



On mean field games in infinite dimension

Salvatore Federico^a, Fausto Gozzi^{b,*}, Andrzej Święch^c^a *Dipartimento di Matematica, Università "Alma Mater" di Bologna, Piazza di Porta San Donato, Bologna, Italy*^b *Dipartimento di Economia e Finanza, LUISS University, Viale Romania 32, Rome, Italy*^c *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA*

ARTICLE INFO

Article history:

Received 21 November 2024

Available online 28 August 2025

MSC:

35Q89

35R15

49N80

35Q84

49L12

60H15

91A16

Keywords:

Mean field games

PDE in infinite dimensional spaces

Hamilton-Jacobi-Bellman equation

Fokker-Planck equation

Wasserstein space

Stochastic differential equation in infinite dimension

ABSTRACT

We study a Mean Field Games (MFG) system in a real, separable infinite dimensional Hilbert space. The system consists of a second order parabolic type equation, called Hamilton-Jacobi-Bellman (**HJB**) equation in the paper, coupled with a nonlinear Fokker-Planck (**FP**) equation. Both equations contain a Kolmogorov operator. Solutions to the HJB equation are interpreted in the mild solution sense and solutions to the FP equation are interpreted in an appropriate weak sense. We prove well-posedness of the considered MFG system under certain conditions. The existence of a solution to the MFG system is proved using Tikhonov's fixed point theorem in a proper space. Uniqueness of solutions is obtained under typical separability and Lasry-Lions type monotonicity conditions.

© 2025 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

R É S U M É

Nous étudions un système de jeux à champ moyen (Mean Field Games, MFG) dans un espace de Hilbert réel, séparable et de dimension infinie. Le système est composé d'une équation parabolique du second ordre, appelée équation de Hamilton-Jacobi-Bellman (**HJB**), couplée à une équation de Fokker-Planck (**FP**) non linéaire. Les deux équations contiennent un opérateur de Kolmogorov. Les solutions de l'équation HJB sont interprétées au sens des solutions *mild*, tandis que les solutions de l'équation FP sont interprétées dans un sens faible approprié. Nous démontrons l'existence et l'unicité du système MFG sous certaines conditions. L'existence d'une solution est obtenue à l'aide du théorème du point fixe de Tikhonov dans un espace approprié. L'unicité est assurée sous des hypothèses classiques de séparabilité et de monotonie de type Lasry-Lions.

© 2025 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail addresses: s.federico@unibo.it (S. Federico), fgozzi@luiss.it (F. Gozzi), swiech@math.gatech.edu (A. Święch).

1. Introduction

In this paper, we are concerned with a class of Mean Field Games (MFG) problems in a real separable Hilbert space $(H, \langle \cdot, \cdot \rangle)$. The analytical framework is the following. Let $\mathcal{P}_1(H)$ denote the space of Borel probability measures on H having finite first moment and let $T > 0$. We consider the following system of coupled Hamilton-Jacobi-Bellman (**HJB**) and nonlinear Fokker-Planck (**FP**) equations for functions $v : [0, T] \times H \rightarrow \mathbb{R}$ and $m : [0, T] \rightarrow \mathcal{P}_1(H)$,

$$-\partial_t v(t, x) - Lv(t, x) + \mathcal{H}(x, Dv(t, x), m(t)) = 0, \quad v(T, \cdot) = G(\cdot, m(T)), \quad (\text{HJB})$$

$$\partial_t m(t) - L^* m(t) - \operatorname{div} \left(\mathcal{H}_p(x, Dv(t, x), m(t)) m(t) \right) = 0, \quad m(0) = m_0, \quad (\text{FP})$$

where

- (i) $m_0 \in \mathcal{P}_1(H)$, $G : H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}$;
- (ii) L is the operator formally defined on functions $\phi : H \rightarrow \mathbb{R}$ by

$$L\phi(x) = \langle Ax, D\phi(x) \rangle + \frac{1}{2} \operatorname{Tr}[D^2\phi(x)],$$

where $A : D(A) \subseteq H \rightarrow H$ is a closed and densely defined linear operator;

- (iii) L^* is, formally, the adjoint of L and the operator **div** is, formally, the opposite of the adjoint of the gradient: both are used here only for notational convenience as the rigorous definition of a solution to (**FP**) will not involve them, see Section 3 below;
- (iv) $\mathcal{H} : H \times H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}$ and $\mathcal{H}_p$ denotes the Fréchet derivative of $\mathcal{H}$ with respect to the second argument.

We point out that we slightly abuse terminology here, since the first equation is not necessarily a Hamilton-Jacobi-Bellman equation as the function $\mathcal{H}$ can be very general. However, since our primary motivation comes from this case, we will call it an HJB equation.

The theory of MFG originated from the pioneering works of Lasry and Lions, and of Huang, Caines, and Malhamé (see [33–36]). This theory was primarily developed to provide an analytical framework for (stochastic) differential games involving a large number of symmetric players. In the limit, as the number of players approaches infinity and assuming that the actions of individual players do not directly influence the overall population dynamics, the concept of Nash equilibrium leads to the study of a coupled system of PDEs of the type given by equations (**HJB**) and (**FP**). Since the inception of this theory, substantial progress has been made in finite-dimensional settings, both theoretically and practically. Notable references for this research area include the book [13] and the survey paper [11]. However, contributions to the theory of Mean Field Games (MFG) in infinite-dimensional spaces remain surprisingly limited. To our knowledge, the literature consists of only a few essential papers, such as [26], which addresses systemic risk by taking the limit of the N -player model in [14], and recent works [24, 37], both of which focus on the linear-quadratic case. This scarcity is somewhat in contrast with the established theory of stochastic optimal control in infinite-dimensional spaces and its connection with Hamilton-Jacobi-Bellman (**HJB**) equations, which has been well developed over several decades, beginning with foundational works by Barbu and Da Prato (see, e.g., [3]). For a general overview of this subject and its applications, we refer to the recent monograph [22], and see also [20, 39].

In this paper, we aim to bridge this gap by providing a preliminary contribution to the study of MFG in infinite dimensions beyond the linear-quadratic case. Building upon existing results for both HJB and

Fokker-Planck equations in infinite-dimensional settings, we investigate the nonlinear coupled system of MFG equations **(HJB)**–**(FP)**. Our main contribution is the proof of well-posedness of this system under certain conditions. The HJB equation **(HJB)** is interpreted in the mild form, allowing us to use well-known theories of mild solutions to partial differential equations in Hilbert spaces. However, this approach requires stronger assumptions on the operator A to ensure certain smoothing properties of the associated Ornstein-Uhlenbeck transition semigroup.

The Fokker-Planck equation **(FP)** is interpreted in a weak form using a special class of test functions. Existence of a solution to the system **(HJB)**–**(FP)** is established with a typical MFG fixed-point argument via Tikhonov's fixed-point theorem, which requires the existence of a unique weak solution to the linear **(FP)** when the term $\mathcal{H}_p(x, Dv(t, x), m(t))$ is fixed. Fokker-Planck-type equations in Hilbert spaces have been widely studied, with various results available on existence and uniqueness (see [6,8,7,5,18,41,43]). To show that our linear **(FP)** has a weak solution, we follow a common approach of linking such an equation to a stochastic differential equation (SDE) in H and noting that the laws of solutions to this SDE are weak solutions of the linear **(FP)**. We apply results from [17] on SDEs in Hilbert spaces with bounded measurable drift, adapting them to our setting. Uniqueness of weak solutions then follows easily from a result in [7], chosen here for its clarity and accessibility. However, this approach imposes additional restrictions on the operator A , ultimately leading to fairly strict assumptions on A .

Uniqueness of solutions to the MFG system **(HJB)**–**(FP)** is obtained under typical monotonicity and separability conditions on $\mathcal{H}$ (see [11] and [36]). To carry out the uniqueness argument, which is now standard in the finite-dimensional case, we approximate the mild solution of the Kolmogorov equation obtained by subtracting two different solutions to the **(HJB)** equations in the **(HJB)**–**(FP)** system by constructing smooth solutions of approximating Kolmogorov equations that belong to the class of test functions used in the definition of a weak solution for the Fokker-Planck equation. Since this is the first study on the subject, we chose not to address the most general case to avoid obscuring the main ideas with excessive technical detail. A possible generalization could involve a more flexible operator L , in which the term $\text{Tr}[D^2\phi(x)]$, associated with the cylindrical Wiener process W in the underlying SDE, is replaced by $\text{Tr}[QD^2\phi(x)]$ for some self-adjoint operator Q (corresponding to a more general additive noise in the SDE). This could potentially relax the assumptions on A . Another interesting direction would be to interpret equation **(HJB)** in the viscosity sense, which could also include the case of first-order MFG. We plan to investigate this in future work.

Regarding possible applications, it would be interesting to try to relax the assumptions on the operator A to cover potentially interesting cases. In particular, it would be interesting to treat the case of first order operators A , some of which appear in financial and economic applications, such as models of systemic risk with delays (as in [26]) or models of optimal investment with vintage capital (see [23,25] in the case of a single firm with vintage structure; [10] in the case of a continuum of homogenous firms but without vintage structure).

The plan of the paper is the following. In Section 2, we introduce the basic notation and assumptions, presenting preliminary technical results on the well-posedness of the SDE related to the (linear) Fokker-Planck equation, along with some estimates for its solutions. Section 3 is dedicated to the existence and uniqueness of weak solutions for the Fokker-Planck equation (3.1). In Section 4, we define the notion of a mild solution to **(HJB)** and collect some foundational results on mild solutions. In Section 5, we show the existence of solutions to the MFG system **(HJB)**–**(FP)**, while the uniqueness of solutions is established in Section 6. An example that satisfies the manuscript's assumptions is discussed in Section 7. Finally, in Appendix A, we present a result on the compactness of sets in spaces of probability measures on a Hilbert space, and in Appendix B, we discuss essential concepts and results related to Kolmogorov equations in Hilbert spaces that are needed in this paper.

2. Preliminaries and assumptions

We recall that H is a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$. We will write $|\cdot|$ for the norm in H ; if needed we stress H in the notation writing $|\cdot|_H$. We always identify H with its dual space. For $R > 0$, we denote $B_R := \{x \in H : |x| \leq R\}$. Throughout the paper, $\mathbb{N}$ will denote the set of natural numbers without 0.

2.1. Basic spaces and notations

We introduce the notation for various functional and operator spaces used in the paper. Below, by I we denote an interval in $\mathbb{R}$. Given a function $u : I \times H \rightarrow \mathbb{R}, (t, x) \mapsto u(t, x)$, we will write $\partial_t u$ for the partial derivative of u with respect to t and Du, D^2u , respectively, for the first and second order Fréchet derivatives of u with respect to the x -variable. For a bounded function between normed spaces, we denote its sup-norm in the usual way by $\|\cdot\|_\infty$.

Below is a list of basic spaces used in the paper.

- (i) $\mathcal{L}(H)$ is the space of bounded linear operators from H to H .
- (ii) $\mathcal{L}^+(H)$ is the space of bounded linear self-adjoint nonnegative operators from H to H .
- (iii) $\mathcal{L}_1(H)$ is the space of bounded linear trace-class operators from H to H .
- (iv) $\mathcal{L}_1^+(H) = \mathcal{L}^+(H) \cap \mathcal{L}_1(H)$.
- (v) $C_b(H)$ is the space of continuous and bounded functions $u : H \rightarrow \mathbb{R}$.
- (vi) $C_b(I \times H)$ is the space of continuous and bounded functions $u : I \times H \rightarrow \mathbb{R}$.
- (vii) $C_b(I \times H; H)$ is the space of continuous and bounded functions $u : I \times H \rightarrow H$.
- (viii) $C_b(I \times H; \mathcal{L}(H))$ (respectively, $C_b(I \times H; \mathcal{L}_1(H))$) is the space of continuous and bounded functions $u : I \times H \rightarrow \mathcal{L}(H)$ (respectively, $u : I \times H \rightarrow \mathcal{L}_1(H)$).
- (ix) $C_b^{0,1}(I \times H)$ is the space of continuous and bounded functions $u : I \times H \rightarrow \mathbb{R}$, which are Fréchet differentiable with respect to the variable x and such that $Du \in C_b(I \times H; H)$.
- (x) $UC_b(H)$ (respectively, $UC_b(H; H)$, $UC_b(H; \mathcal{L}_1(H))$) is the space of bounded and uniformly continuous functions $u : H \rightarrow \mathbb{R}$ (respectively, $u : H \rightarrow H$, $u : H \rightarrow \mathcal{L}_1(H)$).
- (xi) $UC_b^2(H) = \{u : H \rightarrow \mathbb{R} : u \in UC_b(H), Du \in UC_b(H; H), D^2u \in UC_b(H; \mathcal{L}(H))\}$.
- (xii) Finally,

$$C_b^{1,2}(I \times H) = \{u \in C_b(I \times H) : \partial_t u \in C_b(I \times H), Du \in C_b(I \times H; H), D^2u \in C_b(I \times H; \mathcal{L}(H))\}.$$

We denote by $\mathcal{M}(H)$ the topological vector space of finite signed measures on H endowed with the weak topology, i.e., the locally convex topology on $\mathcal{M}(H)$ induced by the family of seminorms

$$|\mu|_f := \left| \int_H f(x) \mu(dx) \right|, \quad f \in C_b(H). \quad (2.1)$$

By $\mathcal{P}(H)$ we denote the space of Borel probability measures on H . The weak topology defined on it is the one inherited by the weak topology of $\mathcal{M}(H)$ defined above. We denote by $\mathcal{P}_1(H)$ the subset of $\mathcal{P}(H)$ of the probability measures having finite first moment. In $\mathcal{P}_1(H)$ we have the Monge-Kantorovich distance (see, e.g., [2,42])

$$\mathbf{d}_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{H \times H} |x - y| \gamma(dx, dy), \quad (2.2)$$

where $\Gamma(\mu, \nu)$ is the set of all Borel probability measures on $H \times H$ with first and second marginals μ and ν , respectively. Unless differently specified, the space $\mathcal{P}_1(H)$ will be considered endowed with the topology induced by $\mathbf{d}_1$ which makes it a complete metric space. We notice that, if $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $X, Y \in L^1(\Omega; H)$ are such that $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$, then $\gamma = (X \times Y)_\# \mathbf{P} \in \Gamma(\mu, \nu)$, so that

$$\mathbf{d}_1(\mu, \nu) \leq \int_{H \times H} |x - y| \gamma(dx, dy) = \int_{\Omega} |X(\omega) - Y(\omega)| d\mathbf{P}(\omega) = \mathbb{E}[|X - Y|].$$

We will deal with the space

$$\mathcal{S} = C([0, T]; (\mathcal{P}_1(H), \mathbf{d}_1)). \quad (2.3)$$

It is a complete metric space endowed with the metric

$$\rho_\infty(\mu, \nu) = \sup_{t \in [0, T]} \mathbf{d}_1(\mu(t), \nu(t)).$$

We will also use the symbol ρ_∞ to denote the induced metric topology.

2.2. Weighted spaces

Given any $\gamma > 0$, we define

$$\begin{aligned} C_{b,\gamma}([0, T] \times H) &:= \{f: [0, T] \times H \rightarrow \mathbb{R} \text{ s.t. } f \in C_b([0, \tau] \times H) \forall \tau \in (0, T) \\ &\text{and } (t, x) \mapsto (T - t)^\gamma f(t, x) \text{ belongs to } C_b([0, T] \times H)\}. \end{aligned} \quad (2.4)$$

Similarly, we define

$$\begin{aligned} C_{b,\gamma}([0, T] \times H; H) &:= \{w: [0, T] \times H \rightarrow H \text{ s.t. } f \in C_b([0, \tau] \times H; H) \forall \tau \in (0, T) \\ &\text{and } (t, x) \mapsto (T - t)^\gamma f(t, x) \text{ belongs to } C_b([0, T] \times H; H)\}. \end{aligned} \quad (2.5)$$

The spaces above are Banach spaces when endowed with the norm (see e.g. [27,30] and [15])

$$\begin{aligned} \|f\|_{C_{b,\gamma}([0, T] \times H)} &:= \sup_{(t,x) \in [0, T] \times H} (T - t)^\gamma |f(t, x)|, \\ \|f\|_{C_{b,\gamma}([0, T] \times H; H)} &:= \sup_{(t,x) \in [0, T] \times H} (T - t)^\gamma |f(t, x)|_H. \end{aligned}$$

We will often simply write $\|f\|_{C_{b,\gamma}}$. Moreover, we define

$$\begin{aligned} C_{b,\gamma}^{0,1}([0, T] \times H) &:= \left\{ f \in C_b([0, T] \times H) : f \in C_b^{0,1}([0, \tau] \times H) \forall \tau \in (0, T) \right. \\ &\left. \text{and } (t, x) \mapsto (T - t)^\gamma Df(t, x) \text{ belongs to } C_b([0, T] \times H; H) \right\}. \end{aligned} \quad (2.6)$$

The space above is a Banach space when endowed with the norm (see e.g. [27,30] and [15])

$$\|f\|_{C_{b,\gamma}^{0,1}([0, T] \times H)} := \|f\|_\infty + \sup_{(t,x) \in [0, T] \times H} (T - t)^\gamma |Df(t, x)|_H.$$

We will sometimes simply write $\|f\|_{C_{b,\gamma}^{0,1}}$.

2.3. The operator A and the generated semigroup

Concerning the operator A , we assume the following.

Assumption 2.1. $A : D(A) \subseteq H \rightarrow H$ is closed, densely defined, negative and self-adjoint. Moreover there exists $\delta > 0$ such that $(-A)^{-1+\delta} \in \mathcal{L}_1^+(H)$.

Given Assumption 2.1, A^{-1} is compact and self-adjoint. Therefore, there exist an orthonormal basis $\{\mathbf{e}_k\}_{k \in \mathbb{N}}$ and a decreasing sequence $(\lambda_k)_{k \in \mathbb{N}} \subset (-\infty, 0)$ such that $\lambda_n \rightarrow -\infty$, with respect to which A is diagonal:

$$A\mathbf{e}_k = \lambda_k \mathbf{e}_k, \quad k \in \mathbb{N},$$

and

$$\sum_{k \in \mathbb{N}} |\lambda_k|^{-1+\delta} < \infty. \quad (2.7)$$

Moreover, A generates an analytic C_0 -semigroup of contractions $\{e^{tA}\}_{t \geq 0} \subseteq \mathcal{L}(H)$ and

$$e^{tA} \mathbf{e}_k = e^{\lambda_k t} \mathbf{e}_k, \quad k \in \mathbb{N}. \quad (2.8)$$

Remark 2.2.

- (i) An example of A that satisfies the above Assumption 2.1 is the Laplace operator ∂_{xx} in $H = L^2([0, 1])$ with zero Dirichlet boundary condition.
- (ii) It can be verified from (2.7) that Assumption 2.1 implies typical conditions on the operator A ensuring existence and uniqueness of mild solutions of equations in infinite dimensional spaces (see [22, Ch. 4]), that is:

- (a) One has

$$Q_t := \int_0^t e^{sA} e^{sA^*} ds \in \mathcal{L}_1^+(H) \quad \forall t > 0$$

and there exists $\alpha > 0$ such that

$$Q_t^{(\alpha)} := \int_0^t s^{-\alpha} e^{sA} e^{sA^*} ds \in \mathcal{L}_1^+(H) \quad \forall t > 0.$$

- (b) $e^{tA}(H) \subset Q_t^{1/2}(H)$ for every $t > 0$ and there exist $C_0 > 0$ and $\gamma \in (0, 1)$ such that

$$|Q_t^{-1/2} e^{tA}| \leq C_0 t^{-\gamma}, \quad \forall t \geq 0,$$

where $Q_t^{-1/2}$ is the pseudoinverse of $Q_t^{1/2}$ (see [22, Def. B.1]).

2.4. The probabilistic framework and the reference SDE

We now define a probabilistic framework. We consider a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ supporting a sequence $\{\beta_k(t)\}_{k \in \mathbb{N}}$ of independent one dimensional Brownian motions. Given the orthonormal basis $\{\mathbf{e}_k\}_{k \in \mathbb{N}}$ of Assumption 2.1, we consider the cylindrical Wiener process $\{W_t\}_{t \geq 0}$, formally written as

$$W_t = \sum_{k \in \mathbb{N}} \beta_k(t) \mathbf{e}_k.$$

Under Assumption 2.1, the stochastic convolution (see Remark 2.2(ii) and [21, Ch. 5])

$$W_t^A := \int_0^t e^{(t-s)A} dW_s, \quad t \geq 0, \quad (2.9)$$

is a well-defined centered Gaussian process in the Hilbert space H with covariance operator

$$Q_t := \int_0^t e^{2sA} ds \in \mathcal{L}_1^+(H)$$

admitting a continuous version. The process

$$X_t^{0,x} := e^{tA}x + W_t^A, \quad t \geq 0,$$

can be seen as a generalized (called *mild*) solution to the SDE

$$dX_t = AX_t dt + dW_t, \quad X_0 = x \in H.$$

$X_t^{0,x}$ is called the Ornstein-Uhlenbeck process and is associated to the Markov transition semigroup

$$[R_t \varphi](x) = \mathbb{E} \left[\varphi(X_t^{0,x}) \right] = \int_H \varphi(y) \mathcal{N}(e^{tA}x, Q_t)(dy), \quad \varphi \in C_b(H), \quad (2.10)$$

called the Ornstein-Uhlenbeck semigroup. Above, $\mathcal{N}(a, Q)$ is the Gaussian measure in H with mean $a \in H$ and covariance operator $Q \in \mathcal{L}_1^+(H)$ (see, e.g., [20, Ch.2, Sec.3]). We will use it later to define the concept of a mild solution to (HJB).

To construct a weak solution to a linear Fokker-Planck equation we will consider the SDE

$$dX_t = AX_t dt + w(t, X_t) dt + dW_t, \quad \mathcal{L}(X_0) = m_0. \quad (2.11)$$

We will need the following assumption.

Assumption 2.3.

- (i) $\int_H |x|^4 m_0(dx) < \infty$.
- (ii) $w \in C_b((0, T) \times H; H)$.

Theorem 2.4. Let $m_0 \in \mathcal{P}_1(H)$ and let Assumptions 2.1 and 2.3(ii) hold. There exists a unique weak-mild solution to (2.11) in the following sense.

- (i) There exist a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbf{P})$, an $\mathcal{F}_0$ -measurable random variable X_0 with $\mathcal{L}(X_0) = m_0$, a cylindrical Wiener process $(W_t)_{t \in [0, T]}$ with respect to $(\mathcal{F}_t)_{t \in [0, T]}$, and an adapted process $(X_t^w)_{t \in [0, T]}$ such that

$$X_t^w = e^{tA} X_0 + \int_0^t e^{(t-s)A} w(s, X_s^w) ds + W_t^A, \quad \forall t \geq 0;$$

- (ii) If another weak-mild solution $\widehat{X}^w$ exists, then $\mathcal{L}(X_t^w) = \mathcal{L}(\widehat{X}_t^w)$ for every $t \in [0, T]$.

Proof. When w does not depend on t , the result is proved [17, Th. 13] and the proof works in the same way if w depends on t . We just need to explain how our setup fits into the framework of [17].

Given our filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ with the Wiener process W , let $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ be another complete atomless probability space. For any $m_0 \in \mathcal{P}_1(H)$, since $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ is nonatomic, there exists an $\mathcal{F}_1$ -measurable random variable $\xi_0 : \Omega_1 \rightarrow X$ such that $m_0 = \mathcal{L}(\xi_0)$ (see, e.g., [12, Lemma 5.29]). We define a new probability space $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ as follows:

$$\begin{aligned} \Omega_2 &= \Omega \times \Omega_1, \quad \mathbf{P}_2 = \mathbf{P} \otimes \mathbf{P}_1 \text{ on } \mathcal{F} \otimes \mathcal{F}_1, \\ \mathcal{F}_2 &= \text{completion of } \mathcal{F} \otimes \mathcal{F}_1 \text{ with respect to } \mathbf{P}_2 = \mathbf{P} \otimes \mathbf{P}_1. \end{aligned}$$

We then define $(\mathcal{F}_t^2)$ to be the augmentation of $\mathcal{F}_t \otimes \mathcal{F}_1$. The cylindrical Wiener process $\{W_t\}_{t \in [0, T]}$ and the random variable ξ_0 extend naturally to the space $(\Omega_2, \mathcal{F}_2, (\mathcal{F}_t^2)_{t \in [0, T]}, \mathbf{P}_2)$. We denote them by W^2 and X_0 respectively. The random variable $X_0 : \Omega \rightarrow H$ is $\mathcal{F}_0^2$ -measurable, $m_0 = \mathcal{L}(X_0)$, and $\{W_t^2\}_{t \in [0, T]}$ is a cylindrical Wiener process in the filtered probability space $(\Omega_2, \mathcal{F}_2, (\mathcal{F}_t^2)_{t \in [0, T]}, \mathbf{P}_2)$.

Now, following the arguments of [17], the weak-mild solution is constructed starting from the space $(\Omega_2, \mathcal{F}_2, (\mathcal{F}_t^2)_{t \in [0, T]}, \mathbf{P}_2)$ using the Girsanov Theorem by introducing a new cylindrical Wiener process $\{\tilde{W}_t^2\}_{t \in [0, T]}$ and replacing $\mathbf{P}_2$ by $\tilde{\mathbf{P}}_2$, where

$$d\tilde{\mathbf{P}}_2 = M_T d\mathbf{P}_2,$$

and $(M_s)_{s \in [0, T]}$ is an appropriate martingale such that $M_0 \equiv 1$. If $f \in C_b(H)$, we then have

$$\tilde{\mathbf{E}}[f(X_0)] = \mathbf{E}[M_T f(X_0)] = \mathbf{E}[\mathbf{E}[M_T f(X_0) | \mathcal{F}_0^2]] = \mathbf{E}[f(X_0) \mathbf{E}[M_T | \mathcal{F}_0^2]] = \mathbf{E}[f(X_0)].$$

Thus, the law of X_0 in the new probability space is the same. This puts us in the framework of [17] and we can now proceed as in the proof of [17, Th. 13]. $\square$

We provide some estimates for the laws $\mathcal{L}(X_t^w)$ of the solution to (2.11) that will be used afterwards.

Proposition 2.5. Let Assumptions 2.1 and 2.3 hold and let X_t^w be the unique weak-mild solution to (2.11).

- (i) There exists c_0 independent of $\|w\|_\infty$ such that

$$\int_H |x|^4 \mathcal{L}(X_t^w)(dx) \leq c_0 \left(1 + \int_H |x|^4 m_0(dx) + \|w\|_\infty^4 \right), \quad \forall t \in [0, T].$$

(ii) For each $R > 0$, there exists a modulus of continuity $\omega_{m_0, R}$ such that, if $|w|_\infty \leq R$, it holds

$$\mathbf{d}_1(\mathcal{L}(X_t^w), \mathcal{L}(X_s^w)) \leq \omega_{m_0, R}(|t - s|), \quad \forall s, t \in [0, T].$$

Proof. (i) Using the definition of mild solution, the fact that $|e^{tA}| \leq 1$ and the estimate on the stochastic convolution W_t^A [22, Prop. 1.144], we have, for some $c_1 > 0$,

$$\begin{aligned} \mathbb{E}[|X_t^w|^4] &\leq \mathbb{E}\left[\left|e^{tA}X_0 + \int_0^t e^{(t-s)A}w(s, X_s^w)ds + W_t^A\right|^4\right] \\ &\leq c_1(\mathbb{E}[|X_0|^4] + T^4\|w\|_\infty^4 + 1), \quad \forall 0 \leq t \leq T, \end{aligned}$$

which implies (i).

(ii) Let $0 \leq s \leq t \leq T$. Using the definition of mild solution, we have the existence of some $c_1 > 0$, independent of s, t such that

$$\begin{aligned} \mathbf{d}_1(\mathcal{L}(X_t^w), \mathcal{L}(X_s^w)) &\leq \mathbb{E}[|X_t^w - X_s^w|] \leq c_1 \left(\mathbb{E}\left[\left|(e^{(t-s)A} - I)X_0\right|\right] \right. \\ &\quad + \mathbb{E}\left[\left|(e^{(t-s)A} - I) \int_0^s e^{(s-r)A}w(r, X_r^w)dr\right| + \int_s^t \left|e^{(t-r)A}w(r, X_r^w)\right|dr\right] \\ &\quad \left. + \mathbb{E}[|W_t^A - W_s^A|] \right). \end{aligned}$$

We proceed to estimate the terms in the right hand side.

(a) Recalling that $|e^{\alpha A}| \leq 1$ for $\alpha \geq 0$, we have, for every $0 \leq s \leq t \leq T$,

$$\mathbb{E}\left[\int_s^t \left|e^{(t-r)A}w(r, X_r^w)\right|dr\right] \leq (t-s)\|w\|_\infty.$$

(b) Considering that $(e^{tA})_{t \in [0, T]}$ is strongly continuous and using dominated convergence, we have

$$\mathbb{E}\left[\left|(e^{(t-s)A} - I)X_0\right|\right] \rightarrow 0, \quad \text{as } s \rightarrow t.$$

Hence, there exists a modulus of continuity $\omega_{m_0, A}$, independent of w such that, for every $0 \leq s \leq t \leq T$,

$$\mathbb{E}\left[\left|(e^{(t-s)A} - I)X_0\right|\right] \leq \omega_{m_0, A}(|t - s|).$$

(c) Let $R > 0$. Since A generates a semigroup of contractions, for every $0 < \eta < s \leq T$ and w such that $|w|_\infty \leq R$, we have (P-a.s.)

$$\int_0^{s-\eta} e^{(s-r)A}w(r, X_r^w)dr \in e^{\eta A}(B_{TR}) =: K_{\eta, R}.$$

Assumption 2.1 guarantees that the semigroup is compact so $K_{\eta, R}$ is compact. Considering also that the semigroup is strongly continuous, we have for $0 \leq s \leq t \leq T$

$$\sup_{x \in K_{\eta,R}} |(e^{(t-s)A} - I)x| \leq \omega_{\eta,A,R}(t-s), \quad (2.12)$$

for some modulus of continuity $\omega_{\eta,A,R}$. Using (2.12), it follows that, for every $0 \leq \eta \leq s \leq t \leq T$

$$\begin{aligned} & \mathbb{E} \left[\left\| (e^{(t-s)A} - I) \int_0^s e^{(s-r)A} w(r, X_r^w) dr \right\| \right] \\ & \leq \mathbb{E} \left[\left\| (e^{(t-s)A} - I) \int_0^{s-\eta} e^{(s-r)A} w(r, X_r^w) dr \right\| + \left\| (e^{(t-s)A} - I) \int_{s-\eta}^s e^{(s-r)A} w(r, X_r^w) dr \right\| \right] \\ & \leq \omega_{\eta,A,R}(t-s) + 2\eta R. \end{aligned}$$

Now, define

$$\tilde{\omega}_{A,R}(r) := \inf_{0 < \eta \leq T} \left\{ \omega_{\eta,A,R}(r) + 2\eta R \right\}, \quad r > 0.$$

Then, $\tilde{\omega}_{A,R}(0^+) = 0$ and $\tilde{\omega}_{A,R}$ is concave; so, it is a modulus of continuity. Taking now the infimum over $0 < \eta \leq T$ in the inequality above, we get, for every $0 \leq s \leq t \leq T$,

$$\mathbb{E} \left[\left\| (e^{(t-s)A} - I) \int_0^s e^{(s-r)A} w(r, X_r^w) dr \right\| \right] \leq \tilde{\omega}_{A,R}(|t-s|).$$

- (d) The map $[0, T] \rightarrow L^2(\Omega, \mathcal{P})$, $t \mapsto W_t^A$ is continuous (see [21, Th. 5.2]); hence, also the map $[0, T] \rightarrow L^1(\Omega, \mathcal{P})$, $t \mapsto W_t^A$ is continuous. Hence, since $[0, T]$ is compact, there exists a modulus of continuity $\hat{\omega}_A$ such that, for every $0 \leq s \leq t \leq T$,

$$\mathbb{E} [|W_t^A - W_s^A|] \leq \hat{\omega}_A(|t-s|).$$

We now combine all the estimates of points (a)–(d) above to obtain Claim (ii). $\square$

2.5. Assumptions on $\mathcal{H}$ and G

We make the following assumptions about $\mathcal{H}$ and G .

Assumption 2.6.

- (i) The function $\mathcal{H}$ is continuous, the function $\mathcal{H}(\cdot, 0, \mu_0)$ is bounded for some $\mu_0 \in \mathcal{P}_1(H)$, and there exists $C > 0$ such that

$$|\mathcal{H}(x, p, \mu) - \mathcal{H}(x, p', \mu')| \leq C(|p - p'| + \mathbf{d}_1(\mu, \mu')), \quad \forall x \in H, p, p' \in H, \mu, \mu' \in \mathcal{P}_1(H).$$

- (ii) $\mathcal{H}$ is differentiable with respect to the second variable, $\mathcal{H}_p$ is continuous and bounded, and there exists $C > 0$ such that

$$|\mathcal{H}_p(x, p, \mu) - \mathcal{H}_p(x, p', \mu')| \leq C(|p - p'| + \mathbf{d}_1(\mu, \mu')), \quad \forall x \in H, p, p' \in H, \mu, \mu' \in \mathcal{P}_1(H).$$

- (iii) G is continuous and such that the function $\mu \mapsto G(\cdot, \mu)$ is continuous as a map from $\mathcal{P}_1(H)$ to $C_b(H)$.

Remark 2.7. Requiring Lipschitz continuity of $\mathcal{H}$ and $\mathcal{H}_p$ with respect to p and μ variables is a little restrictive. However, Assumption 2.6 still covers many basic cases, especially those having separated structure. An example of $\mathcal{H}$ coming from a stochastic optimal control problem is provided in Section 7.

3. The Fokker-Planck equation: weak solutions

In this section, we study (FP) when the term $\mathcal{H}_p(x, Dv(t, x), m(t))$ is frozen. So, given $w \in C_b((0, T) \times H; H)$, we consider the linear Fokker-Planck equation in the space $\mathcal{P}_1(H)$

$$\begin{cases} \partial_t m(t) - L^* m(t) - \operatorname{div}(w(t, x)m(t)) = 0, \\ m(0) = m_0 \in \mathcal{P}_1(H). \end{cases} \quad (3.1)$$

We will interpret (3.1) in a weak sense, following the approach of [7]. For that, we define the set of test functions

$$\begin{aligned} \mathcal{D}_T = \Big\{ \varphi \in C_b^{1,2}([0, T] \times H) : A^* D\varphi(t, x) \in C_b([0, T] \times H; H), \\ D^2 \varphi \in C_b([0, T] \times H; \mathcal{L}_1(H)) \Big\}. \end{aligned} \quad (3.2)$$

To provide a rigorous definition of solution to (3.1), we introduce the operator

$$(L_0 \phi)(x) = \langle x, A^* D\phi(x) \rangle + \frac{1}{2} \operatorname{Tr} [D^2 \phi(x)]. \quad (3.3)$$

We notice that $L_0 \varphi(t, x) := [L_0 \varphi(t, \cdot)](x)$ is well defined for every $\varphi \in \mathcal{D}_T$ and $t \in [0, T]$. The formal operators L^* and div in (3.1) are interpreted similarly as for the Fokker-Planck equation in the case $H = \mathbb{R}^n$ when we apply test functions in $\mathcal{D}_T$. The definition of a *weak solution* to (3.1) given below is similar to the one of [7]; however, we use a larger class of test functions and we assume from the beginning that the solution is continuous in time.

Definition 3.1. A map $m \in \mathcal{S}$ is called a *weak solution* to (3.1) if $m(0) = m_0$ and

$$\begin{aligned} \int_H \varphi(t, x) m(t, dx) - \int_H \varphi(0, x) m(0, dx) \\ = \int_0^t \left(\int_H [\partial_s \varphi(s, x) + L_0 \varphi(s, x) - \langle w(s, x), D\varphi(s, x) \rangle] m(s, dx) \right) ds \end{aligned} \quad (3.4)$$

for every $t \in [0, T]$ and $\varphi \in \mathcal{D}_T$.¹

The connection between (3.1) and (2.11) is provided by the following.

¹ Instead of the requirement that $m \in \mathcal{S}$, in [7] the notion of weak solution $m : [0, T] \rightarrow \mathcal{P}_1(H)$ requires that $t \mapsto m(t)(I)$ is measurable for each $I \in \mathcal{B}(H)$; that is, in the terminology of [7], m is a probability kernel. Our requirement that $m \in \mathcal{S}$ is actually stronger. Indeed, given $I \in \mathcal{B}(H)$ and $m \in \mathcal{S}$, the map $t \mapsto m(t, I)$ is the composition of the continuous map $[0, T] \rightarrow (\mathcal{P}_1(H), d_1)$, $t \mapsto m(t)$ and of the map $(\mathcal{P}_1(H), d_1) \rightarrow \mathbb{R}$, $\mu \mapsto \mu(I)$. The d_1 -topology is stronger than the weak topology of measures; so, to conclude that $m \in \mathcal{S}$ is a probability kernel, it suffices to show that $\mu \mapsto \mu(I)$ is measurable when $\mathcal{P}_1(H)$ is endowed with the Borel σ -algebra induced by the weak topology. This claim follows from the fact that the latter map is upper semicontinuous for I closed (that is standard, as a consequence of one of the equivalent definitions of weak convergence) and by [4, Prop. 7.25]. This makes our notion of solution stronger and enables us to invoke the uniqueness result of [7] in Proposition 3.3.

Proposition 3.2. *Let Assumptions 2.1 and 2.3(ii) hold and let X^w be the unique weak-mild solution to (2.11) provided by Theorem 2.4. Then, $\mathcal{L}(X^w)$ is a weak solution to (2.11).*

Proof. This is just an application of Dynkin's formula for sufficiently regular functions: see, e.g., [22, Prop. 1.169], applied there with

$$b_0(s, x, a_1) = a_1, \quad a_1(s, \omega) = w(s, X_s^w(\omega)), \quad a_2(\cdot) = 0, \quad \Lambda = B_{\|w\|_\infty} \subset H. \quad \square$$

Concerning uniqueness of solutions to (3.1), we have the following.

Proposition 3.3. *Let Assumptions 2.1 and 2.3 hold. Then (3.1) admits at most one weak solution in the class*

$$\mathcal{S}_4 := \left\{ m \in \mathcal{S} : \int_0^T \int_H |x|^4 m(t, dx) dt < \infty \right\}.$$

Proof. The claim follows from [7, Th. 4.1] once we prove that a solution in our sense is also a solution in the sense of [7, Th. 4.1]. To do that we first notice that, in [7, Formula (1.5)], the definition of solution is given using the set test functions (called $\mathcal{E}_A(H)$ there) which is the linear span of the real parts of the functions of the form

$$(t, x) \mapsto \phi(t) e^{i\langle x, h(t) \rangle}, \quad \phi \in C^1([0, T]), \quad h \in C^1([0, T], D(A^*)).$$

Given the above, we observe that such a set is clearly contained in the set $\mathcal{D}_T$ defined in (3.2). Finally, we point out that in [7] the Kolmogorov operator L is defined differently (corresponding to our $\partial_t + L_0 - \langle w, D \rangle$); still, equality (1.5) in [7] corresponds to our (3.4). $\square$

We combine and summarize the above two results in the following theorem.

Theorem 3.4. *Let Assumptions 2.1 and 2.3 hold. Then (3.1) admits a unique weak solution in the class $\mathcal{S}_4$, which coincides with the law of the unique weak-mild solution X^w to (2.11).*

Proof. By Proposition 2.5(i), our solution provided by Proposition 3.2 belongs to the class $\mathcal{S}_4$ and the uniqueness follows from Proposition 3.3. $\square$

4. The HJB equation: mild solutions

In this section we deal with equation (HJB) for a given fixed $m \in \mathcal{S}$. We introduce the concept of a *mild solution* to (HJB) (see [22, Section 4.4.1.2]).

Definition 4.1 (*Mild solution to (HJB)*). Recall (2.4). A *mild solution* to (HJB) is a function $v \in C_{b,\gamma}^{0,1}([0, T] \times H)$ such that, for all $(t, x) \in [0, T] \times H$,

$$v(t, x) = [R_{T-t} G(\cdot, m(T))](x) - \int_t^T [R_{s-t} \mathcal{H}(\cdot, Dv(s, \cdot), m(s))](x) ds.$$

Theorem 4.2. *Let $m \in \mathcal{S}$ and let Assumptions 2.1 and 2.6 hold.*

(i) *There exists a unique mild solution u to (HJB).*

(ii) If $m_n \rightarrow m$ in $\mathcal{S}$, denoting by v_n and v the corresponding mild solutions to (HJB), we have

$$\|Dv_n - Dv\|_{C_{b,\gamma}} \rightarrow 0.$$

Proof. (i) This part is contained in [22, Theorem 4.90].²

(ii) Using the smoothing properties of the Ornstein-Uhlenbeck semigroup (see e.g. [22, Theorem 4.37]), we have the representation of the gradients

$$\begin{aligned} Dv(t, x) &= D[R_{T-t} G(\cdot, m(T))](x) - \int_t^T D[R_{s-t} \mathcal{H}(\cdot, Dv(s, \cdot), m(s))](x) ds. \\ Dv_n(t, x) &= D[R_{T-t} G(\cdot, m_n(T))](x) - \int_t^T D[R_{s-t} \mathcal{H}(\cdot, Dv_n(s, \cdot), m_n(s))](x) ds. \end{aligned}$$

Hence, fixing $t \in [0, T]$ and considering [22, Th. 4.37] and Assumption 2.6, we have, for some $C > 0$,

$$\begin{aligned} \|Dv(t, \cdot) - Dv_n(t, \cdot)\|_\infty &\leq \|D[R_{T-t}(G(\cdot, m(T)) - G(\cdot, m_n(T)))]\|_\infty \\ &\quad + \int_t^T \left\| D \left[R_{s-t} (\mathcal{H}(\cdot, Dv(s, \cdot), m(s)) - \mathcal{H}(\cdot, Dv_n(s, \cdot), m_n(s))) \right] \right\|_\infty ds \\ &\leq C \left((T-t)^{-\gamma} \|G(\cdot, m(T)) - G(\cdot, m_n(T))\|_\infty \right. \\ &\quad \left. + \int_t^T (s-t)^{-\gamma} \left(\|Dv(s, \cdot) - Dv_n(s, \cdot)\|_\infty + \mathbf{d}_1(m(s), m_n(s)) \right) ds \right) \\ &\leq C \left(\eta_n (T-t)^{-\gamma} + \rho_\infty(m, m_n) T^{1-\gamma} + \int_t^T (s-t)^{-\gamma} \|Dv(s, \cdot) - Dv_n(s, \cdot)\|_\infty ds \right), \end{aligned}$$

where, by Assumption 2.6(iii),

$$\eta_n := \|G(\cdot, m(T)) - G(\cdot, m_n(T))\|_\infty \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

By a suitable generalization of Gronwall's Lemma (see [22, Prop. D.30]), we get

$$\|Dv(t, \cdot) - Dv_n(t, \cdot)\|_\infty \leq C(T-t)^{-\gamma} (\eta_n + T\rho_\infty(m, m_n)).$$

Multiplying both sides by $(T-t)^\gamma$ and taking the supremum over $t \in [0, T]$, we get

$$\|Dv - Dv_n\|_{C_{b,\gamma}} \leq C (\eta_n + T\rho_\infty(m, m_n))$$

and the claim follows. $\square$

² More precisely, it is enough to take in that theorem, $m = 0$, $\gamma_G(t) = t^{-\gamma}$, $G = \mathbf{Id}$. We warn that G in this footnote refers to the operator used in [22, Theorem 4.90], which has nothing to do with our function G .

5. The MFG system: existence of solutions

We turn to the analysis of the coupled MFG system (HJB)-(FP). Let Assumptions 2.3(i) and 2.6 hold. Given $m \in \mathcal{S}$, we denote by $v^{(m)}$ the unique mild solution to (HJB). Moreover, recalling the notation X^w used in Section 3, we set

$$X_t^{(m)} := X_t^{\mathcal{H}_p(\cdot, Dv^{(m)}(\cdot, \cdot), m(\cdot))}.$$

In the remainder of this section, $\{\mathbf{e}_k\}_{k \in \mathbb{N}}$ is the orthonormal basis provided by Assumption 2.1.

Lemma 5.1. *Let Assumptions 2.1 and 2.3 holds. For $n \in \mathbb{N}$, set*

$$\alpha_n := (2|\lambda_n|)^{-1}, \quad \beta_n := \int_H \langle x, \mathbf{e}_n \rangle^2 m_0(dx).$$

(i) *We have*

$$\int_H \sup_{t \in [0, T]} \langle e^{tA} x, \mathbf{e}_n \rangle^2 m_0(dx) \leq \beta_n, \quad \forall n \in \mathbb{N}.$$

(ii) *We have*

$$\sup_{t \in [0, T], x \in H} \int_0^t \langle e^{(t-s)A} w(s, x), \mathbf{e}_n \rangle^2 ds \leq \alpha_n \cdot \sup_{s \in (0, T), x \in H} \langle w(s, x), \mathbf{e}_n \rangle^2 \leq \alpha_n \|w\|_\infty^2,$$

for every $n \in \mathbb{N}$.

(iii) *We have*

$$\sup_{t \in [0, T]} \mathbb{E} \langle W_t^A, \mathbf{e}_n \rangle^2 \leq \alpha_n, \quad \forall n \in \mathbb{N}.$$

Proof. (i) In fact,

$$\begin{aligned} \int_H \sup_{t \in [0, T]} \langle e^{tA} x, \mathbf{e}_n \rangle^2 m_0(dx) &= \int_H \sup_{t \in [0, T]} \langle x, e^{tA} \mathbf{e}_n \rangle^2 m_0(dx) \\ &= \int_H \left(\sup_{t \in [0, T]} e^{2\lambda_n t} \right) \langle x, \mathbf{e}_n \rangle^2 m_0(dx) \leq \int_H \langle x, \mathbf{e}_n \rangle^2 m_0(dx) = \beta_n. \end{aligned}$$

(ii) We compute,

$$\begin{aligned} \sup_{t \in [0, T], x \in H} \int_0^t \langle e^{(t-s)A} w(s, x), \mathbf{e}_n \rangle^2 ds &= \sup_{t \in [0, T], x \in H} \int_0^t \langle w(s, x), e^{(t-s)A} \mathbf{e}_n \rangle^2 ds \\ &= \sup_{t \in [0, T], x \in H} \int_0^t e^{2(t-s)\lambda_n} \langle w(s, x), \mathbf{e}_n \rangle^2 ds \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{s \in (0, T), x \in H} \langle w(s, x), \mathbf{e}_n \rangle^2 \cdot \sup_{t \in [0, T]} \int_0^t e^{2(t-s)\lambda_n} ds \\
&\leq \alpha_n \cdot \sup_{s \in (0, T), x \in H} \langle w(s, x), \mathbf{e}_n \rangle^2 \leq \alpha_n \|w\|_\infty^2.
\end{aligned}$$

(iii) Recall that, by Assumption 2.1, we have

$$\text{Cov}(W_t^A) = Q_t = \int_0^t e^{2sA} ds.$$

Then,

$$\sup_{t \in [0, T]} \mathbb{E} \langle W_t^A, \mathbf{e}_n \rangle^2 = \sup_{t \in [0, T]} \langle Q_t \mathbf{e}_n, \mathbf{e}_n \rangle \leq \langle Q_T \mathbf{e}_n, \mathbf{e}_n \rangle = (2|\lambda_n|)^{-1}. \quad \square$$

Set³

$$a_n := 3 \left(\beta_n + \alpha_n + \alpha_n \|\mathcal{H}_p\|_\infty^2 \right).$$

Then, by Assumptions 2.1 and 2.3(i), we have

$$\mathbf{a} = (a_n)_{n \in \mathbb{N}} \in \ell_1. \quad (5.1)$$

Let us consider the map

$$\Psi : \mathcal{S} \rightarrow \mathcal{S}, \quad \Psi(m)(t) := \mathcal{L}(X_t^{(m)}). \quad (5.2)$$

This map is well defined due to Proposition 2.5.

Proposition 5.2. *Let Assumptions 2.1, 2.3(i) and 2.6 hold. Then, for each $m \in \mathcal{S}$, $n \in \mathbb{N}$, and $t \in [0, T]$, we have*

$$\int_H |\langle x, \mathbf{e}_n \rangle|^2 \Psi(m)(t, dx) \leq a_n.$$

Proof. We have, using the equation for $X_t^{(m)}$ and denoting $w(s, x) = \mathcal{H}_p(x, Dv^{(m)}(s, x), m(s))$,

$$\begin{aligned}
&\int_H |\langle x, \mathbf{e}_n \rangle|^2 \Psi(m)(t, dx) = \mathbb{E} \langle X_t^{(m)}, \mathbf{e}_n \rangle^2 \\
&\leq 3 \mathbb{E} \left[\langle e^{tA} X_0, \mathbf{e}_n \rangle^2 + \int_0^t \langle e^{(t-s)A} w(s, X_s^{(m)}), \mathbf{e}_n \rangle^2 ds + \langle W_t^A, \mathbf{e}_n \rangle^2 \right]
\end{aligned}$$

Applying Lemma 5.1(i), we get the following estimate for the first term

$$\mathbb{E} [\langle e^{tA} X_0, \mathbf{e}_n \rangle^2] = \int_H \langle e^{tA} x, \mathbf{e}_n \rangle^2 m_0(dx) \leq \beta_n.$$

³ Note that a_n only depends on the data A , $\mathcal{H}$, and on the initial measure m_0 .

As for the second term, we have by Lemma 5.1(ii)

$$\mathbb{E} \left[\int_0^t \langle e^{(t-s)A} w(s, X_s^{(m)}), \mathbf{e}_n \rangle^2 ds \right] \leq \sup_{x \in H} \int_0^t \langle e^{(t-s)A} w(s, x), \mathbf{e}_n \rangle^2 ds \leq \alpha_n \|\mathcal{H}_p\|_\infty^2.$$

The last term is estimated by Lemma 5.1(iii), which gives

$$\mathbb{E} [\langle W_t^A, \mathbf{e}_n \rangle^2] \leq \alpha_n.$$

The claim now follows from the definition of a_n . $\square$

Let c_0 be the constant of Proposition 2.5. We set

$$\hat{c}_{m_0} := 1 + c_0 \left(1 + \int_H |x|^4 m_0(dx) + \|\mathcal{H}_p\|_\infty^4 \right)$$

and

$$\mathcal{R}_{m_0} := \left\{ \mu \in \mathcal{P}_1(H) : \int_H |x|^4 \mu(dx) \leq \hat{c}_{m_0} \right\}. \quad (5.3)$$

Observe that, by Lemma Appendix A.1 and by [42, Remark 7.13 (ii)], the metric $\mathbf{d}_1$ metrizes on $\mathcal{R}_{m_0}$ the topology of weak convergence in $\mathcal{P}(H)$. In view of Proposition 5.2, we also define

$$\mathcal{Q}_{m_0} := \left\{ \mu \in \mathcal{R}_{m_0} : \int_H \langle x, \mathbf{e}_k \rangle^2 \mu(dx) \leq a_k \right\}.$$

By Proposition Appendix A.2, $\mathcal{Q}_{m_0}$ is $\mathbf{d}_1$ -compact. We now consider a subset of $\mathcal{S}$ of functions which take values in the set $\mathcal{Q}_{m_0}$ and are $\mathbf{d}_1$ -equi-uniformly continuous. More precisely, recalling Proposition 2.5(ii) we define

$$\omega_{m_0} := \omega_{m_0, \|\mathcal{H}_p\|_\infty} \quad (5.4)$$

and consider the subset $\mathcal{C}_{m_0} \subset \mathcal{S}$ defined as

$$\mathcal{C}_{m_0} := \left\{ m : [0, T] \rightarrow \mathcal{Q}_{m_0} : m(0) = m_0, \mathbf{d}_1(m(t), m(s)) \leq \omega_{m_0}(|t - s|) \right\}. \quad (5.5)$$

Lemma 5.3. *Let Assumptions 2.1, 2.3(i) and 2.6 hold. The set $\mathcal{C}_{m_0}$ is convex and compact in $(\mathcal{S}, \rho_\infty)$.*

Proof. *Relative compactness.* Relative compactness follows from Proposition Appendix A.2 and the Arzelà-Ascoli's theorem for functions with values in complete metric spaces.

Closedness. Let $\{m_n\}_{n \in \mathbb{N}} \subset \mathcal{C}_{m_0}$ be such that $m_n \rightarrow m \in \mathcal{S}$. This means that

$$\sup_{[0, T]} \mathbf{d}_1(m_n(t), m(t)) \rightarrow 0.$$

Clearly $m(0) = m_0$. Moreover, $m_n(t)$ converges weakly to $m(t)$ for all $t \in [0, T]$, so

$$\int_H (|x|^4 \wedge K) m(t, dx) = \lim_{n \rightarrow \infty} \int_H (|x|^4 \wedge K) m_n(t, dx) \leq \hat{c}_{m_0}, \quad \forall t \in [0, T], \quad \forall K > 0.$$

Hence, by dominated convergence

$$\int_H |x|^4 m(t, dx) = \lim_{K \rightarrow \infty} \int_H (|x|^4 \wedge K) m(t, dx) \leq \hat{c}_{m_0}.$$

Similarly, one shows that

$$\sup_{t \in [0, T]} \int_H \langle x, e_k \rangle^2 m(t, dx) \leq a_k \quad \forall k \in \mathbb{N}.$$

Next, for $t, s \in [0, T]$, we have

$$\begin{aligned} \mathbf{d}_1(m(t), m(s)) &\leq \mathbf{d}_1(m_n(t), \mu(t)) + \mathbf{d}_1(m_n(s), m(s)) + \mathbf{d}_1(m_n(t), m_n(s)) \\ &\leq \mathbf{d}_1(m_n(t), m(t)) + \mathbf{d}_1(m_n(s), m(s)) + \omega_{m_0}(|t - s|) \end{aligned}$$

Taking the limit as $n \rightarrow \infty$ above, we get

$$\mathbf{d}_1(m(t), m(s)) \leq \omega_{m_0}(|t - s|).$$

Convexity. Let $m_1, m_2 \in \mathcal{C}_{m_0}$, $\lambda \in [0, 1]$, and set $m_\lambda := \lambda m_1 + (1 - \lambda)m_2$. The only nontrivial claim to verify is that, for $t, s \in [0, T]$,

$$\mathbf{d}_1(m_\lambda(t), m_\lambda(s)) \leq \omega_{m_0}(|t - s|).$$

We will use the fact (see, e.g., [42, Rem. 7.5(i)] or [12, Cor. 5.4]) that for $\mu, \nu \in \mathcal{P}_1(H)$

$$\mathbf{d}_1(\mu, \nu) = \sup_{f \in \text{Lip}_1} \int_H f(x)(\mu(dx) - \nu(dx)),$$

where Lip_1 is the set of Lipschitz functions on H with Lipschitz constant 1. Then

$$\begin{aligned} \mathbf{d}_1(m_\lambda(t), m_\lambda(s)) &= \sup_{f \in \text{Lip}_1} \int_H f(x)(m_\lambda(t, dx) - m_\lambda(s, dx)) \\ &\leq \lambda \sup_{f \in \text{Lip}_1} \int_H f(x)(m_1(t, dx) - m_1(s, dx)) + (1 - \lambda) \sup_{f \in \text{Lip}_1} \int_H f(x)(m_2(t, dx) - m_2(s, dx)) \\ &= \lambda \mathbf{d}_1(m_1(t), m_1(s)) + (1 - \lambda) \mathbf{d}_1(m_2(t), m_2(s)) \leq \omega_{m_0}(|t - s|), \end{aligned}$$

the claim. $\square$

In order to use a fixed point theorem to prove existence of solutions to the MFG system, we need to embed properly $\mathcal{C}_{m_0}$ into a topological vector space. We consider the vector space $C([0, T]; \mathcal{M}(H))$, where $\mathcal{M}(H)$ is endowed with the weak topology induced by the family of seminorms (2.1). On the space $C([0, T]; \mathcal{M}(H))$, we define a suitable topology as follows. Define the family of seminorms

$$|m(\cdot)|_f := \sup_{t \in [0, T]} \left| \int_H f(x) m(t, dx) \right|, \quad f \in C_b(H).$$

This family of seminorms induces on $C([0, T]; \mathcal{M}(H))$ a topology, that we denote by $\tau_{\mathbf{uw}}$ and call the *uniform-weak topology*, that makes $C([0, T]; \mathcal{M}(H))$ a locally convex topological vector space.

Lemma 5.4. *Let Assumptions 2.1, 2.3(i) and 2.6 hold. We have the following inclusion:*

$$\mathcal{C}_{m_0} \subset C([0, T]; \mathcal{M}(H)).$$

Moreover, the topology induced by ρ_∞ on $\mathcal{C}_{m_0}$ is the same as the one induced by $\tau_{\mathbf{uw}}$.

Proof. In the following we use the notation $m(\cdot)$ to denote elements of $\mathcal{C}_{m_0}$ in order to avoid confusion.

We first prove the inclusion. Obviously, every function in $\mathcal{C}_{m_0}$ can also be seen as a function with values in $\mathcal{M}(H)$. The inclusion then holds if we prove that every function in $\mathcal{C}_{m_0}$ is continuous as a function from $[0, T]$ to $\mathcal{M}(H)$. Indeed, let $m(\cdot) \in \mathcal{C}_{m_0}$; then, for any $t_0 \in [0, T]$ and any sequence $t_n \rightarrow t_0$ one has $\mathbf{d}_1(m(t_n), m(t_0)) \rightarrow 0$ as $n \rightarrow \infty$. This implies that, in particular, $m(t_n)$ converges weakly to $m(t_0)$ when $n \rightarrow \infty$, which completes the proof.

(i) $\tau_{\mathbf{uw}}$ stronger than ρ_∞ . Let $\bar{m}(\cdot) \in \mathcal{C}_{m_0}$, $\varepsilon > 0$ and consider the ρ_∞ -neighborhood

$$\mathcal{U}(\bar{m}(\cdot)) := \left\{ m(\cdot) \in \mathcal{C}_{m_0} : \sup_{t \in [0, T]} \mathbf{d}_1(m(t), \bar{m}(t)) < \varepsilon \right\}.$$

By the equiuniform $\mathbf{d}_1$ -continuity of the elements of $\mathcal{C}_{m_0}$, there exists $\{t_1, \dots, t_N\} \subset [0, T]$ such that

$$\mathcal{U}_N(\bar{m}(\cdot)) := \left\{ m(\cdot) \in \mathcal{C}_{m_0} : \mathbf{d}_1(m(t_i), \bar{m}(t_i)) < \varepsilon/2, \quad \forall i = 1, \dots, N \right\} \subseteq \mathcal{U}(\bar{m}(\cdot)).$$

Since $\mathbf{d}_1$ metrizes the weak topology in $\mathcal{Q}_{m_0}$, for each $i = 1, \dots, N$ we have the existence of $\delta > 0$ and a family $\{f_{i,1}, \dots, f_{i,M}\} \subset C_b(H)$ such that

$$\begin{aligned} \mathcal{V}_{t_i}(\bar{m}(t_i)) &:= \{\mu \in \mathcal{Q}_{m_0} : |\mu - \bar{m}(t_i)|_{f_{i,j}} < \delta, \quad \forall j = 1, \dots, M\} \\ &\subseteq \{\mu \in \mathcal{Q}_{m_0} : \mathbf{d}_1(\mu, \bar{m}(t_i)) < \varepsilon/2\} =: \mathcal{U}_{t_i}(\bar{m}(t_i)) \end{aligned}$$

Define now the $\tau_{\mathbf{uw}}$ -neighborhood of $\bar{m}(\cdot)$

$$\mathcal{V}(\bar{m}(\cdot)) := \{\mu \in \mathcal{C}_{m_0} : |m(\cdot) - \bar{m}(\cdot)|_{f_{i,j}} < \delta, \quad \forall j = 1, \dots, M, \quad \forall i = 1, \dots, N\}.$$

Then,

$$\begin{aligned} \mathcal{V}(\bar{m}(\cdot)) &\subseteq \{m(\cdot) \in \mathcal{C}_{m_0} : m(t_i) \in \mathcal{V}_{t_i}(\bar{m}(t_i)) \quad \forall i = 1, \dots, N\} \\ &\subseteq \{m(\cdot) \in \mathcal{C}_{m_0} : m(t_i) \in \mathcal{U}_{t_i}(\bar{m}(t_i)), \quad \forall i = 1, \dots, N\} \\ &= \mathcal{U}_N(\bar{m}(\cdot)) \subseteq \mathcal{U}(\bar{m}(\cdot)), \end{aligned}$$

concluding the proof of this part.

(ii) ρ_∞ stronger than $\tau_{\mathbf{uw}}$. First, let us show that, for each $f \in C_b(H)$, the family

$$\left\{ [0, T] \rightarrow \mathbb{R}, \quad t \mapsto \int_H f(x) m(t, dx) \right\}_{m(\cdot) \in \mathcal{C}_{m_0}} \quad (5.6)$$

is equiuniformly continuous; that is, there exists a modulus of continuity ω_f such that

$$\left| \int_H f(x) (m(t, dx) - m(s, dx)) \right| \leq \omega_f(|t - s|), \quad \forall t, s \in [0, T], \quad \forall m \in \mathcal{C}_{m_0}. \quad (5.7)$$

Assume, by contradiction, that this is not true. Then, we may find $\eta > 0$ and suitable sequences $\{s_n\}, \{t_n\} \subset [0, T]$, $\{m_n(\cdot)\} \subset \mathcal{C}_{m_0}$, such that

$$\left| \int_H f(x) (m_n(t_n, dx) - m_n(s_n, dx)) \right| \geq \eta, \quad |s_n - t_n| \rightarrow 0. \quad (5.8)$$

Now, on the one hand, by $\mathbf{d}_1$ -equiuniform continuity of the functions belonging to $\mathcal{C}_{m_0}$, we have

$$\mathbf{d}_1(m_n(s_n), m_n(t_n)) \rightarrow 0; \quad (5.9)$$

on the other hand, by weak compactness of $\mathcal{Q}_{m_0}$, up to passing to subsequences, we have

$$m_n(s_n) \xrightarrow{w} \mu_1, \quad m_n(t_n) \xrightarrow{w} \mu_2 \quad (5.10)$$

for some $\mu_1, \mu_2 \in \mathcal{Q}_{m_0}$. By (5.9), we must have

$$\mu_1 = \mu_2. \quad (5.11)$$

Then, letting $n \rightarrow \infty$ in (5.8) and using (5.10) and (5.11), we get

$$0 < \eta \leq \lim_{n \rightarrow \infty} \left| \int_H f(x) (m_n(t_n, dx) - m_n(s_n, dx)) \right| = 0,$$

a contradiction. This proves (5.6).

Now, let $\bar{m}(\cdot) \in \mathcal{C}_{m_0}$, $\varepsilon > 0$, $\{f_1, \dots, f_N\} \subset C_b(H)$ and consider the $\tau_{\mathbf{uw}}$ -neighborhood

$$\mathcal{V}(\bar{m}(\cdot)) := \left\{ m(\cdot) \in \mathcal{C}_{m_0} : |m(t) - \bar{m}(t)|_{f_i} < \varepsilon \quad \forall t \in [0, T], \quad \forall i = 1, \dots, N \right\}.$$

By (5.6), we have the existence of $\{t_1, \dots, t_M\} \subset [0, T]$ such that

$$\begin{aligned} \mathcal{V}_N(\bar{m}(\cdot)) &:= \left\{ m(\cdot) \in \mathcal{C}_{m_0} : |m(t_j) - \bar{m}(t_j)|_{f_i} < \varepsilon/2, \quad \forall i = 1, \dots, N, \quad \forall j = 1, \dots, M \right\} \\ &\subseteq \mathcal{V}(\bar{m}(\cdot)). \end{aligned}$$

Since $\mathbf{d}_1$ metrizes the weak topology on $\mathcal{Q}_{m_0}$, for each $j = 1, \dots, M$ we have the existence of $\delta_j > 0$ such that

$$\mathbf{d}_1(m(t_j), \bar{m}(t_j)) < \delta_j \Rightarrow |m(t_j) - \bar{m}(t_j)|_{f_i} < \varepsilon/2, \quad \forall i = 1, \dots, N.$$

Let now $\delta := \min\{\delta_1, \dots, \delta_M\} > 0$. Then,

$$\begin{aligned} \mathcal{U}(\bar{m}(\cdot)) &:= \left\{ m(\cdot) \in \mathcal{C}_{m_0} : \sup_{t \in [0, T]} \mathbf{d}_1(m(t), \bar{m}(t)) < \delta \right\} \\ &\subseteq \left\{ m(\cdot) \in \mathcal{C}_{m_0} : \mathbf{d}_1(m(t_j), \bar{m}(t_j)) < \delta \quad \forall j = 1, \dots, M \right\} \\ &\subseteq \mathcal{V}_N(\bar{m}(\cdot)) \subseteq \mathcal{V}(\bar{m}(\cdot)), \end{aligned}$$

concluding the proof. $\square$

We now define a solution to the MFG system (HJB)–(FP).

Definition 5.5. A pair (v, m) is a solution to the MFG system (HJB)–(FP) if v is a mild solution to (HJB) (Definition 4.1) and m is a weak solution of (3.1) with $w(t, x) = \mathcal{H}_p(x, Dv(t, x), m(t))$ (Definition 3.1).

Theorem 5.6. Let Assumptions 2.1, 2.3(i), and 2.6 hold. Then, the MFG system (HJB)–(FP) admits a solution.

Proof. By construction, if m^* is a fixed point of the map Ψ defined in (5.2), then the couple $(v^{(m^*)}, m^*)$ is a solution to the MFG system (HJB)–(FP). We are therefore going to show that Ψ admits a fixed point.

We consider $\mathcal{C}_{m_0}$ embedded in the topological vector space $(C([0, T]; \mathcal{M}(H)), \tau_{\mathbf{uw}})$. By Lemma 5.4 we can indifferently use on $\mathcal{C}_{m_0}$ the $\tau_{\mathbf{uw}}$ or the ρ_∞ topology. In what follows, the topological notions are referred, indifferently, to one of those equivalent topologies. We want to apply Tikhonov's fixed point theorem. First of all, we notice that, by the definitions of $\hat{c}_{m_0}$ and $\mathbf{a}$, using Proposition 2.5 and Proposition 5.2, we have

$$\Psi(\mathcal{C}_{m_0}) \subseteq \mathcal{C}_{m_0}. \quad (5.12)$$

Since $\mathcal{C}_{m_0}$ is compact and convex, and since (5.12) holds, in order to conclude the existence of a fixed point of Ψ in $\mathcal{C}_{m_0}$, we need to show that $\Psi|_{\mathcal{C}_{m_0}}$ is continuous. The remaining part of the proof is devoted to this goal.

Let $\{m_n\} \subset \mathcal{C}_{m_0}$ be such that $m_n \rightarrow m \in \mathcal{C}_{m_0}$. Since $\mathcal{C}_{m_0}$ is compact, from each subsequence (m_{n_k}) , one can extract a sub-subsequence such that $\Psi(m_{n_{k_h}}) \rightarrow \hat{m}$ for some $\hat{m} \in \mathcal{C}_{m_0}$. Set

$$w^{(m)} := \mathcal{H}_p(\cdot, Dv^{(m)}(\cdot, \cdot), m(\cdot)), \quad w^{(m_{n_{k_h}})} := \mathcal{H}_p(\cdot, Dv^{(m_{n_{k_h}})}(\cdot, \cdot), m_{n_{k_h}}(\cdot)).$$

For each $h \in \mathbb{N}$,

$$\begin{aligned} &\int_H \varphi(t, x) \Psi(m_{n_{k_h}})(t, dx) - \int_H \varphi(0, x) m_0(dx) \\ &= \int_0^t \left(\int_H \left[\partial_t \varphi(s, x) + L_0 \varphi(s, x) - \langle w^{(m_{n_{k_h}})}(s, x), D\varphi(s, x) \rangle \right] \Psi(m_{n_{k_h}})(s, dx) \right) ds \end{aligned}$$

for every $t \in (0, T]$ and $\varphi \in D_T$. Considering that, by Theorem 4.2(ii), we have $w^{(m_{n_{k_h}})} \rightarrow w^{(m)}$ pointwise as $h \rightarrow \infty$, we take the limit as $h \rightarrow \infty$ using dominated convergence theorem in the equation above to obtain

$$\begin{aligned} & \int_H \varphi(t, x) \hat{m}(t, dx) - \int_H \varphi(0, x) m_0(dx) \\ &= \int_0^t \left(\int_H \left[\partial_t \varphi(s, x) + L_0 \varphi(s, x) - \langle w^{(m)}(s, x), D\varphi(s, x) \rangle \right] \hat{m}(s, dx) \right) ds. \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} & \int_H \varphi(t, x) \mathcal{L}(X^{(m)})(t, dx) - \int_H \varphi(0, x) m_0(dx) \\ &= \int_0^t \left(\int_H \left[\partial_t \varphi(s, x) + L_0 \varphi(s, x) - \langle w^{(m)}(s, x), D\varphi(s, x) \rangle \right] \mathcal{L}(X^{(m)})(s, dx) \right) ds. \end{aligned}$$

Hence, by Proposition 3.3, we must have $\hat{m} = \mathcal{L}(X^{(m)})$. By the arbitrariness of $k \mapsto n_k$, it follows that

$$\Psi(m_n) \rightarrow \hat{m} = \mathcal{L}(X^{(m)}) = \Psi(m) \quad \text{in } (\mathcal{S}, \rho_\infty),$$

which gives the continuity of Ψ . $\square$

6. The MFG system: uniqueness of solutions

In this section, we establish the uniqueness of solutions for our MFG system **(HJB)**–**(FP)**. It is well known, even in finite-dimensional state spaces (see [11, Rem. 1.4, p.32]), that uniqueness of a solution for this system cannot generally be guaranteed. A typical approach to ensure uniqueness is to impose certain monotonicity conditions on the data. One of the most common monotonicity conditions is the *Lasry-Lions monotonicity condition* (see, for example, [34–36] and [11, Section 1.3.2]). For other monotonicity conditions that ensure uniqueness in finite-dimensional MFG systems, readers can refer to [32,38] and the references therein. Another typical feature of uniqueness results is the separability of the Hamiltonian function $\mathcal{H}$ (as assumed in Assumption 6.1(i) here). Well-posedness of MFG systems without separability has been studied in [38]. Here, we assume separability and use a variant of the Lasry-Lions monotonicity condition from [11, Theorem 1.4]. Our assumptions are as follows.

Assumption 6.1.

- (i) The Hamiltonian $\mathcal{H}$ has a separated form, i.e.

$$\mathcal{H}(x, p, \mu) = \mathcal{H}^0(x, p) - F(x, \mu),$$

for some continuous functions

$$\mathcal{H}^0 : H \times H \rightarrow \mathbb{R}, \quad F : H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}.$$

- (ii) The map $\mathcal{H}^0$ is convex in p .

- (iii) The functions F and G are monotone in $\mu \in \mathcal{P}_1(H)$ in the following sense

$$\int_H [F(x, \mu_1) - F(x, \mu_2)](\mu_1 - \mu_2)(dx) \geq 0 \quad \forall \mu_1, \mu_2 \in \mathcal{P}_1(H),$$

$$\int_H [G(x, \mu_1) - G(x, \mu_2)](\mu_1 - \mu_2)(dx) \geq 0 \quad \forall \mu_1, \mu_2 \in \mathcal{P}_1(H).$$

(iv) One of the following condition is satisfied:

(a) The following is satisfied:

$$\int_H [F(x, \mu_1) - F(x, \mu_2)](\mu_1 - \mu_2)(dx) > 0 \quad \forall \mu_1, \mu_2 \in \mathcal{P}_1(H), \mu_1 \neq \mu_2.$$

(b) The following implications hold:

$$\begin{aligned} \int_H [F(x, \mu_1) - F(x, \mu_2)](\mu_1 - \mu_2)(dx) = 0 &\implies F(\cdot, \mu_1) = F(\cdot, \mu_2), \\ \int_H [G(x, \mu_1) - G(x, \mu_2)](\mu_1 - \mu_2)(dx) = 0 &\implies G(\cdot, \mu_1) = G(\cdot, \mu_2). \end{aligned}$$

(c) The following implication holds:

$$\begin{aligned} \mathcal{H}^0(x, p_1) - \mathcal{H}^0(x, p_2) - \langle \mathcal{H}_p^0(x, p_2), p_1 - p_2 \rangle &= 0 \\ \implies \mathcal{H}_p^0(x, p_1) &= \mathcal{H}_p^0(x, p_2), \quad \forall x \in H, p_1, p_2 \in H \end{aligned}$$

Theorem 6.2. *Let Assumptions 2.3(i), 2.6, and 6.1 hold. Then, there exists a unique solution to the MFG system (HJB)–(FP).*

Proof. *Existence.* Existence of a solution was proved in Theorem 5.6.

Uniqueness. The proof is an adaptation to our Hilbert space setting of a typical proof of uniqueness in finite dimensional spaces (see, e.g., [11, Theorem 1.4]).

Let (v_1, m_1) and (v_2, m_2) be two solutions of our MFG system. We set $\bar{v} = v_1 - v_2$ and $\bar{m} = m_1 - m_2$. Then, $\bar{v}$ solves the integral equation

$$\begin{aligned} \bar{v}(t, x) &= R_{T-t}[G(\cdot, m_1(T)) - G(\cdot, m_2(T))](x) \\ &\quad - \int_t^T R_{s-t}[\mathcal{H}^0(\cdot, Du_1(s, \cdot)) - \mathcal{H}^0(\cdot, Du_2(s, \cdot))](x) ds \\ &\quad - \int_t^T R_{s-t}[F(\cdot, m_1(s)) - F(\cdot, m_2(s))](x) ds, \end{aligned}$$

which means that $\bar{v}$ is a mild solution to the equation

$$\begin{cases} -\partial_t \bar{v} - L\bar{v} + \mathcal{H}^0(x, Dv_1(t, x)) - \mathcal{H}^0(x, Dv_2(t, x)) - (F(x, m_1(t)) - F(x, m_2(t))) = 0, \\ \bar{v}(T, x) = G(x, m_1(T)) - G(x, m_2(T)). \end{cases}$$

On the other hand, observing that $\bar{m}_0 \equiv 0$, we get that $\bar{m}$ satisfies, for every $\varphi \in \mathcal{D}_T$,

$$\begin{aligned} \int_H \varphi(t, x) \bar{m}(t, dx) - \int_H \varphi(0, x) \bar{m}(0, dx) &= \int_0^T \int_H [\partial_t \varphi(s, x) + L_0 \varphi(s, x)] \bar{m}(s, dx) \\ &- \int_0^T \left(\int_H \langle \mathcal{H}_p^0(x, Dv_1(s, x)), D\varphi(s, x) \rangle m_1(s, dx) - \int_H \langle \mathcal{H}_p^0(x, Dv_2(s, x)), D\varphi(s, x) \rangle m_2(s, dx) \right) ds. \end{aligned} \quad (6.1)$$

We now set

$$f(s, x) := \mathcal{H}^0(x, Dv_1(s, x)) - \mathcal{H}^0(x, Dv_2(s, x)) - (F(x, m_1(s)) - F(x, m_2(s))).$$

By Assumption 2.6, by Proposition 2.5, and by the definition of solution to (HJB)-(FP), we have $f \in C_{b,\gamma}([0, T] \times H)$, where the latter space is defined in (2.4). Then, we consider the Kolmogorov equation

$$\begin{cases} \partial_t v(t, x) + Lv(t, x) = f(t, x), \\ v(T, x) = \phi(x) := G(x, m_1(T)) - G(x, m_2(T)). \end{cases}$$

By Theorem Appendix B.4, $\bar{v}$ is also a $\mathcal{K}$ -strong solution of the same equation and moreover, we can choose the approximating data ϕ_n, f_n such that the solutions $\bar{v}_n$ of the approximating Kolmogorov equations

$$\begin{cases} \partial_t \bar{v}_n(t, x) + L_0 \bar{v}_n(t, x) = f_n(t, x), \\ \bar{v}_n(T, x) = \phi_n(x) \end{cases}$$

belong to $\mathcal{D}_T$ for all $n \in \mathbb{N}$. Now, using $\varphi = \bar{v}_n$ in (6.1), we get

$$\begin{aligned} \int_H \bar{v}_n(T, x) \bar{m}(t, dx) &= \int_0^T \left(\int_H [\partial_t \bar{v}_n(s, x) + L_0 \bar{v}_n(s, x)] \bar{m}(s, dx) \right) ds \\ &- \int_0^T \left(\int_H [\langle \mathcal{H}_p^0(x, Dv_1(s, x)), D\bar{v}_n(s, x) \rangle] m_1(s, dx) - \int_H [\langle \mathcal{H}_p^0(x, Dv_2(s, x)), D\bar{v}_n(s, x) \rangle] m_2(s, dx) \right) ds. \end{aligned} \quad (6.2)$$

On the other hand, for every $s \in [0, T]$, we can integrate with respect to the measure $\bar{m}(s)$ the Kolmogorov equations solved by the $\bar{v}_n$'s, to get

$$\int_H (\partial_t \bar{v}_n(s, x) + L_0 \bar{v}_n(s, x)) \bar{m}(s, dx) = \int_H f_n(s, x) \bar{m}(s, dx).$$

Plugging this equality into (6.2), we obtain

$$\begin{aligned} \int_H \bar{v}_n(T, x) \bar{m}(t, dx) - \int_0^T \left(\int_H f_n(s, x) \bar{m}(s, dx) \right) ds &= \\ &- \int_0^T \left(\int_H [\langle \mathcal{H}_p^0(x, Dv_1(s, x)), D\bar{v}_n(s, x) \rangle] m_1(s, dx) - \int_H [\langle \mathcal{H}_p^0(x, Dv_2(s, x)), D\bar{v}_n(s, x) \rangle] m_2(s, dx) \right) ds. \end{aligned} \quad (6.3)$$

We now need to pass to the limit in (6.3) as $n \rightarrow \infty$. By the definition of $\mathcal{K}$ -convergence in $C_{b,\gamma}([0, T] \times H)$ (Definition Appendix B.2), there is $M \in \mathbb{R}$ such that for every $(s, x) \in [0, T] \times H$,

$$\sup_{n \geq 1} |f_n(s, x) - f(s, x)| = (T - t)^{-\gamma} \sup_{n \geq 1} \left[(T - s)^\gamma |f_n(s, x) - f(s, x)| \right] \leq M(T - s)^{-\gamma}.$$

Hence, since $(T - s)^{-\gamma} \in L^1(0, T)$ and $\lim_{n \rightarrow \infty} |f_n(s, x) - f(s, x)| = 0$ for every $(s, x) \in [0, T) \times H$, we obtain

$$\lim_{n \rightarrow \infty} \int_0^T \left(\int_H |f_n(s, x) - f(s, x)| \bar{m}(s, dx) \right) ds = 0$$

by Dominated Convergence Theorem. The convergence of the other terms follows directly by Dominated Convergence Theorem, since the functions involved are uniformly bounded. Therefore, passing to the limit in (6.3), we get

$$\begin{aligned} 0 &= - \int_H \bar{v}(T, x) \bar{m}(t, dx) + \int_0^T \left(\int_H f(s, x) \bar{m}(s, dx) \right) ds \\ &\quad - \int_0^T \left(\int_H [\langle \mathcal{H}_p^0(x, Dv_1(s, x)), D\bar{v}(s, x) \rangle] m_1(s, dx) - \int_H [\langle \mathcal{H}_p^0(x, Dv_2(s, x)), D\bar{v}(s, x) \rangle] m_2(s, dx) \right) ds. \end{aligned} \quad (6.4)$$

Hence, using the definition of f , we get

$$\begin{aligned} 0 &= \int_H \bar{u}(T, x) \bar{m}(t, dx) + \int_0^T \int_H [F(x, m_1(s)) - F(x, m_2(s))] \bar{m}(s, dx) \\ &\quad + \int_0^T \int_H [\mathcal{H}^0(x, Dv_2(s, x)) - \mathcal{H}^0(x, Dv_1(s, x)) - \langle \mathcal{H}_p^0(x, Dv_1(s, x)), -D\bar{v}(s, x) \rangle] m_1(s, dx) \\ &\quad + \int_0^T \int_H [\mathcal{H}^0(x, Dv_1(s, x)) - \mathcal{H}^0(x, Dv_2(s, x)) - \langle \mathcal{H}_p^0(x, Dv_2(s, x)), D\bar{v}(s, x) \rangle] m_2(s, dx). \end{aligned}$$

By Assumption 6.1(ii)-(iii) all terms above are nonnegative, so they all must be equal to 0. We now conclude as follows.

- (a) If case (a) of Assumption 6.1(iv) holds, we have $m_1 = m_2$; then uniqueness of mild solutions of equation (HJB) gives $v_1 = v_2$.
- (b) If case (b) of Assumption 6.1(iv) holds, we first use uniqueness of mild solutions of equation (HJB), which holds since

$$F(x, m_1(s)) = F(x, m_2(s)), \quad G(x, m_1(T)) = G(x, m_2(T)),$$

to obtain $v_1 = v_2$; then, we conclude that $m_1 = m_2$ by uniqueness of weak solutions of (3.4) with $w(s, x) := \mathcal{H}_p^0(x, Dv_1(s, x)) = \mathcal{H}_p^0(x, Dv_2(s, x))$.

- (c) If case (c) of Assumption 6.1(iv) holds, we first obtain

$$\mathcal{H}_p^0(x, Dv_1(s, x)) = \mathcal{H}_p^0(x, Dv_2(s, x)),$$

$m_1(s)$ and $m_2(s)$ a.e. for $s \in [0, T]$. This implies $m_1 = m_2$ by uniqueness of weak solutions of (3.4) with $w(s, x) = \mathcal{H}_p^0(x, Dv_1(s, x))$ (or equivalently with $w(s, x) = \mathcal{H}_p^0(x, Dv_2(s, x))$) and then conclude that $v_1 = v_2$ using uniqueness of mild solutions of equation (HJB). $\square$

7. Examples from stochastic optimal control

A typical application of our framework is when the Hamiltonian $\mathcal{H}$ arises from stochastic optimal control problems. To illustrate that, consider the H -valued stochastic optimal control problem with value function

$$v(t, x) = \inf_{\alpha \in \mathcal{A}_t} \mathbb{E} \left[\int_t^T f(X_s^{t,x,\alpha}, \alpha_s, m(s)) ds + g(X_T^{t,x,\alpha}, m(T)) \right], \quad (7.1)$$

where

$$\mathcal{A}_t = \{ \alpha : [t, T] \times \Omega \rightarrow \Lambda \text{ } (\mathcal{F}_s^t) \text{ -- progressively measurable} \},$$

defined on some reference probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s^t)_{s \geq t}, \mathbb{P}, W)$, where Λ is a complete separable metric space, $m(\cdot) \in \mathcal{S}$, and $X^{t,x,\alpha}$ solves, in mild sense on $[t, T]$, the infinite dimensional SDE

$$dX_s = (AX_s + b(X_s, \alpha_s, m(s)))ds + dW_s, \quad X_t = x. \quad (7.2)$$

We then have

$$\mathcal{H}(x, p, \mu) = \sup_{\alpha \in \Lambda} \{ - \langle b(x, \alpha, \mu), p \rangle - f(x, \alpha, \mu) \}.$$

We make the following assumptions:

- (i) $\Lambda = B_R \subset H$;
- (ii) $b(x, \alpha, \mu) = -\alpha + b_0(x)$, where $b_0 : H \rightarrow H$ is bounded and Lipschitz continuous;
- (iii) $f(x, \alpha, \mu) = f_0(x, \mu) + f_1(|\alpha|)$, where $f_0 : H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}$ is continuous in all variables, $f_0(\cdot, \delta_0)$ is bounded, f_0 is Lipschitz continuous in m with respect to $\mathbf{d}_1$, uniformly in x , and $f_1 \in C_{\text{loc}}^{1,1}(\mathbb{R})$ is symmetric and uniformly convex⁴;
- (iv) $g : H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}$ is continuous and $g(\cdot, \delta_0)$ is bounded.

Then,

$$\begin{aligned} \mathcal{H}(x, p, \mu) &= \sup_{|\alpha| \leq R} \{ \langle \alpha, p \rangle - f_1(|\alpha|) \} - \langle b_0(x), p \rangle - f_0(x, \mu) \\ &=: \mathcal{H}^1(p) - \langle b_0(x), p \rangle - f_0(x, \mu). \end{aligned}$$

In this case

$$\mathcal{H}^0(x, p) = \mathcal{H}^1(p) - \langle b_0(x), p \rangle, \quad F(x, \mu) = f_0(x, \mu), \quad G(x, \mu) = g(x, \mu),$$

and it is easy to see that Assumption 2.6(i) is satisfied. Regarding Assumption 2.6(ii), we compute explicitly

$$\mathcal{H}^1(p) = \begin{cases} (f_1')^{-1}(|p|)|p| - f_1((f_1')^{-1}(|p|)), & \text{if } |p| < f_1'(R), \\ R|p| - f_1(R), & \text{if } |p| \geq f_1'(R). \end{cases}$$

⁴ This means that there exists $\eta > 0$ such that $s \mapsto f_1(s) - \eta s^2$ is convex.

Since $f_1 \in C^{1,1}(\mathbb{R})$ and is uniformly convex, we deduce that $(f'_1)^{-1}$ is Lipschitz continuous, so $\mathcal{H}$ is Lipschitz continuous in p too. Moreover, we have

$$D\mathcal{H}^1(p) = \begin{cases} (f'_1)^{-1}(|p|) \frac{p}{|p|}, & \text{if } |p| < f'_1(R), \\ R \frac{p}{|p|}, & \text{if } |p| \geq f'_1(R), \end{cases}$$

and clearly this function is Lipschitz continuous. Hence, $\mathcal{H}_p^0(x, p) = D\mathcal{H}^1(p) - b_0(x)$ is bounded and Lipschitz continuous in p , uniformly in $x \in H$. This makes our existence result, Theorem 5.6, applicable.

As for uniqueness, examples of functions F satisfying Assumption 6.1(iii) when $H = \mathbb{R}^d$ are for instance in Section 3.4.2 of [13], volume I, and similar examples also work in a real separable Hilbert space H and satisfy the conditions imposed in this paper. We present two examples.

Example 7.1. Let $h_1 : H \times \mathcal{P}_1(H) \rightarrow \mathbb{R}$ and $h_2 : H \times \mathcal{P}_1(H) \rightarrow H$ be Lipschitz continuous and bounded. We define

$$F_1(x, \mu) := h_1(x) \int_H h_1(y) \mu(dy), \quad F_2(x, \mu) := \left\langle h_2(x), \int_H h_2(y) \mu(dy) \right\rangle.$$

We then have

$$\int_H [F_i(x, \mu_1) - F_i(x, \mu_2)] (\mu_1 - \mu_2)(dx) = \left| \int_H h_i(y) (\mu_1 - \mu_2)(dy) \right|^2, \quad i = 1, 2.$$

Hence, Assumption 6.1(iii) and Assumption 6.1(iv)(b) are satisfied for F_1 and F_2 . It is also easy to see that, for every $x \in H$,

$$|F_i(x, \mu_1) - F_i(x, \mu_2)| \leq C_i \mathbf{d}_1(\mu_1, \mu_2), \quad i = 1, 2,$$

where C_i is the Lipschitz constant of h_i . Thus both functions satisfy Assumption 2.6(i).

Example 7.2. Let $\ell : H \times [0, \infty) \rightarrow \mathbb{R}$ be bounded, continuous, and such that $\ell(x, \cdot)$ is strictly increasing for every $x \in H$ and is Lipschitz continuous with Lipschitz constant independent of x . Let $\rho : H \rightarrow [0, \infty)$ be bounded and Lipschitz continuous. Let ν be a positive finite measure on H which is full.⁵ We define

$$F(x, \mu) := \int_H \ell(z, \rho * \mu(z)) \rho(z - x) \nu(dz),$$

where $\rho * m$ denotes the convolution of ρ with μ , that is

$$\rho * \mu(z) = \int_H \rho(z - u) \nu(du).$$

We notice first that this convolution operation is Lipschitz continuous in μ and z . Indeed, let $z, y \in H$ and $\mu_1, \mu_2 \in \mathcal{P}_1(H)$ and let $\gamma \in \Gamma(\mu_1, \mu_2)$. Then,

⁵ That is $\nu(D) > 0$ for every open set D in H . An example of a measure which is full is the nondegenerate Gaussian measure with mean $a \in H$ and covariance operator $Q \in \mathcal{L}_1^+(H)$, see [16, Proposition 1.25] (nondegeneracy means that $\ker Q = \{0\}$).

$$\begin{aligned}
|\rho * \mu_1(z) - \rho * \mu_2(y)| &= \left| \int_H \rho(z-u) \mu_1(du) - \int_H \rho(y-w) \mu_2(dw) \right| \\
&= \left| \int_H (\rho(z-u) - \rho(y-w)) \gamma(du, dw) \right| \leq C \int_H |(z-u) - (y-w)| \gamma(du, dw) \\
&\leq C|z-y| + C \int_H |u-w| \gamma(du, dw).
\end{aligned}$$

Taking the infimum over all $\gamma \in \Gamma(m_1, m_2)$ we thus obtain

$$|\rho * \mu_1(z) - \rho * \mu_2(y)| \leq C(|z-y| + \mathbf{d}_1(\mu_1, \mu_2)).$$

We now have

$$\begin{aligned}
&\int_H [F(x, \mu_1) - F(x, \mu_2)] (\mu_1 - \mu_2)(dx) \\
&= \int_H [\ell(z, \rho * \mu_1(z)) - \ell(z, \rho * \mu_2(z))] \int_H \rho(z-x) (\mu_1(dx) - \mu_2(dx)) \nu(dz) \\
&= \int_H [\ell(z, \rho * \mu_1(z)) - \ell(z, \rho * \mu_2(z))] [\rho * \mu_1(z) - \rho * \mu_2(z)] \nu(dz) \geq 0,
\end{aligned}$$

because $\ell(z, \cdot)$ is increasing. Moreover, if the last expression is equal to 0, using that $\ell(z, \cdot)$ is strictly increasing, we conclude that $\rho * \mu_1(z) - \rho * \mu_2(z) = 0, \mu$ a.e. However, since this function is continuous and μ is full, this implies that $\rho * \mu_1(z) = \rho * \mu_2(z)$ for every $z \in H$. Hence, $F(\cdot, \mu_1) = F(\cdot, \mu_2)$ on H . Therefore, F satisfies Assumption 6.1(iii) and Assumption 6.1(iv)(b). Finally, we compute

$$\begin{aligned}
|F(x, \mu_1) - F(x, \mu_2)| &\leq C \int_H |\ell(z, \rho * \mu_1(z)) - \ell(z, \rho * \mu_2(z))| \mu(dz) \\
&\leq C \int_H |\rho * \mu_1(z) - \rho * \mu_2(z)| \mu(dz) \leq C \mathbf{d}_1(\mu_1, \mu_2)
\end{aligned}$$

which means that F satisfies Assumption 2.6(i).

7.1. An economic model

We illustrate here an application to an economic model in the presence of capital heterogeneity. Inspired by the model presented in [1, Sec. 3(a)], a mean field game model of competition among firms operating in the same sector was studied in [10]. Precisely, in [10], the production capacity of a representative firm is assumed to follow the dynamics

$$dK_s = -\delta K_s ds + \sigma K_s dB_s + I_s ds, \quad K_0 = k > 0,$$

where $\delta, \sigma > 0$, $(I_s)_{s \geq 0}$ is a nonnegative real valued process representing the investment rate of the firm, and B is a standard one dimensional Brownian motion. According to the discussion in [1, Sec. 3(a)], the goal of a representative firm is to solve

$$v(t, k) := \inf_{I.} \mathbb{E} \left[\int_t^T \left(-\pi(K_s, \mathcal{K}(m(s))) + \frac{1}{2} I_s^2 \right) ds \right],$$

where $I.$ belong to a suitable set of admissible control processes,

$$\mathcal{K}(\mu) = \left(\int_{\mathbb{R}} k^\theta \mu(dk) \right)^{1/\theta} \quad \mu \in \mathcal{P}_\theta(\mathbb{R}),$$

with $\theta > 0$ and $\pi : (0, \infty)^2 \rightarrow \mathbb{R}$, $(k, \mathcal{K}) \mapsto \pi(k, \mathcal{K})$ is nondecreasing in k and nonincreasing in $\mathcal{K}$.⁶

This model can be modified as follows. First of all, in order to fit our setting, we assume that the noise is additive⁷ with the diffusion coefficient 1; that is, we look at the dynamics of the form

$$dK_s = -\delta K_s ds + dB_s + I_s ds.$$

We may then enrich the model assuming that, rather than having a large number of single firms, we have a large number of networks of firms, each network being distributed around a circle S^1 . The location of the firm in the representative network is parametrized by $a \in S^1$ and capital K depends on this parameter. Therefore, the representative network's state is represented by a parametrized stochastic process $K_s(a)$, where $s \in [t, T]$ and $a \in S^1$. We may assume, following the modeling of [9], that also the investment rate process $I_s(a)$ is parametrized by the location $a \in S^1$, and that capital $K_s(a)$ naturally evolves in time-space according to a depreciation law and, by taking account of the “spatial” dimension $a \in S^1$, according to a diffusion law; that is, according to the overall law

$$\frac{\partial^2 K_s(a)}{\partial a^2} - \delta K_s(a).$$

Assuming also that the noise depends also on the location a , this leads to the following controlled SPDE:

$$dK_s(a) = \left(\frac{\partial^2 K_s(a)}{\partial a^2} - \delta K_s(a) \right) ds + dB_s(a) + I_s(a) ds, \quad K_0(a) = k(a). \quad (7.3)$$

We consider the following generalization of the optimization problem above:

$$v(t, k(\cdot)) = \inf_{I.(\cdot)} \mathbb{E} \left[\int_t^T \left(\int_{S^1} \left(-\pi(K_s(a), \mathcal{K}(m(s))) + \frac{1}{2} I_s(a)^2 \right) da \right) ds \right], \quad (7.4)$$

where

$$\mathcal{K}(\mu) = \int_H \left(\int_{S^1} f(k(a)) da \right) \mu(dk), \quad \mu \in \mathcal{P}_1(H), \quad (7.5)$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a given Lipschitz and bounded function. Formally setting $W_s = B_s(\cdot)$ and assuming that W is a cylindrical Brownian motion in the Hilbert space H , and formally setting $x = k(\cdot)$, $X_s = K_s(\cdot)$,

⁶ In [10], the case $\theta = 1$ and $\pi(k, \mathcal{K}) = \mathcal{K}^{-\beta} k$, where $\beta > 0$, is considered.

⁷ This is done to fit our additive noise setting. Note that, in this way, the capacity may become negative and this is of course an unpleasant fact from the point of view of applications. It would thus be interesting to investigate, in future works, cases with non-additive noise which may ensure that the infinite dimensional process X lies in the positive cone of $H = L^2(S^1; \mathbb{R})$; see, e.g., [31].

$\alpha_s = -I_s(\cdot)$, the SPDE (7.3) can be viewed in the setting of the present section as a controlled SDE (7.2) in the Hilbert space $H = L^2(S^1; \mathbb{R})$ with the following specifications:

$$A = \frac{\partial^2}{\partial a^2} - \delta, \quad b(x, \alpha) = -\alpha.$$

Similarly, the optimal control problem can be reformulated as (7.1) by setting $g = 0$ and

$$f(x, \alpha, \mu) = \frac{|\alpha|^2}{2} + F(x, \mu),$$

where

$$F(x, \mu) = - \int_{S^1} \pi(x(a), \mathcal{K}(\mu)) da.$$

Imposing the constraint $I_t(\cdot) \in \Lambda = B_R$ and assuming that π is Lipschitz and bounded, the conditions for our existence result (Theorem 5.6) are fulfilled.

Regarding uniqueness, we assume that $\pi(k, \mathcal{K}) = \pi_1(\mathcal{K})f(k)$, where π_1 is Lipschitz continuous, nonnegative, bounded, and strictly decreasing, and f is the same function appearing in (7.5) and required now to be Lipschitz continuous, nonnegative, nondecreasing, and bounded. Then, for every $\mu_1, \mu_2 \in \mathcal{P}_1(H)$, we have

$$\begin{aligned} & \int_H [F(x, \mu_1) - F(x, \mu_2)](\mu_1 - \mu_2)(dx) \\ &= - \int_H \left[\int_{S^1} (\pi(x(a), \mathcal{K}(\mu_1)) - \pi(x(a), \mathcal{K}(\mu_2))) da \right] (\mu_1 - \mu_2)(dx) \\ &= - (\pi_1(\mathcal{K}(\mu_1)) - \pi_1(\mathcal{K}(\mu_2))) \int_H \left(\int_{S^1} f(x(a)) da \right) (\mu_1 - \mu_2)(dx) \\ &= - (\pi_1(\mathcal{K}(\mu_1)) - \pi_1(\mathcal{K}(\mu_2))) (\mathcal{K}(\mu_1) - \mathcal{K}(\mu_2)) \geq 0, \end{aligned}$$

with equality if and only if $\mathcal{K}(\mu_1) = \mathcal{K}(\mu_2)$. This shows that:

- (i) Assumption 6.1(iii) is fulfilled;
- (ii) Since

$$(\pi_1(\mathcal{K}(\mu_1)) - \pi_1(\mathcal{K}(\mu_2))) (\mathcal{K}(\mu_1) - \mathcal{K}(\mu_2)) = 0 \implies \mathcal{K}(\mu_1) = \mathcal{K}(\mu_2),$$

also Assumption 6.1(iv)(b) holds.

The remaining requirements of Assumption 6.1 are fulfilled by construction. Hence, our uniqueness result, Theorem 6.2, applies.

Acknowledgements

Salvatore Federico is a member of the GNAMPA-INdAM group and was supported by the PRIN project 2022BEMMLZ “Stochastic control and games and the role of information”, financed by the Italian Ministry of University and Research. Fausto Gozzi has been supported by the Italian Ministry of University

and Research (MIUR), in the framework of PRIN projects 2017FKHBA8 001 (The Time-Space Evolution of Economic Activities: Mathematical Models and Empirical Applications) and 20223PNJ8K (Impact of the Human Activities on the Environment and Economic Decision Making in a Heterogeneous Setting: Mathematical Models and Policy Implications).

The authors are grateful to Mauro Rosetolato for fruitful discussions on the definition of weak solution to the Fokker-Planck equation.

Appendix A. Compactness in the space of probability measures

We provide a result about compactness of sets in spaces of probability measures on H which is used in the paper.

Lemma Appendix A.1. *Let $L > 0$ and*

$$\mathcal{Q}_L := \left\{ \mu \in \mathcal{P}_1(H) : \int_H |x|^2 \mu(dx) \leq L \right\}.$$

Then

$$\lim_{R \rightarrow \infty} \sup_{\mu \in \mathcal{Q}_L} \int_{B_R^c} |x| \mu(dx) \rightarrow 0.$$

Proof. Since

$$\sup_{\mu \in \mathcal{Q}_L} \int_{B_R^c} |x| \mu(dx) \leq \sup_{\mu \in \mathcal{Q}_L} \frac{1}{R} \int_{B_R^c} |x|^2 \mu(dx) \leq \frac{1}{R} \sup_{\mu \in \mathcal{Q}_L} \int_H |x|^2 \mu(dx) = \frac{L}{R},$$

and the claim follows. $\square$

Proposition Appendix A.2. *Let $S \subset \mathcal{P}_1(H)$. If, for some constant $c_1 > 0$, we have*

$$\sup_{\mu \in S} \int_H |x|^2 \mu(dx) \leq c_1 \quad \text{and} \quad \lim_{N \rightarrow \infty} \sup_{\mu \in S} \sum_{i=N}^{\infty} \int_H \langle x, \mathbf{e}_i \rangle^2 \mu(dx) = 0,$$

then S is relatively compact with respect to $\mathbf{d}_1$.

Proof. By [40, Ch. VI, Th. 2.2],⁸ the set S is tight. The claim follows using [2, Prop. 7.1.5]. $\square$

Appendix B. Classical and strong solutions to Kolmogorov equations

In this section, we define classical, mild, and $\mathcal{K}$ -strong solutions of backward Kolmogorov equations and recall an approximation result used in Section 6. The material is mainly taken from [22, Section B.7] and is presented, for the reader's convenience, in a simplified form needed here. We first consider the following terminal value problem (Kolmogorov equation) in the Hilbert space H :

⁸ In [40, Ch. VI, Th. 2.2] it is stated that the second condition alone guarantees relative compactness of S . This is clearly not true, as can be seen by simply taking as S a non compact sequence concentrated on the line generated by $\mathbf{e}_1$. However the proof of this theorem there remains valid if we add the assumption of uniform boundedness of second moments.

$$\begin{cases} \partial_t u(t, x) + L_0 u(t, x) = f(t, x), & (t, x) \in [0, T] \times H, \\ u(T, x) = \varphi(x), & x \in H, \end{cases} \quad (\text{B.1})$$

where L_0 is given in (3.3), $\varphi \in C_b(H)$ and $f \in C_{b,\gamma}([0, T] \times H)$ (recall that $C_{b,\gamma}([0, T] \times H)$ was defined in (2.4)). Since, as recalled in a footnote in Section 3, the closure (in a suitable weak sense) of the operator L_0 is the generator (again in a suitable weak sense) of the semigroup R_t in (2.10), we can formally rewrite (B.1) in the following mild form:

$$u(t, x) = R_{T-t}[\varphi](x) + \int_t^T R_{s-t}[f(s, \cdot)](x) ds. \quad (\text{B.2})$$

We call the function u above, defined by the right hand side, the *mild solution* of (B.1). Recalling the definition of L_0 in (3.3), we define

$$D(L_0) = \left\{ \phi \in UC_b^2(H) : D^2\phi \in UC_b(H; \mathcal{L}_1(H)), A^*D\phi \in UC_b(H; H) \right\}, \quad (\text{B.3})$$

and endow $D(L_0)$ with the norm

$$\|\phi\|_{D(L_0)} := \|\phi\|_0 + \|D\phi\|_0 + \|A^*D\phi\|_0 + \sup_{x \in H} \|D^2\phi(x)\|_{\mathcal{L}_1(H)}. \quad (\text{B.4})$$

Arguing as in Theorem 2.7 of [19], it can be proved that $D(L_0)$, endowed with the above norm, is a Banach space.⁹ We now recall a notion of classical solution to (B.1) (see [22, Definition B.82]).¹⁰

Definition Appendix B.1. We say that $u \in C_b([0, T] \times H)$ is a *classical solution* to (B.1) in $D(L_0)$ if

$$\begin{cases} u(\cdot, x) \in C^1([0, T]), \quad \forall x \in H, \\ u(t, \cdot) \in D(L_0) \text{ for any } t \in [0, T] \text{ and } \sup_{t \in [0, T]} \|u(t, \cdot)\|_{D(L_0)} < \infty, \\ Du, A^*Du \in C_b([0, T] \times H, H), \quad D^2u \in C_b([0, T] \times H, \mathcal{L}_1(H)) \end{cases} \quad (\text{B.5})$$

and u satisfies (B.1) for every $(t, x) \in [0, T] \times H$.

We stress that Definition Appendix B.1 implicitly implies that $\varphi \in D(L_0)$ and $f \in C([0, T] \times H)$.

To introduce a notion of a strong solution, called $\mathcal{K}$ -strong solution, we first recall a special case of the definition of $\mathcal{K}$ -convergence which is needed in this paper (see [22, Definition B.84]).

Definition Appendix B.2. Let Z be a real Hilbert space. Given $(f_n)_{n \in \mathbb{N}} \subset C_{b,\gamma}([0, T] \times H, Z)$, we say that f_n is $\mathcal{K}$ -convergent to $f \in C_{b,\gamma}[0, T] \times H, Z)$ if

$$\begin{cases} \sup_{n \in \mathbb{N}} \|f_n\|_{C_{b,\gamma}[0, T] \times H, Z} < \infty, \\ \lim_{n \rightarrow \infty} \sup_{(t, x) \in (0, T] \times K} (T - t)^\gamma |f_n(t, x) - f(t, x)| = 0, \end{cases} \quad (\text{B.6})$$

⁹ The definition of the domain of the operator studied in this paper is slightly different, but the arguments of the mentioned reference can be adapted easily. We also notice that the Banach space structure is not essential for our purposes, even if it simplifies the notation.

¹⁰ We also observe that there are other definitions of classical solutions in the literature, see e.g. Section 6.2 of [20, Section 6.2], [28, Definition 4.6], [29, Definition 4.1]. We refer interested readers to [22, Section B.7.1] for a detailed overview on that.

for every compact set $K \subset H$. In this case, we write $\mathcal{K} - \lim_{n \rightarrow \infty} f_n = f$ in $C_{b,\gamma}[0, T] \times H, Z$.

The definition of a $\mathcal{K}$ -strong solution to (B.1) provided below is a special case of a more general definition that can be found in [22, Definition B.85].

Definition Appendix B.3. Let $\gamma \in (0, 1)$. Let $\varphi \in C_b(H)$ and $f \in C_{b,\gamma}([0, T] \times H)$. We say that a function $u \in C_b([0, T] \times H)$ is a $\mathcal{K}$ -strong solution in $D(L_0)$ of (B.1) if $u(t, \cdot)$ is Fréchet differentiable for each $t \in [0, T]$ and there exist sequences $(u_n)_{n \in \mathbb{N}} \subset C_b([0, T] \times H)$, $(\varphi_n)_{n \in \mathbb{N}} \subset D(L_0)$, $(f_n)_{n \in \mathbb{N}} \subset C_{b,\gamma}([0, T] \times H)$ such that:

(i) For every $n \in \mathbb{N}$, u_n is a classical solution in $D(L_0)$ (cf. Definition Appendix B.1) to

$$\begin{cases} w_t = L_0 w + f_n, \\ w(0) = \varphi_n. \end{cases} \quad (\text{B.7})$$

(ii) The following limits hold

$$\begin{cases} \mathcal{K}^- \lim_{n \rightarrow \infty} \varphi_n = \varphi, & \text{in } C_b(H), \\ \mathcal{K}^- \lim_{n \rightarrow \infty} u_n = u, & \text{in } C_b([0, T] \times H), \end{cases}$$

and

$$\begin{cases} \mathcal{K}^- \lim_{n \rightarrow \infty} f_n = f, & \text{in } C_{b,\gamma}([0, T] \times H) \\ \mathcal{K}^- \lim_{n \rightarrow \infty} Du_n = Du & \text{in } C_{b,\gamma}([0, T] \times H, H). \end{cases}$$

We end this section with an approximation result which is a stronger version of [22, Theorem B.95(i)].

Theorem Appendix B.4. $\phi \in C_b(H)$ and $f \in C_{b,\gamma}([0, T] \times H)$. Then the mild solution u (defined in (B.2)) of equation (B.1), is also a $\mathcal{K}$ -strong solution of to (B.1). Moreover the sequence (u_n) can be chosen so that $u_n \in \mathcal{D}_T$ for all $n \in \mathbb{N}$.

Proof. We can apply [22, Theorem B.95 (i)] (in the case when, in the notation there, $m = 0$, $\eta(t) = t^{-\gamma}$), to conclude that u is also a $\mathcal{K}$ -strong solution of the same equation. Moreover, from the proof of [22, Theorem B.95 (i)] one can easily see that the approximating data φ_n , f_n , and consequently the approximating solutions u_n , can be chosen so that $\varphi_n \in D(L_0)$ and $f_n, \partial_t \bar{v}_n \in C_b([0, T] \times H)$ for all $n \in \mathbb{N}$. As a consequence, we get that $u_n \in \mathcal{D}_T$ for all $n \in \mathbb{N}$. $\square$

References

- [1] Y. Achdou, F.J. Buera, J.-M. Lasry, P.-L. Lions, B. Moll, Partial differential equation models in macroeconomics, *Phil. Trans. R. Soc. A* 372 (2014).
- [2] L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Springer Science & Business Media, 2008.
- [3] V. Barbu, G. Da Prato, *Hamilton-Jacobi Equations in Hilbert Spaces*, Pitman Research Notes in Mathematics Series, vol. 86, Longman, Boston, 1983.
- [4] D.P. Bertsekas, S.E. Shreve, *Stochastic Optimal Control – The Discrete Time Case*, Mathematics in Science and Engineering, vol. 139, Academic Press, New York - San Francisco - London, 1978.
- [5] V. Bogachev, G. Da Prato, M. Röckner, Fokker–Planck equations and maximal dissipativity for Kolmogorov operators with time-dependent singular drifts in Hilbert spaces, *J. Funct. Anal.* 256 (4) (2009) 1269–1298.

- [6] V. Bogachev, G. Da Prato, M. Röckner, Existence and uniqueness of solutions for Fokker–Planck equations on Hilbert spaces, *J. Evol. Equ.* 10 (2010) 487–509.
- [7] V. Bogachev, G. Da Prato, M. Röckner, Uniqueness for solutions of Fokker–Planck equations on infinite-dimensional spaces, *Commun. Partial Differ. Equ.* 36 (6) (2011) 925–939.
- [8] V. Bogachev, G. Da Prato, M. Röckner, W. Stannat, Uniqueness of solutions to weak parabolic equations for measures, *Bull. Lond. Math. Soc.* 39 (4) (2007) 631–640.
- [9] R. Boucekkine, G. Fabbri, S. Federico, F. Gozzi, Growth and agglomeration in the heterogeneous space: a generalized AK approach, *J. Econ. Geogr.* 19 (2019) 1287–1318.
- [10] A. Calvia, S. Federico, G. Ferrari, F. Gozzi, A mean-field model of optimal investment, arXiv:2404.02871, 2024.
- [11] P. Cardaliaguet, A. Porretta, *An Introduction to Mean Field Game Theory*, Lecture Notes in Math., vol. 2281, Springer, Cham, 2020.
- [12] R. Carmona, F. Delarue, *Mean Field Games*. Progress in Nonlinear Differential Equations and Their Applications, Springer, Cham, 2017.
- [13] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer, 2018.
- [14] R. Carmona, J.P. Fouque, S.M. Mousavi, L.-H. Sun, Systemic risk and stochastic games with delay, *J. Optim. Theory Appl.* 179 (2018) 366–399.
- [15] S. Cerrai, Weakly continuous semigroups in the space of functions with polynomial growth, *Dyn. Syst. Appl.* 4 (1995) 351–372.
- [16] G. Da Prato, *An Introduction to Infinite-Dimensional Analysis*, Springer Science & Business Media, 2006.
- [17] G. Da Prato, F. Flandoli, E. Priola, M. Röckner, Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift, *Ann. Probab.* 41 (5) (2013) 3306–3344.
- [18] G. Da Prato, F. Flandoli, M. Röckner, Fokker–Planck equations for SPDE with non-trace-class noise, *Commun. Math. Stat.* 1 (2013) 281–304.
- [19] G. Da Prato, J. Zabczyk, Regular densities of invariant measures in Hilbert spaces, *J. Funct. Anal.* 130 (2) (1995) 427–449.
- [20] G. Da Prato, J. Zabczyk, *Second Order Partial Differential Equations in Hilbert Spaces*, London Mathematical Society Lecture Note Series, vol. 293, Cambridge University Press, Cambridge, 2002.
- [21] G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and Its Applications., vol. 152, Cambridge University Press, Cambridge, 2014.
- [22] G. Fabbri, F. Gozzi, A. Święch, *Stochastic Optimal Control in Infinite Dimension*, Probability Theory and Stochastic Modeling, vol. 82, Springer, Switzerland, 2017.
- [23] S. Faggian, F. Gozzi, P.M. Kort, Optimal investment with vintage capital: equilibrium distributions, *J. Math. Econ.* 96 (2021).
- [24] S. Federico, F. Gozzi, D. Ghilli, Linear-quadratic mean field games in Hilbert spaces, arXiv:2402.14935, 2024.
- [25] G. Feichtinger, R.F. Hartl, P.M. Kort, V.M. Veliov, Anticipation effects of technological progress on capital accumulation: a vintage capital approach, *J. Econ. Theory* 126 (1) (2006) 143–164.
- [26] J.-P. Fouque, Z. Zhang, Mean field game with delay: a toy model, *Risks* 6 (3) (2018).
- [27] F. Gozzi, Regularity of solutions of a second order Hamilton–Jacobi equation and application to a control problem, *Commun. Partial Differ. Equ.* 20 (5–6) (1995) 775–826.
- [28] F. Gozzi, Strong solutions for Kolmogorov equation in Hilbert spaces 188 (1997) 163–187.
- [29] F. Gozzi, *Second Order Hamilton–Jacobi Equations in Hilbert Spaces and Stochastic Optimal Control*, PhD thesis, Scuola Normale Superiore, Pisa, 1998.
- [30] F. Gozzi, Second order Hamilton–Jacobi equations in Hilbert spaces and stochastic optimal control, in: G. Da Prato, L. Tubaro (Eds.), *Stochastic Partial Differential Equations and Applications*, in: Lecture Notes in Pure and Applied Mathematics, vol. 227, Dekker, New York, 2002, pp. 255–285.
- [31] F. Gozzi, M. Leocata, A stochastic model of economic growth in time-space, *SIAM J. Control Optim.* 60 (2) (2022) 620–651.
- [32] P.J. Graber, A.R. Mészáros, On monotonicity conditions for mean field games, *J. Funct. Anal.* 285 (9) (2023).
- [33] M. Huang, P.E. Caines, R.P. Malhamé, Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized ϵ -Nash equilibria, *IEEE Trans. Autom. Control* 52 (9) (2007) 1560–1571.
- [34] J.M. Lasry, P.-L. Lions, Jeux à champ moyen. I – le cas stationnaire, *C. R. Math.* 343 (9) (2006) 619–625.
- [35] J.M. Lasry, P.-L. Lions, Jeux à champ moyen. II – horizon fini et contrôle optimal, *C. R. Math.* 343 (10) (2006) 679–684.
- [36] J.M. Lasry, P.-L. Lions, Mean field games, *Jpn. J. Math.* 2 (1) (2007) 229–260.
- [37] H. Liu, D. Firoozi, Hilbert space-valued LQ mean field games: an infinite-dimensional analysis, arXiv:2403.01012, 2024.
- [38] A.R. Mészáros, C. Mou, Mean field games systems under displacement monotonicity, *SIAM J. Math. Anal.* 56 (1) (2024) 529–553.
- [39] M. Nisio, *Stochastic Control Theory*, ISI Lecture Notes., vol. 9, 2015.
- [40] K.R. Parthasarathy, *Probability Measures on Metric Spaces*, Probability and Mathematical Statistics, vol. 3, Academic Press, New York, 1967.
- [41] M. Röckner, R. Zhu, X. Zhu, A note on stochastic semilinear equations and their associated Fokker–Planck equations, *J. Math. Anal. Appl.* 415 (1) (2014) 83–109.
- [42] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, 2003.
- [43] S. Wiesinger, Uniqueness for solutions of Fokker–Planck equations related to singular SPDE driven by Lévy and cylindrical Wiener noise, *J. Evol. Equ.* 13 (2) (2013) 369–394.