Theorem 35. Numerical results for an academical problem are reported in Section 7.

2 Formulation of the mean-field game

2.1 Notation

Let $T > 0$ denote the horizon of the game. We fix d and k in $\mathbb{N}^*$. We denote by $\mathbb{T}^d$ the d-dimensional torus and we set $Q = \mathbb{T}^d \times [0, T]$. For any subsets O and K of $\mathbb{R}^d$, we denote $\mathcal{C}(O; K)$ the set of continuous mappings on O valued in K. In the article, when $K = \mathbb{R}$, we simply denote $\mathcal{C}(O)$ and we make use of this convention for any other functional spaces.

Hölder spaces For any $\alpha \in (0, 1)$, we denote by $\mathcal{C}^\alpha(Q)$ the set of Hölder continuous mappings on Q of exponent α. We denote by $\mathcal{C}^{\alpha,\alpha/2}(Q)$ the set of functions on Q which are Hölder continuous of exponent α with respect to space and Hölder continuous of exponent $\alpha/2$ with respect to time. We denote by $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ the set of functions $u \in \mathcal{C}^\alpha(Q)$ with partial derivatives $\partial_t u$, $\partial_{x_i} u$, and $\partial^2_{x_i,x_j} u$ in $\mathcal{C}^{\alpha,\alpha/2}(Q)$. Finally, we denote by $\mathcal{C}^{2+\alpha}(\mathbb{T}^d)$ the set of α-Hölder continuous functions, such that all partial derivatives up to the order two are α-Hölder continuous.

Sobolev spaces and density space We denote by $W^{n,q}(\mathbb{T}^d)$ the Sobolev space of functions with weak partial derivatives in $L^q(\mathbb{T}^d)$, up to the order n. We set

$$W^{2,1,q}(Q) = W^{1,q}(Q) \cap L^q(0, T; W^{2,q}(\mathbb{T}^d))$$
$$W^{1,0,\infty}(Q) = L^\infty(0, T; W^{1,\infty}(\mathbb{T}^d)).$$

From now on, we fix a real number q such that $q > d + 2$.

Lemma 1. *There exists $\delta \in (0, 1)$ and $C > 0$ such that for all $u \in W^{2,1,q}(Q)$,*

$$\|u\|_{\mathcal{C}^\delta(Q)} + \|\nabla u\|_{\mathcal{C}^\delta(Q;\mathbb{R}^d)} \leq C\|u\|_{W^{2,1,q}(Q)}.$$

Proof. See [51, Lemma II.3.3., page 80 and Corollary, page 342]. □

We introduce the space Θ, which will be used for the control variable of the system. It is defined by

$$\Theta = \left\{ v \in \mathcal{C}(Q; \mathbb{R}^d) \,\middle|\, D_x v \in L^q(Q; \mathbb{R}^{d\times d}) \right\},$$

where $D_x v$ denotes the weak space derivative of v with respect to x. We equip this space with the norm

$$\|v\|_\Theta = \|v\|_{L^\infty(Q;\mathbb{R}^d)} + \|D_x v\|_{L^q(Q;\mathbb{R}^{d\times d})}.$$

The coupling terms (γ, P) of the MFG system of interest will be considered in the space Ξ, defined by

$$\Xi = \big(W^{1,0,\infty}(Q) \cap \mathcal{C}(Q)\big) \times \mathcal{C}(0,T;\mathbb{R}^k).$$

In words, (γ, P) lies in Ξ if and only if γ is continuous in both variables, Lipschitz continuous in x, uniformly in t, and P is continuous. Let $R > 0$, we define the following subset of Ξ:

$$\Xi_R = \Big\{(\gamma, P) \in \Xi \,\Big|\, \|\gamma\|_{W^{1,0,\infty}(Q)} + \|P\|_{L^\infty(0,T;\mathbb{R}^k)} \le R\Big\}.$$

We also define

$$\mathcal{D}_1(\mathbb{T}^d) = \Big\{m \in L^\infty(\mathbb{T}^d) \,\Big|\, m \ge 0,\ \int_{\mathbb{T}^d} m(x)\,\mathrm{d}x = 1\Big\}.$$

Nemytskii operators Given two mappings $g\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and $u\colon \mathcal{X} \to \mathcal{Y}$, we denote by $\boldsymbol{g}[u]\colon \mathcal{X} \to \mathcal{Z}$ the mapping defined by

$$\boldsymbol{g}[u](x) = g(x, u(x)), \quad \forall x \in \mathcal{X},$$

called Nemytskii operator. This notation will for instance be used for the Hamiltonian H: instead of writing $H(x,t,\nabla u(x,t))$, we write $\boldsymbol{H}[\nabla u](x,t)$. Note that $\boldsymbol{H}_p$ will denote the Nemytskii operator associated with the partial derivative of H with respect to p (a similar notation will be used for the other partial derivatives).

Generic constants All along the article, we make use of a generic constant $C > 0$, depending only on the data of the problem. The value of C may increase from an inequality to the next one. In the same way, we will make use of a generic constant $C(R)$ depending only on the data of the problem and some positive real number $R > 0$.

2.2 Mean-field game system

We fix an initial distribution and a terminal cost

$$m_0 \in \mathcal{D}_1(\mathbb{T}^d), \qquad g\colon \mathbb{T}^d \to \mathbb{R},$$

and four maps: a running cost L, an interaction cost f, a price function ϕ and an aggregation term a,

$$\begin{aligned} L&\colon\ Q \times \mathbb{R}^d \to \mathbb{R}, & \phi&\colon\ [0,T] \times \mathbb{R}^k \to \mathbb{R}^d,\\ f&\colon\ Q \times \mathcal{D}_1(\mathbb{T}^d) \to \mathbb{R}, & a&\colon\ Q \to \mathbb{R}^{k\times d}. \end{aligned}$$

For any $(x,t,p) \in Q \times \mathbb{R}^d$, we define the Hamiltonian H by

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

We define two linear operators $A\colon w \in L^1(Q;\mathbb{R}^d) \mapsto A[w] \in L^1(0,T;\mathbb{R}^k)$ and $A^\star\colon P \in L^\infty(0,T;\mathbb{R}^k) \mapsto A^*[P] \in L^\infty(Q;\mathbb{R}^d)$ as follows:

$$A[w](t) = \int_{\mathbb{T}^d} a(x,t)w(x,t)\mathrm{d}x, \qquad A^\star[P](x,t) = a^\star(x,t)P(t), \quad \forall (x,t) \in Q.$$

Note that the function a will be assumed to be bounded. The mean field game system under study is the following:

$$\left\{\begin{array}{lll} \text{(i)} & \left\{\begin{array}{ll} -\partial_t u - \Delta u + \boldsymbol{H}[\nabla u + A^\star P] = \gamma, & (x,t) \in Q, \\ u(x,T) = g(x), & x \in \mathbb{T}^d, \end{array}\right. & \\ \text{(ii)} & v = -\boldsymbol{H}_p[\nabla u + A^\star P], & (x,t) \in Q, \\ \text{(iii)} & \left\{\begin{array}{ll} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, & (x,t) \in Q, \\ m(0,x) = m_0(x), & x \in \mathbb{T}^d, \end{array}\right. & \\ \text{(iv)} & \gamma(x,t) = f(x,t,m(t)), & (x,t) \in Q, \\ \text{(v)} & P(t) = \boldsymbol{\phi}[A[vm]](t), & t \in [0,T], \end{array}\right. \tag{MFG}$$

where the unknown is (m,v,u,γ,P), with $m\colon Q \to \mathbb{R}$, $v\colon Q \to \mathbb{R}^d$, $u\colon Q \to \mathbb{R}$, $\gamma\colon Q \to \mathbb{R}$, and $P\colon [0,T] \to \mathbb{R}^k$. In this model, the coupling terms are the variables γ and P. Given γ and P, the optimal control problem solved by a representative agent is described by (4) and (5); the optimal feedback v is obtained by computing the corresponding value function u, solution the Hamilton-Jacobi-Bellman equation (MFG,i) and then v is obtained with using Equation (MFG,ii). Conversely, the coupling terms γ and P are deduced from v by computing the distribution m of the agents, solution to the Fokker-Planck equation (MFG,iii). The first coupling term γ is deduced from m through (MFG,iv) and the second coupling term P is deduced from m and v through equation (MFG,v).

We will make use the Benamou-Brenier change of variable $w = mv$, where mv denotes the pointwise product of m and v (that is $w(x,t) = m(x,t)v(x,t)$ for a.e. $(x,t) \in Q$).

2.3 Assumptions

We assume that there exist four constants $C_0 > 0$, $C_1 > 0$, $C_2 > 0$, and $\alpha_0 \in (0,1)$ such that the following holds true.

(H1) *Convexity of L.* For any $(x,t) \in Q$, the function $L(x,t,\cdot)$ is strongly convex with modulus $1/C_0$.

(H2) *Lipschitz continuity of L.* For any x and $y \in \mathbb{T}^d$, for any $t \in [0,T]$, and for any $v \in \mathbb{R}^d$,

$$|L(x,t,v) - L(y,t,v)| \leq C_0|x-y|(1+|v|^2).$$

(H3) *Boundedness of L, ϕ, and f.* For any $(x,t) \in Q$, for any $v \in \mathbb{R}^d$, and for any $z \in \mathbb{R}^k$,

$$L(x,t,v) \leq C_0|v|^2 + C_0, \quad |\phi(t,z)| \leq C_0, \quad \text{and} \quad |f(x,t,m)| \leq C_0.$$

(H4) *Regularity assumptions.* The running cost L is differentiable with respect to v and $D_v L$ is differentiable with respect to x and v. The mapping a is differentiable with respect to x. The mapping L, $D_v L$, $D_{vx} L$, $D_{vv} L$, ϕ, a, $D_x a$ are Hölder-continuous on any bounded set. The mappings m_0 and g lie in $C^{2+\alpha_0}(\mathbb{T}^d)$.

(H5) *Regularity of the coupling functions.* For all (x_1, t_1) and (x_2, t_2) in Q and for all m_1 and m_2 in $\mathcal{D}_1(\mathbb{T}^d)$,

$$|f(x_2, t_2, m_2) - f(x_1, t_1, m_1)| \leq C_0\big(|x_2 - x_1| + |t_2 - t_1|^{\alpha_0}\big) + C_1 \|m_2 - m_1\|_{L^2(\mathbb{T}^d)}.$$

For all $t \in [0, T]$ and for all z_1 and z_2 in $\mathbb{R}^k$, $|\phi(t, z_2) - \phi(t, z_1)| \leq C_2 |z_2 - z_1|$.

(H6) *Potential structure.* The map f is monotone with respect to its third variable, that is to say,

$$\int_{\mathbb{T}^d} (f(x, t, m_2) - f(x, t, m_1))(m_2(x) - m_1(x))\, \mathrm{d}x \geq 0,$$

for any m_1 and $m_2 \in \mathcal{D}_1(\mathbb{T}^d)$ and $t \in [0, T]$. We assume that f has a primitive, that is, we assume the existence of a map $F \colon [0, T] \times \mathcal{D}_1(\mathbb{T}^d) \to \mathbb{R}$ such that

$$F(t, m_2) - F(t, m_1) = \int_0^1 \int_{\mathbb{T}^d} f(x, t, sm_2 + (1 - s)m_1)(m_2(x) - m_1(x))\, \mathrm{d}x\, \mathrm{d}s. \tag{1}$$

Moreover, ϕ has a convex potential Φ, that is to say there exists a measurable mapping $\Phi : [0, T] \times \mathbb{R}^k \to \mathbb{R}$, convex and differentiable with respect to its second variable and such that $\phi(t, z) = \nabla_z \Phi(t, z)$ for any $(t, z) \in [0, T] \times \mathbb{R}^k$.

Remark 2. *1. The monotonicity assumption on f implies that*

$$F(t, m_2) \geq F(t, m_1) + \int_{\mathbb{T}^d} f(x, t, m_1)(m_2(x) - m_1(x))\, \mathrm{d}x.$$

Since this inequality holds for any $m_1 \in \mathcal{D}_1(\mathbb{T}^d)$, F is convex with respect to its second variable as the supremum of affine functions.

2. Our assumptions are stronger than those of [9], since we require the boundedness of ϕ and the Lipschitz continuity of f with respect to m, for the $L^2(\mathbb{T}^d)$-norm.

Lemma 3. *The Hamiltonian H is differentiable with respect to p and H_p is differentiable with respect to x and p. Moreover, H, H_p, H_{px}, and H_{pp} are locally Hölder-continuous.*

Proof. See [9, Lemma 1]. □

Theorem 4. *There exists $\alpha \in (0, 1)$ such that the system (MFG) has a unique solution $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$ in $\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \mathcal{C}^{\alpha}(Q; \mathbb{R}^d) \times \mathcal{C}^{2+\alpha, 1+\alpha/2}(Q) \times \mathcal{C}^{\alpha}(Q) \times \mathcal{C}^{\alpha}(0, T; \mathbb{R}^k)$. Moreover $Dv \in \mathcal{C}^{\alpha}(Q; \mathbb{R}^{d \times d})$.*

Proof. Direct application of [9, Theorem 1 and Proposition 2]. □

For the rest of the article, following Theorem 4, we denote by $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$ the unique solution to (MFG) and we set $\bar{w} = \bar{m}\bar{v}$.

3 Generalized conditional gradient and fictitious play

In this section we present the GCG method and we provide an interpretation as a learning method. We state our main convergence result, Theorem 10.

3.1 Mappings

We introduce in this subsection different mappings, which will allow to express in a compact fashion the different mutual dependencies of the variables of the mean field game. The well-posedness of all these mappings will be justified in Section 4.

Fokker-Planck mapping We first define the mapping $\boldsymbol{M}\colon v \in \Theta \mapsto \boldsymbol{M}[v] \in W^{2,1,q}(Q)$ which associates any vector field v to the weak solution to the Fokker-Planck equation

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (vm) &= 0 && (x,t) \in Q, \\ m(0,x) &= m_0(x) && x \in \mathbb{T}^d. \end{aligned} \tag{2}$$

Next we consider the set $\mathcal{R}$ defined by

$$\mathcal{R} = \left\{ \begin{array}{l} (m,w) \in W^{2,1,q}(Q) \times \Theta \,\big|\, \\ \qquad \partial_t m - \Delta m + \nabla \cdot w = 0, \quad m(0,\cdot) = m_0 \\ \qquad \exists v \in L^\infty(Q;\mathbb{R}^d),\ w = mv \end{array} \right\}.$$

Lemma 5. *The set $\mathcal{R}$ is convex. Moreover, given $v \in \Theta$, the pair (m, mv) lies in $\mathcal{R}$, for $m = \boldsymbol{M}[v]$.*

Proof. The proof of convexity is left to the reader. The second part of the lemma is a consequence of Lemma 1 and Lemma 16. □

HJB mapping Given $(\gamma, P) \in \Xi$, we define $\boldsymbol{u}[\gamma, P]$ as the viscosity solution to the following HJB equation:

$$\begin{aligned} -\partial_t u - \Delta u + \boldsymbol{H}[\nabla u + A^\star P] &= \gamma && (x,t) \in Q, \\ u(x,T) &= g(x) && x \in \mathbb{T}^d. \end{aligned} \tag{3}$$

Let us given an interpretation of $\boldsymbol{u}$ as the value function of an optimal control problem. Let $(B_s)_{s\in[0,T]}$ denote a d-dimensional Brownian motion. Let $\mathbb{F}$ denote the filtration generated by the Brownian motion $(B_s)_{s\in[0,T]}$. We denote by $\mathbb{L}^2_{\mathbb{F}}(t,T)$ the set of progressively measurable stochastic processes ν defined on $[t,T]$ and valued in $\mathbb{R}^d$ such that $\mathbb{E}\big[\int_t^T |\nu_s|^2 \mathrm{d}s\big] < +\infty$. Given $(\gamma, P) \in \Xi$, we consider the mapping $J[\gamma,P]\colon Q \times \mathbb{L}^2_{\mathbb{F}}(0,T) \to \mathbb{R}$, defined by

$$J[\gamma,P](x,t,\nu) = \mathbb{E}\Big[\int_t^T \Big(L(X_s,s,\nu_s) + \langle A^\star[P](X_s,s), \nu_s\rangle + \gamma(X_s,s)\Big)\mathrm{d}s + g(X_T)\Big], \tag{4}$$

where $(X_s)_{s\in[t,T]}$ is the solution to $\mathrm{d}X_s = \nu_s\,\mathrm{d}s + \sqrt{2}\,\mathrm{d}B_s$, $X_t = x$. Then, $\boldsymbol{u}[\gamma,P]$ is the value function of the optimal control problem associated with $J[\gamma,P]$, that is to say, for any $(x,t) \in Q$,

$$\boldsymbol{u}[\gamma,P](x,t) = \inf_{\nu \in \mathbb{L}^2_{\mathbb{F}}(0,T)} J[\gamma,P](x,t,\nu). \tag{5}$$

This is a classical result from dynamic programming theory, see [32].

The other mappings Next we introduce three mappings, $\boldsymbol{v}$, $\boldsymbol{m}$, and $\boldsymbol{w}$, defined on Ξ, and such that

$$\boldsymbol{m}[\gamma, P] \in W^{2,1,q}(Q), \qquad \boldsymbol{v}[\gamma, P] \in \Theta, \qquad \text{and} \qquad \boldsymbol{w}[\gamma, P] \in \Theta.$$

For any pair $(\gamma, P) \in \Xi$, they are given by

$$\begin{aligned} \boldsymbol{v}[\gamma, P] &= -\boldsymbol{H}_p[\nabla \boldsymbol{u}[\gamma, P] + A^\star P], \\ \boldsymbol{m}[\gamma, P] &= \boldsymbol{M}\big[\boldsymbol{v}[\gamma, P]\big], \\ \boldsymbol{w}[\gamma, P] &= \boldsymbol{m}[\gamma, P]\boldsymbol{v}[\gamma, P]. \end{aligned}$$

Finally, using Nemytskii operators, we define

$$\begin{aligned} \boldsymbol{\gamma}\colon\ & W^{2,1,q}(Q) \to W^{1,0,\infty}(Q) \cap \mathcal{C}(Q), \\ & m \mapsto f[m], \end{aligned} \qquad \begin{aligned} \boldsymbol{P}\colon\ & \Theta \to \mathcal{C}(0,T;\mathbb{R}^k) \\ & w \mapsto \boldsymbol{\phi}\big[A[w]\big]. \end{aligned}$$

In summary, the mappings $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{m}$, and $\boldsymbol{w}$, derived from the equations (MFG,i-iii), allow to express the behavior of the agents in function of the coupling terms γ and P. Conversely, the mappings $\boldsymbol{\gamma}$ and $\boldsymbol{P}$, derived from the equations (MFG,iv-v), allow to express the coupling terms as a function of the behavior of the agents, described by the variables m and w.

3.2 Optimal control problems

We introduce two optimal control problems involving the Fokker-Planck equation: problem (P), which is the potential formulation associated with the coupled system (MFG), and problem ($\mathcal{P}[\gamma, P]$), which is a PDE formulation for the optimal control problem (5). A key observation is that ($\mathcal{P}[\gamma, P]$) can be seen as a partial linearization of (P). The results announced in this subsection will be proved in Section 5.

The potential problem is given by

$$\inf_{(m,w)\in\mathcal{R}} \mathcal{J}(m,w) \coloneqq \mathcal{J}_1(m,w) + \mathcal{J}_2(m,w), \tag{P}$$

where

$$\begin{aligned} \mathcal{J}_1(m,w) &= \int_Q \tilde{\boldsymbol{L}}[m,w](x,t)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} g(x)m(x,T)\,\mathrm{d}x, \\ \mathcal{J}_2(m,w) &= \int_0^T \big(\boldsymbol{F}[m](t) + \boldsymbol{\Phi}[Aw](t)\big)\mathrm{d}t, \end{aligned} \tag{6}$$

and where $\tilde{\boldsymbol{L}}$ is the Nemytskii operator of the perspective function $\tilde{L}\colon Q\times\mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$ of L, defined by

$$\tilde{L}(x,t,m,w) = \begin{cases} mL\big(x,t,\frac{w}{m}\big), & \text{if } m > 0, \\ 0, & \text{if } m = 0 \text{ and } w = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that $\tilde{L}$ is convex and lower semi-continuous with respect to (m,w) (see [8, Lemma 1.157]).

Lemma 6. *The pair $(\bar{m},\bar{w})$ is the unique solution to the potential problem (P).*

Proof. Lemma 6 is a direct consequence of Corollary 32, proved in Section 5. □

We introduce now the PDE formulation of the stochastic optimal control problem solved by the representative agent. Given a pair $(\gamma, P) \in \Xi$, we consider the criterion

$$\mathcal{Z}[\gamma, P](m, w) = \mathcal{J}_1(m, w) + \Big(\int_Q \gamma(x,t) m(x,t) \,\mathrm{d}x \,\mathrm{d}t + \int_0^T \langle A[w](t), P(t) \rangle \,\mathrm{d}t \Big),$$

and the associated optimal control problem

$$\inf_{(m,w)\in\mathcal{R}} \mathcal{Z}[\gamma, P](m, w). \tag{$\mathcal{P}[\gamma, P]$}$$

Lemma 7. *For any $(\gamma, P) \in \Xi$, the pair $(\boldsymbol{m}[\gamma, P], \boldsymbol{w}[\gamma, P])$ is the unique solution to $(\mathcal{P}[\gamma, P])$.*

Proof. This is a direct consequence of Proposition 31, proved at page 23. □

Take a pair $(\hat{m}, \hat{w}) \in \mathcal{R}$ and set $\hat{\gamma} = \boldsymbol{\gamma}[\hat{m}]$ and $\hat{P} = \boldsymbol{P}[\hat{w}]$. The criterion $\mathcal{Z}[\hat{\gamma}, \hat{P}](\cdot)$ can be seen as a partial linearization of $\mathcal{J}$: while the term $\mathcal{J}_1(m, w)$ is the same in both cost functions, $\int_Q \hat{\gamma} m \,\mathrm{d}x \,\mathrm{d}t$ is a linearization of $\int_0^T \boldsymbol{F}[m] \,\mathrm{d}t$ around $\hat{m}$ and $\int_0^T \langle A[w], \hat{P} \rangle \,\mathrm{d}t$ is a linearization of $\int_0^T \boldsymbol{\Phi}[Aw] \,\mathrm{d}t$ around $\hat{w}$. The connection between $\mathcal{J}$ and $\mathcal{Z}$ is made more precise in the following lemma.

Lemma 8. *Let (m_1, w_1) and (m_2, w_2) be in $\mathcal{R}$. Let $\gamma_1 = \boldsymbol{\gamma}[m_1]$ and let $P_1 = \boldsymbol{P}[w_1]$. Then, there exists a constant $C > 0$, independent of (m_1, w_1) and (m_2, w_2) such that*

$$\int_Q \gamma_1 (m_2 - m_1) \,\mathrm{d}x \,\mathrm{d}t + \int_0^T \langle A[w_2 - w_1], P_1 \rangle \,\mathrm{d}t \le \mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1) \tag{7}$$

and

$$\begin{aligned} \mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1) \le{}& \int_Q \gamma_1 (m_2 - m_1) \,\mathrm{d}x \,\mathrm{d}t + \int_0^T \langle A[w_2 - w_1], P_1 \rangle \,\mathrm{d}t \\ &+ C_1 \int_0^T \|m_2(t,\cdot) - m_1(t,\cdot)\|_{L^2(\mathbb{T}^d)} \|m_2(t,\cdot) - m_1(t,\cdot)\|_{L^1(\mathbb{T}^d)} \,\mathrm{d}t \\ &+ C_2 \int_0^T \Big| \int_{\mathbb{T}^d} a(x,t)(w_2(x,t) - w_1(x,t)) \,\mathrm{d}x \Big|^2 \,\mathrm{d}t. \end{aligned} \tag{8}$$

Proof. Using the definitions of γ_1 and P_1, we obtain

$$\begin{aligned} &\big(\mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1)\big) - \Big(\int_Q \gamma_1 (m_2 - m_1) \,\mathrm{d}x \,\mathrm{d}t + \int_0^T \langle A[w_2 - w_1], P_1 \rangle \,\mathrm{d}t \Big) \\ &\qquad = (a) + (b), \end{aligned}$$

where

$$\begin{aligned} (a) &= \int_0^T \Big(\boldsymbol{F}[m_2](t) - \boldsymbol{F}[m_1](t) - \int_{\mathbb{T}^d} f[m_1](x,t) \big(m_2(x,t) - m_1(x,t)\big) \,\mathrm{d}x \Big) \,\mathrm{d}t, \\ (b) &= \int_0^T \big(\boldsymbol{\Phi}[Aw_2] - \boldsymbol{\Phi}[Aw_1] - \langle \boldsymbol{\phi}[Aw_1](t), Aw_2(t) - Aw_1(t) \rangle \big) \,\mathrm{d}t. \end{aligned}$$

Using Assumption (H6), we obtain that

$$(a) = \int_0^T \int_0^1 \int_{\mathbb{T}^d} \big(f[m_1 + s(m_2 - m_1)] - f[m_1]\big)(m_2 - m_1)\,\mathrm{d}x\,\mathrm{d}s\,\mathrm{d}t.$$

Then Assumptions (H5) and (H6) imply that

$$0 \le (a) \le C_1 \int_0^T \|m_2(t,\cdot) - m_1(t,\cdot)\|_{L^2(\mathbb{T}^d)} \|m_2(t,\cdot) - m_1(t,\cdot)\|_{L^1(\mathbb{T}^d)}\,\mathrm{d}t.$$

We estimate (b) in a similar way and obtain inequalities (7) and (8) easily. □

Corollary 9. *Let (m_1, w_1) and (m_2, w_2) be in $\mathcal{R}$. Let $\gamma_1 = \boldsymbol{\gamma}[m_1]$ and let $P_1 = \boldsymbol{P}[w_1]$. Then,*

$$\mathcal{Z}[\gamma_1, P_1](m_2, w_2) - \mathcal{Z}[\gamma_1, P_1](m_1, w_1) \le \mathcal{J}(m_2, w_2) - \mathcal{J}(m_1, w_1).$$

Proof. This is an immediate consequence of inequality (7) from Lemma 8 and the definitions of $\mathcal{J}$ and $\mathcal{Z}$. □

3.3 GCG algorithm and interpretation as a learning method

Using the partial linearization of $\mathcal{J}$ introduced in the previous subsection, the GCG algorithm yields Algorithm 1.

Algorithm 1 Generalized conditional gradient

Choose $(\bar{m}_0, \bar{w}_0) \in \mathcal{R}$
for $0 \le k < N$ **do**
1. Set $\gamma_k = \boldsymbol{\gamma}[\bar{m}_k]$ and $P_k = \boldsymbol{P}[\bar{w}_k]$.
2. Find the solution to $\mathcal{P}[\gamma_k, P_k]$, that is, define successively:

$$u_k = \boldsymbol{u}[\gamma_k, P_k], \quad v_k = \boldsymbol{v}[\gamma_k, P_k], \quad m_k = \boldsymbol{m}[\gamma_k, P_k], \quad w_k = \boldsymbol{w}[\gamma_k, P_k].$$

3. Choose $\delta_k \in [0, 1]$.
4. Update $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (\bar{m}_k^{\delta_k}, \bar{w}_k^{\delta_k})$, where

$$(\bar{m}_k^{\delta}, \bar{w}_k^{\delta}) = (1 - \delta)(\bar{m}_k, \bar{w}_k) + \delta(m_k, w_k), \quad \forall \delta \in [0, 1]. \tag{9}$$

end for
return $(\bar{m}_N, \bar{w}_N)$.

We can give an interpretation of Algorithm 1 as a learning method. At Step 1, $\bar{m}_k$ and $\bar{w}_k$ are can be seen as predictions of the equilibrium values $\bar{m}$ and $\bar{w}$. The agents use them to make a prediction of the coupling terms, γ_k and P_k. In the second step, they find the corresponding best-response, by solving the HJB equation associated with their optimal control problem. Finally, in Steps 3 and 4, they update their prediction of m and w. In particular, when $\delta_k = 1/(k+1)$, for any $k \in \mathbb{N}$, we are in the setting of the fictitious play, as investigated in [18] (without price interaction).

We denote by ε_k the optimality gap at iterate k, defined by

$$\varepsilon_k = \mathcal{J}(\bar{m}_k, \bar{w}_k) - \inf_{\mathcal{R}} \mathcal{J} = \mathcal{J}(\bar{m}_k, \bar{w}_k) - \mathcal{J}(\bar{m}, \bar{w}).$$

We define the exploitability as follows

$$\begin{aligned}\sigma(m,w) &= \mathcal{Z}[\gamma,P](m,w) - \Big(\inf_{(m',w')\in\mathcal{R}} \mathcal{Z}[\gamma,P](m',w')\Big)\\ &= \mathcal{Z}[\gamma,P](m,w) - \mathcal{Z}[\gamma,P](\boldsymbol{m}[\gamma,P],\boldsymbol{w}[\gamma,P]) \geq 0,\end{aligned}$$

where $\gamma = \boldsymbol{\gamma}[m]$ and $P = \boldsymbol{P}[w]$. The exploitability is the largest decrease in cost that a representative agent can reach by playing its best response, assuming that all other agents play (m,w). At equilibrium, since there is no profitable deviation, the exploitability is null. We denote by σ_k the exploitability at iterate k, given by

$$\sigma_k = \mathcal{Z}[\gamma_k,P_k](\bar{m}_k,\bar{w}_k) - \mathcal{Z}[\gamma_k,P_k](m_k,w_k). \tag{10}$$

3.4 On the choice of stepsizes

The convergence analysis will concern two kinds of stepsizes: predefined stepsizes, whose value only depends on k, and adaptive stepsize, whose value depend on the two pairs $(\bar{m}_k,\bar{w}_k)$ and (m_k,w_k). Following [47], we will investigate three different rules for the determination of adaptive stepsizes.

- **Optimal stepsizes:** Find δ_k such that

$$\delta_k \in \operatorname*{arg\,min}_{\delta\in[0,1]} \mathcal{J}(\bar{m}_k^\delta,\bar{w}_k^\delta), \tag{11}$$

  where $(\bar{m}_k^\delta,\bar{w}_k^\delta)$ is defined as in (9).

- **Quasi-Armijo-Goldstein** condition: fix two parameters $c\in(0,1)$ and $\tau\in(0,1)$. At iterate k, we say that $\delta\in[0,1]$ satisfies the Quasi-Armijo-Goldstein (QAG) condition if

$$\mathcal{J}(\bar{m}_k^\delta,\bar{w}_k^\delta) \leq \mathcal{J}(\bar{m},\bar{w}) - c\delta\sigma_k.$$

  Then δ_k is defined by

$$\delta_k = \tau^{i_k}, \quad i_k = \operatorname{argmin}\Big\{j\in\mathbb{N}\,\Big|\,\tau^j \text{ satisfies the QAG condition}\Big\}. \tag{12}$$

- **Exploitability-based** stepsizes: we take

$$\delta_k = \min\left\{1, \frac{\sigma_k}{2\big(C_1 D_k^{(1)} + C_2 D_k^{(2)}\big)}\right\}, \tag{13}$$

  where C_1 and C_2 are the constants of Assumption (H5) and where

$$\begin{aligned}D_k^{(1)} &= \int_0^T \|m_k(t,\cdot) - \bar{m}_k(t,\cdot)\|_{L^2(\mathbb{T}^d)} \|m_k(t,\cdot) - \bar{m}_k(t,\cdot))\|_{L^1(\mathbb{T}^d)}\,\mathrm{d}t\\ D_k^{(2)} &= \int_0^T \Big|\textstyle\int_{\mathbb{T}^d} a(x,t)(w_2(x,t) - w_1(x,t))\,\mathrm{d}x\Big|^2\,\mathrm{d}t\end{aligned} \tag{14}$$

3.5 Main convergence results and discussion

Theorem 10. *1. At each iteration, assume that $(\delta_k)_{k\in\mathbb{N}}$ satisfy one the three adaptive rules described above: either (11), (12) (for values of τ and c independent of k), or (13). Then, there exist two constants $C > 0$ and $\lambda > 0$ such that*

$$\varepsilon_k \leq C\lambda^k, \quad \forall k \in \mathbb{N}.$$

Moreover, the number i_k of iterations for the QAG condition is bounded by some constant independent of k.

2. There exists a constant $C > 0$, independent of the stepsize choice, such that

$$\varepsilon_{k+1} \leq \frac{\varepsilon_0 \exp\left(C\sum_{j=0}^{k}\delta_j^2\right)}{\exp\left(\sum_{j=0}^{k}\delta_j\right)}, \tag{15}$$

for all $k \in \mathbb{N}$.

The theorem is proved in Subsection 6.2.

Let us first discuss some aspects related to adaptive stepsizes. While the method for computing δ_k in the case of the QAG condition or the exploitability-based formula is explicit, it is impossible to find δ_k that exactly satisfies (11). We suggest to compute an approximation of the optimal stepsize with the golden-section method, which we briefly describe. Denote by $\varphi := (\sqrt{5}+1)/2$ the golden number and choose a tolerance $\kappa \in (0,1)$. At each step $k \in \{0,\dots,N\}$, the learning rate δ_k is computed as follows: Set $(a,d) = (0,1)$ and $(b,c) = (d-(d-a)/\varphi, a+(d-a)/\varphi)$. While $a - d > \kappa$, find

$$\bar{\delta} \in \operatorname*{arg\,min}_{\delta\in\{a,b,c,d\}} \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta). \tag{16}$$

Then set $d = b$ (resp. $d = c$, $a = b$ or $a = c$) if $\bar{\delta} = a$ (resp. $\bar{\delta} = b$, $\bar{\delta} = c$ or $\bar{\delta} = d$). When $a - d \leq \kappa$, stop and set $\delta_k^{\mathrm{GS}} = \bar{\delta}$. Finally, denote δ_k^{QAG} the stepsize determined by (12). As a substitute for (11), one can choose δ_k such that

$$\delta_k \in \operatorname*{arg\,min}_{\delta\in\{\delta_k^{\mathrm{GS}},\delta_k^{\mathrm{QAG}}\}} \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta).$$

In view of the proof of Theorem 10, it is clear that linear convergence is also achieved for this choice of stepsize.

Remark 11. *The practical computation of δ_k requires a bounded number (with respect to k) of evaluations of the cost functional $\mathcal{J}$ in each of the three considered cases, (12) (since i_k is bound), (13) (the formula is explicit), (16) (the number of evaluations can be bounded by some constant depending on κ and φ). Therefore linear convergence is not only achieved with respect to k but also with respect to the number of evaluations of $\mathcal{J}$.*

We discuss now the convergence of $(\varepsilon_k)_{k\in\mathbb{N}}$ for predefined stepsizes. The following lemma covers the case of the fictitious play learning rate ($\delta_k = \frac{1}{k+1}$) and the Frank-Wolfe stepsize ($\delta_k = \frac{2}{k+2}$).

Lemma 12. *Let $p > 0$. Assume that $\delta_k = \frac{p}{k+p}$, for any $k \in \mathbb{N}$. Then, for all $k \in \mathbb{N}$,*

$$\varepsilon_k \leq \frac{\varepsilon_0\, p^p \exp(2Cp)}{(k+p+1)^p}.$$

Proof. We have the following inequalities:

$$\sum_{j=0}^{k} \frac{p}{j+p} \geq \int_0^{k+1} \frac{p}{s+p}\,\mathrm{d}s = p\ln\Big(\frac{k+p+1}{p}\Big)$$
$$\sum_{j=0}^{k} \Big(\frac{p}{j+p}\Big)^2 \leq p^2\Big(\frac{1}{p} + \int_0^{\infty} \frac{1}{(s+p)^2}\,\mathrm{d}s\Big) = 2p.$$

Using inequality (15), we deduce that $\varepsilon_k \leq \varepsilon_0 \exp\big(2Cp - p\ln(k+p) + p\ln(p)\big)$, from which the announced result follows. □

Remark 13. *Lemma 12 shows that rates of convergence of order $\mathcal{O}(k^{-p})$ can be achieved. Yet it is instructive to compute the value of p which yields the best estimate, for a given iteration number k. To this purpose, we define $h(p) = 2Cp - p\ln(k+p+1) + p\ln(p)$. In view of the proof of Lemma 12, it is enough to minimize h with respect to p. We have*

$$h'(p) = 2C - \frac{p}{k+p+1} + \ln\Big(\frac{p}{k+p+1}\Big) + 1.$$

It is easy to verify that $h'(p)$ is increasing over $[0,\infty)$. Using the change of variable $u(p) = \frac{p}{k+p+1}$, we see that $h'(p) = 0$ if and only if $u(p) = \zeta(C)$, where $\zeta(C)$ is the unique solution of the equation

$$(2C+1) - u + \ln(u) = 0 \tag{17}$$

in $(0,1)$. Let us denote by $p(k,C)$ the optimal value of p, we have $p(k,C) = \frac{u(C)(k+1)}{1-u(C)}$. Since $u(C) \in (0,1)$, we deduce from (17) that $\ln(u(C)) \leq -2C$, which finally implies that

$$p(k,C) \leq \frac{\exp(-2C)(k+1)}{1-\exp(-2C)}.$$

This confirms the interest of taking a large value of p when a large number of iterations of k is allowed. However, the optimal value of p decreases exponentially with C. The constant C will be constructed in the proof of Theorem 10, in particular, it is proportional to the Lipschitz moduli C_1 and C_2 appearing in Assumption (H5). This fact severely mitigates the interest of choosing a high value of p, especially as the constant C is cumbersome to compute explicitely.

Lemma 14. *Let the sequence of stepsizes $(\delta_k)_{k\in\mathbb{N}}$ converge to zero. Then, for any $r \in (0,1)$, there exist two constants $C > 0$ and $k_0 \in \mathbb{N}$, both depending on the sequence of stepsizes and r, such that*

$$\varepsilon_{k+1} \leq \frac{C\varepsilon_0}{\exp\big[(1-r)\big(\sum_{j=0}^{k}\delta_j\big)\big]}, \quad \forall k \geq k_0.$$

In particular, if $\sum_{k=0}^{\infty}\delta_k = \infty$, then $\varepsilon_k \underset{k\to\infty}{\longrightarrow} 0$.

Proof. Let $r \in (0,1)$. Let $k_0 \in \mathbb{N}$ be such that $\delta_k \leq \frac{r}{C}$, for any $k \geq k_0$ (and where C is as in (15) holds true). Then for any $k \geq k_0$, we have

$$C\sum_{j=0}^{k}\delta_j^2 \leq C\sum_{j=0}^{k_0-1}\delta_j^2 + \underbrace{C\delta_{k_0}}_{\leq r}\sum_{j=k_0}^{k}\delta_j.$$

It follows that

$$C\sum_{j=0}^{k}\delta_j^2-\sum_{j=0}^{k}\delta_j\le\Big(C\sum_{j=0}^{k_0-1}\delta_j^2+r\sum_{j=0}^{k_0-1}\delta_j\Big)-(1-r)\sum_{j=0}^{k}\delta_j.$$

Applying the exponential function to the above inequality, we obtain the announced result. □

Lemma 14 shows that high convergence rates can be achieved if the sequence $(\delta_k)_{k\in\mathbb{N}}$ decreases slowly, as we already noticed in Lemma 12. However, if the convergence to 0 is slow, the values of k_0 and C may be very large, as was revealed in the proof.

Remark 15. *In the case where $\delta_k=(1+k)^{-\alpha}$, for $\alpha\in(0,1)$. Then*

$$\sum_{j=0}^{k}\delta_j\ge\int_0^{k+1}(s+1)^{-\alpha}\,\mathrm{d}s=\frac{(k+2)^{1-\alpha}-1}{1-\alpha}.$$

Then it is easy to show with Lemma 14 that for any $\alpha'\in(0,\alpha)$, there exist $k_0\in\mathbb{N}$ and $C>0$ (both depending on α) such that

$$\varepsilon_{k+1}\le C\varepsilon_0\exp\left(-\frac{k^{1-\alpha}}{1-\alpha'}\right),\quad\forall k\ge k_0.$$

The best speed of convergence is achieved for values of α close to 0. However, the same phenomenon as the one noticed in Remark 13 arises: the benefit of a small α is likely to be observed only for huge values of k.

4 Properties of the auxiliary mappings

This section is dedicated to the statement of technical lemmas concerning the well-posedness and the stability of auxiliary mappings and to their proofs (given in Subsections 4.3 and 4.4).

4.1 Statement of the results

Lemma 16. *The map $\boldsymbol{M}$ is a bounded linear map from Θ to $W^{2,1,q}(Q)$.*

The proof is given at page 17.

Lemma 17. *Let $R>0$. Let v_1 and $v_2\in\Theta$. Let $m_i=\boldsymbol{M}[v_i]\in W^{2,1,q}(Q)$ for $i\in\{1,2\}$. Assume that $\|v_1\|_\Theta\le R$ and $\|m_2\|_{L^\infty(Q)}\le R$. Then, there exists a constant $C(R)$, independent of v_1 and v_2, such that*

$$\|m_2-m_1\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}\le C(R)\Big(\int_Q|v_2-v_1|^2m_2\,\mathrm{d}x\,\mathrm{d}t\Big)^{1/2}.$$

The proof is given at page 17.

Lemma 18. *Let $R>0$. There exists a constant $C(R)$ such that for all (γ_1,P_1) and (γ_2,P_2) in Ξ_R, the following holds:*

$$\|\boldsymbol{u}[\gamma_2,P_2]-\boldsymbol{u}[\gamma_1,P_1]\|_{L^\infty(Q)}\le C(R)\Big(\|P_2-P_1\|_{L^2(0,T;\mathbb{R}^k)}+\|\gamma_2-\gamma_1\|_{L^\infty(Q)}\Big).$$

The proof will be given at page 19.

Proposition 19. *The map $\boldsymbol{u}$ is well-defined from Ξ to $W^{2,1,q}(Q)$. Moreover, for any $R > 0$, there exists a constant $C(R) > 0$ such that for any $(\gamma, P) \in \Xi_R$, $\|\boldsymbol{u}[\gamma, P]\|_{W^{2,1,q}(Q)} \leq C(R)$.*

The proof will be given at page 20.

Lemma 20. *The maps $\boldsymbol{v}$, $\boldsymbol{m}$, and $\boldsymbol{w}$ are well-defined from Ξ to Θ, $W^{2,1,q}(Q)$, and Θ, respectively. Moreover, for any $R > 0$, there exists $C(R) > 0$ such that for any $(\gamma, P) \in \Xi_R$, it holds:*

$$\|\boldsymbol{v}[\gamma, P]\|_\Theta + \|\boldsymbol{m}[\gamma, P]\|_{W^{2,1,q}(Q)} + \|\boldsymbol{w}[\gamma, P]\|_\Theta \leq C(R).$$

The proof will be given at page 21.

Lemma 21. *The mappings $\boldsymbol{\gamma}$ and $\boldsymbol{P}$ are well-defined. There exists a constant $C > 0$ such that for all $m \in W^{2,1,q}(Q)$ and for all $w \in \Theta$,*

$$\|\gamma\|_{W^{1,0,\infty}(Q)} + \|P\|_{L^\infty(0,T;\mathbb{R}^k)} \leq C,$$

where $\gamma = \boldsymbol{\gamma}[m]$ and $P = \boldsymbol{P}[w]$. There exists a constant $C > 0$ such that for all m_1 and m_2 in $W^{2,1,q}(Q)$ and for all w_1 and w_2 in Θ,

$$\begin{aligned}\|\boldsymbol{\gamma}[m_2] - \boldsymbol{\gamma}[m_1]\|_{L^\infty(Q)} &\leq C\|m_2 - m_1\|_{L^\infty(0,T;L^2(\mathbb{T}^d))},\\ \|\boldsymbol{P}[w_2] - \boldsymbol{P}[w_1]\|_{L^2(0,T;\mathbb{R}^k)} &\leq C\|w_2 - w_1\|_{L^2(Q;\mathbb{R}^d)}.\end{aligned}$$

The proof will be given at page 21.

4.2 Parabolic estimates

In this section we provide estimates for the following parabolic equation:

$$\begin{aligned}\partial_t u - \sigma\Delta u + \langle b, \nabla u\rangle + cu &= h, \qquad (x,t) \in Q,\\ u(x,0) &= u_0(x), \qquad x \in \mathbb{T}^d,\end{aligned} \tag{18}$$

for different assumptions on $b\colon Q \to \mathbb{R}^d$, $c\colon Q \to \mathbb{R}$, $h\colon \to \mathbb{R}$, and $u_0\colon \mathbb{T}^d \to \mathbb{R}$. The proofs of the following results can be found in the Appendix of [9]; they largely rely on [51]. We recall that q is a fixed parameter and $q > d + 2$.

In the next theorem, we consider the Sobolev space $W^{2-2/p,p}(\mathbb{T}^d)$ with a fractional order of derivation, see [51, section II.2] for a definition.

Theorem 22. *For all $R > 0$, there exists $C > 0$ such that for all $u_0 \in W^{2-2/q,q}(\mathbb{T}^d)$, for all $b \in L^q(Q;\mathbb{R}^d)$, for all $c \in L^q(Q)$, and for all $h \in L^q(Q)$ satisfying*

$$\|u_0\|_{W^{2-2/q,q}(\mathbb{T}^d)} + \|b\|_{L^q(Q;\mathbb{R}^d)} + \|c\|_{L^q(Q)} + \|h\|_{L^q(Q)} \leq R,$$

equation (18) has a unique solution u in $W^{2,1,q}(Q)$. Moreover, $\|u\|_{W^{2,1,q}(Q)} \leq C$.

Theorem 23. *There exists $C > 0$ such that for all $u_0 \in W^{2-2/q,q}(\mathbb{T}^d)$ and for all $h \in L^q(Q)$, the unique solution u to (18) (with $b = 0$ and $c = 0$) satisfies the following estimate:*

$$\|u\|_{W^{2,1,q}(Q)} \leq C\big(\|u_0\|_{W^{2-2/q,q}(\mathbb{T}^d)} + \|h\|_{L^q(Q)}\big).$$

Theorem 24. *For all $\beta \in (0,1)$, for all $R > 0$, there exist $\alpha \in (0,1)$ and $C > 0$ such that for all $u_0 \in \mathcal{C}^{2+\beta}(\mathbb{T}^d)$, $b \in \mathcal{C}^{\beta,\beta/2}(Q;\mathbb{R}^d)$, $c \in \mathcal{C}^{\beta,\beta/2}(Q)$ and $h \in \mathcal{C}^{\beta,\beta/2}(Q)$ satisfying*

$$\|u_0\|_{\mathcal{C}^{2+\beta}(\mathbb{T}^d)} + \|b\|_{\mathcal{C}^{\beta,\beta/2}(Q;\mathbb{R}^d)} + \|c\|_{\mathcal{C}^{\beta,\beta/2}(Q)} + \|h\|_{\mathcal{C}^{\beta,\beta/2}(Q)} \leq R,$$

the solution to (18) lies in $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ and satisfies $\|u\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)} \leq C$.

4.3 Fokker-Planck equation

Proof of Lemma 16. Let us write the Fokker-Planck equation in the form of equation (18): $\partial_t m - \Delta m + (\nabla \cdot v)m + \langle v, \nabla m\rangle = 0$. The lemma follows from Theorem 22. □

Proof of Lemma 17. Set $w = v_2 - v_1$ and $\mu = m_2 - m_1$. Then μ is the solution to

$$\begin{aligned} \partial_t \mu - \Delta \mu + \nabla \cdot (v_1 \mu) &= -\nabla \cdot (w m_2), \qquad (x,t) \in Q, \\ \mu(x,0) &= 0, \qquad\qquad\qquad x \in \mathbb{T}^d. \end{aligned}$$

Denote by V the space $W^{1,2}(\mathbb{T}^d)$ and consider the Gelfand triple $(V, L^2(\mathbb{T}^d), V^*)$, where V^* denotes the dual of V. Then μ is solution of a parabolic equation of the form

$$\begin{aligned} \partial_t m(t) + B(t)m(t) &= f(t), \qquad (x,t) \in Q, \\ m(x,0) &= 0, \qquad\quad x \in \mathbb{T}^d, \end{aligned}$$

where $B(t) \in L(V, V^*)$ and $f(t) \in V^*$. For any $m \in V$, we have

$$\begin{aligned} \langle B(t)m, m\rangle_V &= \int_{\mathbb{T}^d} \big(-\Delta m + \nabla \cdot v_1(t) m + \langle v_1(t), \nabla m\rangle\big) m \,\mathrm{d}x \\ &= \int_{\mathbb{T}^d} \big(|\nabla m|^2 - \langle v_1(t), \nabla m\rangle m\big) \mathrm{d}x, \end{aligned}$$

where the second equality is obtained by integration by parts. Using Cauchy-Schwarz inequality and $\|v_1\|_{L^\infty(Q;\mathbb{R}^d)} \leq R$, we obtain the following inequality:

$$\langle B(t)m, m\rangle_V \geq \|\nabla m\|^2_{L^2(\mathbb{T}^d;\mathbb{R}^d)} - C\|\nabla m\|_{L^2(\mathbb{T}^d;\mathbb{R}^d)} \|m\|_{L^2(\mathbb{T}^d)}$$

where the constant C is independent of t (but depends on R). A direct application of Young's inequality yields the existence of C (depending on R) such that

$$\langle B(t)m, m\rangle_V \geq \frac{1}{2}\|m\|^2_V - C\|m\|^2_{L^2(\mathbb{T}^d)}.$$

Thus $B(t)$ is semi-coercive, uniformly in time. With similar techniques, one can show that

$$\langle B(t)m, m'\rangle_V \leq C\|m\|_V \|m'\|_V,$$

for a.e. $t \in (0,T)$ and for all m and m' in V. We can apply [55, Chapter 3, Theorems 1.1 and 1.2], from which we derive

$$\begin{aligned} \|\mu\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} &\leq C\big(\|\mu\|_{L^2(0,T;V)} + \|\partial_t \mu\|_{L^2(0,T;V^*)}\big) \\ &\leq C\|f\|_{L^2(0,T;V^*)} \leq C\|\nabla \cdot (w m_2)\|_{L^2(0,T;V^*)} \\ &\leq C\|w m_2\|_{L^2(Q;\mathbb{R}^d)}. \end{aligned}$$

Finally, since $\|m_2\|_{L^\infty(Q)} \leq R$, we have $\|w m_2\|^2_{L^2(Q;\mathbb{R}^d)} \leq C \int_Q |w|^2 m_2 \,\mathrm{d}x\,\mathrm{d}t$. Combining the two last obtained inequalities, we obtain the announced result. □

4.4 HJB equation

The analysis of the HJB equation relies on its connection with the value function of an optimal control problem, that was introduced in (5). This connection allows first to show a uniform bound for $\boldsymbol{u}[\gamma, P]$.

Lemma 25. *Let $R > 0$ and let $(\gamma, P) \in \Xi_R$. There exists a constant $C(R) > 0$ such that $\|\boldsymbol{u}[\gamma, P]\|_{L^\infty(Q)} \leq C(R)$ and such that u is $C(R)$-Lipschitz continuous with respect to x. Moreover, for any $(x, t) \in Q$,*

$$\boldsymbol{u}[\gamma, P](x, t) = \inf_{\nu \in \mathbb{L}^{2, C(R)}_{\mathbb{F}}(t,T)} J[\gamma, P](x, t, \nu). \tag{19}$$

In the above relation, $\mathbb{L}^{2,C(R)}_{\mathbb{F}}(t, T)$ denotes the set of stochastic processes $\nu \in \mathbb{L}^2_{\mathbb{F}}(t, T)$ such that $\mathbb{E}\big[\int_t^T |\nu_s|^2 \mathrm{d}s\big] \leq C(R)$.

Proof. We first derive a lower bound of L. By assumption (H4), $L(x, t, 0)$ and $L_v(x, t, 0)$ are bounded. It follows then from the strong convexity assumption (Assumption (H1)) that there exists a constant $C > 0$ such that

$$\frac{1}{C}|\nu|^2 - C \leq L(x, t, \nu), \qquad \text{for all } (x, t, \nu) \in Q \times \mathbb{R}^d. \tag{20}$$

Then, for any $(x, s) \in Q$ and for any $\nu \in \mathbb{R}^d$, we have the following estimates:

$$\begin{aligned} L(x, s, \nu) + \langle A^\star[P](x, s), \nu\rangle &\geq \frac{1}{C}|\nu|^2 - \|a\|_{L^\infty(Q;\mathbb{R}^{k\times d})}|P(s)||\nu| - C \\ &\geq \frac{1}{C}(|\nu|^2 - |P(s)|^2 - 1) \geq \frac{1}{C}(|\nu|^2 - 1). \end{aligned}$$

Now we show that $\boldsymbol{u}[\gamma, P]$ is bounded in $L^\infty(Q)$. For any $(x, t) \in Q$, using the above lower bound for the running cost L, the bound $\|\gamma\|_{L^\infty(Q)} \leq R$, together with Assumption (H4) on the terminal cost g, we obtain that

$$\boldsymbol{u}[\gamma, P](x, t) \geq -C(R).$$

In addition, using Assumption (H3) and the fact that that $\|\gamma\|_{L^\infty(Q)} \leq R$, we deduce that

$$\boldsymbol{u}[\gamma, P](x, t) \leq J[\gamma, P](x, t, 0) \leq C(R),$$

from which we conclude that $\|\boldsymbol{u}[\gamma, P]\|_{L^\infty(Q)} \leq C(R)$.

Finally we show equation (19). Let $t \in [0, T]$, let $\varepsilon \in (0, 1)$ and let $\tilde{\nu} \in \mathbb{L}^2_{\mathbb{F}}(t, T)$ be an ε-optimal process. Since g is bounded (Assumption (H4)) and since $(\gamma, P) \in \Xi_R$, we deduce from the above inequality that

$$\begin{aligned} \mathbb{E}\Big[\int_t^T |\tilde{\nu}_s|^2 \mathrm{d}s\Big] &\leq C\Big(\inf_{\nu \in \mathbb{L}^2_{\mathbb{F}}(t,T)} J[\gamma, P](x, t, \nu) + \varepsilon + 1\Big) \\ &\leq C\big(\boldsymbol{u}[\gamma, P](x, t) + 2\big) \leq C, \end{aligned}$$

where the constant C does not depend on t and ε. Thus any ε-optimal process lies in $\mathbb{L}^{2,C}_{\mathbb{F}}(t, T)$, which concludes the proof. □

Using again the interpretation of $\boldsymbol{u}[\gamma, P]$ as the value function of an optimal control problem, we can prove the stability property stated in Lemma 18.

Proof of Lemma 18. Let (γ_1, P_1) and (γ_2, P_2) be in Ξ_R. Let $u_1 = \boldsymbol{u}[\gamma_1, P_1]$ and $u_2 = \boldsymbol{u}[\gamma_2, P_2]$. By Lemma 25, there exists $C > 0$ such that

$$u_2(x,t) - u_1(x,t) = \inf_{\mathbb{L}^{2,C}_{\mathbb{F}}(t,T)} J[\gamma_2, P_2](x,t,\cdot) - \inf_{\mathbb{L}^{2,C}_{\mathbb{F}}(t,T)} J[\gamma_1, P_1](x,t,\cdot),$$

for any $(x,t) \in Q$. We denote $(X^\nu_s)_{s\in[t,T]}$ the solution to the stochastic differential equation $\mathrm{d}X_s = \nu_s \mathrm{d}s + \sqrt{2}\mathrm{d}B_s$ with $X^\nu_t = x$, for any $\nu \in \mathbb{L}^2_{\mathbb{F}}(t,T)$. Then

$$\begin{aligned}|u_2(x,t) - u_1(x,t)| &\leq \sup_{\nu\in\mathbb{L}^{2,C}_{\mathbb{F}}(t,T)} \big|J[\gamma_2,P_2](x,t,\nu) - J[\gamma_1,P_1](x,t,\nu)\big| \\ &\leq \sup_{\nu\in\mathbb{L}^{2,C}_{\mathbb{F}}(t,T)} \mathbb{E}\Big[\int_t^T |\langle A^\star[P_2 - P_1](X^\nu_s, s), \nu_s\rangle| + |(\gamma_2 - \gamma_1)(X^\nu_s, s)|\,\mathrm{d}s\Big].\end{aligned}$$

For any $(x,s) \in Q$ and $\nu \in \mathbb{R}^d$, the Cauchy-Schwarz inequality yields

$$\begin{aligned}|\langle A^\star[P_2 - P_1](x,s), \nu\rangle| &\leq |\langle a(x,s)P_2(s) - P_1(s)||\nu| \\ &\leq \|a\|_{L^\infty(Q;\mathbb{R}^{k\times d})}|P_2(s) - P_1(s)|\,|\nu|.\end{aligned}$$

Using again Cauchy-Schwartz inequality and $\|a\|_{L^\infty(Q;\mathbb{R}^{k\times d})} \leq C$, we finally conclude that

$$|u_2(x,t) - u_1(x,t)| \leq C\Big(\|P_2 - P_1\|_{L^2(0,T;\mathbb{R}^k)} + \|\gamma_2 - \gamma_1\|_{L^\infty(Q)}\Big),$$

as was to be proved. □

We prove Proposition 19 with a density argument. In a nutshell: we prove in Proposition 26 below that the result of Proposition 19 holds true when γ and P are Hölder continuous. Then we pass to the limit, using Lemma 18.

Proposition 26. *Let $R > 0$ and let $\beta \in (0,1)$. For any $(\gamma, P) \in \Xi_R \cap \mathcal{C}^\beta(Q) \times \mathcal{C}^\beta(0,T;\mathbb{R}^k)$, the viscosity solution to the Hamilton-Jacobi-Bellman equation (3) is a classical solution. Moreover, there exists $\alpha \in (0,1)$ such that $\boldsymbol{u}[\gamma,P]$ lies in $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ and there exists a constant $C(R)$, depending only on R, such that $\|\boldsymbol{u}[\gamma,P]\|_{W^{2,1,q}(Q)} \leq C$.*

The proof of Proposition 26 is given at page 20 and relies on a fixed point approach which requires some preparatory work. We introduce the map $\mathcal{T}\colon W^{2,1,q}(Q)\times[0,1] \to W^{2,1,q}(Q)$ which associates to any $u \in W^{2,1,q}(Q)$ and $\tau \in [0,1]$ the classical solution $\tilde{u} = \mathcal{T}[u,\tau]$ to the linear parabolic equation

$$\begin{aligned}-\partial_t\tilde{u} - \Delta\tilde{u} + \tau\boldsymbol{H}[\nabla u + A^\star P] &= \tau\gamma \qquad (x,t) \in Q, \\ \tilde{u}(x,T) &= \tau g(x) \qquad x \in \mathbb{T}^d.\end{aligned}$$

For any $(u,\tau) \in W^{2,1,q}(Q)\times[0,1]$, we have $\tau(\gamma - \boldsymbol{H}[\nabla u + A^\star P]) \in L^\infty(Q)$, by Lemma 3 and Lemma 1. It follows then from Theorem 22 that $\mathcal{T}[u,\tau]$ lies in $W^{2,1,q}(Q)$, proving that $\mathcal{T}$ is well-defined.

Lemma 27. *The mapping $\mathcal{T}$ is continuous and compact. In addition, for all $K > 0$, there exists $\alpha \in (0,1)$ and $C > 0$ depending on K, γ, and P such that $\|u\|_{W^{2,1,q}(Q)} \leq K$ implies $\|\mathcal{T}[u,\tau]\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)} \leq C$.*

Proof. Step 1: Continuity of $\mathcal{T}$. Let $(u_k, \tau_k) \in W^{2,1,q}(Q) \times [0,1]$ be a sequence converging to $(u,\tau) \in W^{2,1,q}(Q) \times [0,1]$. Then $\nabla u_k \to \nabla u$ in $L^\infty(Q;\mathbb{R}^d)$ by Lemma 1. Then $\tau_k(\gamma - \boldsymbol{H}[\nabla u_k + A^\star P]) \to \tau(\gamma - \boldsymbol{H}[\nabla u + A^\star P])$ in $L^\infty(Q;\mathbb{R}^d)$ by continuity of the Hamiltonian (see Lemma 3). Finally the continuity of $\mathcal{T}$ follows from Theorem 23.
Step 2: Compactness of $\mathcal{T}$. Let $K > 0$ and let $(u,\tau) \in W^{2,1,q}(Q) \times [0,1]$ be such that $\|u\|_{W^{2,1,q}(Q)} \leq K$. Combining Lemma 1 and Lemma 3 there exist $\alpha \in (0,1)$ and $C > 0$ such that $\|\gamma - \boldsymbol{H}[\nabla u + A^\star P]\|_{C^\alpha(Q)} \leq C$. Then applying Theorem 24, there exist $\alpha \in (0,1)$ and $C > 0$ such that $\|\mathcal{T}[u,\tau]\|_{\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)} \leq C$. By the Arzela-Ascoli Theorem the centered ball of $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ of radius $C > 0$ is a relatively compact subset of $W^{2,1,q}(Q)$. As a consequence $\mathcal{T}[u,\tau]$ is a compact mapping and the conclusion follows. □

Theorem 28. (Leray-Schauder) *Let X be a Banach space and let $T : X \times [0,1] \to X$ be a continuous and compact mapping. Assume that $T(x,0) = 0$ for all $x \in X$ and assume there exists $C > 0$ such that $\|x\|_X < C$ for all $(x,\tau) \in X \times [0,1]$ such that $T(x,\tau) = x$. Then, there exists $x \in X$ such that $T(x,1) = x$.*

Proof. See [36, Theorem 11.6]. □

Proof of Proposition 26. We prove that under the assumptions of the proposition, the HJB equation has a classical solution in $\mathcal{C}^{2+\alpha,1+\alpha/2}(Q)$ (for some $\alpha \in (0,1)$), which is then necessarily the unique viscosity solution $\boldsymbol{u}[\gamma,P]$. To this purpose, we prove the existence of a solution to the fixed point equation $u = \mathcal{T}[u,1]$. We have $\mathcal{T}[u,0] = 0$ for all $u \in W^{2,1,q}(Q)$. Now let $(u,\tau) \in W^{2,1,q}(Q) \times [0,1]$ be such that $\mathcal{T}[u,\tau] = u$. From Lemma 27, the mapping $\mathcal{T}$ is continuous and compact, in addition u is a classical solution and thus the viscosity solution to the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\partial_t u - \Delta u + \tau \boldsymbol{H}[\nabla u + A^\star P] &= \tau\gamma \qquad (x,t) \in Q, \\ u(x,T) &= \tau g(x) \qquad x \in \mathbb{T}^d, \end{aligned}$$

and can be interpreted as the value function associated to the following stochastic control problem

$$\inf_{\nu \in \mathbb{L}^2_{\mathbb{F}}(0,T)} \tau \mathbb{E}\Big[\int_0^T L(X_s^\tau, s, \nu_s) + \langle A^\star[P](X_s^\tau, s), \nu_s \rangle + \gamma(X_s^\tau, s) \mathrm{d}s + g(X_T^\tau) \Big],$$

where $(X_s^\tau)_{s \in [t,T]}$ is the solution to $\mathrm{d}X_s = \tau \nu_s \mathrm{d}s + \sqrt{2}\mathrm{d}B_s$, $X_0 = Y$. Following [9, Proposition 1, Step 2], there exists a constant $C > 0$, depending only on R, such that $\|u\|_{L^\infty(Q)} + \|\nabla u\|_{L^\infty(Q;\mathbb{R}^d)} \leq C$. Then using Lemma 3 and recalling that $(\gamma, P) \in \Xi_R$, we deduce that $\|\boldsymbol{H}[\nabla u + A^\star P] - \gamma\|_{L^\infty(Q)} \leq C$. It follows that u is the solution to a parabolic PDE with bounded coefficients and thus $\|u\|_{W^{2,1,q}(Q)} \leq C$, by Theorem 22. Again, C only depends on R. Finally, by the Leray-Schauder theorem (Theorem 28), there exists a solution to $u = \mathcal{T}[u,1]$, which is necessarily equal to $\boldsymbol{u}[\gamma,P]$. □

Proof of Proposition 19. Take $(\gamma,P) \in \Xi_R$ and fix $\beta \in (0,1)$. Let (γ_n, P_n) be a sequence in $\Xi_{R+1} \cap \mathcal{C}^\beta(Q) \times \mathcal{C}^\beta(0,T;\mathbb{R}^k)$ such that

$$\|\gamma_n - \gamma\|_{L^\infty(Q)} \longrightarrow 0 \quad \text{and} \quad \|P_n - P\|_{L^2(0,T;\mathbb{R}^k)} \longrightarrow 0.$$

We do not detail the construction of such a sequence, this can be done by convolution. Define $u^n = \boldsymbol{u}[\gamma^n, P^n]$ and $u = \boldsymbol{u}[\gamma, P]$. By Lemma 18, $u_n \to u$ for the L^∞-norm. Moreover, by Proposition 26,

$$\|u^n\|_{W^{2,1,q}(Q)} \leq C(R), \quad \forall n \in \mathbb{N}. \tag{21}$$

Thus, the three sequences $(\partial_t u^n)_{n\in\mathbb{N}}$, $(\Delta u^n)_{n\in\mathbb{N}}$, and $(\nabla u^n)_{n\in\mathbb{N}}$ are bounded in $L^q(Q)$. By the Banach-Alaoglu theorem, the three sequences have at least one accumulation point for the weak topology of $L^q(Q)$. These three accumulation points are necessarily (by definition of weak derivatives) equal to $\partial_t u$, Δu, and ∇u, respectively. Since the L^q-norm is weakly lower semi-continuous, we deduce that $\|u\|_{W^{2,1,q}(Q)} \leq C(R)$, where $C(R)$ is as in (21). This concludes the proof. □

4.5 The other mappings

Proof of Lemma 20. Let $(\gamma, P) \in \Xi_R$. Let $u = \boldsymbol{u}[\gamma, P]$. We already know from Proposition 19 that $\|u\|_{W^{2,1,q}(Q)} \leq C(R)$. Then Lemma 1 implies that u and ∇u are continuous and that $\|u\|_{L^\infty(Q)} \leq C(R)$ and $\|\nabla u\|_{L^\infty(Q;\mathbb{R}^d)} \leq C(R)$. Let $v = \boldsymbol{v}[\gamma, P] = -\boldsymbol{H}_p[\nabla u + A^\star P]$. We have

$$D_x v = -\boldsymbol{H}_{px}[\nabla u + A^\star P] - \boldsymbol{H}_{pp}[\nabla u + A^\star P](D^2_{xx} u + D_x A^\star P).$$

Using the regularity of u, the regularity properties of the Hamiltonian given in Lemma 30, and the regularity assumptions on a (Assumption (H4)), we deduce that $\|v\|_{L^\infty(Q;\mathbb{R}^d)} \leq C(R)$ and that $\|D_x v\|_{L^q(Q;\mathbb{R}^{d\times d})} \leq C(R)$. Moreover, v is continuous.

Next, let $m = \boldsymbol{m}[\gamma, P] = \boldsymbol{M}[v]$. A direct application of Lemma 16 yields that $\|m\|_{W^{2,1,q}(Q)} \leq C$.

Finally, let $w = \boldsymbol{w}[\gamma, P] = mv$. Using again Lemma 1, we obtain that m is continuous and that $\|m\|_{L^\infty(Q)} \leq C(R)$ and $\|\nabla m\|_{L^\infty(Q;\mathbb{R}^d)} \leq C(R)$. We deduce that $w \in \Theta$, with a norm bounded by some constant $C(R)$. The lemma is proved. □

Proof of Lemma 21. The two statements concerning $\boldsymbol{\gamma}$ are directly deduced from Assumptions (H3) and (H5). Let $w \in \Theta$. Recalling the definition of the operator A (page 6), it is easy to see with Assumption (H4) that $Aw \in \mathcal{C}(0, T; \mathbb{R}^d)$. Assumptions (H3) and (H5) ensure then that $\boldsymbol{P}[w] = \boldsymbol{\phi}\big[A[w]\big]$ lies in $\mathcal{C}(0, T; \mathbb{R}^k)$ and that $\|\boldsymbol{P}[w]\|_{L^\infty(0,T;\mathbb{R}^k)} \leq C$.

Let us next consider w_1 and w_2 in Θ. We have

$$\begin{aligned}\|Aw_2 - Aw_1\|_{L^\infty(0,T;\mathbb{R}^k)} &\leq \|a\|_{L^\infty(Q;\mathbb{R}^{k\times d})} \|w_2 - w_1\|_{L^\infty(0,T;L^1(\mathbb{T}^d;\mathbb{R}^d))} \\ &\leq C\|w_2 - w_1\|_{L^2(Q;\mathbb{R}^d)},\end{aligned}$$

by Assumption (H4). Using next the Lipschitz-continuity of ϕ (Assumption (H5)), we obtain that

$$\|\boldsymbol{\phi}[Aw_2] - \boldsymbol{\phi}[Aw_1]\|_{L^2(0,T;\mathbb{R}^k)} \leq C\|w_2 - w_1\|_{L^2(Q;\mathbb{R}^d)},$$

as was to be proved. □

5 Stability results for solutions to stochastic optimal control problems

The main result of this section, Proposition 31, shows that any approximate solution to Problem ($\mathcal{P}[\gamma, P]$) is close to its solution, for suitable norms. This is a key result for achieving linear convergence in the GCG method.

Lemma 29. *Let $(\hat{\gamma}, \hat{P}) \in \Xi$. Let $\hat{m} = \boldsymbol{m}[\hat{\gamma}, \hat{P}]$, $\hat{v} = \boldsymbol{v}[\hat{\gamma}, \hat{P}]$ and let $\hat{w} = \boldsymbol{w}[\hat{\gamma}, \hat{P}]$. There exists a constant $C > 0$ such that for any $(m, w) \in \mathcal{R}$, the following holds:*

$$\mathcal{Z}[\hat{\gamma}, \hat{P}](m, w) - \mathcal{Z}[\hat{\gamma}, \hat{P}](\hat{m}, \hat{w}) \geq \frac{1}{C} \int_Q |v(x,t) - \hat{v}(x,t)|^2 m(x,t) \,\mathrm{d}x \,\mathrm{d}t,$$

where $v \in L^\infty(Q; \mathbb{R}^d)$ is such that $w = mv$.

The proof of Lemma 29 relies on the following inequality.

Lemma 30. *Let $(x,t) \in Q$. Let $v \in \mathbb{R}^d$, $\hat{p} \in \mathbb{R}^d$, $m \geq 0$, and $\hat{m} \geq 0$. Let $\hat{v} = -H_p(x,t,\hat{p})$. Let $C > 0$ be such that $L(x,t,\cdot)$ is strongly convex with modulus $1/C$. Then,*

$$L(x,t,v)m - L(x,t,\hat{v})\hat{m} \geq -H(x,t,p)(m - \hat{m}) - \langle \hat{p}, w - \hat{w} \rangle + \frac{1}{C}|v - \hat{v}|^2 m.$$

Proof. See [9, Proof of Proposition 2]. □

Proof of Lemma 29. Using the definition of $\mathcal{Z}[\hat{\gamma}, \hat{P}]$, we have

$$\begin{aligned}\mathcal{Z}[\hat{\gamma}, \hat{P}](m, w) - \mathcal{Z}[\hat{\gamma}, \hat{P}](\hat{m}, \hat{w}) &= \int_Q \big(L[v]m - L[\hat{v}]\hat{m}\big) \,\mathrm{d}x \,\mathrm{d}t \\ &+ \int_Q \hat{\gamma}(m - \hat{m}) \,\mathrm{d}x \,\mathrm{d}t + \int_Q \langle A^* P, w - \hat{w} \rangle \,\mathrm{d}x \,\mathrm{d}t + \int_{\mathbb{T}^d} g(m(T) - \hat{m}(T)) \,\mathrm{d}x. \quad (22)\end{aligned}$$

We set $\hat{u} = \boldsymbol{u}[\hat{\gamma}, \hat{P}]$. By definition of $\hat{v}$, we have $\hat{v} = -\boldsymbol{H}_p[\nabla \hat{u} + A^\star \hat{P}]$. Applying Lemma 30 with $\hat{p} = \nabla \hat{u} + A^* \hat{P}$, we obtain

$$\begin{aligned}\int_Q \big(L[v]m - L[\hat{v}]\hat{m}\big) \,\mathrm{d}x \,\mathrm{d}t \geq &- \int_Q H[\nabla \hat{u} + A^* \hat{P}](m - \bar{m}) \,\mathrm{d}x \,\mathrm{d}t \\ &- \int_Q \langle \nabla \hat{u} + A^* \hat{P}, w - \bar{w} \rangle \,\mathrm{d}x \,\mathrm{d}t + \frac{1}{C} \int_Q |v - \hat{v}|^2 m \,\mathrm{d}x \,\mathrm{d}t.\end{aligned}$$

We inject the obtained inequality into (22) and we use $\boldsymbol{H}[\nabla \hat{u} + A^* \hat{P}] + \hat{\gamma} = -\partial_t \hat{u} - \Delta \hat{u}$. This yields

$$\begin{aligned}\mathcal{Z}[\hat{\gamma}, \hat{P}](m, w) - \mathcal{Z}[\hat{\gamma}, \hat{P}](\hat{m}, \hat{w}) \geq &\int_Q \big(- \partial_t \hat{u} - \Delta \hat{u}\big)\big(m - \hat{m}\big) \,\mathrm{d}x \,\mathrm{d}t \\ &+ \int_Q \langle -\nabla \hat{u}, w - \hat{w} \rangle \,\mathrm{d}x \,\mathrm{d}t + \int_{\mathbb{T}^d} g(m(T) - \hat{m}(T)) \,\mathrm{d}x + \frac{1}{C} \int_Q |v - \hat{v}|^2 m \,\mathrm{d}x \,\mathrm{d}t.\end{aligned}$$

The first three integrals in the right-hand side cancel out: this can be shown by doing an integration by parts and by using the Fokker-Planck equation satisfied by m and $\hat{m}$. This concludes the proof. □

Proposition 31. *Let $R > 0$ and let $(\hat\gamma, \hat P) \in \Xi_R$. Let $\hat m = \boldsymbol{m}[\hat\gamma, \hat P]$ and let $\hat w = \boldsymbol{w}[\hat\gamma, \hat P]$. Let $(m, w) \in \mathcal{R}$ be such that $\|m\|_{L^\infty(Q)} \le R$. There exists a constant $C(R)$ such that for any $(m, w) \in \mathcal{R}$,*

$$\begin{aligned}\|m - \hat m\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} &\le C(R)\sqrt{\sigma},\\ \|w - \hat w\|_{L^2(Q;\mathbb{R}^d)} &\le C(R)(\sqrt{\sigma} + \sigma),\end{aligned}$$

where $\sigma = \mathcal{Z}[\hat\gamma, \hat P](m, w) - \mathcal{Z}[\hat\gamma, \hat P](\hat m, \hat w)$.

As a consequence of Proposition 31, $(\boldsymbol{m}[\hat\gamma, \hat P], \boldsymbol{w}[\hat\gamma, \hat P])$ is the unique solution to Problem $(\mathcal{P}[\gamma, P])$, with $(\gamma, P) = (\hat\gamma, \hat P)$.

Proof of Proposition 31. Lemma 30 yields $\int_Q |v - \hat v|^2 m \,\mathrm{d}x\,\mathrm{d}t \le C\sigma$. By Lemma 20, we know that $\|\hat v\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \le C$, for some constant depending on R. Applying next Lemma 17, we obtain the estimate of $\|m - \hat m\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}$. The estimate of $\|w - \hat w\|_{L^2(Q;\mathbb{R}^d)}$ follows directly from $w - \hat w = m(v - \hat v) + (m - \hat m)\hat v$. The proposition is proved. □

Corollary 32. *Let $R > 0$. Let $(m, w) \in \mathcal{R}$ be such that $\|m\|_{L^\infty(Q)} \le R$. There exists a constant $C(R)$ such that*

$$\begin{aligned}\|m - \bar m\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} &\le C(R)\sqrt{\varepsilon}\\ \|w - \bar w\|_{L^2(Q;\mathbb{R}^d)} &\le C(R)(\sqrt{\varepsilon} + \varepsilon),\end{aligned}$$

where $\varepsilon = \mathcal{J}(m, w) - \mathcal{J}(\bar m, \bar w)$.

Proof. By definition, $\bar m = \boldsymbol{m}[\bar\gamma, \bar P]$ and $\bar w = \boldsymbol{w}[\bar\gamma, \bar P]$. We obtain the announced estimates by combining Corollary 9 and Proposition 31. □

6 Convergence results

6.1 General results

Proposition 33. *Algorithm 1 is well-posed. It generates sequences in the following sets: $(\bar m_k, \bar w_k) \in \mathcal{R}$, $(\gamma_k, P_k) \in \Xi$, $(u_k, v_k) \in W^{2,1,q}(Q) \times \Theta$, $(m_k, w_k) \in \mathcal{R}$, for all $k \in \mathbb{N}$. Moreover, there exists a constant $C > 0$ such that $(\gamma_k, P_k) \in \Xi_C$ and such that*

$$\begin{aligned}\|u_k\|_{W^{2,1,q}(Q)} + \|v_k\|_\Theta &\le C,\\ \|m_k\|_{W^{2,1,q}(Q)} + \|w_k\|_\Theta &\le C,\\ \|\bar m_k\|_{W^{2,1,q}(Q)} + \|\bar w_k\|_\Theta &\le C,\end{aligned}$$

for all $k \in \mathbb{N}$. Finally, there exists a constant $C > 0$ such that $\varepsilon_k \le C$, for all $k \in \mathbb{N}$.

Proof. The well-posedness of the algorithm and the bounds on $\bar m_k$, $\bar w_k$, γ_k, P_k, u_k, v_k, m_k, and w_k can easily be established by induction and by using Proposition 19, Lemma 20, and Lemma 21. Lemma 5 shows that (m_k, w_k) lies in $\mathcal{R}$, for all $k \in \mathbb{N}$. Using the convexity of $\mathcal{R}$ given in Lemma 5, we deduce by induction that

$(\bar m_k, \bar w_k) \in \mathcal{R}$, for all $k \in \mathbb{N}$. The bounds on m_k and v_k imply the existence of a constant C such $\mathcal{J}(m_k, w_k) \le C$, since

$$\mathcal{J}(m_k, w_k) = \int_Q \boldsymbol{L}[v_k] m_k \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \big(\boldsymbol{F}[m_k] + \boldsymbol{\Phi}[A(m_k v_k)]\big)\mathrm{d}t + \int_{\mathbb{T}^d} g m_k(T)\,\mathrm{d}x.$$

The boundedness of $\mathcal{J}(\bar m_k, \bar w_k)$ follows then by induction, since $\mathcal{J}$ is convex and thus

$$\mathcal{J}(\bar m_{k+1}, \bar w_{k+1}) \le (1-\delta_k)\mathcal{J}(\bar m_k, \bar w_k) + \delta_k \mathcal{J}(m_k, w_k).$$

The proposition is proved. □

Lemma 34. *For all $k \in \mathbb{N}$, it holds that $\varepsilon_k \le \sigma_k$.*

Proof. By Corollary 9, we have for all $(m, w) \in \mathcal{R}$

$$\mathcal{Z}[\gamma_k, P_k](\bar m_k, \bar w_k) - \mathcal{Z}[\gamma_k, P_k](m, w) \le \mathcal{J}(\bar m_k, \bar w_k) - \mathcal{J}(m, w) \le \varepsilon_k.$$

By definition, σ_k is the supremum of the left-hand side with respect to (m, w). The conclusion follows immediately. □

Theorem 35. *There exists $C > 0$ such that for all $k \in \mathbb{N}$,*

$$\begin{aligned}
\|\bar m_k - \bar m\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} + \|\bar w_k - \bar w\|_{L^2(Q;\mathbb{R}^d)} &\le C\sqrt{\varepsilon_k},\\
\|\gamma_k - \bar\gamma\|_{L^\infty(Q)} + \|P_k - \bar P\|_{L^\infty(0,T;\mathbb{R}^k)} &\le C\sqrt{\varepsilon_k},\\
\|u_k - \bar u\|_{L^\infty(Q)} &\le C\sqrt{\varepsilon_k},\\
\|m_k - \bar m_k\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} + \|w_k - \bar w_k\|_{L^2(Q;\mathbb{R}^d)} &\le C\sqrt{\sigma_k}.
\end{aligned}$$

Proof. The first estimate on $\|\bar m_k - \bar m\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}$ and $\|\bar w_k - \bar w\|_{L^2(Q;\mathbb{R}^d)}$ is obtained by combining Corollary 32, the boundedness of ε_k, and the boundedness of $\|m_k\|_{L^\infty(Q)}$. The second estimate on $\|\gamma_k - \bar\gamma\|_{L^\infty(Q)}$ and $\|P_k - \bar P\|_{L^\infty(0,T;\mathbb{R}^k)}$ is obtained by combining the first estimate and Lemma 21. The third estimate on $\|u_k - \bar u\|_{L^\infty(Q)}$ follows from the second one and from Lemma 18. Using Proposition 33 and Lemma 1, there exists $C > 0$ such that $(\gamma_k, P_k) \in \Xi_C$ and $\|\bar m_k\|_{L^\infty(Q)} \le C$, for all $k \in \mathbb{N}$. Moreover, by construction, $m_k = \boldsymbol{m}[\gamma_k, P_k]$ and $w_k = \boldsymbol{w}[\gamma_k, P_k]$ and the pair $(\bar m_k, \bar w_k)$ is σ_k-optimal for the minimization problem of $\mathcal{Z}[\gamma_k, P_k](\cdot)$. Therefore, Proposition 31 applies and yields the last estimate that was to be proved. □

Corollary 36. *Assume that $\varepsilon_k \to 0$. Then v_k uniformly converges to $\bar v$.*

Proof. If $\varepsilon_k \to 0$, then u_k converges to $\bar u$. We know by Proposition 33 that u_k is bounded in $W^{2,1,q}(Q)$; as a consequence of Lemma 1, ∇u_k is bounded in $\mathcal{C}^\alpha(Q;\mathbb{R}^d)$. By Arzela-Ascoli theorem, any subsequence of $(\nabla u_k)_{k\in\mathbb{N}}$ has at least one accumulation point for the uniform topology, which is then necessarily $\nabla \bar u$. Therefore, ∇u_k converges to $\bar u$ for the uniform tolopology. Recalling that $v_k = -\boldsymbol{H}_p[\nabla u_k + A^\star P_k]$, using the convergence of P_k already established, and using the regularity properties of the Hamiltonian, we obtain the announced result. □

Lemma 37. *Let $C_1 > 0$ and $C_2 > 0$ denote the Lipschitz constants of f and ϕ (see Assumption (H5)). Then for any $\delta \in [0,1]$, it holds that*

$$\mathcal{J}(\bar m_k^\delta, \bar w_k^\delta) \le \mathcal{J}(\bar m_k, \bar w_k) - \delta\sigma_k + \big(C_1 D_k^{(1)} + C_2 D_k^{(2)}\big)\delta^2, \tag{23}$$

where $(\bar{m}_k^\delta, \bar{w}_k^\delta)$ *is defined by* (9) *and where* $D_k^{(1)}$ *and* $D_k^{(2)}$ *are defined by* (14)*. Moreover, there exists a constant* $C > 0$ *such that*

$$\mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta) \leq \mathcal{J}(\bar{m}_k, \bar{w}_k) - \delta\sigma_k + C\sigma_k\delta_k^2. \tag{24}$$

Recall that the terms $D_k^{(1)}$ and $D_k^{(2)}$ represent the distance of the current approximate solution to the solution of the (partially) linearized problem. In the classical proof of convergence of the Frank-Wolfe algorithm, one writes a similar estimate to (23), where $D_k^{(1)}$ and $D_k^{(2)}$ are simply bounded by some constant. In order to achieve linear convergence (instead of the classical sublinear rate of convergence), it is crucial to keep these terms and to estimate them with the exploitability σ_k, as will become clear in the next section. This proof technique is largely inspired by [47, Section 4.3].

Proof of Lemma 37. The proof relies on the decomposition $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ introduced (6). First, by convexity of $\tilde{L}$, we have

$$\mathcal{J}_1(\bar{m}_k^\delta, \bar{w}_k^\delta) - \mathcal{J}_1(\bar{m}_k, \bar{w}_k) \leq \delta\big(\mathcal{J}_1(m_k, w_k) - \mathcal{J}_1(\bar{m}_k, \bar{w}_k)\big). \tag{25}$$

Next, using inequality (8) of Lemma 8, we obtain that

$$\begin{aligned}\mathcal{J}_2(\bar{m}_k^\delta, \bar{w}_k^\delta) - \mathcal{J}_2(\bar{m}_k, \bar{w}_k) &\leq \delta \int_Q \gamma_k(m_k - \bar{m}_k)\,\mathrm{d}x\,\mathrm{d}t \\ &+ \delta \int_0^T \langle A[w_k - \bar{w}_k], P_k\rangle\,\mathrm{d}t + C\delta^2\Big(\|m_k - \bar{m}_k\|^2_{L^2(Q)} + \|w_k - \bar{w}_k\|^2_{L^2(Q;\mathbb{R}^d)}\Big).\end{aligned} \tag{26}$$

By definition of $\mathcal{Z}[\gamma_k, P_k]$, we have

$$\begin{aligned}\sigma_k &= \mathcal{Z}[\gamma_k, P_k](\bar{m}_k, \bar{w}_k) - \mathcal{Z}[\gamma_k, P_k](m_k, w_k) \\ &= \mathcal{J}_1(m_k, w_k) - \mathcal{J}_1(\bar{m}_k, \bar{w}_k) \\ &\qquad + \int_Q \gamma_k(m_k - \bar{m}_k)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \langle A[w_k - \bar{w}_k], P_k\rangle\,\mathrm{d}t.\end{aligned} \tag{27}$$

Finally, by Theorem 35 we have

$$\|m_k - \bar{m}_k\|^2_{L^\infty(0,T;L^2(\mathbb{T}^d))} + \|w_k - \bar{w}_k\|^2_{L^2(Q;\mathbb{R}^d)} \leq C\sigma_k. \tag{28}$$

Summing up (25) and (26) and combining the result with (27) and (28), we obtain the announced result. □

Lemma 38. *There exists* $C > 0$ *such that* $\sigma_k \leq C\varepsilon_k$*, for all* $k \in \mathbb{N}$*.*

Proof. For any $\delta \in [0, 1]$, Lemma 37 yields $\mathcal{J}(\bar{m}, \bar{w}) \leq \mathcal{J}(\bar{m}_k, \bar{w}_k) - \delta\sigma_k + C\delta^2\sigma_k$. It follows that

$$0 \leq \varepsilon_k - (1 - C\delta)\delta\sigma_k. \tag{29}$$

Increasing if necessary the value of C in the above inequality, we can assume that $C \geq 1/2$. Taking $\delta = 1/(2C)$, we conclude that $0 \leq \varepsilon - 1/(4C)\sigma_k$, as was to be proved. □

6.2 Proof of the main convergence result

We prove here Theorem 10. We start with the proof of the second part of the theorem, which is almost direct. Combining the upper bound of the cost function proved in Lemma 37 (inequality (24)) with the inequality $\varepsilon_k \leq \sigma_k$ and the bound on the exploitability obtained in Lemma 38, we obtain the following inequality:

$$\varepsilon_{j+1} \leq \big(1 - \delta_j + C\delta_j^2\big)\varepsilon_j, \quad \forall j \in \mathbb{N}.$$

Here the constant C is independent of the choice of stepsize. Multiplying the obtained inequalities for $j = 0, \dots, k$, we obtain

$$\varepsilon_{j+1} \leq \varepsilon_0 \exp\Big(\sum_{j=0}^{k} \ln\big(1 - \delta_j + C\delta_j^2\big)\Big).$$

Using next the inequality $\ln(x) \leq x - 1$, satisfied for any $x > 0$, we obtain the desired result, inequality (15).

Let us consider the case of adaptive stepsizes. Let us fix the iteration number k. It suffices to show that for δ_k satisfying either (11), (12), or (13), there exists a constant $\beta \in (0, 1)$, independent of k, such that $\varepsilon_{k+1} \leq \beta\varepsilon_k$.

Case of the QAG condition. The main idea is to show that the QAG condition is satisfied when δ is smaller than a certain threshold, which is independent of k. Let $\delta > 0$ be such that the condition is not satisfied. Then, using Lemma 37, we obtain that

$$\mathcal{J}(\bar{m}, \bar{w}) - c\delta\sigma_k < \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta) \leq \mathcal{J}(\bar{m}, \bar{w}) - \delta\sigma_k + C\delta^2\sigma_k.$$

Re-arranging, we deduce that $\delta > \bar{\delta} := \frac{(1-c)}{C} > 0$. By contraposition, the QAG condition holds true for any δ such that $\delta \leq \bar{\delta}$.

We can now prove that i_k is finite and uniformly bounded. Two cases can be considered. If $\bar{\delta} \geq 1$, then $i_k = 0$ and $\delta_k = 1$. Otherwise, for any $j \in \mathbb{N}$, we have

$$\tau^j \leq \bar{\delta} \iff j \geq \frac{\ln(\bar{\delta})}{\ln(\tau)}.$$

Therefore, $i_k \leq \lceil \ln(\bar{\delta})/\ln(\tau) \rceil$. If $i_k = 0$, then $\delta_k = 1$. Otherwise, if $i_k > 0$, then $\tau^{i_k - 1}$ does not satisfy the QAG condition and therefore, $\tau^{i_k - i} \geq \bar{\delta}$ and thus $\delta_k = \tau^{i_k} \geq \tau\bar{\delta}$. So, in all cases, we have $\delta_k \geq \delta_{\min} := \min(1, \tau\bar{\delta})$. It follows that

$$\begin{aligned}
\varepsilon_{k+1} &= \mathcal{J}(\bar{m}_{k+1}, \bar{w}_{k+1}) - \mathcal{J}(\bar{m}, \bar{w}) \\
&\leq \big(\mathcal{J}(\bar{m}_k, \bar{w}_k) - \mathcal{J}(\bar{m}, \bar{w})\big) - c\sigma_k\delta_k \\
&\leq \varepsilon_k - c\delta_{\min}\sigma_k \leq (1 - c\delta_{\min})\varepsilon_k. \qquad (30)
\end{aligned}$$

The last inequality was obtained with Lemma 34.

Case of exploitability-based stepsizes. Let us set $a_k = \big(C_1 D_k^{(1)} + C_2 D_k^{(2)}\big)$. By definition, $\delta_k = \min\big(1, \frac{\sigma_k}{2a_k}\big)$. Assume that $\sigma_k \geq 2a_k$, i.e. $a_k \leq \sigma_k/2$. Then $\delta_k = 1$. Inequality (23) in Lemma 37 and Lemma 34 yield

$$\varepsilon_{k+1} \leq \varepsilon_k - \frac{\sigma_k}{2} \leq \frac{1}{2}\varepsilon_k.$$

Now assume that $\sigma_k < 2a_k$. Then $\delta_k = \frac{\sigma_k}{2a_k}$ and inequality (23) yields

$$\varepsilon_{k+1} \leq \varepsilon_k - \frac{\sigma_k^2}{4a_k}.$$

It follows from the last estimate of Theorem 35 that $a_k \leq C\sigma_k$. Therefore, using again Lemma 34,

$$\varepsilon_{k+1} \leq \varepsilon_k - \frac{\sigma_k}{4C} \leq \Big(1 - \frac{1}{4C}\Big)\varepsilon_k,$$

as was to be proved.

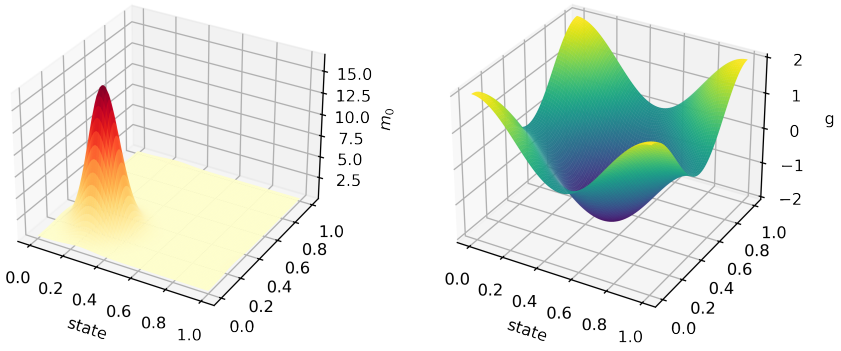
Case of an optimal size. If δ_k minimizes $\mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta)$, then (30) is necessarily satisfied. This concludes the proof of the theorem.

7 Numerical illustration

As a numerical illustration, we solve the mean field game system (MFG) with $f = 0$. In this situation, the agents only interacts through the law of their controls. The associated potential problem has the following form

$$\inf_{(m,w)\in\mathcal{R}} \int_Q \tilde{\boldsymbol{L}}[m,w]\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \boldsymbol{\Phi}[Aw]\,\mathrm{d}t + \int_{\mathbb{T}^d} gm(T)\,\mathrm{d}x.$$

Data and numerical scheme We take $d = 2$, $k = 2$ and $T = 1$ so that $Q = \mathbb{T}^2 \times [0,1]$. The initial measure m_0 is normally distributed on the torus (it is the product of two independent von Mises distributions centered at $1/4$ and is shown in Figure 1a). The terminal condition $g(x) = \sum_{i=1}^2 \cos(2\pi x_i)$ for any $x = (x_1, x_2) \in \mathbb{T}^2$ is shown in Figure 1b.



(a) Initial measure m_0.

(b) Terminal condition g.

We define $L(x,t,v) = \frac{1}{2}|v|^2$, $a(x,t) = \mathrm{Id}$ and $\phi(t,z) = 10z$. Obviously ϕ derives from the potential $\Phi(t,z) = 5(z_1^2 + z_2^2)$. In other words, we have a two-dimensional price variable, and the two price relations write, for $i = 1, 2$ as follows: $P_i = 10\int_{\mathbb{T}^d} v_i(x,t)m_i(x,t)\,\mathrm{d}x\,\mathrm{d}t$. For any control $\nu_t \in \mathbb{L}^2_{\mathbb{F}}(0,1)$, the cost function of a representative agent writes:

$$\mathbb{E}\Big[\int_0^T \frac{1}{2}|\nu_t|^2 + 10\,\langle P(t), \nu_t\rangle\,\mathrm{d}t + g(X_T^\nu)\Big].$$

In this numerical experiment we consider a volatility equal to 0.1 for the controlled stochastic state equation satisfied by X^ν. Algorithm 1 requires to compute the mappings $\boldsymbol{u}$ (i.e. a solution to the Hamilton-Jacobi-Bellman equation) and $\boldsymbol{M}$ (i.e. a solution to the Fokker-Planck equation) at each step. The resolution is done via an explicit finite difference scheme. In the following, we discretize Q with a uniform grid containing 10^2 points in space and 42 points in time.

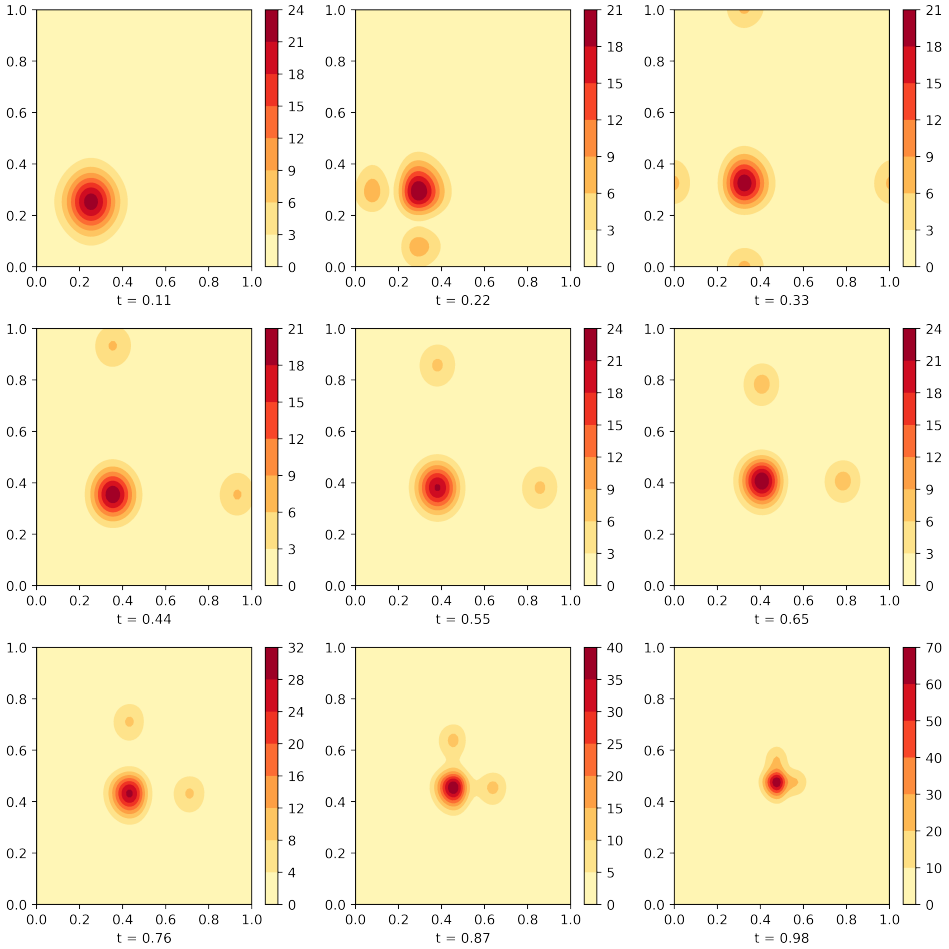


Figure 2: Density of players $\bar{m}$ evaluated at several time steps.

Interpretation and numerical solution When $P = 0$, the optimal trajectories of the agents look like slightly perturbed straight lines (with constant speed), ending in a close neighborhood of the point $(0.5, 0.5)$ (it would be a straight in the deterministic case without diffusion coefficient). In view of the initial condition, located in the "bottom right corner" of the square $[0, 1]^2$, the agents use in this case positive controls v_1 and v_2. However, when P_1 is positive, some agents may try to reach the point $(0.5, 0.5)$ using control with a first coordinate that is negative. Graphically speaking, these agent would cross the left vertical axis ($x_1 = 0$) and "jump" to the right vertical axis ($x_1 = 1$). This strategy is particularly interesting for agents an initial condition x such that x_1 is positive and close to zero. Of course the same reasoning is valid for P_2 positive: some agents would cross the horizontal axis.

The two equilibrium prices are positive, leading to four different kinds of optimal trajectories: those which do not cross any axis, those crossing only the vertical axis, those crossing only the horizontal axis, and those crossing both axes, as can be seen from the graphs of the equilibrium vector field $\bar{v}$ on Figure 3. Thus the initial distribution is split into four groups as shown by Figure 2. The group of agents crossing both axes is actually of very small mass, thus not visible on the graph.

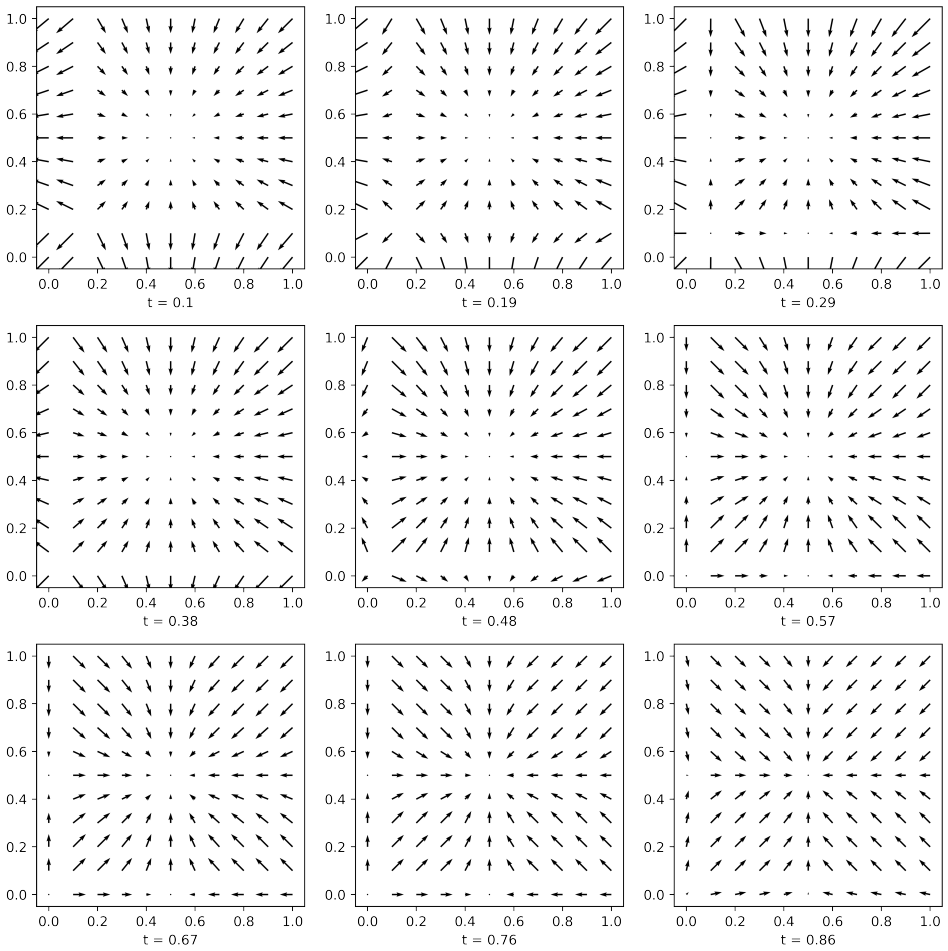


Figure 3: Vector field $\bar{v}$ evaluated at several time steps.

Convergence and execution time The convergence results are reported on Figure 5, for a maximal number $N = 250$ of iterations. We use the exploitability as an indicator of convergence, since it can be evaluated explicitely at each iteration. We recall that $\varepsilon_k \leq \sigma_k \leq C\varepsilon_k$.

- Prescribed stepsizes, with $\delta_k = p/(k+p)$ and $p \in \{1, 2, 5, 10\}$. Figures 5a and 5b show that the convergence of the exploitability with an empirical rate of convergence of order k^{-p}, in accordance with Lemma 12. The value of p which yields the best convergence, for a fixed number of iterations k, increases slowly with respect to k, in accordance with Remark 13.

- Prescribed stepsizes, with $\delta_k = (k+1)^{-\alpha}$, with $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$. See Figure 5c. Similar comments can be done: a better asymptotic rate of convergence is observed for smaller values of α; the value of α which yields the best convergence for a fixed number of iterations k, decreases with respect to k. We have done some tests with smaller values of α, which are not shown on the figure. For $N = 250$ iterations, the performance is severely degraded for such values of α and convergence cannot be observed in a reasonable number of iterations.

- Adaptive stepsizes. See Figure 5d. Optimal stepsizes are approximated with golden section search (see page 13) with tolerances $\kappa = 10^{-3}$ and $\kappa = 10^{-10}$; the QAG condition is implemented with $c = 1/4$ and $\tau = 0.9$. The three methods all yield a linear convergence; the golden section and the exploitability-based methods are particularly fast. Note that in the case of optimal stepsizes, a very precise resolution of problem (11) with tolerance 10^{-10} does not improve the convergence of the method, in comparison with the tolerance 10^{-3}.

Finally we compare the time required for the GCG algorithm to satisfy a precision criterion ($\sigma_k \leq 10^{-3}$ and $\sigma_k \leq 10^{-4}$) for different stepsizes. For the considered example, the time needed to compute the adaptive stepsizes is not significantly longer than the time needed for the resolution of the HJB and the Fokker-Planck equation. Unsurprisingly, the adaptive stepsizes are more efficient than the tested prescribed stepsizes, for the two stopping criteria.

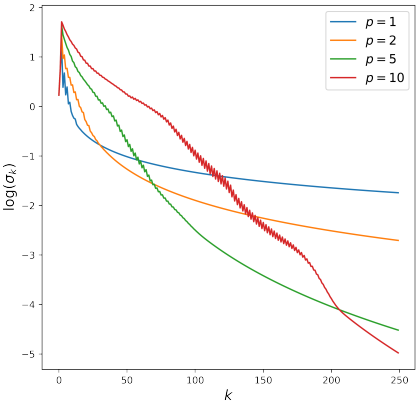
Learning method		$\sigma_k \leq 10^{-3}$	$\sigma_k \leq 10^{-4}$
Prescribed			
$\delta_k = \frac{p}{k+p}$	$p = 1$	345.82	3379.41
	$p = 2$	27.39	85.27
	$p = 5$	11.14	17.0
	$p = 10$	13.93	15.61
$\delta_k = (k+1)^{-\alpha}$	$\alpha = 0.9$	42.46	132.37
	$\alpha = 0.8$	15.62	32.3
	$\alpha = 0.7$	9.46	14.95
	$\alpha = 0.6$	10.3	12.4
Adaptative			
Quasi-Armijo-Goldstein		9.1	13.11
Golden-section $\kappa = 10^{-3}$		2.0	2.2
Golden-section $\kappa = 10^{-10}$		2.54	2.68
Exploitability-based		3.26	3.91

Figure 4: Execution time of generalized conditional gradient

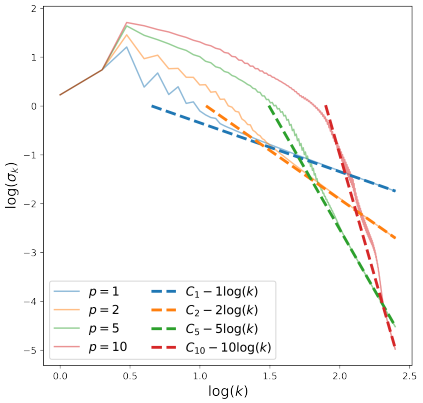
8 Conclusion

The connection between the GCG method and fictitious play investigated in this article is not specific to second-order MFGs and could be established in different settings. Before discussing some possible extensions of our work, let us make some general comments.

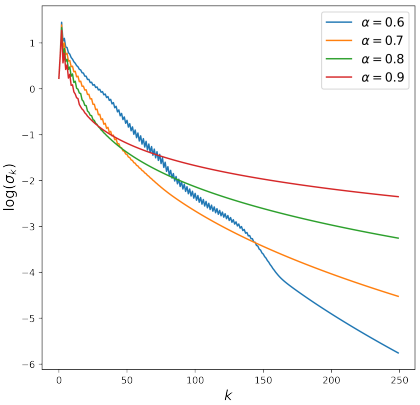
- While we have focused here on a class of MFGs with a convex potential formulation, the case of nonconvex potential MFGs is also of interest. An example of interactions with a nonconvex variational structure, arising from a consensus model, is given in [63]. The article [50] shows that in the case of nonconvex optimization problems, the Frank-Wolfe algorithm converges to stationary points, when suitable stepsizes are utilized. The notion of stationarity involved in [50] should lead to the MFG system associated with the nonconvex potential formulation.



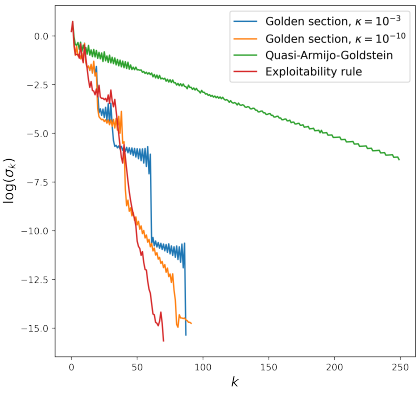
(a) Logarithm of the exploitability for $\delta_k = p/(k+p)$.



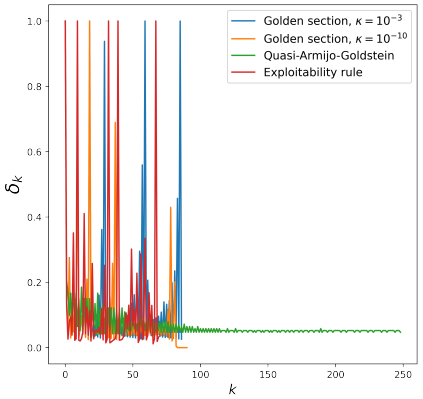
(b) Log-log scale of the exploitablity for $\delta_k = p/(k+p)$.



(c) Logarithm of the exploitability for the prescribed stepsizes $\delta_k = (k+1)^{-\alpha}$.



(d) Logarithm of the exploitability for the adaptative stepsizes.



(e) Golden section search, QAG and exploitability-based stepsizes.

Figure 5: Convergence results for prescribed and adaptative stepsizes.

- The very first assumption to be satisfied in the analysis of the Frank-Wolfe

algorithm is the Lipschitz continuity of the gradient of the cost function. In the case of MFGs, this means that the coupling functions should Lipschitz-continuous in suitable functional spaces. As a consequence, the analysis of the GCG method for potential MFGs with nonsmooth coupling functions (for example, MFGs with a local congestion term) may be particularly difficult.

- In general, the GCG method only has a sublinear rate of convergence. The linear rate of convergence obtained in this article heavily relies on the specific stability analysis which was done in Section 5 for optimal control problems.

A first natural extension of our work concerns the case of an unbounded domain. We expect that the linear convergence can be achieved. At a technical level, one difficulty concerns the boundedness of the distribution m_k. While we have established it with the help of parabolic estimates, we could follow the methodology of [20] to address this more general case. We also think that linear convergence of the GCG method can be established for fully discrete MFGs, as formulated in [10], taking Lipschitz-continuous coupling functions and a strongly convex running cost. Finally, we mention the case of first-order MFGs, in a Lagrangian formulation (as formulated in [18], for example): for this case, we only expect a sublinear rate of convergence for the GCG method.

Finally, let us mention that the fictitious play algorithm is quite similar to the policy iteration method proposed and analyzed in [15] and [16] for MFGs, since this method also relies on iterative resolutions of the HJB and the Fokker-Planck equation. The analysis techniques of our article may bring new insights to the policy iteration method.

References

[1] Yves Achdou, Francisco J. Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Partial differential equation models in macroeconomics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2028):20130397, 2014.

[2] Yves Achdou and Mathieu Laurière. Mean field games and applications: Numerical aspects. *arXiv preprint arXiv:2003.04444*, 2020.

[3] Clémence Alasseur, Imen Ben Tahar, and Anis Matoussi. An extended mean field game for storage in smart grids. *Journal of Optimization Theory and Applications*, 184(2):644–670, 2020.

[4] Dario Bauso, Hamidou Tembine, and Tamer Basar. Opinion dynamics in social networks through mean-field games. *SIAM Journal on Control and Optimization*, 54(6):3225–3257, 2016.

[5] Jean-David Benamou and Guillaume Carlier. Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, Oct 2015.

[6] Jean-David Benamou, Guillaume Carlier, Simone Di Marino, and Luca Nenna. An entropy minimization approach to second-order variational mean-field

games. *Mathematical Models and Methods in Applied Sciences*, 29(08):1553–1583, 2019.

[7] Jean-David Benamou, Guillaume Carlier, and Filippo Santambrogio. Variational mean field games. In *Active Particles, Volume 1*, pages 141–171. Springer, 2017.

[8] J. Frédéric Bonnans. *Convex and Stochastic Optimization*. Universitext Series. Springer-Verlag, 2019.

[9] J. Frédéric Bonnans, Saeed Hadikhanloo, and Laurent Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Applied Mathematics & Optimization*, 83:1431–1464, 2021.

[10] J. Frédéric Bonnans, Pierre Lavigne, and Laurent Pfeiffer. Discrete potential mean field games. *arXiv preprint arXiv:2106.07463*, 2021.

[11] Kristian Bredies, Dirk A Lorenz, and Peter Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *Computational Optimization and Applications*, 42(2):173–193, 2009.

[12] Luis Briceño-Arias, Dante Kalise, Ziad Kobeissi, Mathieu Laurière, A. Mateos González, and Francisco J. Silva. On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings. *ESAIM: Proceedings and Surveys*, 65:330–348, 2019.

[13] Luis M. Briceno-Arias, Dante Kalise, and Francisco J. Silva. Proximal methods for stationary mean field games with local couplings. *SIAM Journal on Control and Optimization*, 56(2):801–836, 2018.

[14] George W. Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

[15] Simone Cacace, Fabio Camilli, and Alessandro Goffi. A policy iteration method for mean field games. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:85, 2021.

[16] Fabio Camilli and Qing Tang. Rates of convergence for the policy iteration method for mean field games systems. *Journal of Mathematical Analysis and Applications*, 512(1):126138, 2022.

[17] Pierre Cardaliaguet, P. Jameson Graber, Alessio Porretta, and Daniela Tonon. Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1287–1317, 2015.

[18] Pierre Cardaliaguet and Saeed Hadikhanloo. Learning in mean field games: the fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017.

[19] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.

[20] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.

[21] Pierre Cardaliaguet, Alpár R. Mészáros, and Filippo Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016.

[22] René Carmona, François Delarue, and Daniel Lacker. Mean field games of timing and models for bank runs. *Applied Mathematics & Optimization*, 76(1):217–260, 2017.

[23] René Carmona and Daniel Lacker. A probabilistic weak formulation of mean field games and applications. *The Annals of Applied Probability*, 25(3):1189–1231, 2015.

[24] Romain Couillet, Samir M. Perlaza, Hamidou Tembine, and Mérouane Debbah. Electrical vehicles in the smart grid: A mean field game analysis. *IEEE Journal on Selected Areas in Communications*, 30(6):1086–1096, 2012.

[25] Kai Cui and Heinz Koeppl. Approximately solving mean field games via entropy-regularized deep reinforcement learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1909–1917. PMLR, 2021.

[26] François Delarue and Athanasios Vasileiadis. Exploration noise for learning linear-quadratic mean field games. *arXiv preprint arXiv:2107.00839*, 2021.

[27] Josu Doncel, Nicolas Gast, and Bruno Gaujal. A mean field game analysis of SIR dynamics with vaccination. *Probability in the Engineering and Informational Sciences*, pages 1–18, 2020.

[28] Joseph 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.

[29] Romuald Elie, Emma Hubert, and Gabriel Turinici. Contact rate epidemic control of covid-19: an equilibrium view. *Mathematical Modelling of Natural Phenomena*, 15:35, 2020.

[30] Romuald Elie, Julien Pérolat, Mathieu Laurière, Matthieu Geist, and Olivier Pietquin. Approximate fictitious play for mean field games. *arXiv preprint arXiv:1907.02633*, 2019.

[31] Olivier Féron, Peter Tankov, and Laura Tinsi. Price formation and optimal trading in intraday electricity markets. *arXiv preprint arXiv:2009.04786*, 2020.

[32] Wendell H. Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

[33] Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[34] Drew Fudenberg and David K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[35] Matthieu Geist, Julien Pérolat, Mathieu Laurière, Romuald Elie, Sarah Perrin, Olivier Bachem, Rémi Munos, and Olivier Pietquin. Concave utility reinforcement learning: the mean-field game viewpoint. *arXiv preprint arXiv:2106.03787*, 2021.

[36] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. springer, 2015.

[37] P. Jameson Graber and Alain Bensoussan. Existence and uniqueness of solutions for Bertrand and Cournot mean field games. *Applied Mathematics & Optimization*, pages 1–25, 2015.

[38] P. Jameson Graber, Vincenzo Ignazio, and Ariel Neufeld. Nonlocal bertrand and cournot mean field games with general nonlinear demand schedule. *Journal de Mathématiques Pures et Appliquées*, 148:150–198, 2021.

[39] P. Jameson Graber and Charafeddine Mouzouni. Variational mean field games for market competition. In *PDE models for multi-agent phenomena*, pages 93–114. Springer, 2018.

[40] P. Jameson Graber and Charafeddine Mouzouni. On mean field games models for exhaustible commodities trade. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:11, 2020.

[41] P. Jameson Graber, Alan Mullenix, and Laurent Pfeiffer. Weak solutions for potential mean field games of controls. *Nonlinear Differential Equations and Applications NoDEA*, 28(5):1–34, 2021.

[42] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer, 2011.

[43] Saeed Hadikhanloo and Francisco 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.

[44] Minyi Huang, Roland P. Malhamé, and Peter 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.

[45] Martin Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *International Conference on Machine Learning*, pages 427–435. PMLR, 2013.

[46] Thomas Kerdreux. *Accelerating conditional gradient methods*. PhD thesis, Université Paris sciences et lettres, 2020.

[47] Karl Kunisch and Daniel Walter. On fast convergence rates for generalized conditional gradient methods with backtracking stepsize. *arXiv preprint arXiv:2109.15217*, 2021.

[48] Aimé Lachapelle, Jean-Michel Lasry, Charles-Albert Lehalle, and Pierre-Louis Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *Mathematics and Financial Economics*, 10(3):223–262, 2016.

[49] Aimé Lachapelle and Marie-Therese Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation research part B: methodological*, 45(10):1572–1589, 2011.

[50] Simon Lacoste-Julien. Convergence rate of frank-wolfe for non-convex objectives. *arXiv preprint arXiv:1607.00345*, 2016.

[51] Olga Aleksandrovna Ladyzhenskaia, Vsevolod Alekseevich Solonnikov, and Nina N Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[52] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[53] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. ii–horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006.

[54] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[55] Jacques-Louis Lions. *Optimal control of systems governed by partial differential equations*, volume 170. Springer Verlag, 1971.

[56] Alpár Richárd Mészáros and Francisco J. Silva. A variational approach to second order mean field games with density constraints: the stationary case. *Journal de Mathématiques Pures et Appliquées*, 104(6):1135–1159, 2015.

[57] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

[58] Julien Perolat, Sarah Perrin, Romuald Elie, Mathieu Laurière, Georgios Piliouras, Matthieu Geist, Karl Tuyls, and Olivier Pietquin. Scaling up mean field games with online mirror descent. *arXiv preprint arXiv:2103.00623*, 2021.

[59] Sarah Perrin, Mathieu Laurière, Julien Pérolat, Matthieu Geist, Romuald Élie, and Olivier Pietquin. Mean field games flock! the reinforcement learning way. *arXiv preprint arXiv:2105.07933*, 2021.

[60] Sarah Perrin, Julien Pérolat, Mathieu Laurière, Matthieu Geist, Romuald Elie, and Olivier Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. *arXiv preprint arXiv:2007.03458*, 2020.

[61] Adam Prosinski and Filippo Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. *Stochastics*, 89(6-7):923–942, 2017.

[62] Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pages 296–301, 1951.

[63] Filippo Santambrogio and Woojoo Shim. A cucker–smale inspired deterministic mean field game with velocity interactions. *SIAM Journal on Control and Optimization*, 59(6):4155–4187, 2021.

[64] Sylvain Sorin. Continuous time learning algorithms in optimization and game theory. *Dynamic Games and Applications*, pages 1–22, 2022.

[65] Weichen Wang, Jiequn Han, Zhuoran Yang, and Zhaoran Wang. Global convergence of policy gradient for linear-quadratic mean-field control/game in continuous time. In *International Conference on Machine Learning*, pages 10772–10782. PMLR, 2021.