



HAL
open science

Generalized conditional gradient and learning in potential mean field games

Pierre Lavigne, Laurent Pfeiffer

► **To cite this version:**

Pierre Lavigne, Laurent Pfeiffer. Generalized conditional gradient and learning in potential mean field games. Applied Mathematics and Optimization, 2023, 88, 10.1007/s00245-023-10056-8 . hal-03341776v3

HAL Id: hal-03341776

<https://hal.science/hal-03341776v3>

Submitted on 17 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generalized conditional gradient and learning in potential mean field games*

Pierre Lavigne[†] Laurent Pfeiffer[‡]

August 17, 2023

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, mean field optimal control, 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 [29] and M. Huang, R. Malhamé, and P. Caines in [24], 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.

*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 a mean field game system of the following form

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

The unknown of the system is (m, v, u, γ, P) , with $m: Q \rightarrow \mathbb{R}$, $v: Q \rightarrow \mathbb{R}^d$, $u: Q \rightarrow \mathbb{R}$, $\gamma: Q \rightarrow \mathbb{R}$, and $P: [0, T] \rightarrow \mathbb{R}^k$. The data of the system (which will be clearly defined latter in Section 2) are the initial distribution m_0 and the terminal cost g . The mapping H denotes the Hamiltonian of the system, f and ϕ are a congestion cost and a price function.

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 (10) and (11); 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). 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 a convex mean field control problem,

$$\inf_{(m,w) \in \mathcal{R}} \mathcal{J}(m, w) := \mathcal{J}_1(m, w) + \mathcal{J}_2(m, w), \quad (\text{P})$$

obtained using the classical Benamou-Brenier transformation. Problem (P) is referred to as the potential problem. Roughly speaking (the precise definition is given in (8)), the feasible set \mathcal{R} is the set of pairs (m, w) satisfying the Fokker-Planck equation $\partial_t m - \Delta m + \nabla \cdot w = 0$ and $m(0) = m_0$ and

$$\begin{aligned} \mathcal{J}_1(m, w) &= \int_Q \tilde{\mathbf{L}}[m, w](x, t) dx dt + \int_{\mathbb{T}^d} g(x)m(x, T) dx, \\ \mathcal{J}_2(m, w) &= \int_0^T (\mathbf{F}[m](t) + \Phi[Aw](t)) dt. \end{aligned} \quad (1)$$

Potential MFGs have been widely investigated, we refer the reader to [4, 13] for interactions through the density m and to [5, 20, 21] for price interactions.

Various methods from convex optimization have been employed to solve potential problems, see [1] 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 [6, 9, 10]. Another approach consists in applying the augmented Lagrangian algorithm to the dual problem of the potential problem, see [2, 4, 6]. Other methods have been investigated such as the Sinkhorn algorithm [3].

Generalized conditional gradient algorithm The aim of the paper is to show that the potential mean field game system (MFG) can be efficiently solved with the *generalized conditional gradient* method (GCG). This method is iterative: at iterate k , given a candidate (\bar{m}_k, \bar{w}_k) , one first solves the following partially linearized problem:

$$\inf_{(m,w) \in \mathcal{R}} \mathcal{J}_1(m, w) + D\mathcal{J}_2[\bar{m}_k, \bar{w}_k](m, w). \quad (2)$$

The derivative $D\mathcal{J}_2[\bar{m}_k, \bar{w}_k]$ will be explicitly defined in the analysis. Denoting by (m_k, w_k) a solution to this problem, the next iterate is defined as $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1 - \delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$ for some stepsize $\delta_k \in [0, 1]$.

The GCG algorithm, first introduced in [7], is an extension of the conditional gradient algorithm, also called Frank-Wolfe algorithm. The conditional gradient method allows to minimize a convex objective function on a convex and compact set. A classical choice of step size is given by $\delta_k = 2/(k + 2)$ (see [25]) which yields the convergence of the objective function in $O(1/k)$.

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 [26], in an infinite-dimensional setting. The main interest of the generalized conditional gradient compared to the Frank-Wolfe algorithm is to use the regularity of the cost function to reach higher convergence rates. This algorithm is thus very suitable for potential mean field games since it is inherently assumed that the potential coupling terms are differentiable.

Learning in mean field games 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. We consider in this article the *fictitious play*, a learning procedure 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. In the context of MFGs, the fictitious play has been investigated in [14, 16, 23, 32]. 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. The convergence of the exploitability has been addressed in [32] 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 fictitious play method can be interpreted as a GCG algorithm. This is a consequence of the fact that the partially linearized problem (2) is equivalent to a standard stochastic optimal control. As will be justified, the unique solution (m_k, w_k) to (2) is obtained by first computing the belief (γ_k, P_k)

$$\gamma_k(x, t) = f(x, t, \bar{m}_k(t)), \quad P_k(t) = \phi(t, A\bar{w}_k(t)),$$

next by solving the PDEs

$$\begin{cases} -\partial_t u_k - \Delta u_k + \mathbf{H}[\nabla u_k + A^* P_k] = \gamma_k, & (x, t) \in Q, \\ \partial_t m_k - \Delta m_k + \nabla \cdot w_k = 0, & (x, t) \in Q, \end{cases}$$

using the same boundary conditions as in (MFG), and finally by computing the best-response $w_k = -\mathbf{H}_p[\nabla u_k + A^* P_k]m_k$. If we further choose $\delta_k = 1/(k+1)$, we recover the fictitious play algorithm introduced in [22].

Exploitability and primal-dual gap The connection between the GCG algorithm and the fictitious play allows us to show that the notion of primal-dual gap *primal-dual gap* and the notion of exploitability of the game are equivalent. This interpretation has already been highlighted in a very recent work [18], 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 investigated in [34] for a general class of potential games.

Contributions The article [14] 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 [18], 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 [14] 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 note that our analysis is restricted to the case of non-degenerate second-order MFGs. At a technical level, the proof of convergence utilizes in a crucial manner techniques from [26]. Let us emphasize that for the methods from convex analysis which we have cited above (Chambolle-Pock, augmented Lagrangian, Sinkhorn), no rate of convergence has been established. To the best of our knowledge, only two references provide a

linear rate of converge for MFGs: [35], which is restricted to linear-quadratic MFGs, and [12], under a restrictive smallness assumption for the coupling function.

Plan of the paper The paper is organized in two main parts. The first part, consisting of Sections 2 to 4, present the GCG algorithm for MFGs. Section 2 is dedicated to the notations, the assumptions and the introduction of the main mappings used all along the article. We then describe the GCG algorithm in Section 3 and we explain its connection with fictitious play. We also state our main convergence results: Theorems 7 and 8. Section 4 present numerical illustrations of the convergence results for an academical problem.

The second part, consisting of Section 5, deals with the analysis of the GCG algorithm. We establish the well-posedness of Algorithm 1 and prove Theorems 7 and 8. Related technical details concerning the Fokker-Planck and the HJB equations are available in Appendix A.

2 Notations, assumptions and mappings

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 $\partial_{x_i, x_j}^2 u$ 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

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

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 [28, Lemma II.3.3., page 80 and Corollary, page 342]. \square

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

$$\Theta = \left\{ v \in \mathcal{C}(Q; \mathbb{R}^d) \mid D_x v \in L^q(Q; \mathbb{R}^{d \times d}) \right\}, \quad (3)$$

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

$$\Xi = (W^{1,0,\infty}(Q) \cap \mathcal{C}(Q)) \times \mathcal{C}(0, T; \mathbb{R}^k). \quad (4)$$

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 :

$$\Xi_R = \left\{ (\gamma, P) \in \Xi \mid \|\gamma\|_{W^{1,0,\infty}(Q)} + \|P\|_{L^\infty(0, T; \mathbb{R}^k)} \leq R \right\}. \quad (5)$$

We also define

$$\mathcal{D}_1(\mathbb{T}^d) = \left\{ m \in L^\infty(\mathbb{T}^d) \mid m \geq 0, \int_{\mathbb{T}^d} m(x) \, dx = 1 \right\}.$$

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

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

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 $\mathbf{H}[\nabla u](x, t)$. Note that \mathbf{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. We will also 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 Assumptions

We define two boundary conditions $m_0 \in \mathcal{D}_1(\mathbb{T}^d)$, $g: \mathbb{T}^d \rightarrow \mathbb{R}$, and four maps: a running cost L , an interaction cost f , a price function ϕ and an aggregation term a ,

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

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 also define the perspective function $\tilde{L}: Q \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ of L ,

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

Note that the function \tilde{L} is convex and lower semi-continuous with respect to (m, w) . Next we define two linear operators $A: w \in L^1(Q; \mathbb{R}^d) \mapsto A[w] \in L^1(0, T; \mathbb{R}^k)$ and $A^*: P \in L^\infty(0, T; \mathbb{R}^k) \mapsto A^*[P] \in L^\infty(Q; \mathbb{R}^d)$ as follows:

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

Note that the function a will be assumed to be bounded.

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 $v \in \mathbb{R}^d$, we have: $|L(x, t, v) - L(y, t, v)| \leq C_0|x - y|(1 + |v|^2)$.
- (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 $\mathcal{D}_1(\mathbb{T}^d)$,

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

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

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

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

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 $F: [0, T] \times \mathcal{D}_1(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

$$F(t, m_2) - F(t, m_1) = \int_0^1 \int_{\mathbb{T}^d} f(x, t, sm_2 + (1-s)m_1)(m_2(x) - m_1(x)) dx ds. \tag{6}$$

Moreover, ϕ has a convex potential Φ , that is to say there exists a measurable mapping $\Phi: [0, T] \times \mathbb{R}^k \rightarrow \mathbb{R}$, 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. An example of coupling term f satisfying the above assumptions can be found in [5, page 8].

2. The monotonicity assumption on f implies that

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

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.

3. If the terminal condition g depends on m , the results of the paper are still valid if g satisfies the same kind of assumptions as f (H3), (H5), and (H6) (boundedness, Lipschitz-continuity, monotonicity, existence of a potential function).

4. Our assumptions are stronger than those of [5], since we require the boundedness of ϕ and the Lipschitz continuity of f w.r.t. m , for the $L^2(\mathbb{T}^d)$ -norm.

Theorem 3. *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 [5, Theorem 1 and Proposition 2]. \square

Lemma 4. *The pair (\bar{m}, \bar{w}) is the unique solution to the potential problem (P).*

Proof. Lemma 4 is a direct consequence of Corollary 22, proved in Section 5.2. \square

For the rest of the article, following Theorem 3, we denote by $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$ the unique solution to (MFG) and we set $\bar{w} = \bar{m}\bar{v}$.

2.3 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 Appendix A.

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

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (vm) &= 0 & (x, t) \in Q, \\ m(0, x) &= m_0(x) & x \in \mathbb{T}^d. \end{aligned} \quad (7)$$

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

$$\mathcal{R} = \left\{ \begin{array}{l} (m, w) \in W^{2,1,q}(Q) \times \Theta \\ \partial_t m - \Delta m + \nabla \cdot w = 0, \quad m(0, \cdot) = m_0 \\ \exists v \in L^\infty(Q; \mathbb{R}^d), w = mv \end{array} \right\}. \quad (8)$$

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

Proof. Let $(m_i, w_i) \in \mathcal{R}$, for $i = 1, 2$. Let $v_i \in L^\infty(Q; \mathbb{R}^d)$ be such that $w_i = m_i v_i$. Let $\theta \in (0, 1)$ and let $(m_\theta, w_\theta) = (1 - \theta)(m_1, w_1) + \theta(m_2, w_2)$. Clearly $\partial_t m_\theta - \Delta m_\theta + \nabla \cdot w_\theta = 0$ and $m_\theta(0, \cdot) = m_0$. Let $C = \max_{i=1,2} \{ \|v_i\|_{L^\infty(Q; \mathbb{R}^d)} \}$. We define

$$\varphi: (m, w) \in W^{2,1,q}(Q) \times \Theta \mapsto \max_{(x,t) \in Q} \|w(x, t)\| - Cm(x, t) \in \mathbb{R}.$$

We have $m_\theta \geq 0$. Define next v_θ as $v_\theta = w_\theta/m_\theta$ on $\{m_\theta > 0\}$, $v_\theta = 0$ on $\{m_\theta = 0\}$. We have $\varphi(m_i, w_i) \leq 0$. By convexity of φ we have $\varphi(m_\theta, w_\theta) \leq 0$, which implies that $\|v_\theta\|_{L^\infty(Q; \mathbb{R}^d)} \leq C$ and that for all $(x, t) \in Q$, if $m_\theta(x, t) = 0$ then $w_\theta(x, t) = 0$. Thus $w_\theta = m_\theta v_\theta$. The convexity of \mathcal{R} follows. The second part of the lemma is a consequence of Lemma 1 and Lemma 13. \square

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

$$\begin{aligned} -\partial_t u - \Delta u + \mathbf{H}[\nabla u + A^*P] &= \gamma & (x, t) \in Q, \\ u(x, T) &= g(x) & x \in \mathbb{T}^d. \end{aligned} \quad (9)$$

Let us give an interpretation of \mathbf{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 $\mathbb{L}_{\mathbb{F}}^2(t, T)$ the set of progressively measurable stochastic processes ν defined on $[t, T]$ and valued in \mathbb{R}^d such that $\mathbb{E}[\int_t^T |\nu_s|^2 ds] < +\infty$. Given $(\gamma, P) \in \Xi$, we consider the mapping $J[\gamma, P]: Q \times \mathbb{L}_{\mathbb{F}}^2(0, T) \rightarrow \mathbb{R}$, defined by

$$J[\gamma, P](x, t, \nu) = \mathbb{E} \left[\int_t^T \left(L(X_s, \nu_s) + \langle A^*[P](X_s, s), \nu_s \rangle + \gamma(X_s, s) \right) ds + g(X_T) \right], \quad (10)$$

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

$$\mathbf{u}[\gamma, P](x, t) = \inf_{\nu \in \mathbb{L}_{\mathbb{F}}^2(0, T)} J[\gamma, P](x, t, \nu). \quad (11)$$

This is a classical result from dynamic programming theory, see [17].

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

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

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

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

Finally, using Nemytskii operators, we define

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

In summary, the mappings \mathbf{u} , \mathbf{v} , \mathbf{m} , and \mathbf{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 γ and \mathbf{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 Generalized conditional gradient and fictitious play

In this section we present the GCG method, provide an interpretation as a learning method, and state our main results.

3.1 Individual criterion and best reply

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

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

and the associated optimal control problem

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

Provided that it exists, any minimizer of the above problem can be interpreted as a best reply of the individual player, given the coupling term (γ, P) . A key observation is that $(\mathcal{P}[\gamma, P])$ can be seen as a partial linearization of (\mathbf{P}) .

Lemma 6. *For any $(\gamma, P) \in \Xi$, $(\mathbf{m}[\gamma, P], \mathbf{w}[\gamma, P])$ is the unique solution to $(\mathcal{P}[\gamma, P])$.*

Proof. This is a direct consequence of Proposition 21, proved at page 20. \square

Take a pair $(\hat{m}, \hat{w}) \in \mathcal{R}$ and set $\hat{\gamma} = \gamma[\hat{m}]$ and $\hat{P} = \mathbf{P}[\hat{w}]$. 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, $\int_Q \hat{\gamma}m dx dt$ is a linearization of $\int_0^T \mathbf{F}[m] dt$ around \hat{m} and $\int_0^T \langle A[w], \hat{P} \rangle dt$ is a linearization of $\int_0^T \Phi[Aw] dt$ around \hat{w} .

3.2 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 $0 \leq k < N$ **do**

1. Set $\gamma_k = \gamma[\bar{m}_k]$ and $P_k = \mathbf{P}[\bar{w}_k]$.

2. Find the solution to $\mathcal{P}[\gamma_k, P_k]$, that is, define successively:

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

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

$$(\bar{m}_k^{\delta}, \bar{w}_k^{\delta}) = (1 - \delta)(\bar{m}_k, \bar{w}_k) + \delta(m_k, w_k), \quad \forall \delta \in [0, 1]. \quad (12)$$

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 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 [14] (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

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

where $\gamma = \boldsymbol{\gamma}[m]$ and $P = \mathbf{P}[w]$. 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

$$\sigma_k = \mathcal{Z}[\gamma_k, P_k](\bar{m}_k, \bar{w}_k) - \mathcal{Z}[\gamma_k, P_k](m_k, w_k). \quad (13)$$

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 [26], we will investigate three different rules for the determination of adaptive stepsizes.

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

$$\delta_k \in \arg \min_{\delta \in [0,1]} \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta), \quad (14)$$

where $(\bar{m}_k^\delta, \bar{w}_k^\delta)$ is defined as in (12).

- **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

$$\delta_k = \tau^{i_k}, \quad i_k = \operatorname{argmin} \left\{ j \in \mathbb{N} \mid \tau^j \text{ satisfies the QAG condition} \right\}. \quad (15)$$

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

$$\delta_k = \min \left\{ 1, \frac{\sigma_k}{2(C_1 D_k^{(1)} + C_2 D_k^{(2)})} \right\}, \quad (16)$$

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

$$\begin{aligned} D_k^{(1)} &= \int_0^T \|m_k(t, \cdot) - \bar{m}_k(t, \cdot)\|_{L^2(\mathbb{T}^d)} \|m_k(t, \cdot) - \bar{m}_k(t, \cdot)\|_{L^1(\mathbb{T}^d)} dt \\ D_k^{(2)} &= \int_0^T \left| \int_{\mathbb{T}^d} a(x, t) (w_2(x, t) - w_1(x, t)) dx \right|^2 dt \end{aligned} \quad (17)$$

3.3 Main convergence results and discussion

Our first result concerns the convergence of the variables of the algorithm. Our second result provides several rates of convergence for the optimality gap, depending on the choice of the learning sequence $(\delta_k)_{k \in \mathbb{N}}$. The two theorems are proved in Subsections 5.3 and 5.4.

Theorem 7. *There exists $C > 0$ such that for all $k \in \mathbb{N}$,*

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

Theorem 8. *1. At each iteration, assume that $(\delta_k)_{k \in \mathbb{N}}$ satisfy one the three adaptive rules described above: either (14), (15) (for values of τ and c independent of k), or (16). 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 $C > 0$ such that whatever the choice of stepsizes $\delta_k \in [0, 1]$, it holds that

$$\varepsilon_{k+1} \leq \frac{\varepsilon_0 \exp\left(C \sum_{j=0}^k \delta_j^2\right)}{\exp\left(\sum_{j=0}^k \delta_j\right)}, \quad \forall k \in \mathbb{N}. \quad (18)$$

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 (14). 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

$$\bar{\delta} \in \arg \min_{\delta \in \{a, b, c, d\}} \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta). \quad (19)$$

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 $\delta_k^{\text{GS}} = \bar{\delta}$. Finally, denote δ_k^{QAG} the stepsize determined by (15). As a substitute for (14), one can choose $\delta_k \in \arg \min_{\delta \in \{\delta_k^{\text{GS}}, \delta_k^{\text{QAG}}\}} \mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta)$. In view of the proof of Theorem 8, it is clear that linear convergence is also achieved for this choice of stepsize.

Remark 9. *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, (15) (since i_k is bounded), (16) (the formula is explicit), (19) (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 10. *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:

$$\begin{aligned} \sum_{j=0}^k \frac{p}{j+p} &\geq \int_0^{k+1} \frac{p}{s+p} ds = p \ln \left(\frac{k+p+1}{p} \right) \\ \sum_{j=0}^k \left(\frac{p}{j+p} \right)^2 &\leq p^2 \left(\frac{1}{p} + \int_0^\infty \frac{1}{(s+p)^2} ds \right) = 2p. \end{aligned}$$

Using inequality (18), we deduce that $\varepsilon_k \leq \varepsilon_0 \exp(2Cp - p \ln(k+p) + p \ln(p))$, from which the announced result follows. \square

Remark 11. *Lemma 10 shows that rates of convergence of order $\mathcal{O}(k^{-p})$ can be achieved. Yet the constant behind the asymptotic rate of convergence, of order $p^p \exp(2Cp)$, increases quickly with p , which mitigates the interest of taking a too large value of p in practice. This is confirmed in the numerical tests of Section 4.*

Lemma 12. *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*

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

In particular, if $\sum_{k=0}^\infty \delta_k = \infty$, then $\varepsilon_k \xrightarrow[k \rightarrow \infty]{} 0$.

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 (18) 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 $C \sum_{j=0}^k \delta_j^2 - \sum_{j=0}^k \delta_j \leq \left(C \sum_{j=0}^{k_0-1} \delta_j^2 + r \sum_{j=0}^{k_0-1} \delta_j \right) - (1-r) \sum_{j=0}^k \delta_j$. Plugging this inequality into (18), we obtain the announced result. \square

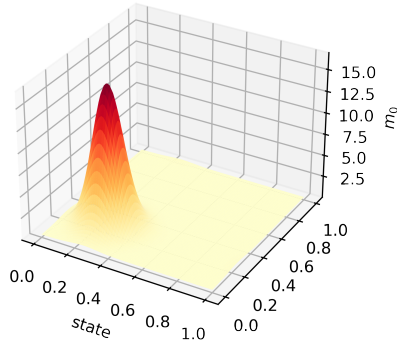
Lemma 12 shows that high convergence rates can be achieved if the sequence $(\delta_k)_{k \in \mathbb{N}}$ decreases slowly, as we already noticed in Lemma 10. However, if the convergence to 0 is slow, the values of k_0 and C may be very large, as was revealed in the proof. This fact is verified for the sequence $\delta_k = (1+k)^{-\alpha}$.

4 Numerical illustration

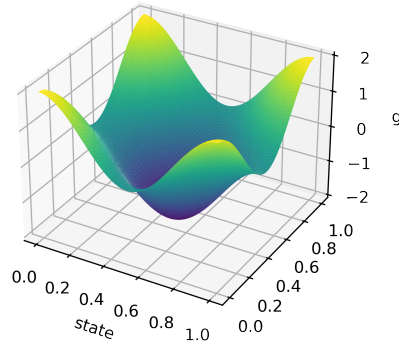
As a numerical illustration, we solve the mean field game system (MFG) with $f = 0$. In this situation, the agents only interact through the law of their controls. The associated potential problem has the following form

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

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 .



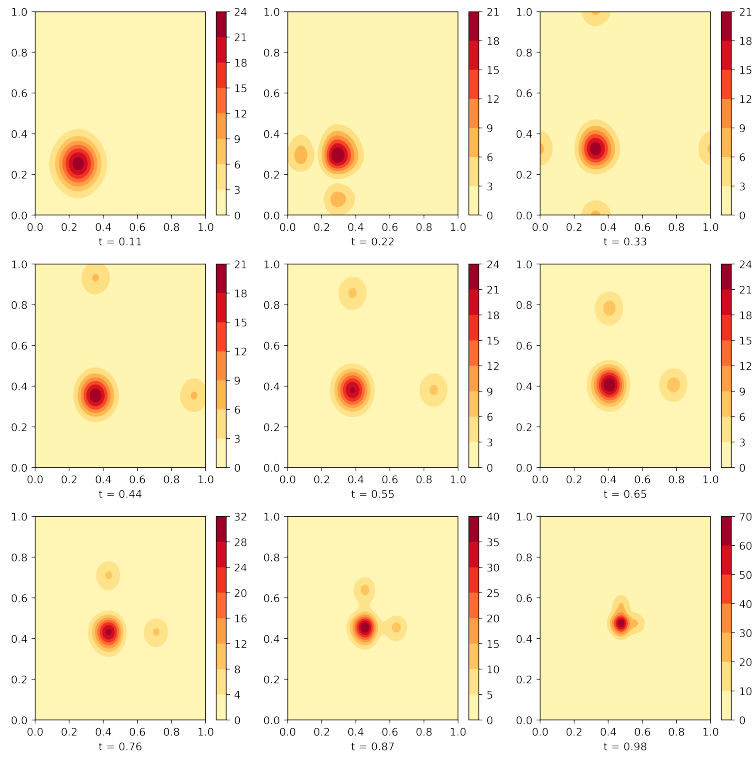
(b) Terminal condition g .

We define $L(x, t, v) = \frac{1}{2}|v|^2$, $a(x, t) = \text{Id}$ 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: $P_i = 10 \int_{\mathbb{T}^d} v_i(x, t) m_i(x, t) dx dt$. For any control $\nu_t \in \mathbb{L}_{\mathbb{F}}^2(0, 1)$, the cost function of a representative agent writes:

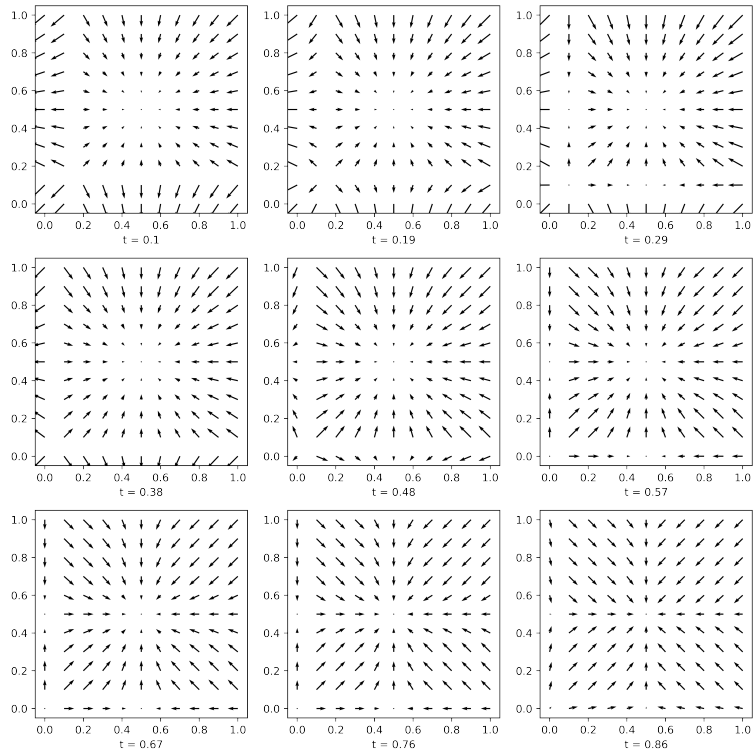
$$\mathbb{E} \left[\int_0^T \frac{1}{2} |\nu_t|^2 + 10 \langle P(t), \nu_t \rangle dt + g(X_T^\nu) \right].$$

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 \mathbf{u} (i.e. a solution to the Hamilton-Jacobi-Bellman equation) and \mathbf{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.

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 line 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



(a) Density of players \bar{m} evaluated at several time steps.



(b) Vector field \bar{v} evaluated at several time steps.

Figure 2: Density and velocity vector field evaluated at several points in time.

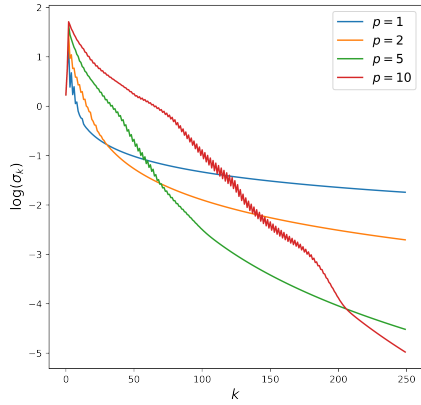
to reach the point $(0.5, 0.5)$ using control with a first coordinate that is negative. Graphically speaking, these agent would cross the left vertical axis ($x_1 = 0$) and “jump” to the right vertical axis ($x_1 = 1$). This strategy is particularly interesting for agents an initial condition x such 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 2b. Thus the initial distribution is split into four groups as shown by Figure 2a. The group of agents crossing both axes is actually of very small mass, thus not visible on the graph.

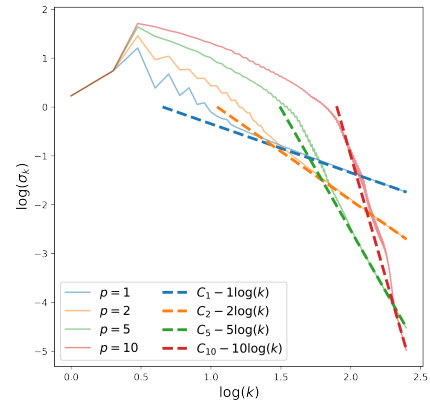
Convergence and execution time The convergence results are reported on Figure 3, for a maximal number $N = 250$ of iterations. We use the exploitability as an indicator of convergence, since it can be evaluated explicitly at each iteration. We recall that $\varepsilon_k \leq \sigma_k \leq C\varepsilon_k$.

- Prescribed stepsizes, with $\delta_k = p/(k + p)$ and $p \in \{1, 2, 5, 10\}$. Figures 3a and 3b show that the convergence of the exploitability with an empirical rate of convergence of order k^{-p} , in accordance with Lemma 10. 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 11.
- Prescribed stepsizes, with $\delta_k = (k + 1)^{-\alpha}$, with $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$. See Figure 3c. 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 3d. Optimal stepsizes are approximated with golden section search (see page 12) with tolerances $\kappa = 10^{-3}$ and $\kappa = 10^{-10}$. The QAG condition is implemented with $c = 1/4$ and $\tau = 0.9$. The choice of c and τ is here arbitrary since there is no general rule for the choice of these parameters (see [31, Chapter 3] for further discussions on the topic). 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 (14) with tolerance 10^{-10} does not improve the convergence of the method, in comparison with the tolerance 10^{-3} .

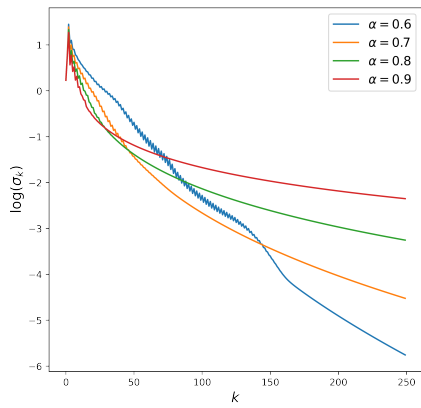
Finally we compare the time required for the GCG algorithm to satisfy a precision criterion ($\sigma_k \leq 10^{-3}$ and $\sigma_k \leq 10^{-4}$) for different stepsizes. For the considered example, the time needed to compute the adaptive stepsizes is not significantly longer than the time needed for 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.



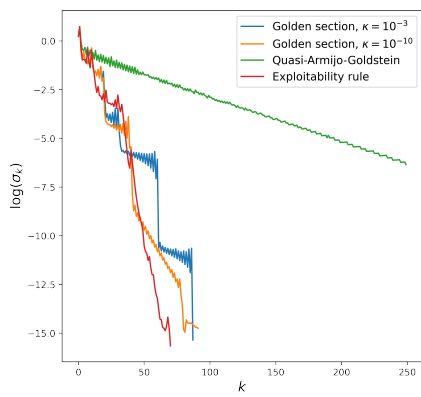
(a) Logarithm of the exploitability for $\delta_k = p/(k + p)$.



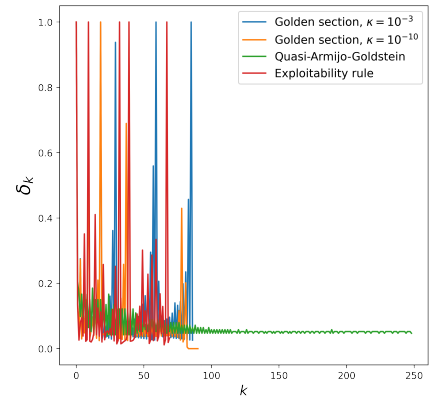
(b) Log-log scale of the exploitability for $\delta_k = p/(k + p)$.



(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 3: Convergence results for prescribed and adaptative stepsizes.

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 in seconds of the generalized conditional gradient

5 Stability and convergence results

This section is dedicated to the analysis of Algorithm 1 and is organized in four subsections. We state in Subsection 5.1 some technical results concerning the well-posedness of the auxiliary mappings introduced in Subsection 2.3. Subsection 5.2 establishes a stability result (Proposition 21), necessary to prove the linear speed of convergence. Subsection 5.3 addresses the well-posedness of Algorithm 1 and the proof of Theorem 7. Subsection 5.4 is dedicated to the proof of Theorem 8.

5.1 Well-posedness of the auxiliary mappings

We provide here technical results, whose proofs can be found in the Appendix. We recall that the sets Θ , Ξ , and Ξ_R have been introduced in (3), (4), and (5).

Lemma 13. *Let $R > 0$. Let $v \in \Theta$ be such that $\|v\|_{\Theta} \leq R$. Then $\mathbf{M}[v]$ is uniquely defined and lies in $W^{2,1,q}(Q)$. Moreover, there exists a constant $C(R) > 0$, independent of v , such that $\|\mathbf{M}[v]\|_{W^{2,1,q}(Q)} \leq C(R)$. Finally, $\mathbf{M}[v](x, t) \geq 0$, for all $(x, t) \in Q$.*

Lemma 14. *Let $R > 0$. Let v_1 and $v_2 \in \Theta$. Let $m_i = \mathbf{M}[v_i] \in W^{2,1,q}(Q)$ for $i \in \{1, 2\}$. Assume that $\|v_1\|_{\Theta} \leq R$ and $\|m_2\|_{L^\infty(Q)} \leq R$. Then, there exists a constant $C(R)$, independent of v_1 and v_2 , such that*

$$\|m_2 - m_1\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} \leq C(R) \left(\int_Q |v_2 - v_1|^2 m_2 \, dx \, dt \right)^{1/2}.$$

Lemma 15. *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:*

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

Proposition 16. *The map \mathbf{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 $(\gamma, P) \in \Xi_R$, $\|\mathbf{u}[\gamma, P]\|_{W^{2,1,q}(Q)} \leq C(R)$.*

Lemma 17. *The maps \mathbf{v} , \mathbf{m} , and \mathbf{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:*

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

Lemma 18. *The mappings γ and \mathbf{P} are well-defined. There exists a constant $C > 0$ such that for all $m \in W^{2,1,q}(Q)$ and for all $w \in \Theta$,*

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

where $\gamma = \gamma[m]$ and $P = \mathbf{P}[w]$. 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 ,

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

5.2 Stability results for stochastic optimal control problems

The main result of this section, Proposition 21 and its Corollary 22, 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 19. *Let $(\hat{\gamma}, \hat{P}) \in \Xi$. Let $\hat{m} = \mathbf{m}[\hat{\gamma}, \hat{P}]$, $\hat{v} = \mathbf{v}[\hat{\gamma}, \hat{P}]$ and let $\hat{w} = \mathbf{w}[\hat{\gamma}, \hat{P}]$. There exists a constant $C > 0$ such that for any $(m, w) \in \mathcal{R}$, the following holds:*

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

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

The proof of Lemma 19 relies on the following inequality.

Lemma 20. *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 [5, Proof of Proposition 2]. □

Proof of Lemma 19. Using the definition of $\mathcal{Z}[\hat{\gamma}, \hat{P}]$, we have

$$\begin{aligned} \mathcal{Z}[\hat{\gamma}, \hat{P}](m, w) - \mathcal{Z}[\hat{\gamma}, \hat{P}](\hat{m}, \hat{w}) &= \int_Q (L[v]m - L[\hat{v}]\hat{m}) \, dx \, dt \\ &+ \int_Q \hat{\gamma}(m - \hat{m}) \, dx \, dt + \int_Q \langle A^* P, w - \hat{w} \rangle \, dx \, dt + \int_{\mathbb{T}^d} g(m(T) - \hat{m}(T)) \, dx. \end{aligned} \quad (20)$$

We set $\hat{u} = \mathbf{u}[\hat{\gamma}, \hat{P}]$. By definition of \hat{v} , we have $\hat{v} = -\mathbf{H}_p[\nabla\hat{u} + A^*\hat{P}]$. Applying Lemma 20 with $\hat{p} = \nabla\hat{u} + A^*\hat{P}$, we obtain

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

We inject the obtained inequality into (20) and we use $\mathbf{H}[\nabla\hat{u} + A^*\hat{P}] + \hat{\gamma} = -\partial_t\hat{u} - \Delta\hat{u}$. This yields

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

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. \square

Proposition 21. *Let $R > 0$ and let $(\hat{\gamma}, \hat{P}) \in \Xi_R$. Let $\hat{m} = \mathbf{m}[\hat{\gamma}, \hat{P}]$ and let $\hat{w} = \mathbf{w}[\hat{\gamma}, \hat{P}]$. Let $(m, w) \in \mathcal{R}$ be such that $\|m\|_{L^\infty(Q)} \leq R$. There exists a constant $C(R)$ such that for any $(m, w) \in \mathcal{R}$,*

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

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 21, $(\mathbf{m}[\hat{\gamma}, \hat{P}], \mathbf{w}[\hat{\gamma}, \hat{P}])$ is the unique solution to Problem $(\mathcal{P}[\hat{\gamma}, \hat{P}])$, with $(\gamma, P) = (\hat{\gamma}, \hat{P})$.

Proof of Proposition 21. Lemma 20 yields $\int_Q |v - \hat{v}|^2 m \, dx \, dt \leq C\sigma$. By Lemma 17, we know that $\|\hat{v}\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C$, for some constant depending on R . Applying next Lemma 14, 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. \square

Corollary 22. *Let $R > 0$. Let $(m, w) \in \mathcal{R}$ be such that $\|m\|_{L^\infty(Q)} \leq R$. There exists a constant $C(R)$ such that*

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

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

Proof. By definition, $\bar{m} = \mathbf{m}[\hat{\gamma}, \bar{P}]$ and $\bar{w} = \mathbf{w}[\hat{\gamma}, \bar{P}]$. We obtain the announced estimates by combining Corollary 25 and Proposition 21. \square

5.3 Well-posedness of Algorithm 1 and proof of Theorem 7.

In this subsection the well-posedness of Algorithm 1 is established. The proof of Theorem 7 is a direct consequence of the well-posedness and the stability results established in the previous subsection.

Proposition 23. *Algorithm 1 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*

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

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

Proof. The well-posedness of the algorithm $\bar{m}_k, \bar{w}_k, \gamma_k, P_k, u_k, v_k, m_k$, and w_k can easily be established by induction. Lemma 18 shows the existence of a constant C_0 , independent of k such that $(\gamma_k, P_k) \in \Xi_{C_0}$, for any $k \in \mathbb{N}$. Applying next Proposition 16 and Lemma 17 with $R = C_0$, we deduce that $\|u_k\|_{W^{2,1,q}(Q)} + \|v_k\|_{\Theta} \leq C$ and $\|m_k\|_{W^{2,1,q}(Q)} + \|w_k\|_{\Theta} \leq C$, for some constant C independent of k . The inequality $\|\bar{m}_k\|_{W^{2,1,q}(Q)} + \|\bar{w}_k\|_{\Theta} \leq C$ follows then immediately from the triangle inequality. 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 $\mathcal{J}(m_k, w_k) \leq C$, since

$$\mathcal{J}(m_k, w_k) = \int_Q \mathbf{L}[v_k] m_k \, dx \, dt + \int_0^T (\mathbf{F}[m_k] + \Phi[A(m_k v_k)]) \, dt + \int_{\mathbb{T}^d} g m_k(T) \, dx.$$

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

$$\mathcal{J}(\bar{m}_{k+1}, \bar{w}_{k+1}) \leq (1 - \delta_k) \mathcal{J}(\bar{m}_k, \bar{w}_k) + \delta_k \mathcal{J}(m_k, w_k).$$

The proposition is proved. \square

Proof of Theorem 7 The first estimate concerning $\|\bar{m}_k - \bar{m}\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}$ and $\|\bar{w}_k - \bar{w}\|_{L^2(Q;\mathbb{R}^d)}$ is obtained by combining Corollary 22, 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 18. The third estimate on $\|u_k - \bar{u}\|_{L^\infty(Q)}$ follows from the second one and from Lemma 15. Using Proposition 23 and Lemma 1, there exists $C > 0$ such that $(\gamma_k, P_k) \in \Xi_C$ and $\|\bar{m}_k\|_{L^\infty(Q)} \leq C$, for all $k \in \mathbb{N}$. Moreover, by construction, $m_k = \mathbf{m}[\gamma_k, P_k]$ and $w_k = \mathbf{w}[\gamma_k, P_k]$ 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 21 applies and yields the last estimate that was to be proved.

5.4 Proof of Theorem 8

In this subsection we prove Theorem 8, leveraging techniques from [26].

Lemma 24. *Let (m_1, w_1) and (m_2, w_2) be in \mathcal{R} . Let $\gamma_1 = \gamma[m_1]$ and let $P_1 = \mathbf{P}[w_1]$. Then, there exists a constant $C > 0$, independent of (m_1, w_1) and (m_2, w_2) such that*

$$\int_Q \gamma_1(m_2 - m_1) dx dt + \int_0^T \langle A[w_2 - w_1], P_1 \rangle dt \leq \mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1) \quad (21)$$

and

$$\begin{aligned} \mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1) &\leq \int_Q \gamma_1(m_2 - m_1) dx dt + \int_0^T \langle A[w_2 - w_1], P_1 \rangle dt \\ &+ C_1 \int_0^T \|m_2(t, \cdot) - m_1(t, \cdot)\|_{L^2(\mathbb{T}^d)} \|m_2(t, \cdot) - m_1(t, \cdot)\|_{L^1(\mathbb{T}^d)} dt \\ &+ C_2 \int_0^T \left| \int_{\mathbb{T}^d} a(x, t)(w_2(x, t) - w_1(x, t)) dx \right|^2 dt. \end{aligned} \quad (22)$$

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

$$\begin{aligned} &(\mathcal{J}_2(m_2, w_2) - \mathcal{J}_2(m_1, w_1)) - \left(\int_Q \gamma_1(m_2 - m_1) dx dt + \int_0^T \langle A[w_2 - w_1], P_1 \rangle dt \right) \\ &= (a) + (b), \end{aligned}$$

where

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

Using Assumption (H6), we obtain that

$$(a) = \int_0^T \int_0^1 \int_{\mathbb{T}^d} (f[m_1 + s(m_2 - m_1)] - f[m_1])(m_2 - m_1) dx ds dt.$$

Then Assumptions (H5) and (H6) imply that

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

We estimate (b) in a similar way and obtain inequalities (21) and (22) easily. \square

Corollary 25. *Let (m_1, w_1) and (m_2, w_2) be in \mathcal{R} . Let $\gamma_1 = \gamma[m_1]$ and let $P_1 = \mathbf{P}[w_1]$. Then,*

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

Proof. This is an immediate consequence of inequality (21) from Lemma 24 and the definitions of \mathcal{J} and \mathcal{Z} . \square

Lemma 26. *For all $k \in \mathbb{N}$, it holds that $\varepsilon_k \leq \sigma_k$.*

Proof. By Corollary 25, we have for all $(m, w) \in \mathcal{R}$

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

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

Lemma 27. *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*

$$\mathcal{J}(\bar{m}_k^\delta, \bar{w}_k^\delta) \leq \mathcal{J}(\bar{m}_k, \bar{w}_k) - \delta\sigma_k + (C_1 D_k^{(1)} + C_2 D_k^{(2)})\delta^2, \quad (23)$$

where $(\bar{m}_k^\delta, \bar{w}_k^\delta)$ is defined by (12) and where $D_k^{(1)}$ and $D_k^{(2)}$ are defined by (17). Moreover, there exists a constant $C > 0$ such that

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

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 [26, Section 4.3].

Proof of Lemma 27. The proof relies on the decomposition $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ introduced in (1). First, by convexity of \tilde{L} , we have

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

Next, using inequality (22) of Lemma 24, we obtain that

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

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

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

Finally, by Theorem 7 we have

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

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

Lemma 28. *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 27 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

$$0 \leq \varepsilon_k - (1 - C\delta)\delta\sigma_k. \quad (29)$$

Increasing if necessary the value of C , 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. \square

Proof of Theorem 8 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 27 (inequality (24)) with the inequality $\varepsilon_k \leq \sigma_k$ and the bound on the exploitability obtained in Lemma 28, we obtain the following inequality:

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

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

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

Using next the inequality $\ln(x) \leq x - 1$, satisfied for any $x > 0$, we obtain the desired result, inequality (18).

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 (14), (15), or (16), 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 27,

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

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

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

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 τ^{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

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

The last inequality was obtained with Lemma 26.

Case of exploitability-based stepsizes. Let us set $a_k = (C_1 D_k^{(1)} + C_2 D_k^{(2)})$. By definition, $\delta_k = \min(1, \frac{\sigma_k}{2a_k})$. Assume that $\sigma_k \geq 2a_k$, i.e. $a_k \leq \sigma_k/2$. Then $\delta_k = 1$. Inequality (23) in Lemma 27 and Lemma 26 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 $\delta_k = \frac{\sigma_k}{2a_k}$. By (23), $\varepsilon_{k+1} \leq \varepsilon_k - \frac{\sigma_k^2}{4a_k}$. It follows from the last estimate of Theorem 7 that $a_k \leq C\sigma_k$. Therefore, by Lemma 26,

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

as was to be proved.

Case of an optimal stepsize. 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.

6 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 [33]. In this situation the GCG algorithm might not converge. In [22], the authors only show that any convergent sub-sequence generated by the fictitious play converges to an equilibrium. This problem can be tackled, assuming that solutions are stable [8]. Also very few results are available concerning the conditional gradient in the nonconvex setting. The article [27] 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 [27] should lead to the MFG system associated with the nonconvex potential formulation.
- 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 5.2 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 [15] 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 [6], 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 [14], 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 [11] and [12] 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] Y. Achdou and M. Laurière. Mean field games and applications: Numerical aspects. *Mean Field Games: Cetraro, Italy 2019*, pages 249–307, 2020.
- [2] J.-D. Benamou and G. Carlier. Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, Oct 2015.
- [3] J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna. An entropy minimization approach to second-order variational mean-field games. *Mathematical Models and Methods in Applied Sciences*, 29(08):1553–1583, 2019.
- [4] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active Particles, Volume 1*, pages 141–171. Springer, 2017.
- [5] J.F. Bonnans, S. Hadikhanloo, and L. Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Appl. Math. Optim.*, 83:1431–1464, 2021.
- [6] J.F. Bonnans, P. Lavigne, and L. Pfeiffer. Discrete potential mean field games: duality and numerical resolution. *Mathematical Programming*, pages 1–38, 2023.
- [7] K. Bredies, D.A. Lorenz, and P. Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *Computational Optimization and Applications*, 42(2):173–193, 2009.
- [8] A. Briani and P. Cardaliaguet. Stable solutions in potential mean field game systems. *Nonlinear Differential Equations Appl.*, 25(1):1, 2018.
- [9] L. Briceño-Arias, D. Kalise, Z. Kobeissi, M. Laurière, A.M. González, and F. 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.
- [10] L. Briceño-Arias, D. Kalise, and F. Silva. Proximal methods for stationary mean field games with local couplings. *SIAM J. Control Optim.*, 56(2):801–836, 2018.

- [11] S. Cacace, F. Camilli, and A. Goffi. A policy iteration method for mean field games. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:85, 2021.
- [12] F. Camilli and Q. Tang. Rates of convergence for the policy iteration method for mean field games systems. *Journal of Mathematical Analysis and Applications*, 512(1):126–138, 2022.
- [13] P. Cardaliaguet, P.J. Graber, A. Porretta, and D. Tonon. Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differential Equations Appl.*, 22(5):1287–1317, 2015.
- [14] P. Cardaliaguet and S. Hadikhanloo. Learning in mean field games: the fictitious play. *ESAIM:COCV*, 23(2):569–591, 2017.
- [15] P. Cardaliaguet and C.-A. Lehalle. Mean field game of controls and an application to trade crowding. *Math. Financ. Econ.*, 12(3):335–363, 2018.
- [16] R. Elie, J. Pérolat, M. Laurière, M. Geist, and O. Pietquin. Approximate fictitious play for mean field games. *arXiv preprint arXiv:1907.02633*, 2019.
- [17] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.
- [18] M. Geist, J. Pérolat, M. Laurière, R. Elie, S. Perrin, O. Bachem, R. Munos, and O. Pietquin. Concave utility reinforcement learning: the mean-field game viewpoint. *arXiv preprint arXiv:2106.03787*, 2021.
- [19] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Springer, 2015.
- [20] P.J. Graber and C. Mouzouni. Variational mean field games for market competition. In *PDE models for multi-agent phenomena*, pages 93–114. Springer, 2018.
- [21] P.J. Graber, A. Mullenix, and L. Pfeiffer. Weak solutions for potential mean field games of controls. *Nonlinear Differential Equations Appl.*, 28(5):1–34, 2021.
- [22] S. Hadikhanloo. Learning in anonymous nonatomic games with applications to first-order mean field games. *arXiv preprint arXiv:1704.00378*, 2017.
- [23] S. Hadikhanloo and F.J. Silva. Finite mean field games: Fictitious play and convergence to a first order continuous mean field game. *J. Math. Pures Appl.*, 132:369 – 397, 2019.
- [24] M. Huang, R.P. Malhamé, and P.E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.
- [25] M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *International Conference on Machine Learning*, pages 427–435. PMLR, 2013.

- [26] K. Kunisch and D. Walter. On fast convergence rates for generalized conditional gradient methods with backtracking stepsize. *Numerical Algebra, Control and Optimization*, pages 0–0, 2022.
- [27] S. Lacoste-Julien. Convergence rate of frank-wolfe for non-convex objectives. *arXiv preprint arXiv:1607.00345*, 2016.
- [28] O.A. Ladyzhenskaia, V.A. Solonnikov, and N.N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Math. Soc., 1988.
- [29] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.
- [30] J.-L. Lions. *Optimal control of systems governed by partial differential equations*, volume 170. Springer Verlag, 1971.
- [31] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer, 1999.
- [32] S. Perrin, J. Pérolat, M. Laurière, M. Geist, R. Elie, and O. Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. *arXiv preprint arXiv:2007.03458*, 2020.
- [33] F. Santambrogio and W. Shim. A cucker–smale inspired deterministic mean field game with velocity interactions. *SIAM J. Control Optim.*, 59(6):4155–4187, 2021.
- [34] S. Sorin. Continuous time learning algorithms in optimization and game theory. *Dynamic Games and Applications*, pages 1–22, 2022.
- [35] W. Wang, J. Han, Z. Yang, and Z. 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.

Statement and Declarations

This work was supported by a public grant as part of the Investissement d’avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH. The authors have no relevant financial or non-financial interests to disclose.

A Appendix: Regularity of the auxiliary mappings

This appendix contains the proofs of the technical lemmas of Subsection 5.1.

A.1 Parabolic estimates

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

$$\begin{aligned} \partial_t u - \sigma \Delta u + \langle b, \nabla u \rangle + cu &= h, & (x, t) \in Q, \\ u(x, 0) &= u_0(x), & x \in \mathbb{T}^d, \end{aligned} \tag{31}$$

for different assumptions on $b: Q \rightarrow \mathbb{R}^d$, $c: Q \rightarrow \mathbb{R}$, $h: \rightarrow \mathbb{R}$, and $u_0: \mathbb{T}^d \rightarrow \mathbb{R}$. The proofs of the following results can be found in the Appendix of [5]; they largely rely on [28]. 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 [28, section II.2] for a definition.

Theorem 29. *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 (31) has a unique solution u in $W^{2,1,q}(Q)$. Moreover, $\|u\|_{W^{2,1,q}(Q)} \leq C$.

Theorem 30. *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 (31) (with $b = 0$ and $c = 0$) satisfies the following estimate:*

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

Theorem 31. *For all $\beta \in (0, 1)$, for all $R > 0$, there exist $\alpha \in (0, 1)$ and $C > 0$ such that for all $u_0 \in C^{2+\beta}(\mathbb{T}^d)$, $b \in C^{\beta, \beta/2}(Q; \mathbb{R}^d)$, $c \in C^{\beta, \beta/2}(Q)$ and $h \in C^{\beta, \beta/2}(Q)$ satisfying $\|u_0\|_{C^{2+\beta}(\mathbb{T}^d)} + \|b\|_{C^{\beta, \beta/2}(Q; \mathbb{R}^d)} + \|c\|_{C^{\beta, \beta/2}(Q)} + \|h\|_{C^{\beta, \beta/2}(Q)} \leq R$, the solution to (31) lies in $C^{2+\alpha, 1+\alpha/2}(Q)$ and satisfies $\|u\|_{C^{2+\alpha, 1+\alpha/2}(Q)} \leq C$.*

A.2 Fokker-Planck equation

Proof of Lemma 13. Let us write the Fokker-Planck equation in the form of equation (31): $\partial_t m - \Delta m + (\nabla \cdot v)m + \langle v, \nabla m \rangle = 0$. The first part of lemma follows from Theorem 29. The nonnegativity of $M[v]$ is proved in [5, Lemma 3]. \square

Proof of Lemma 14. Set $w = v_2 - v_1$ and $\mu = m_2 - m_1$. Then μ is the solution to

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

Set $V = W^{1,2}(\mathbb{T}^d)$ 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

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

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

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

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:

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

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

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

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 [30, Chapter 3, Theorems 1.1 and 1.2], from which we derive

$$\begin{aligned} \|\mu\|_{L^\infty(0, T; L^2(\mathbb{T}^d))} &\leq C(\|\mu\|_{L^2(0, T; V)} + \|\partial_t \mu\|_{L^2(0, T; V^*)}) \\ &\leq C\|f\|_{L^2(0, T; V^*)} \leq C\|\nabla \cdot (wm_2)\|_{L^2(0, T; V^*)} \\ &\leq C\|wm_2\|_{L^2(Q; \mathbb{R}^d)}. \end{aligned}$$

Finally, since $\|m_2\|_{L^\infty(Q)} \leq R$, we have $\|wm_2\|_{L^2(Q; \mathbb{R}^d)}^2 \leq C \int_Q |w|^2 m_2 \, dx \, dt$. Combining the two last obtained inequalities, we obtain the announced result. \square

A.3 HJB equation

Lemma 32. *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 [5, Lemma 1]. \square

The analysis of the HJB equation relies on its connection with the value function of an optimal control problem, that was introduced in (11). This connection allows first to show a uniform bound for $\mathbf{u}[\gamma, P]$.

Lemma 33. *Let $R > 0$ and let $(\gamma, P) \in \Xi_R$. There exists a constant $C(R) > 0$ such that $\|\mathbf{u}[\gamma, P]\|_{L^\infty(Q)} \leq C(R)$ and such that u is $C(R)$ -Lipschitz continuous with respect to x . Moreover, for any $(x, t) \in Q$,*

$$\mathbf{u}[\gamma, P](x, t) = \inf_{\nu \in \mathbb{L}_{\mathbb{F}}^{2, C(R)}(t, T)} J[\gamma, P](x, t, \nu). \quad (32)$$

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

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

$$\frac{1}{C}|\nu|^2 - C \leq L(x, t, \nu), \quad \text{for all } (x, t, \nu) \in Q \times \mathbb{R}^d. \quad (33)$$

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

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

Now we show that $\mathbf{u}[\gamma, P]$ is bounded in $L^\infty(Q)$. For any $(x, t) \in Q$, using the above 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 $\mathbf{u}[\gamma, P](x, t) \geq -C(R)$. In addition, using Assumption (H3) and the fact that $\|\gamma\|_{L^\infty(Q)} \leq R$, we deduce that

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

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

Finally we show equation (32). Let $t \in [0, T]$, let $\varepsilon \in (0, 1)$ and let $\tilde{\nu} \in \mathbb{L}_{\mathbb{F}}^2(t, T)$ 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

$$\begin{aligned} \mathbb{E}\left[\int_t^T |\tilde{\nu}_s|^2 ds\right] &\leq C\left(\inf_{\nu \in \mathbb{L}_{\mathbb{F}}^2(t, T)} J[\gamma, P](x, t, \nu) + \varepsilon + 1\right) \\ &\leq C(\mathbf{u}[\gamma, P](x, t) + 2) \leq C, \end{aligned}$$

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

Proof of Lemma 15. Let (γ_1, P_1) and (γ_2, P_2) be in Ξ_R . Let $u_1 = \mathbf{u}[\gamma_1, P_1]$ and $u_2 = \mathbf{u}[\gamma_2, P_2]$. By Lemma 33, there exists $C > 0$ such that

$$u_2(x, t) - u_1(x, t) = \inf_{\nu \in \mathbb{L}_{\mathbb{F}}^{2,C}(t, T)} J[\gamma_2, P_2](x, t, \nu) - \inf_{\nu' \in \mathbb{L}_{\mathbb{F}}^{2,C}(t, T)} J[\gamma_1, P_1](x, t, \nu'),$$

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

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

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

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

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

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

as was to be proved. \square

We prove Proposition 16 with a density argument. In a nutshell: we prove in Proposition 34 below that the result of Proposition 16 holds true when γ and P are Hölder continuous. Then we pass to the limit, using Lemma 15.

Proposition 34. *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 (9) is a classical solution. Moreover, there exists $\alpha \in (0, 1)$ such that $\mathbf{u}[\gamma, P]$ lies in $\mathcal{C}^{2+\alpha, 1+\alpha/2}(Q)$ and there exists a constant $C(R)$, depending only on R , such that $\|\mathbf{u}[\gamma, P]\|_{W^{2,1,q}(Q)} \leq C$.*

The proof of Proposition 34 is given at page 32 and relies on a fixed point approach which requires some preparatory work. We introduce the map $\mathcal{T}: W^{2,1,q}(Q) \times$

$[0, 1] \rightarrow W^{2,1,q}(Q)$ 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

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

For any $(u, \tau) \in W^{2,1,q}(Q) \times [0, 1]$, we have $\tau(\gamma - \mathbf{H}[\nabla u + A^*P]) \in L^\infty(Q)$, by Lemma 32 and Lemma 1. It follows then from Theorem 29 that $\mathcal{T}[u, \tau]$ lies in $W^{2,1,q}(Q)$, proving that \mathcal{T} is well-defined.

Lemma 35. *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 $\nabla u_k \rightarrow \nabla u$ in $L^\infty(Q; \mathbb{R}^d)$ by Lemma 1. Then $\tau_k(\gamma - \mathbf{H}[\nabla u_k + A^*P]) \rightarrow \tau(\gamma - \mathbf{H}[\nabla u + A^*P])$ in $L^\infty(Q; \mathbb{R}^d)$ by continuity of the Hamiltonian (see Lemma 32). Finally \mathcal{T} is continuous, by Theorem 30.

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 32 there exist $\alpha \in (0, 1)$ and $C > 0$ such that $\|\gamma - \mathbf{H}[\nabla u + A^*P]\|_{\mathcal{C}^\alpha(Q)} \leq C$. Then applying Theorem 31, 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. \square

Theorem 36. (Leray-Schauder) *Let X be a Banach space and let $T : X \times [0, 1] \rightarrow X$ 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 [19, Theorem 11.6]. \square

Proof of Proposition 34. 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 $\mathbf{u}[\gamma, P]$. 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 35, 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

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

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

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

where $(X_s^T)_{s \in [t, T]}$ is the solution to $dX_s = \tau \nu_s ds + \sqrt{2} dB_s$, $X_0 = Y$. Following [5, 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 32 and recalling that $(\gamma, P) \in \Xi_R$, we deduce that $\|\mathbf{H}[\nabla u + A^*P] - \gamma\|_{L^\infty(Q)} \leq C$. 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 29. Again, C only depends on R . Finally, by the Leray-Schauder theorem (Theorem 36), there exists a solution to $u = \mathcal{T}[u, 1]$, which is necessarily $\mathbf{u}[\gamma, P]$. \square

Proof of Proposition 16. 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 $\|\gamma_n - \gamma\|_{L^\infty(Q)} \rightarrow 0$ and such that $\|P_n - P\|_{L^2(0, T; \mathbb{R}^k)} \rightarrow 0$. We do not detail the construction of such a sequence, this can be done by convolution. Define $u^n = \mathbf{u}[\gamma^n, P^n]$ and $u = \mathbf{u}[\gamma, P]$. By Lemma 15, $u_n \rightarrow u$ for the L^∞ -norm. Moreover, by Proposition 34,

$$\|u^n\|_{W^{2,1,q}(Q)} \leq C(R), \quad \forall n \in \mathbb{N}. \quad (34)$$

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 (34). This concludes the proof. \square

A.4 The other mappings

Proof of Lemma 17. Let $(\gamma, P) \in \Xi_R$. Let $u = \mathbf{u}[\gamma, P]$. We already know from Proposition 16 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 $v = \mathbf{v}[\gamma, P] = -\mathbf{H}_p[\nabla u + A^*P]$. We have

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

Using the regularity of u , the regularity properties of the Hamiltonian given in Lemma 20, 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 $m = \mathbf{m}[\gamma, P] = \mathbf{M}[v]$. A direct application of Lemma 13 yields that $\|m\|_{W^{2,1,q}(Q)} \leq C$. Finally, let $w = \mathbf{w}[\gamma, P] = mv$. Using again Lemma 1, we obtain that m is continuous and that $\|m\|_{L^\infty(Q)} \leq C(R)$ and $\|\nabla m\|_{L^\infty(Q; \mathbb{R}^d)} \leq C(R)$. Then $w \in \Theta$, with a norm bounded by some constant $C(R)$. The lemma is proved. \square

Proof of Lemma 18. The two statements concerning γ are directly deduced from Assumptions (H3) and (H5). Let $w \in \Theta$. Recalling the definition of the operator A (page 7), 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 $\mathbf{P}[w] = \phi[A[w]]$ lies in $\mathcal{C}(0, T; \mathbb{R}^k)$ and that $\|\mathbf{P}[w]\|_{L^\infty(0, T; \mathbb{R}^k)} \leq C$. Let us next consider w_1 and w_2 in Θ . We have

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

by Assumption (H4). Using next the Lipschitz-continuity of ϕ (Assumption (H5)), we obtain that $\|\phi[Aw_2] - \phi[Aw_1]\|_{L^2(0, T; \mathbb{R}^k)} \leq C \|w_2 - w_1\|_{L^2(Q; \mathbb{R}^d)}$, as was to be proved. \square