

On the Asymptotic Nature of First Order Mean Field Games

Markus Fischer & Francisco J. Silva

Applied Mathematics & Optimization

ISSN 0095-4616

Appl Math Optim

DOI 10.1007/s00245-020-09711-1



Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.



On the Asymptotic Nature of First Order Mean Field Games

Markus Fischer¹ · Francisco J. Silva²

© The Author(s) 2020

Abstract

For a class of finite horizon first order mean field games and associated N -player games, we give a simple proof of convergence of symmetric N -player Nash equilibria in distributed open-loop strategies to solutions of the mean field game in Lagrangian form. Lagrangian solutions are then connected with those determined by the usual mean field game system of two coupled first order PDEs, and convergence of Nash equilibria in distributed Markov strategies is established.

Keywords Mean field games · Lagrangian form · Deterministic dynamics · Nash equilibrium · Distributed strategies

Mathematics Subject Classification 49N70 · 60B10 · 91A06 · 91A13

1 Introduction

The purpose of this article is to illustrate a simple way of establishing convergence of open-loop Nash equilibria in the case of first-order non-stationary Mean Field Games (MFGs). Introduced by Lasry and Lions and, independently, by Huang, Malhamé and Caines about fifteen years ago (cf. [33,36]), mean field games are limit models for non-cooperative symmetric N -player differential games as the number of players N tends to infinity; see, for instance, the lecture notes [13] and the recent two-volume work [18]. The notion of solution usually adopted for the prelimit models is that of a Nash equilibrium. A standard way of making the connection with the limit model

✉ Markus Fischer
fischer@math.unipd.it

Francisco J. Silva
francisco.silva@unilim.fr

¹ Dipartimento di Matematica “Tullio Levi-Civita”, Università degli Studi di Padova, via Trieste, 63, 35121 Padova, Italia

² Institut de recherche XLIM-DMI, UMR-CNRS 7252, Faculté des Sciences et Techniques, Université de Limoges, 87060 Limoges, France

rigorous is to show that a solution of the mean field game yields approximate Nash equilibria for the N -player games, with approximation error vanishing as $N \rightarrow \infty$. In the opposite direction, one aims to prove that a sequence of N -player Nash equilibria converges, as N tends to infinity, to the mean field game limit.

When Nash equilibria are considered in stochastic open-loop strategies, then their convergence is well understood and can be established under mild conditions; see [28] and [34], both for finite horizon games with general, possibly degenerate, Brownian dynamics. The convergence analysis is much harder when Nash equilibria are defined over Markov feedback strategies with full state information.

A first result in this setting was given by Gomes, Mohr, and Souza [29] for continuous time games with finite state space. There, convergence of Markovian Nash equilibria is proved, but only if the time horizon is small enough. A breakthrough was achieved by Cardaliaguet, Delarue, Lasry, and Lions in [15]. In the setting of games with non-degenerate Brownian dynamics, possibly including common noise, convergence to the mean field game limit is established there for arbitrary time horizon provided the so-called master equation associated with the mean field game possesses a unique sufficiently regular solution. In this case, the convergence analysis can be refined, yielding not only convergence of minimal costs and propagation of chaos for the Nash equilibrium state trajectories, but also fluctuation and large deviations results for the associated empirical measures; see Cecchin and Pelino [19] and, independently, Bayraktar and Cohen [7] for finite state games and the papers by Delarue, Lacker, and Ramanan [23,24] for Brownian dynamics without or with common noise.

Well-posedness of the master equation implies uniqueness of solutions for the mean field game. But also the situation where the mean field game possesses multiple solutions, while the N -player Nash equilibria in full Markov feedback strategies are still uniquely determined, occurs. In this case, the convergence problem is in part open. The most general result appears to be the recent preprint [35] by Lacker. There, for a class of games with non-degenerate Brownian dynamics, it is shown that all limit points of the N -player Nash equilibria are concentrated on weak solutions of the mean field game; these are more general than randomizations of ordinary (“strong”) solutions of the mean field game. In the three recent works [25,38] and [20], the authors present case studies, giving criteria for characterizing those mean field game solutions that can be attained as limits of feedback Nash equilibria with full state information.

Here, we consider a much simpler situation: The underlying dynamics are deterministic with direct control of players’ states; randomness enters only through the players’ initial positions, which are assumed to be independently and identically distributed. Thanks to the deterministic dynamics without explicit interaction, players can directly control their entire state trajectories. Thus, the set of strategies (or actions) of each player consists of trajectory-valued functions depending on time and player’s own position. We call these actions *distributed open-loop strategies*. By considering randomizations of these strategies, obtaining what we call *randomized distributed open-loop strategies*, which are a kind of mixed strategies, we obtain existence of symmetric Nash equilibria in this new class of actions (through a standard fixed point argument), as well as precompactness of sequences of such equilibria. Convergence to the mean field game equilibrium in Lagrangian form (see e.g. [8,10,16,17,37] for the notion of Lagrangian equilibrium and also [30] for a related formulation) along

weakly converging subsequences of Nash equilibria is then established in Theorem 3.1 by showing that the variance of the associated empirical measures (evaluated by integrating over test functions from a countable family) tends to zero as $N \rightarrow \infty$. Thus, to obtain convergence, we essentially prove a law of large numbers for a triangular array $(Y_i^N)_{i \in \{1, \dots, N\}, N \in \mathbb{N}}$ where Y_1^N, \dots, Y_N^N are independent and identically distributed with common marginal law that however depends on N . Let us point out that after obtaining a suitable compactness property for the set of admissible strategies, the convergence result in Theorem 3.1 can actually be deduced from the more general results in [28] or [34]. The main purpose of this result is therefore to give a simple proof of convergence, which takes advantage of the structure of the dynamics, the deterministic nature of the underlying differential games, and the choice of distributed open-loop controls as admissible strategies. Existence or convergence of N -player Nash equilibria in feedback strategies with full state information are beyond the scope of the arguments presented here. The method of our proof is reminiscent of the techniques employed in nonatomic game theory (see [31,39,40]) and could also be useful in justifying the asymptotic nature of more sophisticated *deterministic* MFGs (see e.g. [26,37] for minimal-time MFGs and Remark 3.1(ii) for the state constrained MFG considered in [10]).

In the second part of this article, we consider additional second order assumptions on the data and we assume that the initial distribution is absolutely continuous with respect to the Lebesgue measure. In this framework, and following [16], we link the notion of Lagrangian MFG equilibrium to the well known PDE characterization of mean field games in terms of two coupled first order partial differential equations, namely a backward first order Hamilton-Jacobi-Bellman equation and a forward continuity equation; see equation (MFG) in Sect. 4.1 below. Under the stronger assumptions mentioned above, for each $N \in \mathbb{N}$, any symmetric randomized distributed open-loop Nash equilibrium of the N -player game can be identified with a symmetric distributed open-loop Nash equilibrium (non-randomized). Moreover, the strategies associated to such equilibria can be described by controls which are feedback with respect to the individual states. We call these actions *distributed Markov strategies*. The latter are determined by the solutions of a coupled system of first order PDEs analogous to the mean field game PDE system; see Eq. (MFG_N) in Sect. 4.2. This is in contrast with N -player Nash equilibria defined over Markov feedback strategies with full state information, which are determined through a system of N coupled PDEs. Building on the equivalence of characterizations and the convergence result in Theorem 3.1, we also establish in Theorem 4.1 the convergence of solutions of (MFG_N) to solutions of (MFG), as well as the convergence of the Nash equilibria in distributed Markov strategies.

The rest of this paper is organized as follows. In Sect. 2, we introduce the N -player games together with some notation and our standing assumptions. Existence of Nash equilibria in randomized distributed open-loop strategies is verified in Proposition 2.1. In Sect. 3, the associated mean field game is introduced in Lagrangian form; see Definition 3.1 and Proposition 3.1. We show convergence of symmetric N -player randomized distributed open-loop Nash equilibria to the mean field game limit in Theorem 3.1. Section 4 links, under additional assumptions, the mean field game in Lagrangian form with the mean field game PDE system. Similarly, N -player Nash equilibria in

distributed Markov strategies are described in terms of a PDE system analogous to the limit system, but dependent on the number of players N . The corresponding convergence results are given in Theorem 4.1 and Corollary 4.1, respectively.

2 The N -Player Game

Before introducing the N -player deterministic differential game that we are interested in, let us first fix some notations. Given a Polish space (X, d_X) , we denote by $\mathcal{P}(X)$ the set of probability measures over X . If (Y, d_Y) is another Polish space and $\varphi : X \rightarrow Y$ a Borel measurable function, then the *push-forward* $\varphi\#m \in \mathcal{P}(Y)$ of a measure $m \in \mathcal{P}(X)$ by φ is defined by

$$\varphi\#m(A) := m(\varphi^{-1}(A)) \quad \forall A \in \mathcal{B}(Y), \tag{2.1}$$

where $\mathcal{B}(Y)$ denotes the σ -algebra of Borel sets in Y . Given $p \in [1, +\infty)$ we denote by $\mathcal{P}_p(X)$ the set of probability measures μ over X such that $\int_X d_X(x, x_0)^p d\mu(x) < +\infty$ for some $x_0 \in X$. The set $\mathcal{P}_p(X)$ is endowed with the Monge-Kantorovich distance $d_p : \mathcal{P}_p(X) \times \mathcal{P}_p(X) \rightarrow [0, \infty)$ defined by

$$\begin{aligned} d_p(\mu_1, \mu_2) &:= \inf \left\{ \left(\int_{X \times X} d_X^p(x, y) d\gamma(x, y) \right)^{\frac{1}{p}} \mid \gamma \in \mathcal{P}(X \times X), \right. \\ &\quad \left. \pi_1\#\gamma = \mu_1, \pi_2\#\gamma = \mu_2 \right\}, \end{aligned}$$

where $\pi_i : X \times X \rightarrow X$ ($i = 1, 2$) is the projection on the i -th coordinate, that is, $\pi_i(x_1, x_2) := x_i$. It is well known (see e.g. [41]) that in the particular case $p = 1$ we have

$$d_1(\mu_1, \mu_2) = \sup \left\{ \int_X f(x) d(\mu_1 - \mu_2)(x) \mid f \in \text{Lip}_1(X) \right\}, \tag{2.2}$$

where $\text{Lip}_1(X)$ denotes the set of Lipschitz functions on X with Lipschitz constant equal to one.

We will fix as data some functions $\ell : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$, $\Phi : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ and a probability measure $m_0 \in \mathcal{P}(\mathbb{R}^d)$. The functions ℓ, Φ will represent running and terminal costs, respectively, while m_0 will be an initial state distribution. We will assume that:

(A1) (i) The functions ℓ and Φ are continuous. Moreover, the following properties hold true:

- (i.1) For every $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$, the function $\ell(\cdot, x, \mu)$ is convex.
- (i.2) There exist $q > 1, \underline{c}_\ell > 0, \bar{c}_\ell > 0$ and $C_\ell > 0$ such that

$$\underline{c}_\ell |\alpha|^q - C_\ell \leq \ell(\alpha, x, \mu) \leq \bar{c}_\ell |\alpha|^q + C_\ell \quad \forall \alpha \in \mathbb{R}^d, x \in \mathbb{R}^d, \mu \in \mathcal{P}_1(\mathbb{R}^d), \tag{2.3}$$

where we denote by $|\cdot|$ the Eulidean norm in \mathbb{R}^d .

- (i.3) The function Φ is bounded from below.
- (ii) The support of m_0 , denoted by $\text{supp}(m_0)$, is a compact subset of \mathbb{R}^d .
 A simple example of a function ℓ satisfying the requirements of **(A1)** is given by

$$\ell(\alpha, x, \mu) := \frac{1}{2}|\alpha|^2 + \left| x - \int_{\mathbb{R}^d} y \mu(dy) \right| \wedge C,$$

where C is some positive constant. Notice that the mapping $\mathcal{P}_1(\mathbb{R}^d) \ni \mu \mapsto \int_{\mathbb{R}^d} y \mu(dy) \in \mathbb{R}^d$ is continuous w.r.t. the topology induced by the distance d_1 .

Choose $T > 0$, the finite time horizon, and set $\Gamma := C([0, T]; \mathbb{R}^d)$, the space of continuous \mathbb{R}^d -valued trajectories on $[0, T]$. The space Γ is naturally endowed with the topology of uniform convergence. Let $W^{1,q}([0, T]; \mathbb{R}^d)$ denote the Sobolev space of \mathbb{R}^d -valued absolutely continuous functions that possess first order weak sense partial derivatives in $L^q((0, T); \mathbb{R}^d)$.

In order to introduce the game that we will consider, assume first that there are N individuals ($N \geq 2$), which, from now on, will be called *players*, positioned at $x_1, \dots, x_N \in \mathbb{R}^d$ at time $t = 0$. For $i \in \{1, \dots, N\}$, player i chooses a trajectory $\gamma_i \in \mathcal{A}(x_i)$, where

$$\mathcal{A}(x) := \left\{ \gamma \in \Gamma \mid \gamma \in W^{1,q}([0, T]; \mathbb{R}^d) \text{ and } \gamma(0) = x \right\}, \quad x \in \mathbb{R}^d.$$

Given these initial positions, the cost $j_i^N : \prod_{j=1}^N \mathcal{A}(x_j) \rightarrow \mathbb{R}$ for player i is defined by

$$j_i^N(\gamma_1, \dots, \gamma_N) := \int_0^T \ell(\dot{\gamma}_i(t), \gamma_i(t), (\gamma_j(t))_{j \neq i}) dt + \Phi(\gamma_i(T), (\gamma_j(T))_{j \neq i}), \tag{2.4}$$

where, for notational convenience, we have denoted

$$\begin{aligned} &\ell(\dot{\gamma}_i(t), \gamma_i(t), (\gamma_j(t))_{j \neq i}) \\ &:= \ell\left(\dot{\gamma}_i(t), \gamma_i(t), \frac{1}{N-1} \sum_{j \neq i} \delta_{\gamma_j(t)}\right), \text{ with the same convention for } \Phi. \end{aligned}$$

Assumption **(A1)** implies that j_i^N is well-defined. Note that defining

$$j^N(\gamma_1; (\gamma_j)_{j=2}^N) := \int_0^T \ell(\dot{\gamma}_1(t), \gamma_1(t), (\gamma_j(t))_{j=2}^N) dt + \Phi(\gamma_1(T), (\gamma_j(T))_{j=2}^N), \tag{2.5}$$

we have that $j_i^N(\gamma_1, \dots, \gamma_N) = j^N(\gamma_i; (\gamma_j)_{j \neq i})$.

In the game that we will consider, the initial position of each player is random, independent of the initial positions of the other players and with the same law m_0 . In this new context, we define the set of *distributed open-loop strategies* of the players as

$$\mathcal{A} := \left\{ \gamma : \mathbb{R}^d \rightarrow \Gamma \mid \gamma \text{ is Borel measurable and } \gamma(x) \in \mathcal{A}(x) \text{ for all } x \in \text{supp}(m_0) \right\}.$$

For notational simplicity, for $\gamma \in \mathcal{A}$, we will write $\gamma^x := \gamma(x)$. Given a profile of actions $(\gamma_1, \dots, \gamma_N) \in \mathcal{A}^N$, it is natural to define the cost that it induces on player i as the mean with respect to the initial conditions of costs having the form (2.4). Namely, the *cost function* $J_i^N : \mathcal{A}^N \rightarrow \mathbb{R}$ for player i is defined as

$$J_i^N(\gamma_1, \dots, \gamma_N) := \int_{(\mathbb{R}^d)^N} J_i^N(\gamma_1^{x_1}, \dots, \gamma_N^{x_N}) \otimes_{j=1}^N dm_0(x_j).$$

Recalling (2.5) and defining

$$J^N(\gamma_1; (\gamma_j)_{j=2}^N) := \int_{(\mathbb{R}^d)^N} J^N(\gamma_1^{x_1}; (\gamma_j^{x_j})_{j=2}^N) \otimes_{j=1}^N dm_0(x_j), \tag{2.6}$$

we have $J_i^N(\gamma_1, \dots, \gamma_N) = J^N(\gamma_i; (\gamma_j)_{j \neq i})$ for all $i = 1, \dots, N$. In particular, the costs are *symmetric*, i.e. for every permutation $\sigma : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$, we have that

$$J_i^N(\gamma_1, \dots, \gamma_N) = J_{\sigma(i)}^N(\gamma_{\sigma(1)}, \dots, \gamma_{\sigma(N)}).$$

Let us recall the classical notion of Nash equilibrium when applied to the game defined by the N players, the action set \mathcal{A} (which is the same for all players) and the individual costs $(J_i)_{i=1}^N$.

Definition 2.1 We say that $(\gamma_{1,N}, \dots, \gamma_{N,N}) \in \mathcal{A}^N$ is a *Nash equilibrium* in distributed open-loop strategies if

$$J^N(\gamma_{i,N}; (\gamma_{j,N})_{j \neq i}) \leq J^N(\gamma; (\gamma_{j,N})_{j \neq i}) \quad \forall \gamma \in \mathcal{A}, \quad i = 1, \dots, N. \tag{2.7}$$

The existence of a Nash equilibrium for this symmetric game is not immediate because of the lack of compactness of the set \mathcal{A} . The following simple result shows that the time derivatives of the strategies in a Nash equilibrium configuration (provided that it exists) enjoy a uniform boundedness property in $L^q((0, T); \mathbb{R}^d)$. This fact will allow us to reduce the set of action strategies in Definition 2.1.

Lemma 2.1 Assume that (A1) holds and that $(\gamma_{1,N}, \dots, \gamma_{N,N}) \in \mathcal{A}^N$ is a Nash equilibrium. Then, there exists a constant $C > 0$, independent of N , such that

$$\int_0^T |\dot{\gamma}_{i,N}^x(t)|^q dt \leq C \quad \text{for } m_0\text{-a.e. } x \in \mathbb{R}^d, \quad i = 1, \dots, N. \tag{2.8}$$

Proof Fix $i \in \{1, \dots, N\}$ and define $j_i : W^{1,q}([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ as

$$j_i(\gamma) = \int_0^T \ell_i(\dot{\gamma}(t), \gamma(t), t) dt + \Phi_i(\gamma(T)),$$

where $\ell_i : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ and $\Phi_i : \mathbb{R}^d \rightarrow \mathbb{R}$ are defined by

$$\begin{aligned} \ell_i(\alpha, x, t) &:= \int_{\mathbb{R}^{(N-1) \times d}} \ell \left(\alpha, x, (\gamma_{j,N}^{x_j})_{j \neq i} \right) \otimes_{j \neq i} dm_0(x_j), \\ \Phi_i(x) &:= \int_{\mathbb{R}^{(N-1) \times d}} \Phi \left(x, (\gamma_{j,N}^{x_j}(T))_{j \neq i} \right) \otimes_{j \neq i} dm_0(x_j). \end{aligned} \tag{2.9}$$

By Fubini’s theorem, and (2.7), for all $\gamma \in \mathcal{A}$ we have

$$\begin{aligned} J^N(\gamma_{i,N}; (\gamma_{j,N})_{j \neq i}) &= \int_{\mathbb{R}^d} j_i(\gamma_{i,N}^x) dm_0(x) \leq J^N(\gamma; (\gamma_{j,N})_{j \neq i}) \\ &= \int_{\mathbb{R}^d} j_i(\gamma^x) dm_0(x). \end{aligned} \tag{2.10}$$

For $x \in \mathbb{R}^d$ define

$$\mathcal{S}_i^N(x) := \operatorname{argmin} \{j_i(\gamma) \mid \gamma \in \mathcal{A}(x)\} \subseteq \Gamma.$$

Assumption **(A1)** implies that j_i is lower semicontinuous, w.r.t. the weak topology in $W^{1,q}([0, T]; \mathbb{R}^d)$ (see e.g. [22, Corollary 3.24]). Using this fact, the direct method in the Calculus of Variations and the first inequality in (2.3), we obtain that $\mathcal{S}_i^N(x) \neq \emptyset$ for all $x \in \mathbb{R}^d$.

Claim: The set-valued map $\mathbb{R}^d \ni x \rightrightarrows \mathcal{S}_i^N(x) \in 2^\Gamma$ takes closed values and is upper semicontinuous, i.e. for all closed sets $M \subseteq \Gamma$ we have that $\{x \in \mathbb{R}^d \mid \mathcal{S}_i^N(x) \cap M \neq \emptyset\}$ is closed.

Let us assume for a moment that the claim is true. Then, as a consequence of its second statement, the set valued map \mathcal{S}_i^N is Borel measurable and, hence, by the Kuratowski-Ryll-Nardzewski Selection Theorem (see e.g. [1, Theorem 18.13]) we have the existence of a Borel measurable function $\mathbb{R}^d \ni x \mapsto \hat{\gamma}(x) \in \Gamma$ such that $\hat{\gamma}(x) \in \mathcal{S}_i^N(x)$ for all $x \in \mathbb{R}^d$. Since $\hat{\gamma} \in \mathcal{A}$, relation (2.10) implies that $\gamma_{i,N}^x \in \mathcal{S}_i^N(x)$ for m_0 -almost every $x \in \mathbb{R}^d$. Taking $x \in \mathbb{R}^d$ such that $\gamma_{i,N}^x \in \mathcal{S}_i^N(x)$, we have that $j_i(\gamma_{i,N}^x) \leq j_i(\bar{\gamma}^x)$, where $\bar{\gamma}^x(t) := x$ for all $t \in [0, T]$. Assumption **(A1)** implies that $j_1(\bar{\gamma}^x)$ is bounded by a constant, which is uniform for $x \in \operatorname{supp}(m_0)$ and independent of N . Using this fact, the first inequality in (2.3) easily yields (2.8).

It remains to prove the claim. First note that if (γ_n^x) is a sequence in $\mathcal{S}_i^N(x)$ converging to γ^x uniformly in $[0, T]$, then, by the first inequality in (2.3), the sequence $(\dot{\gamma}_n^x)$ is bounded in $L^q([0, T]; \mathbb{R}^d)$. If g is a weak limit point of $(\dot{\gamma}_n^x)$ in $L^q([0, T]; \mathbb{R}^d)$, then, passing to the limit along a subsequence in the relation $\gamma_n^x(t) = x + \int_0^t \dot{\gamma}_n^x(s) ds$ for all $t \in [0, T]$, we get that $\gamma^x \in W^{1,q}([0, T]; \mathbb{R}^d)$, with $\dot{\gamma}^x = g$, and, hence, the whole sequence $(\dot{\gamma}_n^x)$ converges weakly to $\dot{\gamma}^x$ in $L^q([0, T]; \mathbb{R}^d)$. Using this fact and the weak lower semicontinuity of j_i in $W^{1,q}([0, T]; \mathbb{R}^d)$, we obtain that $\gamma^x \in \mathcal{S}_i^N(x)$ and, hence, $\mathcal{S}_i^N(x)$ is closed in Γ . In order to show that \mathcal{S}_i^N is upper semicontinuous, let M be a closed subset of Γ and let (x_n) be a sequence in $\{x \in \mathbb{R}^d \mid \mathcal{S}_i^N(x) \cap M \neq \emptyset\}$ converging to some $\bar{x} \in \mathbb{R}^d$. Then, by definition, there exists $\gamma_n^{x_n} \in \mathcal{S}_i^N(x_n) \cap M$.

Arguing as before, using the first inequality in (2.3), we obtain that $(\dot{\gamma}_n^{x_n})$ is bounded in $L^q([0, T]; \mathbb{R}^d)$. This implies that, up to some subsequence, $\dot{\gamma}_n^{x_n}$ converge uniformly to some $\dot{\gamma}^{\bar{x}} \in W^{1,q}([0, T]; \mathbb{R}^d)$ and $\dot{\gamma}^{x_n} \rightharpoonup \dot{\gamma}^{\bar{x}}$ weakly in $L^q([0, T]; \mathbb{R}^d)$. Using that M is closed, w.r.t. the uniform convergence, we get that $\dot{\gamma}^{\bar{x}} \in M$. On the other hand, noticing that

$$j_i(\dot{\gamma}_n^{x_n}) \leq j_i(\dot{\gamma}^{x_n}) \quad \forall \dot{\gamma}^{x_n} \in \mathcal{A}(x_n),$$

and the fact that any $\dot{\gamma} \in \mathcal{A}(\bar{x})$ satisfies $\dot{\gamma} - \bar{x} + x_n \in \mathcal{A}(x_n)$, the weak lower semicontinuity of j_i yields that $j_i(\dot{\gamma}^{\bar{x}}) \leq j_i(\dot{\gamma}) \quad \forall \dot{\gamma} \in \mathcal{A}(\bar{x})$, i.e. $\dot{\gamma}^{\bar{x}} \in \mathcal{S}_i^N(\bar{x}) \cap M$, and, hence, $\bar{x} \in \{x \in \mathbb{R}^d \mid \mathcal{S}_i^N(x) \cap M \neq \emptyset\}$, which implies that the latter set is closed. \square

Now, we focus our attention on the existence of Nash equilibria for the described game. Using Lemma 2.1, a reformulation of the set of admissible strategies and cost functionals of the N -players game will be useful. Let $C > 0$ be given by Lemma 2.1 and define the set

$$Q_C := \left\{ \dot{\gamma} \in W^{1,q}([0, T]; \mathbb{R}^d) \mid \int_0^T |\dot{\gamma}(t)|^q dt \leq C, \quad \dot{\gamma}(0) \in \text{supp}(m_0) \right\}.$$

Since $\text{supp}(m_0)$ is compact, Hölder’s inequality yields the existence of a compact set $K_C \subseteq \mathbb{R}^d$ such that

$$\dot{\gamma}(t) \in K_C \quad \text{for all } t \in [0, T], \quad \dot{\gamma} \in Q_C. \tag{2.11}$$

Using this fact and arguing as in the proof of Lemma 2.1, we have that Q_C is compact as a subset of Γ , that is, when it is endowed with the topology of uniform convergence.

Given $\dot{\gamma}_i \in \mathcal{A}$, let us set $m_i := \dot{\gamma}_i \# m_0 \in \mathcal{P}(\Gamma)$. By (2.1), for any profile of strategies $(\dot{\gamma}_i)_{i=1}^N \in \mathcal{A}^N$, the cost for player i is given by

$$J^N(\dot{\gamma}_i; (\dot{\gamma}_j)_{j \neq i}) = \int_{\Gamma^N} j^N(\tilde{\gamma}_i; (\tilde{\gamma}_j)_{j \neq i}) \otimes_{j=1}^N dm_j(\tilde{\gamma}_j). \tag{2.12}$$

This expression for the cost motivates a relaxation of the game with strategies in \mathcal{A} by considering strategies taking values in $\mathcal{P}(\Gamma)$. Let us define the set \mathcal{A}_{rel} of randomized distributed open-loop strategies by

$$\mathcal{A}_{rel} := \{m \in \mathcal{P}(\Gamma) \mid e_0 \# m = m_0, \quad \text{supp}(m) \subseteq Q_C\}, \tag{2.13}$$

where, for each $t \in [0, T]$, $e_t : \Gamma \rightarrow \mathbb{R}^d$ is defined by $e_t(\dot{\gamma}) := \dot{\gamma}(t)$.

Lemma 2.2 *The set \mathcal{A}_{rel} is convex and compact as a subset of $\mathcal{P}(\Gamma)$.*

Proof Convexity follows directly from the definition. On the other hand, since $\mathcal{A}_{rel} \subseteq \mathcal{P}(Q_C)$ and $\mathcal{P}(Q_C)$ is compact as a subset of $\mathcal{P}(\Gamma)$ (because Q_C is a compact subset of Γ), it suffices to check that \mathcal{A}_{rel} is closed in $\mathcal{P}(\Gamma)$, but this follows directly from [3, Proposition 5.1.8] and the fact that Q_C is closed. \square

Remark 2.1 For later use, note that if $m \in \mathcal{A}_{rel}$, then $[0, T] \ni t \rightarrow e_t \# m \in \mathcal{P}_1(\mathbb{R}^d)$ is well-defined and, by (2.2), belongs to $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. Moreover, since $\text{supp}(m) \subseteq Q_C$, we easily check that there exists $C' > 0$, independent of $m \in \mathcal{A}_{rel}$, such that

$$d_1(e_t \# m, e_s \# m) \leq C'|t - s|^{\frac{1}{q'}} \quad \forall s, t \in [0, T].$$

Therefore, by (2.11), [3, Proposition 7.1.5] and Lemma 2.2, the set $\{[0, T] \ni t \rightarrow e_t \# m \in \mathcal{P}_1(\mathbb{R}^d) \mid m \in \mathcal{A}_{rel}\}$ is compact in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$.

Motivated by (2.12), we introduce the new relaxed game which has \mathcal{A}_{rel} as set of strategies for each player and, given a strategy profile $(m_j)_{j=1}^N \subseteq \mathcal{A}_{rel}^N$, the cost for player i is given by

$$J_{rel}^N(m_i; (m_j)_{j \neq i}) := \int_{Q_C^N} j^N(\gamma_i; (\gamma_j)_{j \neq i}) \otimes_{j=1}^N dm_j(\gamma_j).$$

Note that the this game is still symmetric. In this framework, a profile of strategies $(m_{1,N}, \dots, m_{N,N}) \in \mathcal{A}_{rel}^N$ is called a *Nash equilibrium* in randomized distributed open-loop strategies if

$$J_{rel}^N(m_{i,N}; (m_{j,N})_{j \neq i}) \leq J_{rel}^N(m; (m_{j,N})_{j \neq i}) \quad \forall m \in \mathcal{A}_{rel}, \quad i = 1, \dots, N. \quad (2.14)$$

Proposition 2.1 below establishes existence of Nash equilibria in this class of strategy profiles. The corresponding trajectory distributions actually have product form. The convergence result of Sect. 3 will apply to this kind of equilibrium. Notice that non-symmetric Nash equilibria might exist as well.

Proposition 2.1 *Under assumption (A1), there exists at least one Nash equilibrium in randomized distributed open-loop strategies having the form $(m_N, \dots, m_1) \in \mathcal{A}_{rel}^N$.*

Proof It suffices to show the existence of a fixed point of the following set-valued map

$$\mathcal{A}_{rel} \ni \mu \rightrightarrows \mathcal{S}^N(\mu) := \text{argmin} \left\{ J_{rel}^N(\mu', \mu, \dots, \mu) \mid \mu' \in \mathcal{A}_{rel} \right\} \subseteq \mathcal{A}_{rel}.$$

For $\mu \in \mathcal{A}_{rel}$, $\alpha \in \mathbb{R}^d$, $x \in \mathbb{R}^d$, $t \in [0, T]$ and $\gamma \in Q_C$, let us set

$$\begin{aligned} \ell_\mu(\alpha, x, t) &:= \int_{Q_C^{N-1}} \ell \left(\alpha, x, (\gamma_j(t))_{j=1}^{N-1} \right) \otimes_{j=1}^{N-1} d\mu(\gamma_j), \\ \Phi_\mu(x) &:= \int_{Q_C^{N-1}} \Phi \left(x, (\gamma_j(T))_{j=1}^{N-1} \right) \otimes_{j=1}^{N-1} d\mu(\gamma_j), \\ j_\mu(\gamma) &:= \int_0^T \ell_\mu(\dot{\gamma}(t), \gamma(t), t) dt + \Phi_\mu(\gamma(T)), \end{aligned}$$

and notice that, by Fubini's theorem, for every $\mu' \in \mathcal{A}_{rel}$ we have

$$J_{rel}^N(\mu', \mu, \dots, \mu) = \int_{Q_C} j_\mu(\gamma) d\mu'(\gamma). \quad (2.15)$$

Assumption **(A1)** and [22, Theorem 3.23] imply that $Q_C \ni \gamma \mapsto j_\mu(\gamma) \in \mathbb{R}$ is lower semicontinuous. Moreover, by **(A1)**, j_μ is bounded from below. Therefore, by [3, Lemma 5.1.7] and (2.15), for every $\mu \in \mathcal{A}_{rel}$ the map $\mathcal{A}_{rel} \ni \mu' \rightarrow J_{rel}^N(\mu', \mu, \dots, \mu) \in \mathbb{R}$ is lower semicontinuous. Thus, the compactness of \mathcal{A}_{rel} yields that $\mathcal{S}^N(\mu) \neq \emptyset$ for all $\mu \in \mathcal{A}_{rel}$. A similar argument implies that $\mathcal{S}^N(\mu)$ is closed for every $\mu \in \mathcal{A}_{rel}$. Notice also that $\mathcal{S}^N(\mu)$ is convex for all $\mu \in \mathcal{A}_{rel}$. Let us show that \mathcal{S}^N is upper semicontinuous. Consider a closed set $M \subseteq \mathcal{A}_{rel}$ and a sequence (μ_n) in $\{\mu \in \mathcal{A}_{rel} \mid \mathcal{S}^N(\mu) \cap M \neq \emptyset\}$ converging to some $\bar{\mu} \in \mathcal{A}_{rel}$. By definition, there exists $v_n \in M$ such that

$$J_{rel}^N(v_n; \mu_n, \dots, \mu_n) \leq J_{rel}^N(\mu'; \mu_n, \dots, \mu_n) \quad \forall \mu' \in \mathcal{A}_{rel}. \tag{2.16}$$

Since M is compact (because \mathcal{A}_{rel} is compact), there exists $\bar{v} \in M$ such that, up to some subsequence, $v_n \rightarrow \bar{v}$ narrowly. By [9, Theorem 3.2] and **(A1)** for all $\mu' \in \mathcal{A}_{rel}$ we have that $J_{rel}^N(\mu'; \mu_n, \dots, \mu_n)$ converges to $J_{rel}^N(\mu'; \bar{\mu}, \dots, \bar{\mu})$. On the other hand, by **(A1)** and [22, Theorem 3.23], we obtain that $Q_C^N \ni (\gamma_1, \dots, \gamma_N) \rightarrow j^N(\gamma_1; (\gamma_j)_{j \geq 2}) \in \mathbb{R}$ is bounded from below and lower semicontinuous. Thus, using [3, Lemma 5.1.7] again, we obtain $J_{rel}^N(\bar{v}; \bar{\mu}, \dots, \bar{\mu}) \leq \liminf_{n \rightarrow \infty} J_{rel}^N(v_n; \mu_n, \dots, \mu_n)$. Altogether, passing to the limit in (2.16), we get

$$J_{rel}^N(\bar{v}; \bar{\mu}, \dots, \bar{\mu}) \leq J_{rel}^N(\mu'; \bar{\mu}, \dots, \bar{\mu}) \quad \forall \mu' \in \mathcal{A}_{rel},$$

i.e. $\bar{v} \in \mathcal{S}^N(\bar{\mu}) \cap M$, which implies the closedness $\{\mu \in \mathcal{A}_{rel} \mid \mathcal{S}^N(\mu) \cap M \neq \emptyset\}$ and the upper semicontinuity of \mathcal{S}^N . Using the properties above, the existence of a fixed point for \mathcal{S}^N follows from the Kakutani-Fan-Glicksberg fixed-point theorem (see e.g. [1, Corollary 17.55]). □

Corollary 2.1 *Let $(m_N, \dots, m_N) \in \mathcal{A}_{rel}^N$ be a Nash equilibrium for the game defined by the cost J_{rel}^N and the set of strategies \mathcal{A}_{rel} . Define $\hat{\ell} : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ and $\hat{\Phi} : \mathbb{R}^d \rightarrow \mathbb{R}$ by*

$$\begin{aligned} \hat{\ell}(\alpha, x, t) &:= \int_{Q_C^{N-1}} \ell \left(\alpha, x, (\gamma_j(t))_{j=2}^N \right) \otimes_{j=2}^N dm_N(\gamma_j), \\ \hat{\Phi}(x) &:= \int_{Q_C^{N-1}} \Phi \left(x, (\gamma_j(T))_{j=2}^N \right) \otimes_{j=2}^N dm_N(\gamma_j), \end{aligned} \tag{2.17}$$

and assume that for m_0 -almost every $x \in \mathbb{R}^d$ the optimization problem

$$\inf \left\{ \int_0^T \hat{\ell}(\dot{\gamma}(t), \gamma(t), t) dt + \hat{\Phi}(\gamma(T)) \mid \gamma \in Q_C, \gamma(0) = x \right\}, \tag{2.18}$$

admits a unique solution. Then, there exists $\gamma_N \in \mathcal{A}$ such that $m_N = \gamma_N \# m_0$. Moreover, γ_N is m_0 -uniquely determined.

In particular, $(\gamma_N, \dots, \gamma_N) \in \mathcal{A}^N$ is a Nash equilibrium in distributed open-loop strategies.

Proof Define $\widehat{j} : Q_C \rightarrow \mathbb{R}$ by

$$\widehat{j}(\gamma) := \int_0^T \widehat{\ell}(\dot{\gamma}(t), \gamma(t), t) dt + \widehat{\Phi}(\gamma(T)). \tag{2.19}$$

By definition of Nash equilibrium and Fubini’s theorem, for all $m \in \mathcal{A}_{rel}$ we have that

$$\begin{aligned} J_{rel}^N(m_N; m_N, \dots, m_N) &= \int_{Q_C} \widehat{j}(\gamma) dm_N(\gamma) \leq J_{rel}^N(m; m_N, \dots, m_N) \\ &= \int_{Q_C} \widehat{j}(\gamma) dm(\gamma). \end{aligned} \tag{2.20}$$

By the disintegration theorem (see e.g. [3, Theorem 5.3.1]), there exists a Borel family $\{m_N^x \mid x \in \mathbb{R}^d\} \subseteq \mathcal{P}(Q_C)$, such that $m_N^x(\{\gamma \in Q_C \mid \gamma(0) = x\}) = 1$, for m_0 -almost every $x \in \mathbb{R}^d$, and

$$\int_{Q_C} \widehat{j}(\gamma) dm_N(\gamma) = \int_{\mathbb{R}^d} \int_{Q_C} \widehat{j}(\gamma) dm_N^x(\gamma) dm_0(x). \tag{2.21}$$

Define the set-valued function $\widehat{\mathcal{S}}^N : \mathbb{R}^d \rightarrow 2^\Gamma$ by

$$\widehat{\mathcal{S}}^N(x) := \operatorname{argmin} \left\{ \widehat{j}(\gamma) \mid \gamma \in \mathcal{A}(x), \int_0^T |\dot{\gamma}(t)|^q \leq C \right\}.$$

Arguing as in the proof of Lemma 2.1, we have the existence of a Borel measurable selection $\mathbb{R}^d \ni x \mapsto \gamma_N^x \in \widehat{\mathcal{S}}^N(x)$, which, by assumption, is m_0 -uniquely determined. Moreover, by definition, $\gamma_N \in \mathcal{A}$. Now, let us define $\widehat{m} := \gamma_N \# m_0 \in \mathcal{A}_{rel}$. Then, by (2.20), taking $m = \widehat{m}$, and (2.21), we have that

$$\int_{\mathbb{R}^d} \left[\int_{Q_C} \widehat{j}(\gamma) dm_N^x(\gamma) - \widehat{j}(\gamma_N^x) \right] dm_0(x) \leq 0.$$

Since the integrand in the expression above is non-negative, by definition of γ_N^x , we deduce that $\widehat{j}(\gamma_N^x) = \int_{Q_C} \widehat{j}(\gamma) dm_N^x(\gamma)$ for m_0 -a.e. $x \in \mathbb{R}^d$, and, hence, $\widehat{j}(\gamma) = \widehat{j}(\gamma_N^x)$ for m_N^x -a.e. $\gamma \in Q_C$. Since, by assumption, $\widehat{\mathcal{S}}^N(x) = \{\gamma_N^x\}$ for m_0 -a.e. $x \in \mathbb{R}^d$, we deduce that $m_N^x = \delta_{\gamma_N^x}$ for m_0 -a.e. $x \in \mathbb{R}^d$, hence $m_N = \gamma_N \# m_0$. The result follows. \square

Remark 2.2 An example of application of Corollary 2.1 is provided in Sect. 4.2 below.

3 Convergence to a Mean Field Game Equilibrium

In this section, we study the limit behavior, as $N \rightarrow \infty$, of symmetric Nash equilibria in randomized distributed open-loop strategies. The existence of such Nash equilibria

is ensured by Proposition 2.1. We begin by defining the limit object, i.e. the MFG equilibrium. Then we will prove that any cluster point of the sequence (m_N) is a MFG equilibrium.

Let us define $J : W^{1,q}([0, T]; \mathbb{R}^d) \times \mathcal{P}_1(\Gamma) \rightarrow \mathbb{R}$ by

$$J(\gamma, m) := \int_0^T \ell(\dot{\gamma}(t), \gamma(t), e_t \# m) dt + \Phi(\gamma(T), e_T \# m).$$

It is straightforward to check that $m \in \mathcal{P}_1(\Gamma)$ implies that $e_t \# m \in \mathcal{P}_1(\mathbb{R}^d)$ for all $t \in [0, T]$, which implies that J is well-defined.

Following the terminology in [37], we consider next the notion of *Lagrangian MFG equilibrium* (see e.g. [8, 10, 16, 17]).

Definition 3.1 We say that $m_* \in \mathcal{P}_1(\Gamma)$ is a Lagrangian MFG equilibrium if $e_0 \# m_* = m_0$ and

$$\begin{aligned} \text{supp}(m_*) &\subseteq \left\{ \gamma \in W^{1,q}([0, T]; \mathbb{R}^d) \mid J(\gamma, m_*) \right. \\ &\quad \left. \leq J(\gamma', m_*) \ \forall \gamma' \in W^{1,q}([0, T]; \mathbb{R}^d), \ \gamma'(0) = \gamma(0) \right\}. \end{aligned} \tag{3.1}$$

Reasoning as in the proof of Lemma 2.1, assumption (A1) implies that if m_* is a Lagrangian MFG equilibrium, then $\text{supp}(m_*) \subseteq Q_C$ and, hence, $m_* \in \mathcal{A}_{rel}$. Thus, m_* is a Lagrangian MFG if and only if $m_* \in \mathcal{A}_{rel}$, $e_0 \# m_* = m_0$ and

$$\text{supp}(m_*) \subseteq \left\{ \gamma \in Q_C \mid J(\gamma, m_*) \leq J(\gamma', m_*) \ \forall \gamma' \in Q_C, \ \gamma'(0) = \gamma(0) \right\}. \tag{3.2}$$

We still denote by J the restriction of J to $Q_C \times \mathcal{A}_{rel}$ and recall that Q_C , endowed with the topology of uniform convergence, is a compact set. For later use, let us state the following simple results.

Lemma 3.1 *Suppose that (A1) holds. Then, the following assertions hold true:*

- (i) *The relative topology on \mathcal{A}_{rel} , as a subset of $\mathcal{P}_1(\Gamma)$, coincides with the topology induced by the narrow convergence.*
- (ii) *The function J is lower semicontinuous in $Q_C \times \mathcal{A}_{rel}$.*
- (iii) *For all $\gamma \in Q_C$, the function $J(\gamma, \cdot)$ is bounded, uniformly in γ , and continuous in \mathcal{A}_{rel} .*

Proof Assertion (i) follows from the fact that both topologies coincide on $\mathcal{P}(Q_C)$, because Q_C is a compact subset of Γ . In particular, if $m_n \rightarrow m$ narrowly in \mathcal{A}_{rel} , then for all $t \in [0, T]$ we have that

$$e_t \# m_n \rightarrow e_t \# m \quad \text{in } \mathcal{P}_1(\mathbb{R}^d). \tag{3.3}$$

Thus, assertion (ii) follows from (3.3), (A1) and the proof of [22, Theorem 3.23]. Assertion (iii) follows directly from (2.3), (3.3) and dominated convergence. \square

If m_* is a Lagrangian MFG equilibrium, we will denote by $\{m_*^x \mid x \in \mathbb{R}^d\}$ the m_0 -uniquely determined Borel family of probability measures on Q_C satisfying that $m_*^x(Q_C \setminus \mathcal{A}(x)) = 0$ and $dm_*(\gamma) = dm_*^x(\gamma) \otimes dm_0(x)$. The existence of such a family is ensured by the disintegration theorem (see e.g. [3, Theorem 5.3.1]). We have the following equivalent characterization of a Lagrangian MFG equilibrium.

Proposition 3.1 *The following assertions are equivalent:*

- (i) *The measure $m_* \in \mathcal{A}_{rel}$ is a Lagrangian MFG equilibrium.*
- (ii) *For m_0 -a.e. $x \in \mathbb{R}^d$ we have that*

$$\text{supp}(m_*^x) \subseteq \mathcal{S}(x) := \text{argmin} \left\{ J(\gamma', m^*) \mid \gamma' \in \mathcal{A}(x), \int_0^T |\dot{\gamma}'(t)|^q \leq C \right\}. \tag{3.4}$$

- (iii) *The following inequality holds true:*

$$\int_{Q_C} J(\gamma, m_*) dm_*(\gamma) \leq \int_{Q_C} J(\gamma, m_*) dm(\gamma) \text{ for all } m \in \mathcal{A}_{rel}. \tag{3.5}$$

Proof Let us prove the equivalence between (i) and (ii). Let $m_* \in \mathcal{A}_{rel}$ be a Lagrangian MFG equilibrium. If (3.4) does not hold, there exists $A \in \mathcal{B}(\mathbb{R}^d)$, with $m_0(A) > 0$, such that $m_*^x(\mathcal{S}(x)^c) > 0$ for all $x \in A$. Define the set $E := \{\gamma \in Q_C \mid \gamma(0) \in A, \gamma \in \mathcal{S}(\gamma(0))^c\} = e_0^{-1}(A) \cap \{\gamma \in Q_C \mid \gamma \in \mathcal{S}(\gamma(0))^c\}$. Arguing as in the proof of the claim in Lemma 2.1, the set $\{\gamma \in Q_C \mid \gamma \in \mathcal{S}(\gamma(0))\}$ is closed in Q_C , which implies that $E \in \mathcal{B}(Q_C)$. Since $m_*(E) = \int_A m_*^x(\mathcal{S}(x)^c) dm_0(x) > 0$, we obtain a contradiction with (3.2). Conversely, suppose that (ii) holds and that m_* is not a Lagrangian MFG equilibrium. Then, defining $E' := \{\gamma \in Q_C \mid \gamma \in \mathcal{S}(\gamma(0))^c\}$, which is an open set and, hence, belongs to $\mathcal{B}(Q_C)$, we have that

$$0 < m_*(E') = \int_{\mathbb{R}^d} \int_{Q_C} \mathbb{I}_{E'}(\gamma) dm_*^x(\gamma) dm_0(x),$$

which is impossible because (3.4) implies that the r.h.s. above is equal to 0.

Let us now prove that (ii) \Leftrightarrow (iii). Notice that

$$\int_{Q_C} J(\gamma, m_*) dm_*(\gamma) = \int_{\mathbb{R}^d} \int_{Q_C} J(\gamma, m_*) dm_*^x(\gamma) dm_0(x). \tag{3.6}$$

Analogously, given $m \in \mathcal{A}_{rel}$, we disintegrate it w.r.t. m_0 and write $dm(\gamma) = dm^x(\gamma) \otimes dm_0(x)$, where $\{m^x \mid x \in \mathbb{R}^d\}$ is a m_0 -uniquely determined Borel family of probability measures on Q_C such that $m^x(Q_C \setminus \mathcal{A}(x)) = 0$ for m_0 -a.e. $x \in \mathbb{R}^d$. Thus,

$$\int_{Q_C} J(\gamma, m_*) dm(\gamma) = \int_{\mathbb{R}^d} \int_{Q_C} J(\gamma, m_*) dm^x(\gamma) dm_0(x). \tag{3.7}$$

If (ii) holds, then for m_0 -a.e. $x \in \mathbb{R}^d$ and m_*^x -a.e. $\gamma \in Q_C$ we have

$$J(\gamma, m_*) \leq J(\gamma', m^*) \quad \forall \gamma' \in \mathcal{A}(x) \cap Q_C. \tag{3.8}$$

Integrating both sides of (3.8), first with respect to $dm^x(\gamma')$ and then with respect to $dm_*^x(\gamma)$, and using (3.6)-(3.7) we obtain (3.5). Conversely, using the notations introduced above, suppose that (3.5) holds and let $\hat{\gamma} \in \mathcal{A}$ be a Borel measurable selection of \mathcal{S} (the existence of such selection can be justified arguing exactly as in the proof of Lemma 2.1). Then, taking the measure $m \in \mathcal{A}_{rel}$ defined by $dm(\gamma) = d\delta_{\hat{\gamma}^x(\gamma)} \otimes dm_0(x)$ in (3.5) and using (3.6), we deduce that

$$\int_{Q_C} J(\gamma, m_*) dm_*^x(\gamma) = J(\gamma^x, m_*) \quad \text{for } m_0\text{-a.e. } x \in \mathbb{R}^d,$$

and, hence, m_*^x -almost every γ belongs to $\mathcal{S}(x)$. The conclusion follows. □

Now, consider the symmetric N -player game defined in Sect. 2, with randomized distributed open-loop strategies, and let $(m_N, \dots, m_N) \in \mathcal{A}_{rel}^N$ be a symmetric equilibrium. Our main result in this section, stated in the next theorem, shows that any limit point of this sequence is a Lagrangian MFG equilibrium.

Theorem 3.1 *Suppose that (A1) holds. Then any limit point m_* of $(m_N)_{N \in \mathbb{N}}$ (there exist at least one) is a Lagrangian MFG equilibrium. Moreover, if $(m_{N_k})_{k \in \mathbb{N}}$ is a subsequence of $(m_N)_{N \in \mathbb{N}}$ converging to m_* , then $\sup_{t \in [0, T]} d_1(e_t \# m_{N_k}, e_t \# m_*) \rightarrow 0$ as $k \rightarrow \infty$.*

Proof We only prove the first assertion, since, having this result, the second one follows directly from Remark 2.1. For $N \geq 2$, let Y_1^N, \dots, Y_N^N be independent and identically distributed (i.i.d.) Q_C -valued random variables with common distribution m_N defined on some probability space $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$. We denote by \mathbf{E}_N the expectation with respect to \mathbf{P}_N . For $i \in \{1, \dots, N\}$, let $\mu^{N,i}$ denote the (random) empirical measure of Y_1^N, \dots, Y_N^N excluding Y_i^N , that is,

$$\mu^{N,i}(\omega) := \frac{1}{N-1} \sum_{j \neq i} \delta_{Y_j^N(\omega)} \in \mathcal{P}(Q_C) \quad \forall \omega \in \Omega_N.$$

Notice that Y_i^N and $\mu^{N,i}$ are independent for every i , while $\mu^{N,1}, \dots, \mu^{N,N}$ are identically distributed (not independent in general) with common distribution depending on N . Moreover, for every $i \in \{1, \dots, N\}$,

$$J_{rel}^N(m_N; m_N, \dots, m_N) = \mathbf{E}_N \left[J \left(Y_i^N, \mu^{N,i} \right) \right]. \tag{3.9}$$

Let $(m_{N_k})_{k \in \mathbb{N}}$ be a converging subsequence of $(m_N)_{N \in \mathbb{N}}$ with limit m_* for some $m_* \in \mathcal{A}_{rel}$. The existence of such a subsequence follows from the compactness of \mathcal{A}_{rel} . Let us prove that $(\mu^{N_k,1})_{k \in \mathbb{N}}$ converges in distribution to the deterministic limit m_* . Let $\mathcal{T} \subset C_b(Q_C)$ be countable and measure determining (or separating). Thus, \mathcal{T} is a countable collection of bounded continuous functions on Q_C such that two probability measures $\nu, \tilde{\nu} \in \mathcal{P}(Q_C)$ are equal whenever $\int \psi d\nu = \int \psi d\tilde{\nu}$ for all $\psi \in \mathcal{T}$. Observe that \mathcal{T} can be chosen countable since Q_C is a Polish space, hence separable, under the supremum norm topology.

For $\psi \in \mathcal{T}$ set

$$m_\psi^N := \int_{Q_C} \psi(\gamma) dm_N(\gamma), \quad v_\psi^N := \mathbf{E}_N \left[\left(\int_{Q_C} \psi(\gamma) d\mu^{N,1}(\gamma) - m_\psi^N \right)^2 \right].$$

By construction and symmetry, for every $i \in \{1, \dots, N\}$,

$$m_\psi^N = \mathbf{E}_N \left[\int_{Q_C} \psi(\gamma) d\mu^{N,i}(\gamma) \right] = \mathbf{E}_N \left[\psi(Y_i^N) \right],$$

$$v_\psi^N = \mathbf{E}_N \left[\left(\frac{1}{N-1} \sum_{j \neq i} \psi(Y_j^N) - m_\psi^N \right)^2 \right].$$

Since Y_1^N, \dots, Y_N^N are independent and the functions in \mathcal{T} are bounded, it follows that

$$v_\psi^N = \frac{1}{(N-1)^2} \sum_{j=2}^N \mathbf{E}_N \left[\left(\psi(Y_j^N) - m_\psi^N \right)^2 \right] \xrightarrow{N \rightarrow \infty} 0 \quad \text{for every } \psi \in \mathcal{T}. \tag{3.10}$$

Moreover, considering the subsequence $(m_{N_k})_{k \in \mathbb{N}}$, we have that

$$m_\psi^{N_k} \xrightarrow{k \rightarrow \infty} \int_{Q_C} \psi(\gamma) dm_*(\gamma) =: m_\psi^* \quad \text{for every } \psi \in \mathcal{T}. \tag{3.11}$$

Since $\mathcal{P}(Q_C)$ is compact, the set $\mathcal{P}(\mathcal{P}(Q_C))$ is also compact. Thus, there exists a subsequence $(N_{k_l})_{l \in \mathbb{N}}$ such that $(\mu^{N_{k_l},1})_{l \in \mathbb{N}}$ converges in distribution to μ for some $\mathcal{P}(Q_C)$ -valued random variable μ defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We denote by \mathbf{E} the expectation under \mathbf{P} . By convergence in distribution and (3.11), we obtain

$$\lim_{l \rightarrow \infty} v_\psi^{N_{k_l}} = \mathbf{E} \left[\left(\int_{Q_C} \psi(\gamma) d\mu(\gamma) - m_\psi^* \right)^2 \right] \quad \text{for every } \psi \in \mathcal{T}.$$

Indeed, the mapping $\nu \mapsto \left(\int_{Q_C} \psi(\gamma) d\nu(\gamma) - m_\psi^* \right)^2$ is bounded and continuous on $\mathcal{P}(Q_C)$ since ψ is bounded and continuous on Q_C , while by definition of m_ψ^N and v_ψ^N ,

$$v_\psi^N = \mathbf{E}_N \left[\left(\int_{Q_C} \psi(\gamma) d\mu^{N,1}(\gamma) - m_\psi^* \right)^2 \right] - \left(m_\psi^* - m_\psi^N \right)^2 \quad \text{for all } N \in \mathbb{N}.$$

On the other hand, thanks to (3.10), $\lim_{l \rightarrow \infty} v_{\psi}^{N_{kl}} = 0$. It follows that for every $\psi \in \mathcal{T}$,

$$\int_{Q_C} \psi(\gamma) d\mu(\gamma) = m_{\psi}^* = \int_{Q_C} \psi(\gamma) dm_*(\gamma) \quad \mathbf{P}\text{-almost surely.}$$

Since \mathcal{T} is countable, we have the existence of $A \in \mathcal{F}$ such that $\mathbf{P}(A) = 1$ and for every $\omega \in A$,

$$\int_{Q_C} \psi(\gamma) d\mu_{\omega}(\gamma) = \int_{Q_C} \psi(\gamma) dm_*(\gamma) \quad \text{for all } \psi \in \mathcal{T}.$$

Since \mathcal{T} is measure determining, we find that

$$\mu_{\omega} = m_* \quad \text{for } \mathbf{P}\text{-almost all } \omega \in \Omega.$$

As we can always choose converging (sub-)subsequences, we deduce that $(\mu^{N_k,1})_{k \in \mathbb{N}}$ converges in distribution to m_* . By independence of $Y_1^{N_k}$ and $\mu^{N_k,1}$, and using the fact that sequences of product measures converge weakly if and only if the sequences of marginals converge weakly [9, Theorem 3.2], we find that

$$\left(Y_1^{N_k}, \mu^{N_k,1} \right) \# \mathbf{P}_{N_k} = m_{N_k} \otimes \left(\mu^{N_k,1} \right) \# \mathbf{P}_{N_k} \xrightarrow{k \rightarrow \infty} m_* \otimes \delta_{m_*}.$$

This implies, thanks to (3.9), Lemma 3.1(ii) and a version of Fatou’s lemma [3, Lemma 5.1.7], that

$$\begin{aligned} & \liminf_{k \rightarrow \infty} J_{rel}^{N_k}(m_{N_k}; m_{N_k}, \dots, m_{N_k}) \\ &= \liminf_{k \rightarrow \infty} \mathbf{E}_{N_k} \left[J \left(Y_1^{N_k}, \mu^{N_k,1} \right) \right] \\ &\geq \int_{Q_C} J(\gamma, m_*) dm_*(\gamma). \end{aligned} \tag{3.12}$$

Let $m \in \mathcal{A}_{rel}$. By Lemma 3.1(iii) and dominated convergence, it follows that

$$\begin{aligned} & J_{rel}^{N_k}(m; m_{N_k}, \dots, m_{N_k}) \\ &= \int_{Q_C} \mathbf{E}_{N_k} \left[J \left(\gamma, \mu^{N_k,1} \right) \right] dm(\gamma) \xrightarrow{k \rightarrow \infty} \int_{Q_C} J(\gamma, m_*) dm(\gamma). \end{aligned} \tag{3.13}$$

Passing to the limit in the Nash equilibrium inequality

$$J_{rel}^{N_k}(m_{N_k}; m_{N_k}, \dots, m_{N_k}) \leq J_{rel}^{N_k}(m; m_{N_k}, \dots, m_{N_k}),$$

and using (3.12)-(3.13), we obtain that (3.5) holds. The result now follows from Proposition 3.1. □

Remark 3.1 (i) In particular, under assumption **(A1)**, Theorem 3.1 ensures the existence of at least one Lagrangian MFG equilibrium.

(ii) The previous approach can also be applied to deal with MFGs with state constraints (see [10]). Let us sketch the main changes. Given a nonempty bounded open set \mathcal{O} , suppose that $\text{supp}(m_0) \subseteq \bar{\mathcal{O}}$ and redefine \tilde{Q}_C , \tilde{A}_{rel} and \tilde{J}_{rel}^N by

$$\begin{aligned} \tilde{Q}_C(x) &:= \left\{ \gamma \in W^{1,q}([0, T]; \mathbb{R}^d) \mid \gamma(0) = x, \gamma(t) \in \bar{\mathcal{O}} \text{ for all } t \right. \\ &\quad \left. \in [0, T], \text{ and } \int_0^T |\dot{\gamma}(t)|^q dt \leq C \right\}, \\ \tilde{A}_{rel} &:= \left\{ m \in \mathcal{P}(\Gamma) \mid e_0 \# m = m_0, \text{ supp}(m) \subseteq \tilde{Q}_C \right\}, \\ \tilde{J}_{rel}^N(m_i; (m_j)_{j \neq i}) &:= \int_{\tilde{Q}_C} j^N(\gamma_i; (\gamma_j)_{j \neq i}) \\ &\quad \times \otimes_{j=1}^N dm_j(\gamma_j) \text{ for all } (m_1, \dots, m_N) \in \tilde{\mathcal{A}}_{rel}^N, \end{aligned}$$

where we recall that j^N is defined in (2.5). Consider the symmetric N -player game defined by the action space \tilde{A}_{rel} and the cost \tilde{J}_{rel}^N . Then, arguing exactly as in the proof of Proposition 2.1, we get the existence of a symmetric Nash equilibrium (m_N, \dots, m_N) . Now, let us define

$$\tilde{J}(\gamma, m) := \int_0^T \ell(\dot{\gamma}(t), \gamma(t), e_t \# m) dt + \Phi(\gamma(T), e_T \# m) \quad \forall \gamma \in \tilde{Q}_C, m \in \tilde{A}_{rel}.$$

Arguing as in the proof of Theorem 3.1, as $N \rightarrow \infty$, and up to some subsequence, we get that m_N converges to some $m_* \in \tilde{A}_{rel}$ satisfying

$$\int_{\tilde{Q}_C} \tilde{J}(\gamma, m_*) dm_*(\gamma) \leq \int_{\tilde{Q}_C} \tilde{J}(\gamma, m_*) dm(\gamma) \quad \text{for all } m \in \tilde{A}_{rel}. \tag{3.14}$$

Finally, assuming that \mathcal{O} has a C^2 boundary, it is easy to see from [10, Proposition 3.1] and the proof of the implication (iii) \Rightarrow (ii) in Proposition 3.1 above that m_* is a constrained MFG equilibrium in the sense of [10, Definition 3.1].

4 The First Order Mean Field Game system

In this section, we first discuss, following [16], the relation between the notion of Lagrangian MFG equilibrium and the first order PDE system introduced by Lasry and Lions in [36, Section 2.5] for some particular data. Next, in Sect. 4.2, we consider symmetric randomized distributed open-loop Nash equilibria for the N -player game, which, thanks to Corollary 2.1 and Assumption **(A2)** below, can be identified with symmetric distributed open-loop Nash equilibria (non-randomized). Arguing as in the MFG limit, we connect these equilibria with a first order PDE system which is similar

to the one appearing in the limit case. Consequently, analytic techniques can also be used in order study the limit behavior of these equilibria as the number of players tends to infinity.

Let $L : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, $f, g : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ and $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ satisfying that

(A2)(i) The function L belongs to $C^2(\mathbb{R}^d \times \mathbb{R}^d)$, is bounded from below and (i.1) there exist $C_L > 0, \bar{L} > 0$ such that

$$L(\alpha, x) \leq \bar{L}|\alpha|^2 + C_L \quad \forall \alpha, x \in \mathbb{R}^d. \tag{4.1}$$

(i.2) There exist $c_L, c'_L > 0$ such that

$$\begin{aligned} \partial_{\alpha, \alpha}^2 L(\alpha, x)(\alpha', \alpha') &\geq c_L |\alpha'|^2 \forall \alpha, \alpha' \in \mathbb{R}^d, x \in \mathbb{R}^d, \\ |\partial_x L(\alpha, x)| &\leq c'_L (1 + |\alpha|^2) \forall \alpha, x \in \mathbb{R}^d. \end{aligned} \tag{4.2}$$

(ii) The functions f and g are continuous. Moreover, for every $m \in \mathcal{P}_1(\mathbb{R}^d)$ the functions $f(\cdot, m)$ and $g(\cdot, m)$ belong to $C^2(\mathbb{R}^d)$ and there exists a constant $C_{f,g} > 0$ such that

$$\sup_{m \in \mathcal{P}_1(\mathbb{R}^d)} \{ \|f(\cdot, m)\|_{C^2} + \|g(\cdot, m)\|_{C^2} \} \leq C_{f,g},$$

where, for $h = f, g$, we have set

$$\|h(\cdot, m)\|_{C^2} := \sup_{x \in \mathbb{R}^d} \left\{ |h(x, m)| + \sum_{i=1}^d |\partial_{x_i} h(x, m)| + \sum_{i,j=1}^d |\partial_{x_i x_j} h(x, m)| \right\}.$$

(iii) The measure m_0 is absolutely continuous w.r.t. to the Lebesgue measure \mathcal{L}^d , with density still denoted by m_0 , and has a compact support.

A typical example of a function L satisfying **(A2)(i)** is given by $\mathbb{R}^d \times \mathbb{R}^d \ni (\alpha, x) \rightarrow L(\alpha, x) := b_1(x)|\alpha|^2 + b_2(x)$, where, for $i = 1, 2$, $b_i \in C^2(\mathbb{R}^d)$, b_i is Lipschitz, and there exist constants $\underline{b}_1 > 0, \underline{b}_2 \in \mathbb{R}$ and $\bar{b}_i > 0$ such that $\underline{b}_i \leq b_i \leq \bar{b}_i$.

Remark 4.1 (i) Assumption **(A2)(i)** above implies the convexity of $L(\cdot, x)$ and the existence of $\underline{L}, C'_L, c''_L > 0$ such that

$$\begin{aligned} L(\alpha, x) &\geq \underline{L}|\alpha|^2 - C'_L \quad \forall \alpha, x \in \mathbb{R}^d, \\ |\partial_\alpha L(\alpha, x)| &\leq c''_L (1 + |\alpha|) \quad \forall \alpha, x \in \mathbb{R}^d. \end{aligned} \tag{4.3}$$

(ii) For $x \in \mathbb{R}^d$ let us denote by $H(\cdot, x)$ the convex conjugate of $L(\cdot, x)$. The bound in (i.1) and the first bound in (4.3) imply the existence of constants \underline{H}, \bar{H} and $C_H > 0$ such that

$$\underline{H}|\xi|^2 - C_H \leq H(\xi, x) \leq \bar{H}|\xi|^2 + C_H \quad \forall \xi, x \in \mathbb{R}^d. \tag{4.4}$$

(iii) By the first estimate in (4.2) we have that $\partial_\xi H(\xi, x)$ is characterized as the unique solution $\alpha(\xi, x)$ of the optimization problem $\max_{\alpha \in \mathbb{R}^d} \{\xi \cdot \alpha - L(\alpha, x)\}$. As a consequence of this fact and the first relation in (4.3), we obtain the existence of $c_H > 0$ such that

$$|\partial_\xi H(\xi, x)| \leq c_H(1 + |\xi|) \quad \forall \xi, x \in \mathbb{R}^d. \tag{4.5}$$

Moreover, from the convexity of $L(x, \cdot)$, for all $\xi, x \in \mathbb{R}^d$ we have that $\partial_\xi H(\xi, x)$ is the unique solution to

$$\partial_\alpha L(\partial_\xi H(\xi, x), x) = \xi. \tag{4.6}$$

Using the relation above, the regularity $L \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$, the first estimate in (4.2) and the implicit function theorem, we obtain that $\mathbb{R}^d \times \mathbb{R}^d \ni (\xi, x) \rightarrow \partial_\xi H(\xi, x) \in \mathbb{R}^d$ belongs to $C^1(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$. Using this fact, we get that $\mathbb{R}^d \times \mathbb{R}^d \ni (\xi, x) \mapsto \partial_x H(\xi, x) = -\partial_x L(\partial_\xi H(\xi, x), x) \in \mathbb{R}^d$ is also of class C^1 . As a consequence H is of class C^2 .

Let us define $\ell : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ and $\Phi : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ by

$$\ell(\alpha, x, \mu) := L(-\alpha, x) + f(x, \mu) \quad \text{and} \quad \Phi(x, \mu) := g(x, \mu). \tag{4.7}$$

Clearly, Assumption (A2) implies that ℓ, Φ and m_0 satisfy (A1).

4.1 Lagrangian MFG Equilibria and the MFG PDE System

As pointed out in [16], under (A2) the existence of a Lagrangian equilibrium for the MFG problem defined by ℓ, Φ and m_0 , is equivalent the existence of a solution (u, ρ) of the following PDE system, which was first introduced in [36],

$$\left. \begin{aligned} -\partial_t u + H(\nabla u, x) &= f(x, \rho(t)) \quad \text{in } \mathbb{R}^d \times (0, T), \\ \partial_t \rho - \operatorname{div}(\partial_\xi H(\nabla u, x)\rho) &= 0 \quad \text{in } \mathbb{R}^d \times (0, T), \\ u(\cdot, T) = g(\cdot, \rho(T)), \quad \rho(0) &= m_0 \text{ in } \mathbb{R}^d. \end{aligned} \right\} \quad (MFG)$$

In the system above, $u : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ is a solution to the first equation, with the associated terminal condition, if it is globally Lipschitz, locally semi-concave with respect to its first argument (see [12, Section 2]), uniformly in $t \in [0, T]$, and the equation holds in the viscosity sense (see e.g. [4, Chapter III, Section 3]). In (MFG), ∇u denotes a Borel measurable selection of the set-valued map

$$\begin{aligned} &\mathbb{R}^d \times [0, T] \ni (x, t) \mapsto D_x^+ u(x, t) \\ &:= \left\{ p \in \mathbb{R}^d \mid \limsup_{x' \rightarrow x} \frac{u(x', t) - u(x, t) - p \cdot (x' - x)}{|x' - x|} \leq 0 \right\} \subseteq \mathbb{R}^d. \end{aligned}$$

The existence of such measurable selection follows from the fact that the above set-valued map has a closed graph (thanks to the semi-concavity property of u , see e.g. [12, Proposition 3.3.4]). Moreover, since u is Lipschitz, $\nabla u(x, t)$ is uniformly bounded in $(x, t) \in \mathbb{R}^d \times [0, T]$.

The function $\rho : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ is a solution to the second equation, with the associated initial condition, if $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, and the equation is satisfied in the sense of distributions, i.e. for all $\phi \in C^\infty(\mathbb{R}^d)$ with compact support we have

$$\int_{\mathbb{R}^d} \phi(x) d\rho(t)(x) = \int_{\mathbb{R}^d} \phi(x) dm_0(x) - \int_0^t \int_{\mathbb{R}^d} \partial_\xi H(\nabla u(x, s), x) \cdot \nabla \phi(x) d\rho(s)(x) ds. \tag{4.8}$$

Note that by the previous considerations, the second term in the right-hand-side of (4.8) is well-defined.

Any pair (u, ρ) satisfying (MFG) is called an *equilibrium of the first order MFG problem*.

For the sake of completeness, let us provide the main arguments that justify the equivalence between both notions of equilibria. Let $m_* \in \mathcal{A}_{rel}$ be a Lagrangian MFG equilibrium. Let us define $\rho : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ by $\rho(t) := e_t \# m_*$ for all $t \in [0, T]$, and the value function $u : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ by

$$u(x, t) := \inf \left\{ \int_t^T \ell(\dot{\gamma}(s), \gamma(s), \rho(s)) ds + \Phi(\gamma(T), \rho(T)) \mid \gamma \in W^{1,2}([t, T]; \mathbb{R}^d) \text{ and } \gamma(t) = x \right\}, \tag{4.9}$$

for all $(x, t) \in \mathbb{R}^d \times [0, T]$. Since $m_* \in \mathcal{A}_{rel}$, we have that $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. Using this fact, assumption (A2) and [5, Proposition 1.1 and Remark 1.1], we obtain that u is a viscosity solution of

$$\begin{aligned} -\partial_t u + H(\nabla u, x) &= f(x, \rho(t)) \text{ in } \mathbb{R}^d \times (0, T), \\ u(\cdot, T) &= g(\cdot, \rho(T)) \text{ in } \mathbb{R}^d. \end{aligned} \tag{4.10}$$

Moreover, by [21, Theorem 2.1], the Hamilton-Jacobi-Bellman equation above admits a comparison principle, which implies that u is its unique viscosity solution. We will need the following result, whose proof follows from standard arguments.

Lemma 4.1 *Under (A2) we have:*

(i) *For every $(x, t) \in \mathbb{R}^d \times [0, T]$ the set $\mathcal{S}(x, t)$ of paths $\gamma_* \in W^{1,2}([t, T]; \mathbb{R}^d)$ such that $\gamma_*(t) = x$ and*

$$u(x, t) = \int_t^T \ell(\dot{\gamma}_*(s), \gamma_*(s), \rho(s)) ds + \Phi(\gamma_*(T), \rho(T)) \tag{4.11}$$

is non-empty. Moreover, there exists a constant $C > 0$, independent of (x, t) , such that

$$\sup_{s \in [t, T], \gamma_* \in \mathcal{S}(x, t)} |\dot{\gamma}_*(s)| \leq C. \tag{4.12}$$

(ii) *The value function u is globally Lipschitz.*

(iii) The value function u is locally semi-concave w.r.t. to the space variable, uniformly in $t \in [0, T]$. More precisely, for any compact set $K \subseteq \mathbb{R}^d$ there exists a constant C_K , independent of t , such that for every $\lambda \in [0, 1]$, $x, y \in K$, such that the segment $[x, y]$ is contained in K , the following inequality holds

$$\lambda u(x, t) + (1 - \lambda)u(y, t) \leq u(\lambda x + (1 - \lambda)y) + C_K \frac{\lambda(1 - \lambda)}{2} |x - y|^2. \tag{4.13}$$

Proof The proof being standard, we only sketch the main ideas. The fact that $\mathcal{S}(x, t)$ is non-empty follows directly from **(A2)** and the direct method in the Calculus of Variations. Moreover, by **(A2)** and arguing as in the proof of Lemma 2.1 we obtain the existence of $c > 0$, independent of (x, t, ρ) , such that

$$u(x, t) = \inf \left\{ \int_t^T \ell(\dot{\gamma}(s), \gamma(s), \rho(s)) ds + \Phi(\gamma(T), \rho(T)) \mid \gamma \in W^{1,2}([0, T]; \mathbb{R}^d), \gamma(t) = x, \text{ and } \int_0^T |\dot{\gamma}(s)|^2 ds \leq c \right\}. \tag{4.14}$$

Using this fact, the Euler-Lagrange equation associated to any element $\gamma_* \in \mathcal{S}(x, t)$, the second estimate in (4.2) and arguing as in the proof of [12, Theorem 6.2.5], we easily obtain (4.12), which proves assertion (i). In order to prove (ii), notice that (i) implies that the value function can also be written as

$$u(x, t) = \inf \left\{ \int_t^T \ell \left(\alpha(s), x + \int_t^s \alpha(s') ds', \rho(s) \right) ds + \Phi \left(x + \int_t^T \alpha(s') ds', \rho(T) \right) \mid \alpha \in \hat{A}_C \right\}, \tag{4.15}$$

where $\hat{A}_C := \{ \alpha \in L^\infty([0, T]; \mathbb{R}^d) \mid \|\alpha\|_{L^\infty} \leq C \}$. Using the estimate $|\inf_{\alpha \in \hat{A}_C} A(\alpha) - \inf_{\alpha \in \hat{A}_C} B(\alpha)| \leq \sup_{\alpha \in \hat{A}_C} |A(\alpha) - B(\alpha)|$ for any functions $A, B : L^\infty([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$, expression (4.7), the uniform Lipschitz property for f and g in **(A2)**(ii), and the second estimate in (4.2), we easily obtain that $u(\cdot, t)$ is globally Lipschitz, with a Lipschitz constant which is independent of $t \in [0, T]$. Similarly, using (4.15) and the estimate (4.1), we get that $u(x, \cdot)$ is globally Lipschitz, with a Lipschitz constant which is independent of $x \in \mathbb{R}^d$. Assertion (ii) follows. Finally, assertion (iii) follows directly from [12, Theorem 6.4.1]. \square

Now, let us consider the set-valued map

$$\mathbb{R}^d \ni x \mapsto \mathcal{S}(x) := \operatorname{argmin} \left\{ \int_0^T \ell(\dot{\gamma}(t), \gamma(t), \rho(t)) dt + \Phi(\gamma(T), \rho(T)) \mid \gamma \in W^{1,2}([0, T]; \mathbb{R}^d), \gamma(0) = x \right\}.$$

Since Lemma 4.1(ii) implies that $u(\cdot, 0)$ is a.e. differentiable, [12, Corollary 6.4.10] yields that for a.e. $x \in \mathbb{R}^d$ we have $\mathcal{S}(x) = \{\tilde{\gamma}^x\}$ for some $\tilde{\gamma}^x \in \mathcal{A}(x)$. Now, as in the proof of Proposition 3.1, let $\gamma_* \in \mathcal{A}$ be a Borel measurable selection of \mathcal{S} . Then, for a.e. $x \in \mathbb{R}^d$ we have that $\gamma_*^x = \tilde{\gamma}^x$. Thus, Proposition 3.1(ii) yields $m_*^x = \delta_{\gamma_*^x}$ for a.e. $x \in \text{supp}(m_0)$ and, hence, $m_* = \gamma_* \# m_0$. In particular, $\rho(t) = \gamma_*^{(\cdot)}(t) \# m_0$ for all $t \in [0, T]$.

On the other hand, by [12, Theorems 6.4.9, 6.3.3 and 6.4.8], for a.e. $x \in \mathbb{R}^d$, we have

$$\dot{\gamma}_*^x(t) = -\partial_\xi H(\nabla u(\gamma_*^x(t), t), \gamma_*^x(t)) \quad \forall t \in (0, T), \quad \gamma_*^x(0) = x, \quad (4.16)$$

where we underline that u is differentiable w.r.t. to its first argument at the point $(\gamma_*^x(t), t)$ if $t \in (0, T)$ (see [12, Theorem 6.4.7]). Denoting still by ∇u a measurable selection of $(x, t) \mapsto D_x^+ u(x, t)$, for every $\phi \in C^\infty(\mathbb{R}^d)$ with compact support and $t \in [0, T]$, we have

$$\begin{aligned} & \int_{\mathbb{R}^d} \phi(x) d\rho(t)(x) \\ &= \int_{\mathbb{R}^d} \phi(\gamma_*^x(t)) dm_0(x) \\ &= \int_{\mathbb{R}^d} \phi(x) dm_0(x) - \int_{\mathbb{R}^d} \int_0^t \partial_\xi H(\nabla u(\gamma_*^x(s), s), \gamma_*^x(s)) \nabla \phi(\gamma_*^x(s)) ds dm_0(x), \\ &= \int_{\mathbb{R}^d} \phi(x) dm_0(x) - \int_0^t \int_{\mathbb{R}^d} \partial_\xi H(\nabla u(x, s), x) \nabla \phi(x) d\rho(s)(x) ds, \end{aligned}$$

which implies that ρ satisfies (4.8) and, hence, the couple (u, ρ) solves (MFG) . Notice that under **(A2)** a Lagrangian MFG equilibrium m_* exists (see Remark 3.1) and, hence, the previous arguments show, in particular, the existence of at least one solution (u, ρ) to (MFG) .

Conversely, if (u, ρ) solves (MFG) , then the first equation therein implies that u and ρ are still related by (4.9). By the second equation in (MFG) and [3, Theorem 8.2.1], there exists a probability measure $m_* \in \mathcal{P}(\Gamma)$ such that $\rho(t) = e_t \# m_*$ for all $t \in [0, T]$ and, considering the disintegration $dm_*(\gamma) = dm_*^x(\gamma) \otimes dm_0(x)$, for a.e. $x \in \text{supp}(m_0)$ the support of the measure m_*^x is contained in the set of solutions of (4.16). By Lemma 4.1 and arguing as in the proof of [13, Lemma 4.11], we have that every solution to (4.16) solves the optimization problem in the r.h.s. of (4.9) with $t = 0$. Thus, by Proposition 3.1(ii) we obtain that m_* is a Lagrangian MFG equilibrium. Notice also that $\mathcal{S}(x)$ being a singleton for a.e. $x \in \mathbb{R}^d$, the previous argument shows, in particular, that $[0, T] \ni t \mapsto \gamma_*^{(\cdot)}(t) \# m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ is the unique solution in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ of the continuity equation

$$\partial_t \rho - \text{div}(\partial_\xi H(\nabla u, x)\rho) = 0 \quad \text{in } \mathbb{R}^d \times (0, T), \quad \rho(0) = m_0 \quad \text{in } \mathbb{R}^d. \quad (4.17)$$

Remark 4.2 In the state constrained case (see Remark 3.1(ii)), the equivalence between the Lagrangian notion of equilibrium and the solution of a PDE system has been thoroughly investigated in [11].

In addition to the relation between Lagrangian MFG equilibria and the solutions of (MFG), assumption (A2) has also consequences on the regularity of the time marginals $\{\rho(t) \mid t \in [0, T]\}$ as the following result shows.

Proposition 4.1 *In addition to (A2)(iii), assume that $m_0 \in L^p(\mathbb{R}^d)$ for some $p \in (1, +\infty]$ and let (u, ρ) be a solution to (MFG). Then, the following assertions hold true:*

(i) *There exists $c_1 > 0$, independent of $t \in [0, T]$, such that $\text{supp}(\rho(t)) \subseteq B(0, c_1)$ for all $t \in [0, T]$.*

(ii) *For all $t \in [0, T]$ the measure $\rho(t)$ is absolutely continuous w.r.t. the Lebesgue measure. Moreover, the density of $\rho(t)$, that we will still denote by $\rho(t)$, belongs to $L^p(\mathbb{R}^d)$ and there exists a constant $c_2 > 0$, independent of $p \in (1, +\infty]$ and $t \in [0, T]$, such that*

$$\|\rho(t)\|_{L^p} \leq c_2 \|m_0\|_{L^p}. \tag{4.18}$$

Proof Assertion (i) follows directly from the formula $\rho(t) = \gamma_*^{(\cdot)}(t) \# m_0$, where $\gamma_*^x \in \mathcal{S}(x)$ for all $x \in \mathbb{R}^d$, Lemma 4.1(i) and the fact that $\text{supp}(m_0)$ is compact. In order to prove (ii), let $\beta \in C^\infty(\mathbb{R}^d)$, non-negative, with support contained in the unit ball and such that $\int_{\mathbb{R}^d} \beta(x) dx = 1$. For $\varepsilon > 0$, let us define $\beta_\varepsilon(x) := \varepsilon^{-d} \beta(x/\varepsilon)$, $u_\varepsilon(x, t) := (\beta_\varepsilon * u(\cdot, t))(x)$ and consider the equation

$$\partial_t \rho_\varepsilon - \text{div}(\partial_\xi H(\nabla u_\varepsilon, x) \rho_\varepsilon) = 0 \quad \text{in } \mathbb{R}^d \times (0, T), \quad \rho_\varepsilon(0) = m_0 \quad \text{in } \mathbb{R}^d. \tag{4.19}$$

For every $x \in \mathbb{R}^d$, let us define $\gamma_\varepsilon^x \in C^1([0, T]; \mathbb{R}^d)$ as the unique solution to

$$\dot{\gamma}_\varepsilon^x(t) = -\partial_\xi H(\nabla u_\varepsilon(\gamma_\varepsilon^x(t), t), \gamma_\varepsilon^x(t)) \quad \forall t \in (0, T), \quad \gamma_\varepsilon^x(0) = x. \tag{4.20}$$

By [3, Proposition 8.1.8], Eq. (4.19) admits a unique solution in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, which is given by $\rho_\varepsilon(t) := \gamma_\varepsilon^{(\cdot)}(t) \# m_0$ for all $t \in [0, T]$. Moreover, by a standard change of variable argument (see e.g. [2, Section 2]), for every $t \in [0, T]$ we have that $\rho_\varepsilon(t)$ is absolutely continuous, with density given by

$$\rho_\varepsilon(y, t) = \frac{m_0\left([\gamma_\varepsilon^{(\cdot)}(t)]^{-1}(y)\right)}{\left|\det\left(Y\left([\gamma_\varepsilon^{(\cdot)}(t)]^{-1}(y), t\right)\right)\right|} \quad \text{for a.e. } y \in \mathbb{R}^d,$$

where, for each $x \in \mathbb{R}^d$, $Y(x, t) = \exp\left(-\int_0^t L_\varepsilon(x, s) ds\right)$ with

$$\begin{aligned} L_\varepsilon(x, s) &:= D_x \left[\partial_\xi H(\nabla u_\varepsilon(\cdot, s), \cdot) \right] (\gamma_\varepsilon^x(s)) \\ &= \partial_{\xi, \xi}^2 H(\nabla u_\varepsilon(\gamma_\varepsilon^x(s), s), \gamma_\varepsilon^x(s)) \partial_{x, x}^2 u_\varepsilon(\gamma_\varepsilon^x(s), s) \\ &\quad + \partial_{\xi, x}^2 H(\nabla u_\varepsilon(\gamma_\varepsilon^x(s), s), \gamma_\varepsilon^x(s)). \end{aligned} \tag{4.21}$$

Let us assume that $p \in (1, +\infty)$. By a change of variable again, we obtain that

$$\begin{aligned} \|\rho_\varepsilon(t)\|_{L^p}^p &= \int_{\mathbb{R}^d} m_0^p(x) |\det(Y(x, t))|^{1-p} dx \\ &= \int_{\text{supp}(m_0)} m_0^p(x) |\det(Y(x, t))|^{1-p} dx. \end{aligned} \tag{4.22}$$

Now, for all $x \in \mathbb{R}^d$ and $t \in [0, T]$, we have (see [2, Section 2, estimate (2.4)])

$$\begin{aligned} |\det(Y(x, t))|^{1-p} &\leq \exp\left((p-1) \int_0^t \left\| [\text{div}(-\partial_\xi H(\nabla u_\varepsilon(\cdot, s), \cdot))]_- \right\|_{L^\infty} ds\right) \\ &\leq \exp\left(p \int_0^t \left\| [\text{Tr}(L_\varepsilon(x, t))]_+ \right\|_{L^\infty} dt\right), \end{aligned} \tag{4.23}$$

where $[a]_- := \max\{0, -a\}$, $[a]_+ = a + [a]_-$, for any $a \in \mathbb{R}$, and $\text{Tr}(L_\varepsilon(x, t))$ denotes the trace of the matrix $L_\varepsilon(x, t)$. On the other hand, Lemma 4.1(ii), and the definition of u_ε , imply that $\nabla u_\varepsilon(x, t)$ is bounded, uniformly in $(x, t) \in \mathbb{R}^d \times [0, T]$ and $\varepsilon > 0$. Moreover, by Lemma 4.1(iii), the compactness of $\text{supp}(m_0)$ and the definition of $u_\varepsilon(\cdot, t)$ again, we can assume that $u_\varepsilon(\cdot, t)$ is uniformly semiconcave in a bounded open set \mathcal{O} containing $\text{supp}(m_0)$, i.e. $u_\varepsilon(\cdot, t)$ satisfies (4.13) for all $x, y \in \mathcal{O}$, with C_K replaced by \tilde{c} , for some \tilde{c} independent of t and ε small enough. By [12, Proposition 1.1.3], we have that $\partial_{xx}^2 u_\varepsilon(x, t) - \tilde{c}I_d$ negative semidefinite for all $(x, t) \in \mathcal{O} \times [0, T]$ and, hence, using that $\partial_{\xi, \xi} H(\xi, x)$ is positive semidefinite for all $\xi \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$, by (4.21) we get that $\text{Tr}(L_\varepsilon(x, t))$ is bounded from above by a constant which is independent of $\varepsilon > 0$ small enough, $x \in \mathcal{O}$, and $t \in [0, T]$. Thus, by (4.23) and taking the power $1/p$ in (4.22), there exists $c_2 > 0$, independent of ε, t and p , such that

$$\|\rho_\varepsilon(t)\|_{L^p} \leq c_2 \|m_0\|_{L^p} \quad \forall t \in [0, T]. \tag{4.24}$$

The previous estimate shows the existence of $\tilde{\rho} \in L^\infty([0, T]; L^p(\mathbb{R}^d))$ and a sequence $(\rho_{\varepsilon_n})_{n \in \mathbb{N}}$ such that, as $n \rightarrow \infty$, $\varepsilon_n \rightarrow 0$ and $\rho_{\varepsilon_n} \rightarrow \tilde{\rho} \in L^\infty([0, T]; L^p(\mathbb{R}^d))$ in the weak* topology. By dominated convergence, we have that $\partial_\xi H(\nabla u_{\varepsilon_n}, \cdot) \rightarrow \partial_\xi H(\nabla u, \cdot)$ in $L^1([0, T]; L^s(\mathbb{R}^d))$ for any $s \in [1, +\infty)$. As a consequence, $\tilde{\rho}$ satisfies estimate (4.24) and, passing to the limit in (4.19), we get that the measure $[0, T] \ni t \rightarrow \tilde{\rho}(t)\mathcal{L}^d \in L^p(\mathbb{R}^d)$ satisfies (4.8). Using that $[0, T] \ni t \mapsto \rho(t) \in \mathcal{P}_1(\mathbb{R}^d)$ is the unique solution to (4.17) in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, [3, Lemma 8.1.2] implies that $\tilde{\rho}(t)\mathcal{L}^d = \rho(t)$ for a.e. $t \in [0, T]$. Thus, for a.e. $t \in [0, T]$, $\rho(t)$ is absolutely continuous w.r.t. to the Lebesgue measure and estimate (4.18) holds for its density. Using this fact and that $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, the previous statement is valid in the whole time interval $[0, T]$, which proves (ii) when $p < +\infty$. Since c_2 does not depend on p , assertion (ii) for $p = \infty$ follows by taking the limit in (4.18) when $p \rightarrow \infty$. \square

Remark 4.3 Similar regularization techniques have been recently employed in [26], in order to establish L^p -estimates for the time evolving distributions describing equilibria in optimal-exit MFGs. Let us also point out that the absolute continuity of $\rho(t)$, for

all $t \in [0, T]$, has already been established in [30, Section 5], under more general assumptions than **(A2)**, but without providing the L^p -estimate (4.18).

4.2 The N -Player Equilibria: Associated Time Marginals and Value Functions

Let us consider the game with N players defined in Sect. 2 with ℓ and Φ given by (4.7). Let $(m_N, \dots, m_N) \in \mathcal{A}_{rel}^N$ be a symmetric equilibrium in randomized distributed open-loop strategies for the N -player game. Note that if for $h = f, g$ we define

$$h_N(x, \mu) := \int_{(\mathbb{R}^d)^{N-1}} h \left(x, \frac{1}{N-1} \sum_{j=2}^N \delta_{x_j} \right) \otimes_{j=2}^N d\mu(x_j) \quad \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_1(\mathbb{R}^d), \tag{4.25}$$

we have that f_N and g_N satisfy the assumptions for f and g in **(A2)**(ii) (with the same constant $C_{f,g}$). As a consequence of this fact, the results in [12, Chapter 6] and Corollary 2.1 we obtain the existence of $\gamma_N \in \mathcal{A}$ such that (m_N, \dots, m_N) with $m_N = \gamma_N \# m_0$ can be identified with the non-randomized symmetric equilibrium in distributed open-loop strategies given by $(\gamma_N, \dots, \gamma_N) \in \mathcal{A}^N$. Furthermore, setting $\rho_N(t) := e_t \# m_N = \gamma_N^{(\cdot)}(t) \# m_0$ for all $t \in [0, T]$, we have that $\gamma_N^x \in \mathcal{S}^N(x) := \mathcal{S}^N(x, 0)$, where

$$\begin{aligned} \mathcal{S}^N(x, t) := \operatorname{argmin} & \left\{ \int_t^T [L(-\dot{\gamma}(s), \gamma(s)) + f_N(\gamma(s), \rho_N(s))] ds \right. \\ & \left. + g_N(\gamma(T), \rho_N(T)) \mid \gamma \in W^{1,2}([t, T]; \mathbb{R}^d) \text{ and } \gamma(t) = x \right\} \forall x \in \mathbb{R}^d, t \in [0, T]. \end{aligned}$$

Remark 4.4 (i) Recall that the representation $m_N = \gamma_N \# m_0$ is only m_0 -uniquely determined. In particular, if $\gamma'_N \in \mathcal{A}$ is different from γ_N but coincides with it on a set A such that $m_0(A) = 1$, then we also have that $m_N = \gamma'_N \# m_0$. For the sake of simplicity, we have chosen to represent always m_N via a measurable selection γ_N of the set-valued map \mathcal{S}^N . Notice that **(A2)** and the results in [12, Chapter 6] imply that γ_N^x is uniquely defined for a.e. $x \in \mathbb{R}^d$.

(ii) Exactly as in the limit case (see Lemma 4.1(i)), we have the existence of a constant $C > 0$, independent of (x, t) and $N \in \mathbb{N}$, such that

$$\sup_{s \in [t, T], \gamma_* \in \mathcal{S}^N(x, t)} |\dot{\gamma}_*(s)| \leq C \quad \forall x \in \mathbb{R}^d, t \in [0, T], N \in \mathbb{N}. \tag{4.26}$$

As a consequence, there exists a compact set $K_C \subseteq \mathbb{R}^d$ such that $\gamma_N^x(t) \in K_C$ for all $N \in \mathbb{N}, x \in \operatorname{supp}(m_0)$ and $t \in [0, T]$. In particular, the representation $\rho_N(t) = e_t \# m_N = \gamma_N^{(\cdot)}(t)$ implies that $\operatorname{supp}(\rho_N(t)) \subseteq K_C$ for all $N \in \mathbb{N}$ and $t \in [0, T]$.

Let us define $u_N : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ by

$$\begin{aligned}
 u_N(x, t) := \inf \left\{ \int_t^T [L(-\dot{\gamma}(s), \gamma(s)) + f_N(\gamma(s), \rho_N(s))] ds \right. \\
 \left. + g_N(\gamma(T), \rho_N(T)) \mid \gamma \in W^{1,2}([t, T]; \mathbb{R}^d) \text{ and } \gamma(t) = x \right\} \\
 \forall x \in \mathbb{R}^d, t \in [0, T].
 \end{aligned}
 \tag{4.27}$$

Remark 4.5 Mimicking the proofs of Lemma 4.1 and [12, Theorem 6.4.1] we obtain that u_N is globally Lipschitz and locally semi-concave. Moreover, the Lipschitz and local semi-concavity constants are independent of N .

Arguing as in the previous subsection, the pair (u_N, ρ_N) solves

$$\left. \begin{aligned}
 -\partial_t u_N + H(\nabla u_N, x) &= f_N(x, \rho_N(t)) \quad \text{in } \mathbb{R}^d \times (0, T), \\
 \partial_t \rho_N - \operatorname{div}(\partial_\xi H(\nabla u_N, x)\rho_N) &= 0 \quad \text{in } \mathbb{R}^d \times (0, T), \\
 u_N(\cdot, T) = g_N(\cdot, \rho_N(T)), \quad \rho_N(0) &= m_0 \text{ in } \mathbb{R}^d.
 \end{aligned} \right\} \quad (MFG_N)$$

Conversely, associated to any solution (u_N, ρ_N) we have the existence of a symmetric equilibrium in distributed open-loop strategies $(\gamma_N, \dots, \gamma_N) \in \mathcal{A}^N$ for the N -player game. Moreover, using again the results in [12, Chapter 6], any $\gamma_N \in \mathcal{A}$ defining such equilibrium satisfies

$$\dot{\gamma}_N^x(t) = -\partial_\xi H(\nabla u_N(\gamma_N^x(t), t), \gamma_N^x(t)) \quad \forall t \in (0, T), \quad \gamma_N^x(0) = x, \tag{4.28}$$

for a.e. $x \in \operatorname{supp}(m_0)$. Thus, we can think of the r.h.s. above as an optimal control which is feedback with respect to the individual states. We call $(\gamma_N, \dots, \gamma_N)$ a Nash equilibrium in *distributed Markov strategies* for the N -player game. As a consequence of the previous discussion, such equilibria exist for all $N \in \mathbb{N}$ provided that **(A2)** holds true.

Let us consider a sequence $(\gamma_N)_{N \in \mathbb{N}}$ of elements in \mathcal{A} defining Nash equilibria in distributed Markov strategies for the N -player games. Theorem 3.1 yields the existence of a Lagrangian equilibrium $m_* \in \mathcal{A}_{rel}$ and a subsequence $(\gamma_{N_k})_{k \in \mathbb{N}}$ such that $\gamma_{N_k} \# m_0 \rightarrow m_*$ as $k \rightarrow \infty$. Let $\gamma_* \in \mathcal{A}$ be such that $m_* = \gamma_* \# m_0$ and $\gamma_*^x \in \mathcal{S}(x)$ for all $x \in \mathbb{R}^d$, i.e. an equilibrium in distributed Markov strategies for the MFG.

Our aim now is to study the convergence of the associated time marginals ρ_{N_k} to ρ , the convergence of the associated value functions u_{N_k} to u and, finally, the convergence of $\gamma_{N_k}^x$ to γ_*^x for a.e. $x \in \mathbb{R}^d$.

We will need the following preliminary result.

Lemma 4.2 Assume **(A2)**(ii) and let $K \subseteq \mathbb{R}^d$ be a nonempty compact set. Consider a sequence of measures $(\mu_k)_{k \in \mathbb{N}} \subseteq \mathcal{P}_1(\mathbb{R}^d)$ such that $\operatorname{supp}(\mu_k) \subseteq K$, for all $k \in \mathbb{N}$, and, as $k \rightarrow \infty$, $\mu_k \rightarrow \mu$ for some $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. Then, for any sequence $(x_k)_{k \in \mathbb{N}}$ and $x \in \mathbb{R}^d$ such that $x_k \rightarrow x$, we have

$$h(x, \mu) = \lim_{k \rightarrow \infty} h_k(x_k, \mu_k), \tag{4.29}$$

where $h = f, g$ and h_k is defined by (4.25).

Proof Notice that (A2)(ii) implies that

$$|h_k(x_k, \mu_k) - h(x, \mu_k)| \leq C_{f,g}|x - x_k|. \tag{4.30}$$

Now, let Y_1^k, \dots, Y_k^k be independent and identically distributed K -valued random variables, defined on some probability space $(\Omega_k, \mathcal{F}_k, \mathbf{P}_k)$, with common distribution μ_k . Using that $\mathcal{P}(\mathcal{P}(K))$ is compact and arguing as in the proof of Theorem 3.1, we obtain that, as $k \rightarrow \infty$, the $\mathcal{P}_1(\mathbb{R}^d)$ -valued random sequence $\left(\frac{1}{k-1} \sum_{j=2}^k \delta_{Y_j^k}\right)_{k \in \mathbb{N}}$ converges in distribution to the deterministic measure μ . Since (4.30) can be written as

$$\left| h_k(x_k, \mu_k) - \mathbf{E}_k \left(h \left(x, \frac{1}{k-1} \sum_{j=2}^k \delta_{Y_j^k} \right) \right) \right| \leq C_{f,g}|x - x_k|,$$

relation (4.29) follows by letting $k \rightarrow \infty$. □

Theorem 4.1 Assume that (A2) holds. Then, the following assertions hold true:

- (i) The sequence $(\rho_{N_k})_{k \in \mathbb{N}}$ converges to ρ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$.
- (ii) The sequence $(u_{N_k})_{k \in \mathbb{N}}$ converges to u uniformly on compact subsets of $\mathbb{R}^d \times [0, T]$.
- (iii) For a.e. $x \in \mathbb{R}^d$, the sequence $(\gamma_{N_k}^x)_{k \in \mathbb{N}}$ converges to γ_*^x uniformly in $[0, T]$ and $(\dot{\gamma}_{N_k}^x)_{k \in \mathbb{N}}$ converges to $\dot{\gamma}_*^x$ in the weak* topology in $L^\infty([0, T]; \mathbb{R}^d)$.

Proof Assertion (i) follows directly from Theorem 3.1. Note that (A2) implies that $(u_{N_k})_{k \in \mathbb{N}}$ is a sequence of uniformly bounded functions on $\mathbb{R}^d \times [0, T]$. Let us fix $(x, t) \in \mathbb{R}^d \times [0, T]$. The definition of u_{N_k} , Remark 4.4(ii) and Lemma 4.2 imply that

$$\limsup_{k \rightarrow \infty} u_{N_k}(x, t) \leq u(x, t). \tag{4.31}$$

Let $\gamma_{N_k}^{x,t} \in \mathcal{S}^N(x, t)$ and $\gamma^{x,t} \in C([t, T]; \mathbb{R}^d)$ be a cluster point of $(\gamma_{N_k}^{x,t})_{k \in \mathbb{N}}$, with respect to the uniform convergence. The existence of $\gamma^{x,t}$ is ensured by (4.26) and the Arzelà-Ascoli theorem. Up to the extraction of a subsequence, we can assume that $\liminf_{k \rightarrow \infty} u_{N_k}(x, t) = \lim_{k \rightarrow \infty} u_{N_k}(x, t)$ and $\lim_{k \rightarrow \infty} \gamma_{N_k}^{x,t} = \gamma^{x,t}$ in $C([t, T]; \mathbb{R}^d)$. Using estimate (4.26) again, we get that $\dot{\gamma}^{x,t}$ exists and $\dot{\gamma}_{N_k}^{x,t} \rightarrow \dot{\gamma}^{x,t}$ in the weak* topology in $L^\infty([0, T]; \mathbb{R}^d)$. By the weak lower semi-continuity of the cost functional we obtain

$$\begin{aligned} u(x, t) &\leq \int_t^T [L(-\dot{\gamma}^{x,t}(s), \gamma^{x,t}(s)) + f(\gamma^{x,t}(s), \rho(s))] ds + g(\gamma^{x,t}(T), \rho(T)) \\ &\leq \liminf_{k \rightarrow \infty} u_{N_k}(x, t). \end{aligned}$$

Thus, by (4.31) we get the pointwise convergence

$$\lim_{k \rightarrow +\infty} u_{N_k}(x, t) = u(x, t) \quad \forall (x, t) \in \mathbb{R}^d \times [0, T],$$

and hence, using that u_{N_k} is Lipschitz continuous, with a Lipschitz constant which is independent of k , assertion (ii) follows from the Arzelà-Ascoli theorem. Finally, (iii) is a consequence of the previous analysis with $t = 0$ and the fact that [12, Corollary 6.4.10] implies that for a.e. $x \in \mathbb{R}^d$ the optimization problem associated with $u(x, 0)$ admits a unique solution. \square

Recall that, as in the case of (MFG) , to each solution (u_N, ρ_N) of (MFG_N) we can associate a symmetric equilibrium $(m_N, \dots, m_N) \in \mathcal{A}_{rel}^N$ of the N -player game. As a consequence of this fact, Theorems 3.1 and 4.1, we have the following result.

Corollary 4.1 *Let $((u_N, \rho_N))_{N \in \mathbb{N}}$ be a sequence of solutions to (MFG_N) ($N \in \mathbb{N}$). Then, there exists a solution $(u, \rho) \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ to (MFG) such that, up to some subsequence, $u_N \rightarrow u$ uniformly over compact subsets of $\mathbb{R}^d \times [0, T]$ and $\rho_N \rightarrow \rho$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$.*

Remark 4.6 If $h = f, g$ satisfies

$$\int_{\mathbb{R}^d} (h(x, \mu) - h(x, \mu')) \, d(\mu - \mu')(x) \geq 0 \quad \forall \mu, \mu' \in \mathcal{P}_1(\mathbb{R}^d),$$

then the solution (u, ρ) to (MFG) is unique (see [36] and [32, Corollary 5.2]). Since any Lagrangian equilibrium m_* can be represented by $\gamma_* \# m_0$, where $\gamma_*^x \in \mathcal{S}(x)$ is uniquely determined for a.e. $x \in \mathbb{R}^d$, the Lagrangian equilibrium must also be unique. In this case, the results in Theorem 4.1 hold for the entire sequence (u_N, ρ_N, γ_N) and the result in Corollary 4.1 holds for the entire sequence (u_N, ρ_N) .

Finally, let us point out that the convergence result in Corollary 4.1 can also be established directly, without appealing to Theorem 3.1, under a stronger regularity assumption than **(A2)(ii)**. Indeed, assume that, in addition to **(A2)**, $m_0 \in L^p(\mathbb{R}^d)$ for some $p \in (1, +\infty]$. If (u_N, ρ_N) solves (MFG_N) , then by [3, Theorem 8.2.1] and the results in [12, Chapter 6], we must have that $\rho_N(t) = \gamma_N^{(\cdot)}(t) \# m_0$, for some $\gamma_N \in \mathcal{A}$ such that $\gamma_N^x \in \mathcal{S}^N(x)$ for all $x \in \text{supp}(m_0)$. Arguing as in the proof of Lemma 4.1 we get the existence of $C' > 0$ such that

$$\sup_{x \in \text{supp}(m_0), t \in [0, T]} |\dot{\gamma}_N^x(t)| \leq C' \quad \forall N \in \mathbb{N}.$$

Therefore, there exists $C > 0$ and a compact set $K \subseteq \mathbb{R}^d$, both independent of N , such that $\gamma_N^x(t) \in K$ for a.e. $x \in \text{supp}(m_0)$ and all $t \in [0, T]$. In particular, $\text{supp}(\rho_N(t)) \subseteq K$ and $d_1(\rho_N(s), \rho_N(t)) \leq C|s - t|$ for all $s, t \in [0, T]$ and $N \in \mathbb{N}$. This implies the existence of $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ such that, up to some subsequence, $\rho_N \rightarrow \rho$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ as $N \rightarrow \infty$. Since Lemma 4.2 implies that $f_N(\cdot, \rho_N(\cdot))$ converges uniformly to $f(\cdot, \rho(\cdot))$ on compact subsets of $\mathbb{R}^d \times [0, T]$, standard stability results for viscosity solutions of Hamilton-Jacobi-Bellman equations imply that, up to some subsequence, $u_N \rightarrow u$ uniformly on compact subsets of $\mathbb{R}^d \times [0, T]$, u being the unique viscosity solution to (4.10). In particular, u_N being locally semi-concave with respect to the space variable, uniformly in N , for all $t \in [0, T]$ we have that $\nabla u_N(x, t) \rightarrow$

$\nabla u(x, t)$ for a.e. $x \in \mathbb{R}^d$. Using that u_N is globally Lipschitz, uniformly in N , by dominated convergence we deduce that

$$\partial_\xi H(\nabla u_N(\cdot, \cdot), \cdot) \rightarrow \partial_\xi H(\nabla u(\cdot, \cdot), \cdot) \text{ in } L^1([0, T]; L^s(\mathbb{R}^d)) \text{ for any } s \in [1, +\infty). \quad (4.32)$$

On the other hand, using again the uniform local semiconcavity of $u_N(\cdot, t)$ and arguing as in the proof Proposition 4.1 we get the existence of $c_3 > 0$, independent of N , such that

$$\|\rho_N(t)\|_{L^p} \leq c_3 \|m_0\|_{L^p} \quad \forall t \in [0, T]. \quad (4.33)$$

Using this bound, we obtain that $\rho(t)$ is absolutely continuous w.r.t. the Lebesgue measure for a.e. $t \in [0, T]$, and its density, denoted likewise by $\rho(t)$, satisfies $\|\rho(t)\|_{L^p} \leq c_3 \|m_0\|_{L^p}$ for a.e. $t \in [0, T]$. Since $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, the previous bound implies that for all $t \in [0, T]$, the measure $\rho(t)$ is absolutely continuous w.r.t. the Lebesgue measure and the estimate $\|\rho(t)\|_{L^p} \leq c_3 \|m_0\|_{L^p}$ holds. Moreover, using (4.32)–(4.33) we can pass to the limit in the second equation (MFG_N) to obtain that the pair (u, ρ) solves (MFG).

Acknowledgements The second author was partially supported by the ANR project MFG ANR-16-CE40-0015-01 and by the FMJH Program PGMO, reference Project PGMO VarPDEMFG. The first author acknowledges partial support through the research projects “Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games” of the Fondazione CaRiPaRo and “Large Scale Random Structures” of the Italian Ministry of Education, Universities and Research (PRIN 20155PAWZB-004). Both authors thank two anonymous Referees for their critique and comments, which helped improve the manuscript.

Funding Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis. A Hitchhiker’s Guide. Springer, Berlin (2006)
2. Ambrosio, L.: Transport equation and Cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, (2008)
3. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space Of Probability measures. Lecture notes in Mathematics ETH Zürich, 2nd edn. Birkhäuser Verlag, Basel (2008)
4. Bardi, M., Capuzzo Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Birkhäuser, Basel (1996)
5. Bardi, M., Da Lio, F.: On the Bellman equation for some unbounded control problems. *NoDEA Nonlinear Differ. Equ. Appl.* **4**(4), 491–510 (1997)

6. Bardi, M., Priuli, F.S.: Linear-quadratic N -person and mean-field games with ergodic cost. *SIAM J. Control Optim.* **52**(5), 3022–3052 (2014)
7. Bayraktar, E., Cohen, A.: Analysis of a finite state many player game using its master equation. *SIAM J. Control Optim.* **56**(5), 3538–3568 (2018)
8. Benamou, J.-D., Carlier, G., Santambrogio, F.: Variational mean field games. In *Active particles. Vol. I. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 141–171. Birkhäuser/Springer, Cham (2017)
9. Billingsley, P.: Convergence of probability measures. Wiley, New York-London-Sydney (1968)
10. Cannarsa, P., Capuani, R.: Existence and uniqueness for mean field games with state constraints. In *PDE models for multi-agent phenomena.*, Springer INdAM Ser. Vol. 28, pages 49–71, Springer Cham (2018)
11. Cannarsa, P., Capuani, R., Cardaliaguet, P.: Mean field games with state constraints: from mild to pointwise solutions of the PDE system. *arXiv preprint arXiv:1812.11374* (2018)
12. Cannarsa, P., Sinestrari, C.: Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, Basel (2004)
13. Cardaliaguet, P.: Notes on Mean Field Games: from P.-L. Lions' lectures at Collège de France. *Lecture Notes given at Tor Vergata* (2010)
14. Cardaliaguet, P.: On the convergence of open loop Nash equilibria in mean field games with a local coupling. In *Trends in Control Theory and Partial Differential Equations*, Springer Indam Series, pages 105–122. Springer, Cham (2019)
15. Cardaliaguet, P., Delarue, F., Lasry, J.-M., Lions, P.-L.: The master equation and the convergence problem in mean field games. *arXiv:1509.02505 [math.AP]* (2015)
16. Cardaliaguet, P., Hadikhanloo, S.: Learning in mean field games: the fictitious play. *ESAIM Control Optim. Calc. Var.* **23**(2), 569–591 (2017)
17. Cardaliaguet, P., Mészáros, A., Santambrogio, F.: First order mean field games with density constraints: pressure equals price. *SIAM J. Control Optim.* **54**(5), 2672–2709 (2016)
18. Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications, volumes 83 and 84 of Probability Theory and Stochastic Modelling. Springer, Cham (2018)
19. Cecchin, A., Pelino, G.: Convergence, fluctuations and large deviations for finite state mean field games via the Master Equation. *Stochastic Process. Appl.* **129**(11), 4510–4555 (2019)
20. Cecchin, A., Dai Pra, P., Fischer, M., Pelino, G.: On the convergence problem in Mean Field Games: a two state model without uniqueness. *SIAM J. Control Optim.* **57**(4), 2443–2466 (2019)
21. Da Lio, F., Ley, O.: Convex Hamilton-Jacobi equations under superlinear growth conditions on data. *Appl. Math. Optim.* **63**(3), 309–339 (2011)
22. Dacorogna, B.: Direct Methods in the Calculus of Variations. Applied Mathematical Sciences, vol. 78. Springer-Verlag, Berlin (1980)
23. Delarue, F., Lacker, D., Ramanan, K.: From the master equation to mean field game limit theory: a central limit theorem. *Electron. J. Probab.* **24**, 51 (2019)
24. Delarue, F., Lacker, D., Ramanan, K.: From the master equation to mean field game limit theory: large deviations and concentration of measure. *Ann. Probab.* **48**(1), 211–263 (2020)
25. Delarue, F., Foguen Tchuendom, R.: Selection of equilibria in a linear quadratic mean field game. *Stochastic Process. Appl.* **130**(2), 1000–1040 (2020)
26. Dweik, S., Mazanti, G.: Sharp semi-concavity in a non-autonomous control problem and L^p estimates in an optimal exit MFG. *Nonlinear Differ. Equ. Appl.* **27**, 11 (2020)
27. Feleqi, E.: The derivation of ergodic mean field game equations for several populations of players. *Dyn. Games Appl.* **3**(4), 523–536 (2013)
28. Fischer, M.: On the connection between symmetric N -player games and mean field games. *Ann. Appl. Probab.* **27**(2), 757–810 (2017)
29. Gomes, D.A., Mohr, J., Souza, R.R.: Continuous time finite state mean field games. *Appl. Math. Optim.* **68**(1), 99–143 (2013)
30. Gomes, D.A., Voskanyan, V.K.: Extended deterministic mean-field games. *SIAM J. Control Optim.* **54**(2), 1030–1055 (2016)
31. Green, E.J.: Continuum and finite-player noncooperative models of competition. *Econometrica* **52**(4), 975–993 (1984)
32. Hadikhanloo, S.: Learning in anonymous nonatomic games with applications to first-order mean field games. *arXiv preprint arXiv:1704.00378* (2017)

33. Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6**(3), 221–252 (2006)
34. Lacker, D.: A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Relat. Fields* **165**(3), 581–648 (2016)
35. Lacker, D.: On the convergence of closed loop Nash equilibria to the mean field game limit. *Ann. Appl. Probab.* **30**(4), 1693–1761 (2020)
36. Lasry, J.-M., Lions, P.-L.: Mean field games. *Jpn. J. Math.* **2**, 229–260 (2007)
37. Mazanti, G., Santambrogio, F.: Minimal-time mean field games. Preprint, (2018)
38. Nutz, M., San Martin, J., Tan, X.: Convergence to the mean field game limit: a case study. *Ann. Appl. Probab.* **30**(1), 259–286 (2020)
39. Rashid, S.: Equilibrium points of non-atomic games: asymptotic results. *Econ. Lett.* **12**(1), 7–10 (1983)
40. Schmeidler, D.: Equilibrium points of nonatomic games. *J. Stat. Phys.* **7**(4), 295–300 (1973)
41. Villani, C.: *Topics in Optimal Transportation*. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence, RI (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.