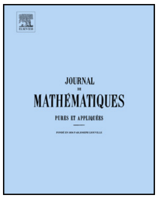
J. Math. Pures Appl. 184 (2024) 190–217

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

journal homepage: www.elsevier.com/locate/matpur



Minimal solutions of master equations for extended mean field games

Chenchen Mou[a,*], Jianfeng Zhang[b]

[a] *Department of Mathematics, City University of Hong Kong, Hong Kong Special Administrative Region of China*
[b] *Department of Mathematics, University of Southern California, CA, USA*

ARTICLE INFO

Article history:
Received 2 March 2023
Available online 1 March 2024

MSC:
35Q89
49N80
35D40
60H30
91A16
93E20

Keywords:
Mean field games
Master equation
Weak solution
Viscosity solution
Comparison principle

ABSTRACT

In an extended mean field game the vector field governing the flow of the population can be different from that of the individual player at some mean field equilibrium. This new class strictly includes the standard mean field games. It is well known that, without any monotonicity conditions, mean field games typically contain multiple mean field equilibria and the wellposedness of their corresponding master equations fails. In this paper, a partial order for the set of probability measure flows is proposed to compare different mean field equilibria. The minimal and maximal mean field equilibria under this partial order are constructed and satisfy the flow property. The corresponding value functions, however, are in general discontinuous. We thus introduce a notion of weak-viscosity solutions for the master equation and verify that the value functions are indeed weak-viscosity solutions. Moreover, a comparison principle for weak-viscosity semi-solutions is established and thus these two value functions serve as the minimal and maximal weak-viscosity solutions in appropriate sense. In particular, when these two value functions coincide, the value function becomes the unique weak-viscosity solution to the master equation. The novelties of the work persist even when restricted to the standard mean field games.

RÉSUMÉ

Dans les jeux à champ moyen étendu, le champ vectoriel gouvernant le flot de la population peut être différent de celui du joueur individuel pour certaines solutions d'équilibre du jeu. Il s'agit d'une classe de jeux à champ moyen qui inclut strictement les jeux à champ moyen standard. Il est bien connu que, sans des conditions de monotonie, les jeux à champ moyen admettent généralement plusieurs solutions d'équilibre, et que les équations maîtresses correspondantes ne seraient pas bien posées. Dans cet article, nous introduisons un ordre partiel sur l'ensemble des flot de mesures de probabilité pour comparer différentes solutions d'équilibre. Sous cet ordre partiel, les équilibres de champ moyen minimaux et maximaux sont construits, et ils satisfont la propriété du flot. Cependant, les fonctions de valeur correspondantes sont généralement discontinues. Nous introduisons donc une notion de solution de viscosité faible pour l'équation maîtresse et vérifions que les fonctions de valeur sont effectivement des solutions de viscosité faible. De plus, un

---

* Corresponding author.
E-mail addresses: chencmou@cityu.edu.hk (C. Mou), jianfenz@usc.edu (J. Zhang).

https://doi.org/10.1016/j.matpur.2024.02.002

principe de comparaison pour les semi-solutions de viscosité faible est établi, et ainsi ces deux fonctions de valeur servent de solutions de viscosité faible minimale et maximale dans un sens approprié. En particulier, lorsque ces deux fonctions de valeur coïncident, la fonction de valeur devient l'unique solution de viscosité faible de l'équation maîtresse. Restreint à la classe des jeux à champ moyen standard, les résultats demeurent nouveaux.

1. Introduction

In this paper we consider the following extended mean field game system: given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\begin{aligned} &\partial_t \nu(t,x) - \frac{1}{2}\mathrm{tr}\big(\partial_{xx}\nu(t,x)\big) + div\big(\nu(t,x)\widehat{b}(x,\partial_x v(t,x),\nu_t)\big) = 0, \quad \nu_0 = \mu;\\ &\partial_t v(t,x) + \frac{1}{2}\mathrm{tr}(\partial_{xx} v(t,x)) + H(x,\partial_x v(t,x),\nu_t) = 0, \quad v(T,x) = G(x,\nu_T). \end{aligned} \tag{1.1}$$

The master equation, see (2.12) below, is to characterize its decoupling field V in the sense that

$$v(t,x) = V(t,x,\nu_t). \tag{1.2}$$

The standard mean field game and its master equation correspond to the special case:

$$\widehat{b}(x,p,\mu) = \partial_p H(x,p,\mu). \tag{1.3}$$

Initiated independently by Caines-Huang-Malhamé [12] and Lasry-Lions [37], mean field games (MFGs, for short) have received very strong attention and is by now a well-established theory for the study of the asymptotic behavior of stochastic differential games with a large number of players interacting in certain symmetric way. We refer to the monographs Carmona-Delarue [17,18] and the lecture note Cardaliaguet-Porretta [14] for a complete introduction of recent progresses on the subject.

Extended MFGs were first introduced by Lions-Souganidis [40] to study a more general class of MFGs where the vector field governing the flow of the population might be different from that of the individual player at some mean field equilibrium (MFE, for short). Their motivation comes from two folds. Firstly, the homogenization limit of a class of oscillatory classical MFGs is in general not a classical MFG but an extended MFG. Secondly, extended MFGs arise naturally in the optimal transportation-type control problems. More precisely, the Euler-Lagrange systems of optimal transportation-type control problems are in general not of the classical MFG type but of the extended MFG type. A new and meaningful monotonicity condition was proposed in [40] to study the wellposedness of extended MFG systems, and their wellpoedness results were further extended in Munõz [43]. In particular, the proposed monotonicity condition ensures the uniqueness of MFE of extended MFGs.

It should be noted that [40,43] consider extended MFG systems with local coupling, that is, the data $G, H, \widehat{b}$ depend on $\nu(t,x)$, rather than ν_t. We instead study extended MFGs with nonlocal coupling, as in (1.1), via the master equation (2.12). Our motivation for studying such extended MFGs comes from the study of MFGs with a major player. These games consist of a major player and infinite many homogeneous minor players where the major player can have a significant impact on the minor players while all the minor players as a whole can have an impact on the major player. In this case, the value function of the major player will take the form

$$V_0(t, X_t^0, \mathcal{L}_{X_t|\mathcal{F}_t^{X^0}}), \tag{1.4}$$

where X^0 and X stand for the major player's state and the representative minor player's state, respectively. In particular, the measure variable $\mathcal{L}_{X_t|\mathcal{F}_t^{X^0}}$ is not the law of the major player's state X_t^0. This is exactly in the spirit of the extended MFG. The local (in time) wellposedness of the MFG systems for MFGs with a major player has been established in Cardaliaguet-Cirant-Porretta [13]. Its global wellposedness has not been studied in the literature, to the best of our knowledge, and we shall address it in an accompanying paper.

In the literature of standard MFGs, the global wellposedness of master equations requires the uniqueness of MFE, typically under certain monotonicity conditions. See, e.g., Bertucci [6], Bertucci-Cecchin [7], Cardaliaguet-Delarue-Lasry-Lions [15], Cardaliaguet-Souganidis [16], Carmona-Delarue [18], Chassagneux-Crisan-Delarue [20], Lions [38], Mou-Zhang [41], for the well-known Lasry-Lions monotonicity condition; Ahuja [1], Bensoussan-Graber-Yam [2,3], Gangbo-Meszaros [30], Gangbo-Meszaros-Mou-Zhang [31] for the displacement monotonicity condition; and Mou-Zhang [42] for the anti-monotonicity condition. We emphasize that, all these monotonicity conditions require the measure variable to be the law of the state process, and thus fail automatically for value functions in the form (1.4). The works Graber-Meszaros [32,33] proposed a new type of monotonicity condition, which does not have this constraint. We should mention the very recent work Bertucci-Lasry-Lions [10] concerning master equations for extended MFGs with nonlocal coupling as in the present paper. It shows that the master equation admits at most one global solution which is Lipschitz continuous in the measure variable. However, the existence of such a solution requires additional structural conditions and remains open. Moreover, there are studies on master equations for finite state extended MFGs, see e.g. Bertucci [5] and Bertucci-Lasry-Lions [8,9]. We shall investigate the existence of global classical solutions of master equations for extended MFGs in another accompanying paper.

In this paper we focus on extended MFGs and their master equations, with possibly multiple MFEs. Our main idea is to introduce a partial order $\preceq$ for the set of probability measure flows, in the spirit of stochastic dominance. This allows us to compare different MFEs, and we shall construct the minimal/maximal MFE for extended MFGs under this partial order, following the Knaster–Tarski fixed point theorem. To be precise, we shall construct MFEs $\underline{\nu}$ and $\overline{\nu}$ such that:

$$\underline{\nu} \preceq \nu^* \preceq \overline{\nu}, \quad \text{for all MFE } \nu^*. \tag{1.5}$$

For this purpose, we shall assume the data $G, H, \widehat{b}$ are monotone in μ under the partial order $\preceq$. We emphasize that this type of monotonicity under $\preceq$ has a completely different nature from the various monotonicity conditions mentioned in the previous paragraph. Our approach is strongly inspired by Dianetti-Ferrari-Fischer-Nendel [26,27] and Dianetti [25] which obtained (1.5) under the same partial order for standard MFGs. A similar idea has also been applied previously to investigate MFGs of optimal stopping, see Carmona-Delarue-Lacker [19] and Bertucci [4].

We next establish the flow property of the minimal/maximal MFEs, which is crucial for studying the dynamic value function and the master equation. That is, let $\underline{\nu}^{t,\mu}$ denote the minimal MFE for the extended MFG on $[t, T]$ with initial distribution μ. Then, for any $t_0 < t_1$,

$$\underline{\nu}_t^{t_0,\mu} = \underline{\nu}_t^{t_1,\underline{\nu}_{t_1}^{t_0,\mu}}, \quad t \geq t_1. \tag{1.6}$$

This implies the following value function is time consistent:

$$\underline{V}(t_0, x, \mu) = v(t_0, x), \tag{1.7}$$

where v solves the backward PDE in (1.1) with $\nu = \underline{\nu}^{t_0,\mu}$. This function $\underline{V}$ is smooth in x, but is typically discontinuous in (t, μ), as we will see in Section 8 below. So a classical solution theory for the master equation is not viable under our conditions.

We thus turn to weak solutions, by adapting the notion of weak-viscosity solution proposed in our previous paper [41]. We shall show that, by introducing $\overline{V}$ associated to the maximal MFE, both $\underline{V}$ and $\overline{V}$ are weak-viscosity solutions of the master equation (2.12). Moreover, for any weak-viscosity solution V, the spatial derivative $\partial_x V$ always stays between $\partial_x \underline{V}$ and $\partial_x \overline{V}$ component wise. In this sense, $\underline{V}$ and $\overline{V}$ can be viewed as the minimal and maximal weak-viscosity solution of the master equation. In particular, the weak-viscosity solution is unique if and only if $\underline{V} = \overline{V}$. We would like to note that, the very recent work Lions-Seeger [39] has used the same approach to establish the global well-posedness for linear and nonlinear finite dimensional transport equations with coordinate-wise increasing velocity fields, and the theory has also been applied to study MFGs in a finite state space.

We note that our consideration of $\underline{\nu}$ and $\overline{\nu}$ can be viewed as a special selection of MFEs. In the literature there have been other selection criteria for standard MFGs with multiple MFEs, see e.g. Delarure-Foguen Tchuendom [24], Cecchin-Dai Pra-Fisher-Pelino [21], Cecchin-Delaure [22,23]. In [24], three methods of selection, including the minimal cost, zero noise limit, N-player limit selections, are considered for the linear quadratic MFGs. In particular, in this case the master equation is reduced to a one dimensional PDE and the MFE selected by the last two methods provides an entropy solution to this PDE. Similar results have been obtained for two-state MFGs in [21]. In [22,23] the authors established the global wellposedness of master equations for potential MFGs with multiple MFEs. The potential game structure allows to link the MFG to a mean field control problem in the sense that the selected MFE for the MFG is an optimal strategy for the control problem. We would also like to mention that Iseri-Zhang [36] takes a different approach by investigating the set value of MFGs, namely the set of game values over all MFEs, which satisfies the dynamic programming principle. Again our $\underline{V}$ and $\overline{V}$ can be viewed as the minimal and maximal (in terms of $\partial_x V$ instead of V) elements of the set value.

The rest of the paper is organized as follows. In Section 2 we introduce the problem, the main results, and the assumptions. In Section 3 we investigate the backward PDE in (1.1) for given ν. In Section 4 we construct the minimal MFE for the extended MFG. In Section 5 we study the basic properties of the value function $\underline{V}$. In Section 6 we establish the weak-viscosity solution theory. In Section 7 we present the results concerning the maximal MFE and its corresponding value function $\overline{V}$; the results under an alternative set of monotonicity condition under the partial order; as well as the extension of the current results to extended MFGs with a common noise. Finally in Section 8 we solve an example explicitly, which in particular shows that $\underline{V}$ is discontinuous in (t, μ).

ACKNOWLEDGEMENTS. The research of CM was supported in part by Hong Kong RGC Grants ECS 21302521, GRF 11311422 and GRF 11303223. The research of JZ was supported in part by NSF grants DMS-1908665 and DMS-2205972.

2. The setting and the main results

Throughout the paper, we fix a finite time horizon $[0, T]$ and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, on which is defined a d-dimensional Brownian motion B. For any $p \geq 1$, let $\mathcal{P}_p(\mathbb{R}^d)$ denote the set of probability measures on $\mathbb{R}^d$ with finite p-th moment, equipped with the p-Wasserstein distance W_p. We assume $\mathcal{F}_0$ is rich enough to support any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and $\mathcal{F}_t := \mathcal{F}_0 \vee \mathcal{F}_t^B$. For any $p \geq 1$, $\mathcal{G} \subset \mathcal{F}$, and $\mu \in \mathcal{P}_p(\mathbb{R}^d)$, denote by $\mathbb{L}^p(\mathcal{G})$ the set of $\mathcal{G}$-measurable and p-integrable random variables ξ; and $\mathbb{L}^p(\mathcal{G}; \mu)$ the set of those $\xi \in \mathbb{L}^p(\mathcal{G})$ with $\mathcal{L}_\xi = \mu$. For any $t_0 \in [0, T]$, denote $B_t^{t_0} := B_t - B_{t_0}$, $t \in [t_0, T]$, and $\mathbb{F}^{t_0} := \{\mathcal{F}_t\}_{t_0 \leq t \leq T}$. Moreover, we denote $\mathbf{0} := (0, \cdots, 0)$ and $\mathbf{1} := (1, \cdots, 1)$ with appropriate dimensions.

2.1. The extended mean field game

First, given $t_0 \in [0, T]$ and $\nu \in C([t_0, T]; \mathcal{P}_2(\mathbb{R}^d))$, consider the following parabolic PDE on $[t_0, T]$:

$$\begin{aligned}&\partial_t v(\nu;t,x)+\frac{1}{2}\mathrm{tr}(\partial_{xx}v(\nu;t,x))+H(x,\partial_x v(\nu;t,x),\nu_t)=0,\\&v(\nu;T,x)=G(x,\nu_T).\end{aligned}\tag{2.1}$$

Under certain technical conditions on H, G as we will specify later, the above PDE has a unique classical solution $v(\nu;\cdot,\cdot)$. Next, given $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0})$, consider the following SDE on $[t_0, T]$:

$$X_t^{t_0,\xi,\nu}=\xi+\int_{t_0}^{t}\widehat{b}\big(X_s^{t_0,\xi,\nu},\partial_x v(\nu;s,X_s^{t_0,\xi,\nu}),\nu_s\big)ds+B_t^{t_0}.\tag{2.2}$$

It is clear that the mapping $\xi \mapsto \mathcal{L}_{X^{t_0,\xi,\nu}}$ is law invariant. We then define the Nash field Φ for the extended MFG as follows: for any $(t_0,\mu) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0};\mu)$,

$$\Phi(t_0,\mu,\nu):=\{\mathcal{L}_{X_t^{t_0,\xi,\nu}}\}_{t_0\le t\le T},\quad \forall \nu\in C([t_0,T];\mathcal{P}_2(\mathbb{R}^d)).\tag{2.3}$$

Definition 2.1. For any $(t_0,\mu)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, we say $\nu^*\in C([t_0,T];\mathcal{P}_2(\mathbb{R}^d))$ is a mean field equilibrium (MFE) at (t_0,μ) if it is a fixed point of the Nash field $\Phi(t_0,\mu,\cdot)$:

$$\Phi(t_0,\mu,\nu^*)=\nu^*.\tag{2.4}$$

Remark 2.2. (i) The typical case is that H is a Hamiltonian and thus (2.1) is the HJB equation:

$$H(x,p,\mu):=\inf_{a\in\mathbb{R}}h(x,p,\mu,a),\tag{2.5}$$

where

$$h(x,p,\mu,a):=p\cdot b_0(x,a,\mu)+f(x,a,\mu).$$

In this case, as in the standard theory we have a representation formula for v:

$$\begin{aligned}&X_t^{0,\nu;t_0,x,\alpha}=x+\int_{t_0}^{t}b_0(X_s^{0,\nu;t_0,x,\alpha},\alpha(s,X_s^{0,\nu;t_0,x,\alpha}),\nu_s)ds+B_t^{t_0};\\&J(\nu;t_0,x,\alpha):=\mathbb{E}\Big[g(X_T^{0,\nu;t_0,x,\alpha},\nu_T)\\&\qquad\qquad+\int_{t_0}^{T}f(X_s^{0,\nu;t_0,x,\alpha},\alpha(s,X_s^{0,\nu;t_0,x,\alpha}),\nu_s)ds\Big];\\&v(\nu;t_0,x):=\inf_{\alpha\in\mathcal{A}_{t_0}}J(\nu;t_0,x,\alpha)\end{aligned}\tag{2.6}$$

where $\mathcal{A}_{t_0}$ denotes the appropriate set of admissible controls $\alpha:[t_0,T]\times\mathbb{R}^d\to\mathbb{R}^d$.
(ii) In the case in which the Hamiltonian H has a minimizer $a^*=\phi(x,p,\mu)$, namely

$$H(x,p,\mu)=h(x,p,\mu,\phi(x,p,\mu)).\tag{2.7}$$

By (2.5) one can easily check that

$$\begin{aligned}b_0(x,\phi(x,p,\mu),\mu)&=\partial_p H(x,p,\mu),\\f(x,\phi(x,p,\mu),\mu)&=H(x,p,\mu)-p\cdot\partial_p H(x,p,\mu).\end{aligned}\tag{2.8}$$

(iii) Assuming (2.7) holds true, one typical case of $\widehat{b}$ is: for some appropriate function b,

$$\widehat{b}(x,p,\mu) = b(x, \phi(x,p,\mu), \mu).$$

When $b = b_0$ or $\widehat{b}(x,p,\mu) = \partial_p H(x,p,\mu)$, the extended MFG becomes a standard MFG.

2.2. The master equation

When there is a unique MFE for each $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, denoted as $(\alpha^*(t_0,\mu;\cdot), \nu^*(t_0,\mu))$. Then the game problem leads to the following value function:

$$V(t_0,x,\mu) := J(\nu^*(t_0,\mu); t_0, x, \alpha^*(t_0,\mu;\cdot)) \quad \text{for any } x \in \mathbb{R}^d. \tag{2.9}$$

Recall the extended MFG (2.1), (2.2), (2.3), and (2.4). In light of (2.6) and (2.8) we introduce the following FBSDE system (the system does not require the structure in Remark 2.2 (i) though):

$$\begin{aligned} X_t^{0,*} &= x + \int_{t_0}^{t} \partial_p H(X_s^{0,*}, \partial_x V(s, X_s^{0,*}, \nu_s^*), \nu_s^*) ds + B_t^{t_0}; \\ X_t^* &= \xi + \int_{t_0}^{t} \widehat{b}(X_s^*, \partial_x V(s, X_s^*, \nu_s^*), \nu_s^*) ds + B_t^{t_0}; \\ Y_t^* &= G(X_T^{0,*}, \nu_T^*) - \int_t^T Z_s^* dB_s \\ &\quad + \int_t^T \Big[H(\cdot) - \partial_x V(s, X_s^{0,*}, \nu_s^*) \cdot \partial_p H(\cdot) \Big] \Big(X_s^{0,*}, \partial_x V(s, X_s^{0,*}, \nu_s^*), \nu_s^* \Big) ds; \\ &\text{where} \quad \nu_t^* := \mathcal{L}_{X_t^*}. \end{aligned} \tag{2.10}$$

In particular, we have

$$Y_t^* = V\big(t, X_t^{0,*}, \nu_t^*\big) = V\big(t, X_t^{0,*}, \mathcal{L}_{X_t^*}\big). \tag{2.11}$$

By applying the Itô's formula (cf. [11,20]) and comparing it with (2.10), we derive the master equation:

$$\partial_t V + \frac{1}{2}\mathrm{tr}(\partial_{xx} V) + H(x, \partial_x V, \mu) + \mathcal{M}V = 0, \quad V(T,x,\mu) = G(x,\mu), \tag{2.12}$$

where

$$\begin{aligned} \mathcal{M}V(t,x,\mu) := \mathrm{tr}\Big(\int_{\mathbb{R}^d} \Big[&\partial_\mu V(t,x,\mu,\tilde{x}) \widehat{b}^\top(\tilde{x}, \partial_x V(t,\tilde{x},\mu), \mu) \\ &+ \frac{1}{2} \partial_{\tilde{x}} \partial_\mu V(t,x,\mu,\tilde{x}) \Big] \mu(d\tilde{x}) \Big). \end{aligned}$$

Note that we may alternatively view V as the decoupling field of the following FBSDE system:

$$\begin{aligned}
&\mathcal{X}_t^{0,*} = x + B_t^{t_0};\\
&\mathcal{X}_t^* = \xi + \int_{t_0}^t \widehat{b}(X_s^*, \partial_x V(s, X_s^*, \nu_s^*), \nu_s^*)ds + B_t^{t_0}, \quad \text{where} \quad \nu_t^* := \mathcal{L}_{\mathcal{X}_t^*};\\
&\mathcal{Y}_t^* = G(\mathcal{X}_T^{0,*}, \nu_T^*) + \int_t^T H\big(\mathcal{X}_s^{0,*}, \partial_x V(s, \mathcal{X}_s^{0,*}, \nu_s^*), \nu_s^*\big)ds - \int_t^T \mathcal{Z}_s^* dB_s;\\
&\text{in the sense} \quad \mathcal{Y}_t^* = V(t, \mathcal{X}_t^{0,*}, \nu_t^*).
\end{aligned} \tag{2.13}$$

Moreover, V also serves as the decoupling field of the extended MFG system, see (1.1) and (1.2).

The main feature here is that the measure variable ν_t^* in (2.11) is the law of X_t^*, rather than that of $X_t^{0,*}$. Consequently, the $\mathcal{MV}$ above involves the term $\partial_\mu V \widehat{b}^\top$, instead of $\partial_\mu V b_0^\top = \partial_\mu V \partial_p H^\top$ as in the standard master equations. This feature appears naturally in MFG with a major player, which is the main motivation of this paper and will be the subject of an accompanying paper. We also refer to [40] for more applications of extended MFGs.

However, in general there could be multiple MFEs, which lead to multivalued functions. Our goal in this paper is to construct the minimal/maximal MFE and to verify that their value functions satisfy the master equation, in the sense of weak-viscosity solutions introduced in [41].

2.3. The main results

The main results of this paper build on the following partial order $\preceq$ (or alternatively $\succeq$).

Definition 2.3. For a generic dimension n and for $i = 1, 2$,
(i) for any $x^i = (x_1^i, \cdots, x_n^i) \in \mathbb{R}^n$, we say that $x^1 \preceq x^2$ if $x_j^1 \le x_j^2$ for all $j = 1, \cdots, n$;
(ii) for any $\mu_i \in \mathcal{P}_2(\mathbb{R}^n)$, we say that $\mu_1 \preceq \mu_2$ if there exist $\xi^i \in \mathbb{L}^2(\mathcal{F}_0; \mu_i)$ s.t. $\xi^1 \preceq \xi^2$ $\mathbb{P}$-a.s.;
(iii) for any $\nu^i \in C([t_0, T]; \mathcal{P}_2(\mathbb{R}^n))$, we say that $\nu^1 \preceq \nu^2$ if $\nu_t^1 \preceq \nu_t^2$ for all $t \in [t_0, T]$.

We note that $\mu_1 \preceq \mu_2$ is equivalent to the stochastic dominance. We say $x^1 \succeq x^2$ if $x^2 \preceq x^1$, and a function $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is increasing (resp. decreasing) if $\varphi(x^1) \preceq \varphi(x^2)$ whenever $x^1 \preceq$ (resp. $\succeq$) x^2. Similarly we define the monotonicity of functions on $\mathcal{P}_2(\mathbb{R}^d)$ and $C([t_0, T]; \mathcal{P}_2(\mathbb{R}^d))$.

We first have the following simple proposition.

Proposition 2.4. *Assume $\varphi \in \mathcal{C}^1(\mathcal{P}_2(\mathbb{R}^d))$, namely it has a continuous Lions derivative $\partial_\mu \varphi$. Then φ is increasing if and only if $\partial_\mu \varphi(\mu, x) \succeq \mathbf{0}$ for all $(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$.*

Proof. We first prove the if part. Assume $\partial_\mu \varphi \succeq \mathbf{0}$. Let $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ be such that $\mu_1 \preceq \mu_2$, i.e. there exist $\xi^i \in \mathbb{L}^2(\mathcal{F}_0; \mu_i)$, $i = 1, 2$, such that $\xi^1 \preceq \xi^2$ $\mathbb{P}$-a.s. Then

$$\varphi(\mu_2) - \varphi(\mu_1) = \int_0^1 \mathbb{E}\Big[\partial_\mu \varphi\big(\mathcal{L}_{\xi^1 + \theta(\xi^2 - \xi^1)}, \xi^1 + \theta(\xi^2 - \xi^1)\big) \cdot (\xi^2 - \xi^1)\Big] d\theta \ge 0.$$

We next prove the only if part. Assume φ is increasing. For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\xi \in \mathbb{L}^2(\mathcal{F}_0; \mu)$, and $\eta \in \mathbb{L}^2(\mathcal{F}_0)$ such that $\eta \succeq \mathbf{0}$, we have

$$0 \le \lim_{\varepsilon \downarrow 0} \frac{\varphi(\mathcal{L}_{\xi + \varepsilon\eta}) - \varphi(\mu)}{\varepsilon} = \mathbb{E}\Big[\partial_\mu \varphi(\mu, \xi) \cdot \eta\Big].$$

By the arbitrariness of $\eta \succeq \mathbf{0}$, this implies that $\partial_\mu \varphi(\mu, \xi) \succeq \mathbf{0}$, $\mathbb{P}$-a.s. That is $\partial_\mu \varphi(\mu, x) \succeq \mathbf{0}$, for μ-a.e. x. Since $\partial_\mu \varphi$ is continuous, we see that $\partial_\mu \varphi(\mu, x) \succeq \mathbf{0}$ for all (μ, x). □

Remark 2.5. As we saw in [31], a smooth function U on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ satisfies the Lasry-Lions monotonicity condition if and only if: for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\xi \in \mathbb{L}^2(\mathcal{F}_0;\mu)$, $\eta \in \mathbb{L}^2(\mathcal{F}_0)$,

$$\mathbb{E}\Big[\langle \partial_{x\mu} U(\xi,\mu,\tilde{\xi})\eta,\ \tilde{\eta}\rangle\Big] \geq 0. \tag{2.14}$$

We note that (2.14) is always under expectation, while in Proposition 2.4 we require $\partial_\mu \varphi(\mu,x) \succeq \mathbf{0}$ pointwisely. In this sense we are considering pointwise monotonicity in this paper. We shall remark that (2.14) and the pointwise monotonicity of $\partial_x U(x,\cdot)$ do not imply each other.

Our main results consist of two parts, under the conditions specified in the next subsection.

- First, given $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, we will construct the minimal MFE $\underline{\nu}^{t,\mu}$ and the maximal MFE $\overline{\nu}^{t,\mu}$ at (t,μ), in the sense that for any other MFE ν^* at (t,μ) it holds:

  $$\underline{\nu}^{t,\mu} \preceq \nu^* \preceq \overline{\nu}^{t,\mu}.$$

- Next, we define the dynamic value functions

  $$\underline{V}(t,x,\mu) := v(\underline{\nu}^{t,\mu};t,x), \quad \overline{V}(t,x,\mu) := v(\overline{\nu}^{t,\mu};t,x).$$

  We shall show that they are weak-viscosity solutions of the master equation (2.12) such that $\partial_x \underline{V}$ and $\partial_x \overline{V}$ satisfy certain minimal/maximal property.

Since the analyses are similar, in the paper we will focus only on $\underline{\nu}^{t,\mu}$ and $\underline{V}(t,x,\mu)$, and we will present the results concerning $\overline{\nu}^{t,\mu}$ and $\overline{V}(t,x,\mu)$ in Section 7.1 below.

2.4. The assumptions

We first introduce some technical assumptions on the coefficients, which are more or less standard in the literature. Denote, for any $R > 0$,

$$O_R := \{p \in \mathbb{R}^d : |p| < R\}, \quad \forall R > 0. \tag{2.15}$$

Assumption 2.6. (i) $G \in C^0(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ and $H \in C^0(\mathbb{R}^{2d} \times \mathcal{P}_2(\mathbb{R}^d))$ are functions satisfying $G(\cdot,\mu) \in C^2(\mathbb{R}^d)$ and $H(\cdot,\cdot,\mu) \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$ for each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$;
(ii) there exist constants L_0^G, L_0^H, and $L^H(R)$ for each $R > 0$, such that

$$\begin{gathered}
|\partial_x G(x,\mu)|, |\partial_{xx} G(x,\mu)| \leq L_0^G, \quad \forall (x,p,\mu);\\
|\partial_x H(x,p,\mu)| \leq L_0^H[1+|p|], \quad \forall (x,p,\mu);\\
|\partial_p H|, |\partial_{xx} H|, |\partial_{xp} H|, |\partial_{pp} H| \leq L^H(R) \quad \text{on } \mathbb{R}^d \times O_R \times \mathcal{P}_2(\mathbb{R}^d);
\end{gathered}$$

(iii) for each $R > 0$ and any compact set $K \subset \mathcal{P}_2(\mathbb{R}^d)$, $\partial_x G, \partial_{xx} G$ are uniformly continuous in (x,μ) on $\mathbb{R}^d \times K$, and $\partial_x H$, $\partial_p H$, $\partial_{xx} H$, $\partial_{xp} H$, $\partial_{pp} H$ are uniformly continuous in (x,p,μ) on $\mathbb{R}^d \times O_R \times K$.

Assumption 2.7. Assume that $\widehat{b}(\cdot,\cdot,\mu) \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$ for each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and for each $R > 0$ and any compact set $K \subset \mathcal{P}_2(\mathbb{R}^d)$, $\widehat{b}$, $\partial_x \widehat{b}, \partial_p \widehat{b}$ are bounded with bound $L^{\widehat{b}}(R)$ and $\widehat{b}$ is uniformly continuous in μ on $\mathbb{R}^d \times O_R \times K$.

The following pointwise monotonicity condition under partial order $\preceq$ is crucial.

Assumption 2.8. (i) $\partial_x G$ is increasing in (x, μ);
(ii) $\partial_x H$ is increasing in (x, μ), $\partial_p H$ is increasing in (p, μ), and $\partial_{x_i p_j} H \geq 0$ for all $i \neq j$ (which is slightly weaker than that $\partial_p H$ is increasing in x);
(iii) $\widehat{b}$ is increasing in (p, μ) and $\partial_{x_j} \widehat{b}_i \geq 0$ for all $i \neq j$.

Alternatively, we may replace the above assumption with the following monotonicities.

Assumption 2.9. (i) $\partial_x G$ is decreasing in (x, μ);
(ii) $\partial_x H$ is decreasing in (x, μ), $\partial_p H$ is increasing in (p, μ), and $\partial_{x_i p_j} H \geq 0$ for all $i \neq j$;
(iii) $\widehat{b}$ is decreasing in p, increasing in μ, and $\partial_{x_i} \widehat{b}_j \geq 0$ for all $i \neq j$.

In the paper we will focus only on the analyses under Assumption 2.8. The corresponding results under Assumption 2.9 are essentially the same, with obvious changes, so we will present them in Section 7.2 without proofs.

2.5. Some preliminary comparison results

In this subsection we present two well known comparison results for multidimensional SDEs and BSDEs, which will play an important role in the paper. The proofs are rather standard, and we refer to [35] for further discussions on the BSDE case.

Lemma 2.10. *Consider the following two n-dimensional SDE systems: for $k = 1, 2$,*

$$X_t^{k,i} = \xi_k^i + \int_0^t b_k^i(s, X_s^k)ds + B_t^i, \quad i = 1, \cdots, n, \tag{2.16}$$

where $\xi_k^i \in \mathbb{L}^2(\mathcal{F}_0)$ and $b_k^i : [0, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}$ is $\mathbb{F}$-progressively measurable. Assume
(i) for $k = 1, 2$, b_k is uniformly Lipschitz continuous in x and

$$\mathbb{E}[\int_0^T |b_k(t, 0)|^2 dt] < \infty;$$

(ii) b_1^i (or b_2^i) is increasing in x_j for any $i \neq j$, and $\xi_1 \preceq \xi_2$ and $b_1 \preceq b_2$.
Then $X_t^1 \preceq X_t^2$, $0 \leq t \leq T$, $\mathbb{P}$-a.s.

Lemma 2.11. *Consider the following two n-dimensional BSDE systems: for $k = 1, 2$,*

$$Y_t^{k,i} = \xi_k^i + \int_t^T f_k^i(s, Y_s^k, Z_s^{k,i})ds - \int_t^T Z_s^{k,i} \cdot dB_s, \quad i = 1, \cdots, n, \tag{2.17}$$

where $\xi_k^i \in \mathbb{L}^2(\mathcal{F}_T)$ and $f_k^i : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^d$ is $\mathbb{F}$-progressively measurable. Assume
(i) for $k = 1, 2$, f_k is uniformly Lipschitz continuous in (y, z) and

$$\mathbb{E}[\int_0^T |f_k(t, 0, 0)|^2 dt] < \infty;$$

(ii) f_1^i (or f_2^i) is increasing in y_j for any $i \neq j$, and $\xi_1 \preceq \xi_2$ and $f_1 \preceq f_2$.
Then $Y_t^1 \preceq Y_t^2$, $0 \le t \le T$, $\mathbb{P}$-a.s.

3. The PDE (2.1)

In this section we focus on the properties of the solution v for the PDE (2.1). The following lemma is more or less standard. For the sake of completeness, we sketch a proof here. In particular, our probabilistic arguments will remain valid for the common noise case which will be discussed in Section 7.3 below.

Lemma 3.1. *Let Assumption 2.6 hold.*
(i) For any given $\nu \in C([0,T];\mathcal{P}_2(\mathbb{R}^d))$, the equation (2.1) admits a unique classical solution v, and there exist constants $C_1, C_2 > 0$, depending on T, d, L_0^G, L_0^H, and the function L^H, but independent of ν, such that

$$|\partial_x v| \le C_1 \quad \text{and} \quad |\partial_{xx} v| \le C_2; \tag{3.1}$$

(ii) for any compact set $K \subset \mathcal{P}_2(\mathbb{R}^d)$, there exists a modulus of continuity function ρ_K such that: for any $\nu, \nu^1, \nu^2 \in C([0,T];\mathcal{P}_2(\mathbb{R}^d))$ satisfying $\nu_t, \nu_t^1, \nu_t^2 \in K$ for all t,

$$|\partial_x v(\nu^1;t,x) - \partial_x v(\nu^2;t,x)| \le \rho_K\Big(\sup_{t\le s\le T} W_2(\nu_s^1,\nu_s^2)\Big); \tag{3.2}$$

$$|\partial_x v(\nu;t_1,x) - \partial_x v(\nu;t_2,x)| \le \rho_K(t_2-t_1), \quad \forall 0 \le t_1 < t_2 \le T. \tag{3.3}$$

Proof. First it follows from [31, Proposition 6.1] that the following function $v(\nu;t,x)$ satisfies (3.1): denoting $X_s^{t,x} := x + B_s^t$, $t \le s \le T$,

$$\begin{aligned} &v(\nu;t,x) := Y_t^{t,x,\nu}, \quad \text{where} \\ &Y_s^{t,x,\nu} = G(X_T^{t,x},\nu_T) + \int_s^T H(X_r^{t,x}, Z_r^{t,x,\nu}, \nu_r)dr \\ &\qquad - \int_s^T Z_r^{t,x,\nu} \cdot dB_r, \quad t \le s \le T. \end{aligned} \tag{3.4}$$

In particular, we have

$$|Z_s^{t,x,\nu}| = |\partial_x v(\nu;s,X_s^{t,x})| \le C_1. \tag{3.5}$$

We note that the assumptions in the statement of [31, Proposition 6.1] involve the derivatives of G and H with respect to μ as well, but they are never used in that proof.

We next prove (3.2). Fix K and let ρ_K^0 denote the common modulus of continuity function of $\partial_x G, \partial_{xx} G$ on $\mathbb{R}^d \times K$ and that of $\partial_x H, \partial_p H, \partial_{xx} H, \partial_{xp} H$, $\partial_{pp} H$ on $\mathbb{R}^d \times O_{C_1} \times K$ for the C_1 in (3.1) or (3.5). By standard arguments we have

$$\partial_x v(\nu;t,x) = \nabla_x Y_t^{t,x,\nu}, \quad \partial_{xx} v(\nu;t,x) = \nabla_{xx}^2 Y_t^{t,x,\nu}, \tag{3.6}$$

where $\nabla_x Y^{t,x,\nu} \in \mathbb{R}^d$ and $\nabla_{xx}^2 Y^{t,x,\nu} \in \mathbb{R}^{d\times d}$ satisfy the following linear BSDEs on $[t,T]$:

$$\begin{aligned}\nabla_{x_i} Y^{t,x,\nu}_s = \partial_{x_i} G(X^{t,x}_T, \nu_T) - \int_s^T \nabla_{x_i} Z^{t,x,\nu}_r \cdot dB_r \\ + \int_s^T [\partial_{x_i} H + \partial_p H \nabla_{x_i} Z^{t,x,\nu}_r](X^{t,x}_r, \nabla_x Y^{t,x,\nu}_r, \nu_r) dr,\end{aligned} \tag{3.7}$$

$$\begin{aligned}&\nabla_{x_i x_j} Y^{t,x,\nu}_s = \partial_{x_i x_j} G(X^{t,x}_T, \nu_T) - \int_s^T \nabla_{x_i x_j} Z^{t,x,\nu}_r \cdot dB_s \\ &+ \int_s^T \Big[\partial_{x_i x_j} H + \sum_{k=1}^d [\partial_{x_i p_k} H \nabla_{x_j x_k} Y^{t,x,\nu}_r + \partial_{x_j p_k} H \nabla_{x_i x_k} Y^{t,x,\nu}_r] \\ &\qquad + \sum_{k,l=1}^d [\nabla_{x_j x_k} Y^{t,x,\nu}_r \partial_{p_k p_l} H \nabla_{x_i x_l} Y^{t,x,\nu}_r] \\ &\qquad + \partial_p H \nabla_{x_i x_j} Z^{t,x,\nu}_r\Big](X^{t,x}_r, \nabla_x Y^{t,x,\nu}_r, \nu_r) dr.\end{aligned} \tag{3.8}$$

Here we used the fact that $Z^{t,x,\nu}_r = \partial_x v(\nu; r, X^{t,x}_r) = \nabla_x Y^{t,x,\nu}_r$. Recall (3.5) again, then we may rewrite (3.7) as:

$$\begin{aligned}\nabla_{x_i} Y^{t,x,\nu}_s = \partial_{x_i} G(X^{t,x}_T, \nu_T) - \int_s^T \nabla_{x_i} Z^{t,x,\nu}_r \cdot dB_r \\ + \int_s^T \big[\partial_{x_i} H + \partial_p H(-C_1 \vee \nabla_{x_i} Z^{t,x,\nu}_r \wedge C_1)\big](X^{t,x}_r, \nabla_x Y^{t,x,\nu}_r, \nu_r) dr,\end{aligned}$$

where the truncation is in the component wise sense. Note that the generator of the above BSDE is Lipschitz continuous. Then, by the standard BSDE estimates (cf. [46, Chapter 4]) we can easily obtain (3.2). Similarly, we can show that $\partial_x v$ and $\partial_{xx} v$ are uniformly continuous in x, with a possibly different modulus of continuity function ρ.

Moreover, for any $t_1 < t_2$, note that $\nabla_x Y^{t_1,x,\nu}_{t_2} = \partial_x v(\nu; t_2, X^{t_1,x}_{t_2})$ and thus, by (3.7),

$$\begin{aligned}&\partial_x v(\nu; t_1, x) = \nabla_x Y^{t_1,x,\nu}_{t_1} \\ &= \partial_x v(\nu; t_2, X^{t_1,x}_{t_2}) + \int_{t_1}^{t_2} [\partial_x H + \partial_p H \nabla_x Z^{t,x,\nu}_r](X^{t,x}_r, Z^{t,x,\nu}_r, \nu_r) dr \\ &\quad - \int_{t_1}^{t_2} \nabla_x Z^{t,x,\nu}_r \cdot dB_r.\end{aligned}$$

Then, noting that $|\nabla_x Z^{t,x,\nu}_r| = |\partial_{xx} v(\nu, r, X^{t,x}_r)| \le C_2$, one can easily prove (3.3), for a possibly different ρ_K. Similarly $\partial_x v$ and $\partial_{xx} v$ are also uniformly continuous in t. Moreover, since G and H are continuous, by (3.4) one can easily show that v is also continuous in t. Then by (3.4) clearly $v(\nu; \cdot, \cdot)$ is the unique classical solution of (2.1). □

Proposition 3.2. *Under Assumptions 2.6 and 2.8 (i)-(ii), $\partial_x v$ is increasing in (x, ν).*

Proof. First we may rewrite (3.8) as: omitting ${}^{t,x,\nu}$ for notational simplicity,

$$\begin{aligned}\nabla_{x_ix_j}Y_s &= \partial_{x_ix_j}G(X_T,\nu_T) - \int_s^T \nabla_{x_ix_j}Z_r\cdot dB_r\\ &+\int_s^T\Big[f_0\big(r,(\nabla_{x_kx_l}Y_r)_{(k,l)\neq(i,j)}\big)+\Gamma_r\nabla_{x_ix_j}Y_r\\ &\qquad +\partial_pH(X_r,\nabla_xY_r,\nu_r)\nabla_{x_ix_j}Z_r\Big]dr,\end{aligned}\tag{3.9}$$

where

$$\begin{aligned}\Gamma_r &:= \Big[\partial_{x_ip_i}H+\partial_{x_jp_j}H+\sum_{l\neq j}\partial_{p_ip_l}H\nabla_{x_ix_l}Y_r\\ &\qquad +\sum_{k\neq i}\partial_{p_kp_j}H\nabla_{x_jx_k}Y_r\Big](X_r,\nabla_xY_r,\nu_r),\\ f_0\big(r,(y_{k,l})_{(k,l)\neq(i,j)}\big) &:= \Big[\partial_{x_ix_j}H+\sum_{k\neq i}\partial_{x_ip_k}Hy_{jk}+\sum_{k\neq j}\partial_{x_jp_k}Hy_{ik}\\ &\qquad +\sum_{k\neq i,l\neq j}\partial_{p_kp_l}H[(-C_2)\vee y_{j,k}\wedge C_2][(-C_2)\vee y_{i,l}\wedge C_2]\Big](X_r,\nabla_xY_r,\nu_r).\end{aligned}$$

Here the constant C_2 is from (3.1) and we used (3.6). We may view (3.9) as a d^2-dimensional BSDE system, with index (i,j) and solution $\{(\nabla_{x_ix_j}Y,\nabla_{x_ix_j}Z)\}_{(i,j)}$, where Γ is viewed as a given coefficient. We next introduce two d^2-dimensional BSDE systems, again with index (i,j):

$$\begin{aligned}Y_s^{1,(i,j)} &= -\int_s^T Z_r^{1,(i,j)}\cdot dB_r+\int_s^T\Big[\Gamma_rY_r^{1,(i,j)}+\partial_pH(X_r,\nabla_xY_r,\nu_r)Z_r^{1,(i,j)}\Big]dr;\\ Y_s^{2,(i,j)} &= \partial_{x_ix_j}G(X_T,\nu_T)-\int_s^T Z_r^{2,(i,j)}\cdot dB_r\\ &+\int_s^T\Big[f_0\big(r,\{(Y_r^{2,(k,l)})^+\}_{(k,l)\neq(i,j)}\big)+\Gamma_rY_r^{2,(i,j)}+\partial_pH(X_r,\nabla_xY_r,\nu_r)Z_r^{2,(i,j)}\Big]dr.\end{aligned}$$

By Assumption 2.8 (i)-(ii), we have for all (i,j) and $r\in[t,T]$

$$\partial_{x_ix_j}G(X_T,\nu_T)\geq 0,\quad f_0\big(r,\{(y_{k,l})^+\}_{(k,l)\neq(i,j)}\big)\geq 0.$$

Note that f_0 is increasing in $\{(y_{k,l})^+\}_{(k,l)\neq(i,j)}$, and it is obvious that $Y_s^{1,(i,j)}\equiv 0$. Then it follows from Lemma 2.11 that $Y_s^{2,(i,j)}\geq Y_s^{1,(i,j)}=0$, and thus

$$f_0\big(r,\{(Y_r^{2,(k,l)})^+\}_{(k,l)\neq(i,j)}\big)=f_0\big(r,\{Y_r^{2,(k,l)}\}_{(k,l)\neq(i,j)}\big).$$

This implies that $\big\{Y^{2,(i,j)},Z^{2,(i,j)}\big\}_{(i,j)}$ satisfies BSDE system (3.9). Then $\partial_{x_ix_j}v(\nu;t,x)=\nabla_{x_ix_j}Y_t=Y_t^{2,(i,j)}\geq 0$. That is, ∂_xv is increasing in x.

Similarly, given $\nu^1,\nu^2\in C([0,T];\mathcal{P}_2(\mathbb{R}^d))$ such that $\nu^1\preceq\nu^2$, omit ^{t,x} and denote, for $\theta\in[0,1]$,

$$\begin{aligned}\bar\nabla_{x_i}Y_s &:= \nabla_{x_i}Y_s^{\nu^2}-\nabla_{x_i}Y_s^{\nu^1},\quad \bar\nabla_{x_i}Z_s:=\nabla_{x_i}Z_s^{\nu^2}-\nabla_{x_i}Z_s^{\nu^1},\\ \nabla_xY_s^\theta &:= (1-\theta)\nabla_xY_s^{\nu^2}+\theta\nabla_xY_s^{\nu^1}.\end{aligned}$$

Note that $\nabla_{x_i} Z = \big(\nabla_{x_i x_1} Y, \cdots, \nabla_{x_i x_d} Y\big)^\top$. By (3.7) we have

$$\begin{aligned}\bar\nabla_{x_i} Y_s &= [\partial_{x_i} G(X_T, \nu_T^2) - \partial_{x_i} G(X_T, \nu_T^1)] - \int_s^T \bar\nabla_{x_i} Z_r \cdot dB_r \\ &\quad + \int_s^T \Big[\bar\gamma_r + \bar f_0\big(r, \{\bar\nabla_{x_j} Y_r\}_{j\neq i}\big) + \bar\Gamma_r \bar\nabla_{x_i} Y_r \\ &\qquad + \partial_p H(X_r, \nabla_x Y_r^{\nu^1}, \nu_r^1) \bar\nabla_{x_i} Z_r\Big] dr,\end{aligned} \tag{3.10}$$

where

$$\begin{aligned}\bar\Gamma_r &:= \int_0^1 \Big[\partial_{x_i p_i} H + \sum_{k=1}^d \partial_{p_i p_k} H \nabla_{x_i x_k} Y_r^{\nu^2}\Big](X_r, \nabla_x Y^\theta, \nu^1) d\theta, \\ \bar f_0\big(r, \{y_j\}_{j\neq i}\big) &:= \sum_{j\neq i} \int_0^1 \Big[\partial_{x_i p_j} H + \sum_{k=1}^d \partial_{p_j p_k} H \nabla_{x_i x_k} Y_r^{\nu^2}\Big](X_r, \nabla_x Y^\theta, \nu^1) d\theta \; y_j \\ \bar\gamma_r &:= [\partial_{x_i} H(X_r, \nabla_x Y_r^{\nu^2}, \nu_r^2) - \partial_{x_i} H(X_r, \nabla_x Y_r^{\nu^2}, \nu_r^1)] \\ &\quad + \sum_{k=1}^d [\partial_{p_k} H(X_r, \nabla_x Y_r^{\nu^2}, \nu_r^2) - \partial_{p_k} H(X_r, \nabla_x Y_r^{\nu^2}, \nu_r^1)] \nabla_{x_i x_k} Y_r^{\nu^2}.\end{aligned}$$

Note that $\nabla_{x_i x_k} Y_r^{\nu^2} = \partial_{x_i x_k} v(\nu^2; r, X_r) \geq 0$. Then, by Assumption 2.8 (i)-(ii) we see that $\bar f_0$ is increasing in $\{y_j\}_{j\neq i}$ and, for all i and $r \in [t, T]$,

$$[\partial_{x_i} G(X_T, \nu_T^2) - \partial_{x_i} G(X_T, \nu_T^1)] \geq 0, \quad \bar\gamma_r \geq 0.$$

Now compare (3.10) with the following d-dimensional linear BSDE system:

$$\begin{aligned}\bar Y_s^i &= \int_s^T \Big[\bar f_0\big(r, \{Y_r^j\}_{j\neq i}\big) + \bar\Gamma_r \bar Y_r^i + \partial_p H(X_r, \nabla_x Y_r^{\nu^1}, \nu_r^1) \bar Z_r^i\Big] dr \\ &\quad - \int_s^T \bar Z_r^i \cdot dB_r.\end{aligned} \tag{3.11}$$

It follows from Lemma 2.11 again that $\bar\nabla_{x_i} Y_s \geq \bar Y_s^i$ for all i. From (3.11) it is obvious that $\bar Y_s^i \equiv 0$. Then

$$\partial_{x_i} v(\nu^2; t, x) - \partial_{x_i} v(\nu^1; t, x) = \bar\nabla_{x_i} Y_t = \bar Y_s^i \geq 0.$$

That is, $\partial_x v$ is increasing in ν. □

4. The minimal MFE

In this section we construct the minimal MFE for the extended MFG. We first establish the pointwise monotonicity of the Nash field Φ.

Theorem 4.1. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then for any $t_0 \in [0, T]$, $\Phi(t_0, \cdot, \cdot)$ is increasing in (μ, ν).*

Proof. Let $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\nu^1, \nu^2 \in C([t_0, T]; \mathcal{P}_2(\mathbb{R}^d))$ be such that $\mu_1 \preceq \mu_2$, $\nu^1 \preceq \nu^2$, and $\xi_1 \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu_1)$, $\xi_2 \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu_2)$ be such that $\xi_1 \le \xi_2$. For $k = 1, 2$, we have

$$X_t^{t_0,\xi_k,\nu^k} = \xi_k + \int_{t_0}^{t} \widehat{b}(X_s^{t_0,\xi_k,\nu^k}, \partial_x v(\nu^k; s, X_s^{t_0,\xi_k,\nu^k}), \nu_s^k) ds + B_t^{t_0}.$$

Denote $b_k(s,x) := \widehat{b}(x, \partial_x v(\nu^k; s, x), \nu_s^k)$, $k = 1, 2$. By Lemma 3.1 b_k satisfies Lemma 2.10 (i). Moreover, by Assumption 2.8 (iii) and Proposition 3.2 we see that $b_1 \preceq b_2$ and

$$\partial_{x_j} b_k^i(s,x) = \Big[\partial_{x_j} \widehat{b}^i + \partial_p \widehat{b}^i \cdot \partial_{x_j x} v\Big](x, \partial_x v(\nu^k; s, x), \nu_s^k) \ge 0, \quad i \neq j.$$

Since $\xi_1 \preceq \xi_2$, then by Lemma 2.10 we have $X_t^{t_0,\xi_1,\nu^1} \preceq X_t^{t_0,\xi_2,\nu^2}$, $t_0 \le t \le T$, $\mathbb{P}$-a.s. This implies that $\Phi(t_0, \mu_1, \nu^1) \preceq \Phi(t_0, \mu_2, \nu^2)$. $\square$

We now construct the minimal MFE by Picard iteration, following the standard procedure in Knaster-Tarski fixed point theorem. Fix $(t_0, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu)$. Recall Assumption 2.7 and (3.1), we set

$$\underline{X}_t^{t_0,\xi,0} := \xi - L^{\widehat{b}}(C_1)\mathbf{1} + B_t^{t_0}, \quad \overline{X}_t^{t_0,\xi,0} := \xi + L^{\widehat{b}}(C_1)\mathbf{1} + B_t^{t_0}, \tag{4.1}$$

and, for $n = 0, \cdots$,

$$\underline{X}_t^{t_0,\xi,n+1} = \xi + \int_{t_0}^{t} \widehat{b}(\underline{X}_s^{t_0,\xi,n+1}, \partial_x v(\mathcal{L}_{\underline{X}^{t_0,\xi,n}}; s, \underline{X}_s^{t_0,\xi,n+1}), \mathcal{L}_{\underline{X}_s^{t_0,\xi,n}}) ds + B_t^{t_0}. \tag{4.2}$$

We then have the first main result of the paper.

Theorem 4.2. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then for any* $(t_0, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ *and* $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu)$, *there exists a process* $\underline{X}^{t_0,\xi}$ *on* $[t_0, T]$ *such that*
(i) $\underline{X}_t^{t_0,\xi,n} \preceq \underline{X}_t^{t_0,\xi,n+1}$, $\forall n, t$, $\mathbb{P}$*-a.s. with*

$$\lim_{n\to\infty} \mathbb{E}[\sup_{t_0 \le t \le T} |\underline{X}_t^{t_0,\xi,n} - \underline{X}_t^{t_0,\xi}|^2] = 0;$$

(ii) $\underline{\nu}^{t_0,\mu} := \mathcal{L}_{\underline{X}^{t_0,\xi}}$ *is an MFE of the extended MFG at* (t_0, μ)*;*
(iii) for any MFE ν^* *of the extended MFG at* (t_0, μ)*, we have* $\underline{\nu}^{t_0,\mu} \preceq \nu^*$*. That is,* $\underline{\nu}^{t_0,\mu}$ *is the minimal MFE.*

Proof. For notational simplicity we omit ${}^{t_0,\xi}$ and ${}^{t_0,\mu}$.

First, by Assumption 2.7 and (3.1),

$$\widehat{b}(\underline{X}_s^1, \partial_x v(\mathcal{L}_{\underline{X}^0}; s, \underline{X}_s^1), \mathcal{L}_{\underline{X}_s^0}) \succeq -L^{\widehat{b}}(C_1)\mathbf{1}.$$

Then $\underline{X}_t^0 \preceq \underline{X}_t^1$, $t_0 \le t \le T$, $\mathbb{P}$-a.s. and thus $\mathcal{L}_{\underline{X}^0} \preceq \mathcal{L}_{\underline{X}^1}$. Applying Theorem 4.1 repeatedly, we see that $\underline{X}^n$ is increasing in n, and thus we may define $\underline{X} := \lim_{n\to\infty} \underline{X}^n$. Moreover, following similar arguments one can easily see that $\underline{X}_t^n \preceq \overline{X}_t^0$, $t_0 \le t \le T$, $\mathbb{P}$-a.s. for all n. Then it follows from the dominated convergence theorem that $\lim_{n\to\infty} \mathbb{E}[|\underline{X}_t^n - \underline{X}_t|^2] = 0$, for any t.

Next, by Assumption 2.7 and (3.1) we see that $\widehat{b}(\cdot, \partial_x v(\cdot), \cdot)$ is bounded by $L^{\widehat{b}}(C_1)$. Then it follows from [45, Lemma 4.1] that the set $\cup_{n\geq 1}\{\mathcal{L}_{\underline{X}_t^n}\}_{0\leq t\leq T}$ is precompact. Now send $n\to\infty$ in (4.2), by the desired continuity of $\widehat{b}$ in Assumption 2.7 and that of $\partial_x v$ in Lemma 3.1, we have

$$\underline{X}_t = \xi + \int_{t_0}^{t} \widehat{b}(\underline{X}_s, \partial_x v(\mathcal{L}_{\underline{X}}; s, \underline{X}_s), \mathcal{L}_{\underline{X}_s})ds + B_t^{t_0}. \tag{4.3}$$

This implies that $\underline{\nu} := \mathcal{L}_{\underline{X}}$ is an MFE of the extended MFG at (t_0, μ). Moreover, compare this with (4.2), one can easily see that $\lim_{n\to\infty}\mathbb{E}[\sup_{t_0\leq t\leq T}|\underline{X}_t^n - \underline{X}_t|^2] = 0$.

Finally, for any MFE ν^* of the extended MFG at (t_0,μ), consider the related SDE system:

$$X_t^* = \xi + \int_{t_0}^{t} \widehat{b}(X_s^*, \partial_x v(\nu^*; s, X_s^*), \nu_s^*)ds + B_t^{t_0}. \tag{4.4}$$

Since ν^* is an MFE, we have $\nu^* = \mathcal{L}_{X^*}$. Again since $\widehat{b}(X_s^*, \partial_x v(\nu^*; s, X_s^*), \nu_s^*) \succeq -L^{\widehat{b}}(C_1)\mathbf{1}$, we have $\underline{X}_t^0 \preceq X_t^*$, $t_0 \leq t \leq T$, $\mathbb{P}$-a.s. Applying Theorem 4.1 repeatedly, we see that $\underline{X}_t^n \preceq X_t^*$, $t_0\leq t\leq T$, $\mathbb{P}$-a.s. for all n. Then $\underline{X}_t \preceq X_t^*$, $t_0 \leq t \leq T$, $\mathbb{P}$-a.s. and thus $\underline{\nu} \preceq \nu^*$. □

We conclude this section with the following crucial flow property.

Proposition 4.3. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then, for any* $(t_0,\mu) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$,

$$\underline{\nu}_t^{t_0,\mu} = \underline{\nu}_t^{t_1, \underline{\nu}_{t_1}^{t_0,\mu}}, \quad \textit{for all } t_0 \leq t_1 \leq t \leq T. \tag{4.5}$$

Proof. Let $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu)$. Then $\underline{\nu}_t^{t_0,\mu} = \mathcal{L}_{\underline{X}_t^{t_0,\xi}}$, $\forall t\geq t_0$, where $\underline{X}^{t_0,\xi}$ satisfies (4.3). Note that

$$\underline{X}_t^{t_0,\xi} = \underline{X}_{t_1}^{t_0,\xi} + \int_{t_1}^{t} \widehat{b}(\underline{X}_s^{t_0,\xi}, \partial_x v(\mathcal{L}_{\underline{X}^{t_0,\xi}}; s, \underline{X}_s^{t_0,\xi}), \mathcal{L}_{\underline{X}_s^{t_0,\xi}})ds + B_t^{t_1}, \quad t\geq t_1.$$

We see that $\underline{\nu}^{t_0,\mu}$ is an MFE of the extended MFG at $(t_1, \mathcal{L}_{\underline{X}_{t_1}^{t_0,\xi}}) = (t_1, \underline{\nu}_{t_1}^{t_0,\mu})$. Then by Theorem 4.2 (iii) we have $\underline{\nu}_t^{t_1,\underline{\nu}_{t_1}^{t_0,\mu}} \preceq \underline{\nu}_t^{t_0,\mu}$, for all $t\geq t_1$.

On the other hand, for the Picard iteration in (4.1) and (4.2), by Theorem 4.2 (i) we have $\underline{X}_{t_1}^{t_0,\xi,n} \preceq \underline{X}_{t_1}^{t_0,\xi} =: \xi_1$, for all n. By (4.1) it is clear that $\underline{X}_t^{t_0,\xi,0} \preceq \underline{X}_t^{t_1,\xi_1,0}$ for all $t\geq t_1$. Note that

$$\underline{X}_t^{t_0,\xi,1} = \underline{X}_{t_1}^{t_0,\xi,0} + \int_{t_1}^{t} \widehat{b}(\underline{X}_s^{t_0,\xi,1}, \partial_x v(\mathcal{L}_{\underline{X}^{t_0,\xi,0}}; s, \underline{X}_s^{t_0,\xi,1}), \mathcal{L}_{\underline{X}_s^{t_0,\xi,0}})ds + B_t^{t_1}.$$

Since $\underline{X}_{t_1}^{t_0,\xi,1} \preceq \xi_1$, by Theorem 4.1 we see that $\underline{X}_t^{t_0,\xi,1} \preceq \underline{X}_t^{t_1,\xi_1,1}$, $t\geq t_1$, $\mathbb{P}$-a.s. Repeat the arguments, we obtain $\underline{X}_t^{t_0,\xi,n} \preceq \underline{X}_t^{t_1,\xi_1,n}$. Send $n\to\infty$, by Theorem 4.2 (i) we have $\underline{X}_t^{t_0,\xi} \preceq \underline{X}_t^{t_1,\xi_1}$, $t\geq t_1$, $\mathbb{P}$-a.s. That is, $\underline{\nu}_t^{t_0,\mu} \preceq \underline{\nu}_t^{t_1,\underline{\nu}_{t_1}^{t_0,\mu}}$, for all $t\geq t_1$. Then we must have the equality. □

5. The corresponding value function

In this section we investigate the dynamic value function corresponding to the minimal MFE:

$$\underline{V}(t,x,\mu) := v(\underline{\nu}^{t,\mu}; t, x). \tag{5.1}$$

The following properties are immediate.

Proposition 5.1. *Let Assumptions 2.6, 2.7, and 2.8 hold.*
(i) For any $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$, $\underline{V}(t, \cdot, \mu) \in C^2(\mathbb{R}^d)$ *with* $|\partial_x \underline{V}| \le C_1$ *and* $|\partial_{xx} \underline{V}| \le C_2$ *for the* C_1, C_2 *in (3.1);*
(ii) for any $t \in [0, T]$, $\partial_x \underline{V}(t, \cdot, \cdot)$ *is increasing in* (x, μ).

Proof. (i) is a direct consequence of Lemma 3.1 (i).

(ii) Assume $x_1 \preceq x_2$, $\mu_1 \preceq \mu_2$ and let $\xi_i \in \mathbb{L}(\mathcal{F}_{t_0}, \mu_i)$, $i = 1, 2$, be such that $\xi_1 \preceq \xi_2$. Then $\underline{X}_t^{t_0,\xi_1,0} \preceq \underline{X}_t^{t_0,\xi_2,0}$ for all $t_0 \le t \le T$. Apply Theorem 4.1 repeatedly, we have $\underline{X}_t^{t_0,\xi_1,n} \preceq \underline{X}_t^{t_0,\xi_2,n}$, $t_0 \le t \le T$, $\mathbb{P}$-a.s. for all n. Then $\underline{X}_t^{t_0,\xi_1} \preceq \underline{X}_t^{t_0,\xi_2}$, $t_0 \le t \le T$, $\mathbb{P}$-a.s. and hence $\underline{\nu}^{t_0,\mu_1} \preceq \underline{\nu}^{t_0,\mu_2}$. Since $\partial_x \underline{V}(t, x, \mu) = \partial_x v(\underline{\nu}^{t,\mu}; t, x)$, then it follows from Proposition 3.2 that $\partial_x \underline{V}(t, x_1, \mu_1) \preceq \partial_x \underline{V}(t, x_2, \mu_2)$. □

However, as we will see in Section 8 below, in general $\underline{V}$ is discontinuous in (t, μ). At below we show that $\partial_x \underline{V}$ is lower semi-continuous in μ in the following sense.

Definition 5.2. (i) Let $\mu_n, \mu \in \mathcal{P}_2(\mathbb{R}^d)$, $n \ge 1$. We say that $\mu_n \uparrow \mu$ (resp. $\mu_n \downarrow \mu$) if $\mu_n \preceq$ (resp. $\succeq$) μ_{n+1} for all n and $\lim_{n\to\infty} W_2(\mu_n, \mu) = 0$;
(ii) we say a function $U : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ is lower semi-continuous (resp. upper semi-continuous) if $\liminf\limits_{n\to\infty} U(\mu_n) \succeq U(\mu)$ (resp. $\limsup\limits_{n\to\infty} U(\mu_n) \preceq U(\mu)$) whenever $\lim\limits_{n\to\infty} W_2(\mu_n, \mu) = 0$.

Here lim inf and lim sup are taken component wise. We then have the semi-continuity of $\underline{V}$ in (t, μ).

Proposition 5.3. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then*
(i) for any $(t_k, \mu_k) \to (t, \mu)$, *we have* $\liminf_{n\to\infty} \partial_x \underline{V}(t_k, x, \mu_k) \succeq \partial_x \underline{V}(t, x, \mu)$, *i.e.* $\partial_x \underline{V}$ *is lower semi-continuous in* (t, μ). *Moreover, if* $\mu_k \uparrow \mu$, *then* $\lim_{k\to\infty} \partial_x \underline{V}(t, x, \mu_k) = \partial_x \underline{V}(t, x, \mu)$;
(ii) for any $x \in \mathbb{R}^d$ *and* $\nu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$, *the mapping* $t \mapsto \partial_x \underline{V}(t, x, \nu_t)$ *is lower semi-continuous, and in particular it is Borel measurable.*

Proof. (i) Fix x and let $(t_k, \mu_k) \to (t, \mu)$, with $\xi_k \in \mathbb{L}^2(\mathcal{F}_{t_k}; \mu_k)$, $\xi \in \mathbb{L}^2(\mathcal{F}_t; \mu)$. Denote $\varepsilon_k := |t_k - t| + W_2(\mu_k, \mu)$ and $\hat{t}_k := t_k \vee t$. Then, by Proposition 3.2 and (3.3) we have

$$\begin{aligned}\partial_x \underline{V}(t_k, x, \mu_k) &= \partial_x v(\underline{\nu}^{t_k,\mu_k}; t_k, x) \succeq \partial_x v(\mathcal{L}_{\underline{X}^{t_k,\xi_k,n}}; t_k, x)\\ &\succeq \partial_x v(\mathcal{L}_{\underline{X}^{t_k,\xi_k,n}_{[\hat{t}_k,T]}}; \hat{t}_k, x) - \rho(\varepsilon_k).\end{aligned}$$

Recall (4.1) and (4.2). It is clear that $\sup_{\hat{t}_k \le s \le T} W_2(\mathcal{L}_{\underline{X}_s^{t_k,\xi_k,0}}, \mathcal{L}_{\underline{X}_s^{t,\xi,0}}) \le \varepsilon_k + \sqrt{\varepsilon_k}$. Similarly to the arguments in Theorem 4.2, we may utilize the locally uniform regularity in Assumption 2.7 with $R = C_1$ and with appropriate compact set K. Then, by Lemma 3.1 and stability of SDEs, one can easily show that there exists a modulus of continuity function ρ_1 such that $\sup_{\hat{t}_k \le s \le T} W_2(\mathcal{L}_{\underline{X}_s^{t_k,\xi_k,1}}, \mathcal{L}_{\underline{X}_s^{t,\xi,1}}) \le \rho_1(\varepsilon_k)$. Moreover, by Lemma 3.1 and (4.2) again, we can show by induction on n that there exists a modulus of continuity function ρ_n such that $\sup_{\hat{t}_k \le s \le T} W_2(\mathcal{L}_{\underline{X}_s^{t_k,\xi_k,n}}, \mathcal{L}_{\underline{X}_s^{t,\xi,n}}) \le \rho_n(\varepsilon_k)$. Then, by (3.2) and (3.3) we have, for each n, k,

$$\begin{aligned}\partial_x \underline{V}(t_k, x, \mu_k) &\succeq \partial_x v(\mathcal{L}_{\underline{X}^{t,\xi,n}_{[\hat{t}_k,T]}}; \hat{t}_k, x) - \rho(\rho_n(\varepsilon_k)) - \rho(\varepsilon_k)\\ &\succeq \partial_x v(\mathcal{L}_{\underline{X}^{t,\xi,n}}; t, x) - \rho(\rho_n(\varepsilon_k)) - 2\rho(\varepsilon_k).\end{aligned}$$

Send $k \to \infty$, we have $\liminf_{k\to\infty} \partial_x \underline{V}(t_k, x, \mu_k) \succeq \partial_x v(\mathcal{L}_{\underline{X}^{t,\xi,n}}; t, x)$. Now send $n \to \infty$, by (3.2) again we have

$$\liminf_{k\to\infty} \partial_x \underline{V}(t_k, x, \mu_k) \succeq \partial_x v(\mathcal{L}_{\underline{X}^{t,\xi}}; t, x) = \partial_x \underline{V}(t, x, \mu).$$

Moreover, if $\mu_k \uparrow \mu$, by Proposition 5.1 we have $\partial_x \underline{V}(t, x, \mu_k) \preceq \partial_x \underline{V}(t, x, \mu)$, then the above inequality implies $\lim_{k\to\infty} \partial_x \underline{V}(t, x, \mu_k) = \partial_x \underline{V}(t, x, \mu)$.

(ii) For $t_k \to t$, since $\nu_{t_k} \to \nu_t$, then $\liminf_{k\to\infty} \partial_x \underline{V}(t_k, x, \nu_{t_k}) \succeq \partial_x \underline{V}(t, x, \nu_t)$. This proves the claimed lower semi-continuity, which implies further the Borel measurability. □

Definition 5.4. Let $\mathcal{C}^2$ denote the set of functions $V : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ satisfying:
(i) $V(t, \cdot, \mu) \in C^2(\mathbb{R}^d)$ for each (t, μ), and $\partial_x V, \partial_{xx} V$ are uniformly bounded;
(ii) for any $x \in \mathbb{R}^d$ and $\nu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$, the mapping $t \mapsto \partial_x V(t, x, \nu_t)$ is Borel measurable.

Then it is clear that $\underline{V} \in \mathcal{C}^2$. The following lemma will be important in the next section.

Lemma 5.5. *Let Assumptions 2.7 and 2.8 (iii) hold and $V \in \mathcal{C}^2$. Assume further that $\partial_x V$ is increasing in μ and lower or upper semi-continuous in μ. Then, for any $(t_0, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0}; \mu)$, the following McKean-Vlasov SDE has a strong solution:*

$$X_t^{t_0,\xi} = \xi + \int_{t_0}^{t} \widehat{b}\big(X_s^{t_0,\xi}, \partial_x V(s, X_s^{t_0,\xi}, \mathcal{L}_{X_s^{t_0,\xi}}), \mathcal{L}_{X_s^{t_0,\xi}}\big) ds + B_t^{t_0}. \tag{5.2}$$

Equivalently, the following Fokker-Planck equation has a weak solution $\nu(t, x)$:

$$\begin{aligned} &\partial_t \nu(t, x) - \frac{1}{2}\mathrm{tr}(\partial_{xx}\nu(t, x)) + div(\nu(t, x)\widehat{b}(x, \partial_x V(t, x, \nu_t), \nu_t)) = 0, \\ &\nu_{t_0} = \mu. \end{aligned} \tag{5.3}$$

Proof. We shall only prove the case that $\partial_x V$ is lower semi-continuous in μ. The upper semi-continuous case can be proved similarly, in the same spirit as we construct the maximal MFE in Subsection 7.1 below.

Recall (4.1) and (4.2). Denote $X^{t_0,\xi,0} := \underline{X}^{t_0,\xi,0}$, with possibly a larger C_1 which is an upper bound of $|\partial_x V|$, and for $n = 0, 1, \cdots$,

$$X_t^{t_0,\xi,n+1} = \xi + \int_{t_0}^{t} \widehat{b}(X_s^{t_0,\xi,n+1}, \partial_x V(s, X_s^{t_0,\xi,n+1}, \mathcal{L}_{X_s^{t_0,\xi,n}}), \mathcal{L}_{X_s^{t_0,\xi,n}}) ds + B_t^{t_0}. \tag{5.4}$$

Since $\partial_x V$ is increasing in μ and by Assumption 2.8 (iii), it is clear that $X^{t_0,\xi,n}$ is increasing in n, and $X_t^{t_0,\xi,n} \le \overline{X}_t^{t_0,\xi,0}$ for all $t \in [t_0, T]$. Then there exists $X^{t_0,\xi}$ such that $\lim_{n\to\infty} \sup_{t_0 \le t \le T} \mathbb{E}[|X_t^{t_0,\xi,n} - X_t^{t_0,\xi}|^2] = 0$. Note that, since $\partial_x V$ is increasing and lower semi-continuous in μ, and $\mathcal{L}_{X^{t_0,\xi,n}} \uparrow \mathcal{L}_{X^{t_0,\xi}}$, as in Proposition 5.3 (i) we have $\lim_{n\to\infty} \partial_x V(t, x, \mathcal{L}_{X_t^{t_0,\xi,n}}) = \partial_x V(t, x, \mathcal{L}_{X_t^{t_0,\xi}})$. Then by sending $n \to \infty$ in (5.4) we see that $X^{t_0,\xi}$ satisfies (5.2). □

6. Weak-viscosity solutions to the master equation

6.1. Viscosity solution to PDE system

Differentiate (2.1) formally in x, we obtain the following system of PDEs: for $i = 1, \cdots, d$,

$$\begin{aligned} &\partial_t u^i(t, x) + \frac{1}{2}\mathrm{tr}(\partial_{xx} u^i(t, x)) + \partial_{x_i} H(x, u(t, x), \nu_t) \\ &\qquad + \partial_p H(x, u(t, x), \nu_t) \cdot \partial_x u^i(t, x) = 0. \end{aligned} \tag{6.1}$$

Definition 6.1. Fix $\nu \in C([0,T] \times \mathcal{P}_2(\mathbb{R}^d))$ and consider $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that both u and $\partial_x u$ are bounded. We say that u is a viscosity subsolution (resp. supersolution, solution) of the PDE system (6.1) if, for each i and for given $u^{-i} := (u^1, \cdots, u^{i-1}, u^{i+1}, \cdots, u^d)$, the function u^i is a viscosity subsolution (resp. supersolution, solution) to the PDE (6.1) for fixed i in the standard sense.

Lemma 6.2. *Let Assumption 2.6 hold true. Fix $\nu \in C([0,T]; \mathcal{P}_2(\mathbb{R}^d))$ and let $v(\nu; \cdot, \cdot)$ be the unique classical solution of the PDE (2.1). Then $u(t,x) := \partial_x v(\nu; t, x)$ is a viscosity solution to the PDE system (6.1).*

Proof. Recall (3.6) and (3.7). Note that $\nabla_x Y_s^{t,x,\nu} = u(s, X_s^{t,x})$. Then, for fixed i, (3.7) becomes:

$$\nabla_{x_i} Y_s^{t,x,\nu} = \partial_{x_i} G(X_T^{t,x}, \nu_T) - \int_s^T \nabla_{x_i} Z_r^{t,x,\nu} \cdot dB_r$$
$$+ \int_s^T [\partial_{x_i} H + \partial_p H \nabla_{x_i} Z_r^{t,x,\nu}](X_r^{t,x}, u^{-i}(r, X_r^{t,x}), \nabla_{x_i} Y_r^{t,x,\nu}, \nu_r) dr.$$

Then by the standard BSDE theory we see that $u^i(t,x) = \nabla_{x_i} Y_t^{t,x,\nu}$ is a viscosity solution to the PDE (6.1) for each fixed i. $\square$

The next comparison principle is more or less standard, see e.g. [34] in slightly different contexts. We nevertheless sketch a proof for completeness.

Lemma 6.3. *Let Assumptions 2.6 and 2.8 (i)-(ii) hold true, and fix $\nu \in C([0,T]; \mathcal{P}_2(\mathbb{R}^d))$. Let u be as in Lemma 6.2, and $\underline{u}$ and $\overline{u}$ be a viscosity subsolution and a viscosity supersolution, respectively, to the PDE system (6.1) in the sense of Definition 6.1. If $\underline{u}(T,x) \preceq \partial_x G(x, \nu_T) \preceq \overline{u}(T,x)$ for all $x \in \mathbb{R}^d$, then $\underline{u} \preceq u \preceq \overline{u}$ on $[0,T] \times \mathbb{R}^d$.*

Proof. We shall prove only $\underline{u} \preceq u$. The inequality $u \preceq \overline{u}$ can be proved similarly.

Fix (t,x) and denote $X_s := x + B_s^t$. For a possibly larger C_1 such that $|\underline{u}| \le C_1$, recall (3.7) and introduce the following linear BSDEs recursively: $\nabla_i Y^0 := C_1$, and for $n \ge 0$,

$$\nabla_i Y_s^{n+1} = \partial_{x_i} G(X_T, \nu_T) - \int_s^T \nabla_i Z_r^{n+1} \cdot dB_r$$
$$+ \int_s^T \Big[\partial_{x_i} H(X_r, \nabla^{-i} Y_r^n, \nabla_i Y_r^{n+1}, \nu_r) + \partial_p H(X_r, \nabla Y_r^n, \nu_r) \cdot \nabla_i Z_r^{n+1}\Big] dr. \tag{6.2}$$

That is, $\nabla Y_s^{n+1} = u_{n+1}(s, X_s)$, where $u_0^i \equiv C_1$, and for $n \ge 0$ and for given u_n, the function u_{n+1}^i is the unique viscosity solution to the following PDE:

$$\begin{aligned} &\partial_t u_{n+1}^i(t,x) + \frac{1}{2}\mathrm{tr}(\partial_{xx} u_{n+1}^i(t,x)) + \partial_{x_i} H(x, u_n^{-i}(t,x), u_{n+1}^i(t,x), \nu_t) \\ &+ \partial_p H(x, u_n(t,x), \nu_t) \cdot \partial_x u_{n+1}^i(t,x) = 0, \quad u_{n+1}^i(T,x) = \partial_{x_i} G(x, \nu_T). \end{aligned} \tag{6.3}$$

Recall (3.7). One can easily show that $\lim_{n\to\infty} \sup_{t \le s \le T} \mathbb{E}[|\nabla Y_s^n - \nabla_x Y_s^{t,x,\nu}|^2] = 0$, and thus $\lim_{n\to\infty} u_n = u$. Moreover, similar (actually easier) to the proof of Proposition 3.2, we can prove by induction on n that u_n is increasing in x for all n. We claim that

$$\underline{u} \preceq u_n, \quad \text{for all } n. \tag{6.4}$$

Then, by sending $n \to \infty$, we obtain $\underline{u} \preceq u$.

To see (6.4), first, since $u_0^i \equiv C_1 \geq \underline{u}^i$, it holds true for $n = 0$. Assume it holds true for n, and we shall verify it for $n + 1$. By Assumption 2.8 (ii) and $\partial_x u_{n+1}^i \succeq \mathbf{0}$, we see that

$$\begin{aligned}&\partial_{x_i} H(x, u_n^{-i}(t,x), u_{n+1}^i(t,x), \nu_t) + \partial_p H(x, u_n(t,x), \nu_t) \cdot \partial_x u_{n+1}^i(t,x)\\ &\geq \partial_{x_i} H(x, \underline{u}^{-i}(t,x), u_{n+1}^i(t,x), \nu_t) + \partial_p H(x, \underline{u}(t,x), \nu_t) \cdot \partial_x u_{n+1}^i(t,x).\end{aligned}$$

Then u_{n+1}^i is a viscosity supersolution of the following PDE:

$$\begin{aligned}&\partial_t u_{n+1}^i(t,x) + \frac{1}{2}\mathrm{tr}(\partial_{xx} u_{n+1}^i(t,x)) + \partial_{x_i} H(x, \underline{u}^{-i}(t,x), u_{n+1}^i(t,x), \nu_t)\\ &+\partial_p H(x, \underline{u}(t,x), \nu_t) \cdot \partial_x u_{n+1}^i(t,x) \leq 0, \quad u_{n+1}^i(T,x) = \partial_{x_i} G(x, \nu_T).\end{aligned} \tag{6.5}$$

Notice that $\underline{u}^i$ is a viscosity subsolution of the above PDE. Then by the standard comparison principle we obtain $\underline{u}^i \leq u_{n+1}^i$. This proves (6.4) for $n + 1$, and hence $\underline{u} \preceq u$. □

6.2. Weak-viscosity solutions to the master equation

We now introduce a notion of weak-viscosity solution to the master equation (2.12), adapted from [41]. Recall Definition 5.4.

Definition 6.4. We say that $V \in \mathcal{C}^2$ is a weak-viscosity subsolution (resp. supersolution, solution) of the master equation (2.12) if, for any $(t_0, \mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, the Fokker-Planck equation (5.3) has a weak solution ν such that the function $u(t,x) := \partial_x V(t,x,\nu_t)$ is a viscosity subsolution (resp. supersolution, solution) to the PDE system (6.1) on $[t_0,T]$ in the sense of Definition 6.1 and satisfies $u(T,x) \preceq$ (resp. $\succeq$, $=$)$\partial_x G(x,\nu_T)$.

We first have the following simple result.

Proposition 6.5. *Let Assumptions 2.6, 2.7, and 2.8 (i)-(ii) hold. Assume $V \in \mathcal{C}^2$ is a weak-viscosity solution of the master equation (2.12). Then, for any $(t_0, \mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, the ν in Definition 6.4 is an MFE of the extended MFG at (t_0, μ).*

Proof. First by Lemma 3.1 let $v(\nu;\cdot,\cdot)$ be the classical solution of the PDE (2.1). Then by Lemma 6.2 $\tilde{u} := \partial_x v(\nu;\cdot,\cdot)$ is a viscosity solution of the PDE system (6.1) in the sense of Definition 6.1 with $\tilde{u}(T,x) = \partial_x G(x,\nu_T)$. Now by Definition 6.4 and the comparison principle in Lemma 6.3, we have $\partial_x v(\nu;t,x) = \partial_x V(t,x,\nu_t)$. This identifies (5.3) and (2.2) with $\nu_t = \mathcal{L}_{X_t^{t_0,\xi,\nu}}$, except that one is in PDE form while the other is in SDE form. Thus $\nu = \Phi(t_0,\mu,\nu)$, namely ν is an MFE at (t_0,μ). □

Remark 6.6. Alternatively, we may call $V \in \mathcal{C}^2$ a weak-viscosity solution of the master equation (2.12) if it is both a weak-viscosity subsolution and a weak-viscosity supersolution of (2.12), where the weak-viscosity subsolution and supersolution are defined in Definition 6.4. Under this alternative definition, we may use one ν for the subsolution property and another different ν (and hence a different u) for the supersolution property. So this is weaker than Definition 6.4, in particular, a weak-viscosity solution in this alternative sense does not necessarily provide an MFE as in Proposition 6.5.

Our second main result of the paper is the following.

Theorem 6.7. *Let Assumptions 2.6, 2.7, and 2.8 hold.*
(i) $\underline{V}$ is a weak-viscosity solution to the master equation (2.12);
(ii) for any weak-viscosity supersolution V to the master equation (2.12), we have

$$\partial_x \underline{V} \;\preceq\; \partial_x V. \tag{6.6}$$

Proof. (i) Fix $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$. By Theorem 4.2 and in particular (4.3) we see that $\underline{\nu}^{t_0,\mu}$ is a weak solution to the Fokker-Planck equation (5.3) with $V = \underline{V}$. Moreover, by Proposition 4.3 we have

$$\underline{u}(t,x) := \partial_x \underline{V}(t,x,\underline{\nu}^{t_0,\mu}_t) = \partial_x v(\underline{\nu}^{t,\underline{\nu}^{t_0,\mu}_t}; t,x) = \partial_x v(\underline{\nu}^{t_0,\mu}; t,x).$$

Then by Lemma 6.2 $\underline{u}$ is a viscosity solution to the PDE system (6.1) with $\nu_t = \underline{\nu}^{t_0,\mu}_t$. Moreover, $\underline{u}(T,x) = \partial_x G(x, \underline{\nu}^{t_0,\mu}_T)$. Therefore, $\underline{V}$ is a weak-viscosity solution to the master equation (2.12).

(ii) Let V be an arbitrary weak-viscosity supersolution to the master equation (2.12). For any $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, let ν, u be as in Definition 6.4. Then, for any $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0};\mu)$, the McKean-Vlasov SDE (5.2) has a strong solution $X^{t_0,\xi}$ with $\nu = \mathcal{L}_{X^{t_0,\xi}}$. Recall (4.1) and (4.2). It is clear that $\underline{X}^{t_0,\xi,0}_t \preceq X^{t_0,\xi}_t$ for all $t \in [t_0,T]$. Denote $\underline{\nu}^{t_0,\mu,0} := \mathcal{L}_{\underline{X}^{t_0,\xi,0}} \preceq \nu$. Note that $\partial_x v(\underline{\nu}^{t_0,\mu,0}; \cdot,\cdot)$ is a viscosity solution to the PDE system (6.1) with $\underline{\nu}^{t_0,\mu,0}$ and by Proposition 3.2 $\partial_x v$ is increasing in x. Then by Assumption 2.8 (ii) one can easily see that $\partial_x v(\underline{\nu}^{t_0,\mu,0}; \cdot,\cdot)$ is a viscosity subsolution to the PDE system (6.1) with ν. Moreover, by Assumption 2.8 (i),

$$\partial_x v(\underline{\nu}^{t_0,\mu,0}; T,x) = \partial_x G(x, \underline{\nu}^{t_0,\mu,0}_T) \preceq \partial_x G(x,\nu_T) \preceq u(T,x).$$

Since u is a viscosity supersolution of this system, then by the comparison principle Lemma 6.3, we have $\partial_x v(\underline{\nu}^{t_0,\mu,0}; t,x) \preceq u(t,x) = \partial_x V(t,x,\nu_t)$ for all (t,x). Denote

$$\underline{b}(t,x) := \widehat{b}(x, \partial_x v(\underline{\nu}^{t_0,\mu,0}; t,x), \underline{\nu}^{t_0,\mu,0}), \quad b(t,x) := \widehat{b}(x, \partial_x V(t,x,\nu_t); \nu_t).$$

By Assumption 2.8 (iii) one can easily see that $\underline{b} \preceq b$, and $\partial_{x_j} \underline{b}^i \geq 0$ for all $i \neq j$. Then, comparing (4.2) and (5.2), it follows from Lemma 2.10 that $\underline{X}^{t_0,\xi,1}_t \preceq X^{t_0,\xi}_t$, $t_0 \leq t \leq T$, $\mathbb{P}$-a.s. Repeat the arguments we can show that $\underline{X}^{t_0,\xi,n}_t \preceq X^{t_0,\xi}_t$, $t_0 \leq t \leq T$, $\mathbb{P}$-a.s. and $\partial_x v(\mathcal{L}_{\underline{X}^{t_0,\xi,n}}; t,x) \preceq u(t,x)$ for all n. Send $n \to \infty$, by Theorem 4.2 and Lemma 3.1 (ii) we see that $\underline{X}^{t_0,\xi}_t \preceq X^{t_0,\xi}_t$, $t_0 \leq t \leq T$, $\mathbb{P}$-a.s. and $\partial_x v(\underline{\nu}^{t_0,\mu}; t,x) \preceq u(t,x)$. Therefore, $\partial_x \underline{V}(t_0,x,\mu) = \partial_x v(\underline{\nu}^{t_0,\mu}; t_0,x) \preceq u(t_0,x) = \partial_x V(t_0,x,\mu)$. Since (t_0,x,μ) is arbitrary, we conclude the proof. □

7. Some extensions

7.1. The maximal case

Similarly to Section 4, we can construct the maximal MFE as follows. Fix $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0};\mu)$. Let $\overline{X}^{t_0,\xi,0}$ be defined by (4.1), and for $n \geq 0$,

$$\overline{X}^{t_0,\xi,n+1}_t = \xi + \int_{t_0}^{t} \widehat{b}(\overline{X}^{t_0,\xi,n+1}_s, \partial_x v(\mathcal{L}_{\overline{X}^{t_0,\xi,n}}; s, \overline{X}^{t_0,\xi,n+1}_s), \mathcal{L}_{\overline{X}^{t_0,\xi,n}_s}) ds + B^{t_0}_t. \tag{7.1}$$

Then, as in Theorem 4.2 and Proposition 4.3, we have the following results.

Theorem 7.1. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then for any $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0};\mu)$, there exists a process $\overline{X}^{t_0,\xi}$ on $[t_0,T]$ such that*
(i) $\overline{X}^{t_0,\xi,n+1}_t \preceq \overline{X}^{t_0,\xi,n}_t$, $\forall n,t$, $\mathbb{P}$-a.s. with

$$\lim_{n\to\infty} \mathbb{E}[\sup_{t_0 \leq t \leq T} |\overline{X}^{t_0,\xi,n}_t - \overline{X}^{t_0,\xi}_t|^2] = 0;$$

(ii) $\overline{\nu}^{t_0,\mu} := \mathcal{L}_{\overline{X}^{t_0,\xi}}$ *is an MFE of the extended MFG at* (t_0,μ) *and satisfies the flow property:*

$$\overline{\nu}^{t_0,\mu}_t = \overline{\nu}_t^{t_1,\overline{\nu}^{t_0,\mu}_{t_1}}, \quad \text{for all } t_0 < t_1 \le t \le T; \tag{7.2}$$

(iii) for any MFE ν^* *of the extended MFG at* (t_0,μ)*, we have* $\overline{\nu}^{t_0,\mu} \succeq \nu^*$*. That is,* $\overline{\nu}^{t_0,\mu}$ *is the maximal MFE.*

We next define

$$\overline{V}(t,x,\mu) := v(\overline{\nu}^{t,\mu}; t,x). \tag{7.3}$$

Theorem 7.2. *Let Assumptions 2.6, 2.7, and 2.8 hold.*
(i) $\overline{V} \in \mathcal{C}^2$, $\partial_x \overline{V}$ *is increasing in* (x,μ) *and upper semi-continuous in* (t,μ)*. Moreover, if* $\mu_k \downarrow \mu$*, then* $\lim_{k\to\infty} \partial_x \overline{V}(t,x,\mu_k) = \partial_x \overline{V}(t,x,\mu)$*;*
(ii) $\overline{V}$ *is a weak-viscosity solution to the master equation (2.12);*
(iii) for any weak-viscosity subsolution V *to the master equation (2.12), we have*

$$\partial_x V \preceq \partial_x \overline{V}. \tag{7.4}$$

The following result is an immediate consequence of Theorems 6.7 and 7.2.

Corollary 7.3. *Let Assumptions 2.6, 2.7, and 2.8 hold. If* $\underline{V} = \overline{V}$ *on* $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$*, then the master equation (2.12) admits a unique weak-viscosity solution* $V := \underline{V} = \overline{V}$*.*

7.2. The decreasing case

In this subsection we replace Assumption 2.8 with Assumption 2.9.

Theorem 7.4. *Let Assumptions 2.6, 2.7, and 2.9 hold true.*
(i) $\partial_x v$ *is decreasing in* (x,ν)*, and* Φ *in increasing in* (μ,ν)*;*
(ii) for any $(t_0,\mu)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$*, there exist MFEs* $\underline{\nu}^{t_0,\mu}$ *and* $\overline{\nu}^{t_0,\mu}$ *of the extended MFG at* (t_0,μ) *such that, for any other MFE* ν^* *of the extended MFG at* (t_0,μ)*, we have* $\underline{\nu}^{t_0,\mu} \preceq \nu^* \preceq \overline{\nu}^{t_0,\mu}$*;*
(iii) the minimal MFE $\underline{\nu}^{t_0,\mu}$ *and the maximal MFE* $\overline{\nu}^{t_0,\mu}$ *satisfy the flow property (4.5) and (7.2).*

Again we define the value functions:

$$\underline{V}(t,x,\mu) := v(\underline{\nu}^{t,\mu}; t,x), \quad \overline{V}(t,x,\mu) := v(\overline{\nu}^{t,\mu}; t,x). \tag{7.5}$$

Theorem 7.5. *Let Assumptions 2.6, 2.7, and 2.9 hold.*
(i) $\underline{V}, \overline{V} \in \mathcal{C}^2$, $\partial_x \underline{V}$ *is decreasing in* (x,μ) *and upper semi-continuous in* (t,μ)*, and* $\partial_x \overline{V}$ *is decreasing in* (x,μ) *and lower semi-continuous in* (t,μ)*;*
(ii) $\underline{V}$, $\overline{V}$ *are weak-viscosity solutions to the master equation (2.12);*
(iii) for any weak-viscosity subsolution V_1 *and weak-viscosity supersolution* V_2 *to the master equation (2.12), we have*

$$\partial_x \underline{V} \succeq \partial_x V_1, \qquad \partial_x \overline{V} \preceq \partial_x V_2. \tag{7.6}$$

(iv) If $\underline{V} = \overline{V}$ *on* $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$*, then the master equation (2.12) admits a unique weak-viscosity solution* $V := \underline{V} = \overline{V}$*.*

7.3. The common noise case

In this subsection we study the extended mean field game with a common noise. We shall only consider the problem under Assumption 2.8. The case under Assumption 2.9 is similar.

Let B^0 be the common noise which is independent of $\mathbb{F}$, $\beta \geq 0$ be a constant and $\hat{\beta}^2 := 1 + \beta^2$. For any $t_0 \in [0,T]$, denote $B_t^{0,t_0} := B_t^0 - B_{t_0}^0$, $t \in [t_0,T]$ and $\mathbb{F}^{0,t_0} := \{\mathcal{F}_t^{B^{0,t_0}}\}_{t_0 \leq t \leq T}$. Let $C(\mathbb{F}^{0,t_0};\mathcal{P}_2(\mathbb{R}^d))$ denote the set of stochastic measure flow $\nu : [t_0,T] \times \Omega \to \mathcal{P}_2(\mathbb{R}^d)$ which is $\mathbb{F}^{0,t_0}$-progressively measurable and continuous in t. Given any $\nu \in C(\mathbb{F}^{0,t_0};\mathcal{P}_2(\mathbb{R}^d))$, consider the following backward stochastic PDE on $[t_0,T]$:

$$\begin{aligned} dv(\nu;t,x) &= z(\nu;t,x) \cdot dB_t^0 \\ &\quad - \Big[\mathrm{tr}\big(\frac{\hat{\beta}^2}{2}\partial_{xx} v(\nu;t,x) + \beta \partial_x z^\top(\nu;t,x)\big) + H(x, \partial_x v(\nu;t,x), \nu_t)\Big] dt, \\ v(\nu;T,x) &= G(x,\nu_T), \end{aligned} \tag{7.7}$$

where the solution pair (v,z) is $\mathbb{F}^{0,t_0}$-progressively measurable. Given $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0})$, we still use $X^{t_0,\xi,\nu}$ to denote the strong solution to the following SDE on $[t_0,T]$:

$$X_t^{t_0,\xi,\nu} = \xi + \int_{t_0}^{t} \widehat{b}\big(X_s^{t_0,\xi,\nu}, \partial_x v(\nu;s,X_s^{t_0,\xi,\nu}), \nu_s\big) ds + B_t^{t_0} + \beta B_t^{0,t_0}. \tag{7.8}$$

Introduce the Nash field Φ on $C(\mathbb{F}^{0,t_0};\mathcal{P}_2(\mathbb{R}^d))$: for any $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2(\mathcal{F}_{t_0};\mu)$,

$$\Phi(t_0,\mu,\nu) := \{\mathcal{L}_{X_t^{t_0,\xi,\nu}|\mathcal{F}_t^{0,t_0}}\}_{t_0 \leq t \leq T}, \quad \forall \nu \in C(\mathbb{F}^{0,t_0};\mathcal{P}_2(\mathbb{R}^d)). \tag{7.9}$$

Fix (t_0,μ), define MFE as a fixed point of $\Phi(t_0,\mu,\cdot)$. Then the corresponding master equation becomes second order:

$$\partial_t V + \frac{1}{2}\mathrm{tr}(\partial_{xx} V) + H(x, \partial_x V, \mu) + \mathcal{M}V = 0, \quad V(T,x,\mu) = G(x,\mu), \tag{7.10}$$

where

$$\begin{aligned} \mathcal{M}V(t,x,\mu) := \mathrm{tr}\Big(\int_{\mathbb{R}^d} \Big[&\frac{\hat{\beta}^2}{2}\partial_{\tilde{x}}\partial_\mu V(t,x,\mu,\tilde{x}) + \beta^2 \partial_x \partial_\mu V(t,x,\mu,\tilde{x}) \\ &+ \partial_\mu V(t,x,\mu,\tilde{x}) \widehat{b}^\top(\tilde{x}, \partial_x V(t,\tilde{x},\mu), \mu) + \frac{\beta^2}{2}\int_{\mathbb{R}^d} \partial_{\mu\mu} V(t,x,\mu,\bar{x},\tilde{x})\mu(d\bar{x})\Big]\mu(d\tilde{x})\Big). \end{aligned}$$

Theorem 7.6. *Let Assumptions 2.6, 2.7, and 2.8 hold true.*
(i) $\partial_x v$ is increasing in (x,ν), and Φ in increasing in (μ,ν);
(ii) for any $(t_0,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, there exist MFEs $\underline{\nu}^{t_0,\mu}$ and $\overline{\nu}^{t_0,\mu}$ of the extended MFG at (t_0,μ) such that $\underline{\nu}^{t_0,\mu} \preceq \nu^* \preceq \overline{\nu}^{t_0,\mu}$ for all other MFE ν^* of the extended MFG at (t_0,μ);
(iii) the minimal MFE $\underline{\nu}^{t_0,\mu}$ and the maximal MFE $\overline{\nu}^{t_0,\mu}$ satisfy the flow property (4.5) and (7.2), respectively, $\mathbb{P}$-a.s.

Here, for any $\nu^i \in C(\mathbb{F}^{0,t_0};\mathcal{P}_2(\mathbb{R}^n))$, $i = 1,2$, the partial order $\nu^1 \preceq \nu^2$ is extended naturally: $\nu_t^1 \preceq \nu_t^2$ for all $t \in [t_0,T]$, a.s. The monotonicity of $\partial_x v$ in (x,ν) is also in obvious sense.

Define the value functions corresponding to the minimal and maximal MFEs respectively:

$$\underline{V}(t,x,\mu) := v(\underline{\nu}^{t,\mu};t,x), \quad \overline{V}(t,x,\mu) := v(\overline{\nu}^{t,\mu};t,x). \tag{7.11}$$

We note that $\underline{V}$ and $\overline{V}$ are $\mathcal{F}^{0,t}_t$-measurable and hence are actually deterministic.

Theorem 7.7. *Let Assumptions 2.6, 2.7, and 2.8 hold. Then* $\underline{V}, \overline{V} \in \mathcal{C}^2$, $\partial_x \underline{V}$ *is increasing in* (x, μ) *and lower semi-continuous in* (t, μ), *and* $\partial_x \overline{V}$ *is increasing in* (x, μ) *and upper semi-continuous in* (t, μ).

We may continue to study weak-viscosity solution of the master equation (7.10) as in Section 6. In this case the PDE (6.1) becomes a backward SPDE (7.7), which can be viewed as a path dependent PDE, see e.g. Zhang [46, Chapter 11]. However, in this case $\underline{u}(t, x, \omega) := \partial_x V(t, x, \underline{\nu}^{t_0,\mu}_t(\omega))$ is in general discontinuous in (t, ω), thus the viscosity theory for path dependent PDEs in Ekren-Touzi-Zhang [28,29] and Zhou [47] cannot be applied here. One possibility is to adapt the viscosity solution for backward SPDEs in Qiu [44], which does not require the regularity in ω. On the other hand, we note that the value function (1.4) for the MFG with a major player will have the same regularity issue, even when there is no common noise. So we shall leave the systematic investigation of this issue to a future research.

8. An example

In this section we solve an example completely. In particular, we shall show that $\underline{V}$ is in general discontinuous in (t, μ). Set $d = 1$ and denote

$$m(\mu) := \int_{\mathbb{R}} x\mu(dx).$$

Consider the example:

$$G(x, \mu) := xm(\mu), \quad H(x, p, \mu) := \frac{p^2}{2},$$
$$\widehat{b}(x, p, \mu) := \widehat{b}(p) := \begin{cases} -2, & p < -2; \\ 2p + \frac{1}{2}p^2, & -2 \le p < 0; \\ 2p - \frac{1}{2}p^2, & 0 \le p < 2; \\ 2, & p \ge 2. \end{cases}$$

One can easily verify that Assumptions 2.6, 2.7, and 2.8 hold true. Moreover, (2.1) becomes

$$\partial_t v(\nu; t, x) + \frac{1}{2}\mathrm{tr}(\partial_{xx} v(\nu; t, x)) + \frac{|\partial_x v(\nu; t, x)|^2}{2} = 0,$$
$$v(\nu; T, x) = G(x, \nu_T) = xm(\nu_T).$$

It admits a unique solution:

$$v(\nu; t, x) = xm(\nu_T) + \frac{1}{2}(T - t)|m(\nu_T)|^2. \tag{8.1}$$

Then $\partial_x v(\nu; t, x) = m(\nu_T)$ and thus (2.2) becomes:

$$X^{t_0,\xi,\nu}_t = \xi + \widehat{b}(m(\nu_T))(t - t_0) + B^{t_0}_t. \tag{8.2}$$

Note that Φ depends on ν only through $m(\nu_T)$. Introduce the following operator:

$$\widehat{\Phi}(t_0, \mu, p) := \mathbb{E}\Big[\xi + \widehat{b}(p)(T - t_0) + B^{t_0}_T\Big] = m(\mu) + \widehat{b}(p)(T - t_0), \ p \in \mathbb{R}. \tag{8.3}$$

One can easily see that ν^* is an MFE at (t_0, μ) if and only if $p^* := m(\nu_T^*)$ is a fixed point of $\widehat{\Phi}$:

$$p^* = \widehat{\Phi}(t_0, \mu, p^*) = m(\mu) + \widehat{b}(p^*)(T - t_0), \tag{8.4}$$

or equivalently,

$$\widehat{b}(p^*) = \frac{p^* - m(\mu)}{T - t_0}.$$

Then, by (8.2) and (8.1), the corresponding MFE and value are:

$$X_t^{t_0,\xi,p^*} = \xi + \widehat{b}(p^*)(t - t_0) + B_t^{t_0}, \quad v(p^*; t, x) = xp^* + \frac{1}{2}(T - t)|p^*|^2. \tag{8.5}$$

Note that one side of (8.4) is piecewise quadratic, and the other side is linear. By elementary calculation we solve (8.4) in four cases. Denote

$$\begin{aligned} \lambda := \frac{1}{T - t_0}, \quad m_1 := 2 - \frac{2}{\lambda}, \quad m_2 := \frac{\lambda}{2} + \frac{2}{\lambda} - 2; \\ \phi_-(\lambda, m) := \sqrt{(\lambda - 2)^2 - 2\lambda m}, \quad \phi_+(\lambda, m) := \sqrt{(\lambda - 2)^2 + 2\lambda m} \end{aligned} \tag{8.6}$$

Case 1. $\lambda \geq 2$. In this case $m_1 > 0$, and there is a unique fixed point p^*:

$$p^* = \begin{cases} m(\mu) - \frac{2}{\lambda}, \quad \text{if } m(\mu) < -m_1; \\ \lambda - 2 - \phi_-(\lambda, m(\mu)), \quad \text{if } -m_1 \leq m(\mu) < 0; \\ 2 - \lambda + \phi_+(\lambda, m(\mu)), \quad \text{if } 0 \leq m(\mu) < m_1; \\ m(\mu) + \frac{2}{\lambda}, \quad \text{if } m(\mu) \geq m_1. \end{cases} \tag{8.7}$$

Case 2. $4 - 2\sqrt{2} < \lambda < 2$. In this case $0 < m_2 < m_1$. We solve the problem in three subcases.
Case 2.1. $|m(\mu)| > m_2$. In this case there is a unique fixed point:

$$p^* = \begin{cases} m(\mu) - \frac{2}{\lambda}, \quad \text{if } m(\mu) < -m_1; \\ \lambda - 2 - \phi_-(\lambda, m(\mu)), \quad \text{if } -m_1 \leq m(\mu) < -m_2; \\ 2 - \lambda + \phi_+(\lambda, m(\mu)), \quad \text{if } m_2 < m(\mu) < m_1; \\ m(\mu) + \frac{2}{\lambda}, \quad \text{if } m(\mu) \geq m_1. \end{cases} \tag{8.8}$$

Case 2.2. $|m(\mu)| = m_2$. In this case there are two fixed points:

$$\begin{aligned} p^* = \lambda - 2 - \phi_-(\lambda, m(\mu)) \quad \text{or} \quad p^* = 2 - \lambda, \quad \text{if} \quad m(\mu) = -m_2; \\ p^* = \lambda - 2 \quad \text{or} \quad p^* = 2 - \lambda + \phi_+(\lambda, m(\mu)), \quad \text{if} \quad m(\mu) = m_2. \end{aligned} \tag{8.9}$$

Case 2.3. $|m(\mu)| < m_2$. In this case there are three fixed points:

$$\begin{aligned} \begin{matrix} p^* = \lambda - 2 - \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = 2 - \lambda \pm \phi_+(\lambda, m(\mu)) \end{matrix}, \quad \text{if} \quad -m_2 < m(\mu) \leq 0; \\ \begin{matrix} p^* = \lambda - 2 \pm \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = 2 - \lambda + \phi_+(\lambda, m(\mu)) \end{matrix}, \quad \text{if} \quad 0 < m(\mu) < m_2. \end{aligned} \tag{8.10}$$

Case 3. $1 < \lambda \leq 4 - 2\sqrt{2}$. In this case $0 < m_1 \leq m_2$. We solve the problem in three subcases.
Case 3.1. $|m(\mu)| > m_2$. In this case there is a unique fixed point:

$$p^* = \begin{cases} m(\mu) - \frac{2}{\lambda}, & \text{if } m(\mu) < -m_2; \\ m(\mu) + \frac{2}{\lambda}, & \text{if } m(\mu) > m_2. \end{cases} \tag{8.11}$$

Case 3.2. $|m(\mu)| = m_2$. In this case there are two fixed points:

$$\begin{array}{c} p^* = m(\mu) - \dfrac{2}{\lambda} \quad \text{or} \quad p^* = 2 - \lambda, \quad \text{if} \quad m(\mu) = -m_2; \\ p^* = \lambda - 2 \quad \text{or} \quad p^* = m(\mu) + \dfrac{2}{\lambda}, \quad \text{if} \quad m(\mu) = m_2. \end{array} \tag{8.12}$$

Case 3.3. $|m(\mu)| < m_2$. In this case there are three fixed points:

$$\begin{array}{ccl} \begin{array}{c} p^* = m(\mu) - \dfrac{2}{\lambda} \\ \text{or} \quad p^* = 2 - \lambda \pm \phi_+(\lambda, m(\mu)) \end{array}, & \text{if} & -m_2 < m(\mu) \leq -m_1; \\ \begin{array}{c} p^* = \lambda - 2 - \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = 2 - \lambda \pm \phi_+(\lambda, m(\mu)) \end{array}, & \text{if} & -m_1 < m(\mu) \leq 0; \\ \begin{array}{c} p^* = \lambda - 2 \pm \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = 2 - \lambda + \phi_+(\lambda, m(\mu)) \end{array}, & \text{if} & 0 < m(\mu) \leq m_1; \\ \begin{array}{c} p^* = \lambda - 2 \pm \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = m(\mu) + \frac{2}{\lambda} \end{array}, & \text{if} & m_1 < m(\mu) < m_2; \end{array} \tag{8.13}$$

Case 4. $0 < \lambda \leq 1$. In this case $0 \leq -m_1 < m_2$. We solve the problem in three subcases.
Case 4.1. $|m(\mu)| > m_2$. In this case there is a unique fixed point:

$$p^* = \begin{cases} m(\mu) - \frac{2}{\lambda}, & \text{if } m(\mu) < -m_2; \\ m(\mu) + \frac{2}{\lambda}, & \text{if } m(\mu) > m_2. \end{cases} \tag{8.14}$$

Case 4.2. $|m(\mu)| = m_2$. In this case there are two fixed points:

$$\begin{array}{c} p^* = m(\mu) - \dfrac{2}{\lambda} \quad \text{or} \quad p^* = 2 - \lambda, \quad \text{if} \quad m(\mu) = -m_2; \\ p^* = \lambda - 2 \quad \text{or} \quad p^* = m(\mu) + \dfrac{2}{\lambda}, \quad \text{if} \quad m(\mu) = m_2. \end{array} \tag{8.15}$$

Case 4.3. $|m(\mu)| < m_2$. In this case there are three fixed points:

$$\begin{array}{ccl} \begin{array}{c} p^* = m(\mu) - \dfrac{2}{\lambda} \\ \text{or} \quad p^* = 2 - \lambda \pm \phi_+(\lambda, m(\mu)) \end{array}, & \text{if} & -m_2 < m(\mu) \leq m_1; \\ \begin{array}{c} p^* = m(\mu) \pm \dfrac{2}{\lambda} \\ \text{or} \quad p^* = 2 - \lambda - \phi_+(\lambda, m(\mu)) \end{array}, & \text{if} & m_1 \leq m(\mu) < 0; \\ \begin{array}{c} p^* = m(\mu) \pm \dfrac{2}{\lambda} \\ \text{or} \quad p^* = \lambda - 2 + \phi_-(\lambda, m(\mu)) \end{array}, & \text{if} & 0 \leq m(\mu) \leq -m_1; \\ \begin{array}{c} p^* = \lambda - 2 \pm \phi_-(\lambda, m(\mu)) \\ \text{or} \quad p^* = m(\mu) + \frac{2}{\lambda} \end{array}, & \text{if} & -m_1 < m(\mu) < m_2. \end{array} \tag{8.16}$$

Put all the cases together, we find that the minimal p^*, denoted as $\underline{p}^{t_0,\mu}$, is:

$$\underline{p}^{t_0,\mu} := \begin{cases} m(\mu) - \dfrac{2}{\lambda}, \quad \text{if} \quad \lambda > 0, m(\mu) \le -m_1; \\ \lambda - 2 - \phi_-(\lambda, m(\mu)), \quad \text{if} \quad \lambda \ge 2, -m_1 \le m(\mu) < 0, \\ \qquad\qquad\qquad\qquad \text{or } 0 < \lambda < 2, -m_1 < m(\mu) \le m_2; \\ 2 - \lambda + \phi_+(\lambda, m(\mu)), \quad \text{if } \lambda \ge 2, 0 \le m(\mu) < m_1, \\ \qquad\qquad\qquad\qquad \text{or } 4 - 2\sqrt{2} < \lambda < 2, m_2 < m(\mu) < m_1; \\ m(\mu) + \dfrac{2}{\lambda}, \quad \text{if } \lambda > 4 - 2\sqrt{2}, m(\mu) \ge m_1, \\ \qquad\qquad\qquad\qquad \text{or } 0 < \lambda \le 4 - 2\sqrt{2}, m(\mu) > m_2. \end{cases} \tag{8.17}$$

By (8.5), we then have the minimal MFE and the corresponding value function:

$$\begin{aligned} \underline{X}_t^{t_0,\xi} &= \xi + \widehat{b}(\underline{p}^{t_0,\mu})(t - t_0) + B_t^{t_0}, \\ \underline{V}(t_0, x, \mu) &= x\underline{p}^{t_0,\mu} + \frac{1}{2}(T - t_0)|\underline{p}^{t_0,\mu}|^2. \end{aligned} \tag{8.18}$$

Similarly, we find that the maximal p^*, denoted as $\overline{p}^{t_0,\mu}$, is:

$$\overline{p}^{t_0,\mu} := \begin{cases} m(\mu) - \dfrac{2}{\lambda}, \quad \text{if} \quad \lambda > 4 - 2\sqrt{2}, m(\mu) < -m_1, \\ \qquad\qquad\qquad\qquad \text{or } 0 < \lambda \le 4 - 2\sqrt{2}, m(\mu) < -m_2; \\ \lambda - 2 - \phi_-(\lambda, m(\mu)), \quad \text{if} \quad \lambda \ge 2, -m_1 \le m(\mu) < 0, \\ \qquad\qquad\qquad\qquad \text{or } 4 - 2\sqrt{2} < \lambda < 2, -m_1 \le m(\mu) < -m_2; \\ 2 - \lambda + \phi_+(\lambda, m(\mu)), \quad \text{if } \lambda \ge 2, 0 \le m(\mu) < m_1, \\ \qquad\qquad\qquad\qquad \text{or } 0 < \lambda < 2, -m_2 \le m(\mu) < m_1; \\ m(\mu) + \dfrac{2}{\lambda}, \quad \text{if } \lambda > 0, m(\mu) \ge m_1; \end{cases} \tag{8.19}$$

and the maximal MFE and the corresponding value function are:

$$\begin{aligned} \overline{X}_t^{t_0,\xi} &= \xi + \widehat{b}(\overline{p}^{t_0,\mu})(t - t_0) + B_t^{t_0}, \\ \overline{V}(t_0, x, \mu) &= x\overline{p}^{t_0,\mu} + \frac{1}{2}(T - t_0)|\overline{p}^{t_0,\mu}|^2. \end{aligned} \tag{8.20}$$

We note that, when $\lambda > 2$, namely $T - t < \frac{1}{2}$, $\underline{p}^{t,\mu}$ is smooth in (t, μ) and actually in this case $\underline{V} = \overline{V}$ is a classical solution of the master equation (2.12). This is consistent with the standard result that the master equation admits a unique classical solution over small time interval.

However, for $4 - 2\sqrt{2} < \lambda < 2$, namely $\frac{1}{2} < T - t < \frac{1}{4-2\sqrt{2}}$, we have

$$\begin{aligned} \lim_{m(\mu)\uparrow m_2} \underline{p}^{t,\mu} &= \lambda - 2 - \phi_-(\lambda, m_2) = \frac{1}{T - t} - 2; \\ \lim_{m(\mu)\downarrow m_2} \underline{p}^{t,\mu} &= 2 - \lambda + \phi_+(\lambda, m_2) = (1 + \sqrt{2})(2 - \frac{1}{T - t}); \end{aligned} \tag{8.21}$$

That is, $\partial_x \underline{V}(t, x, \mu) = \underline{p}^{t,\mu}$ is discontinuous in μ when $\frac{1}{2} < T - t < \frac{1}{4-2\sqrt{2}}$ and $m(\mu) = m_2$.

Similarly, when $m(\mu) = \frac{1}{20}$, we see that $m_2 > m(\mu)$ if $T - t > \frac{5}{8}$ and $m_2 < m(\mu)$ if $T - t > \frac{1}{8}$. Then, by (8.21) we have

$$\lim_{t\uparrow(T-\frac{5}{8})} \underline{p}^{t,\mu} = \frac{1}{T-(T-\frac{5}{8})} - 2 = -\frac{2}{5},$$
$$\lim_{t\downarrow(T-\frac{5}{8})} \underline{p}^{t,\mu} = (1+\sqrt{2})(2 - \frac{1}{T-(T-\frac{5}{8})}) = \frac{2(1+\sqrt{2})}{5}.$$

That is, $\partial_x \underline{V}(t,x,\mu) = \underline{p}^{t,\mu}$ is discontinuous in t at $t = \frac{1}{8}$ and $m(\mu) = \frac{1}{20}$.

References

[1] S. Ahuja, Wellposedness of mean field games with common noise under a weak monotonicity condition, SIAM J. Control Optim. 54 (2016) 30–48.
[2] A. Bensoussan, P.J. Graber, S.C.P. Yam, Stochastic control on space of random variables, preprint, arXiv:1903.12602.
[3] A. Bensoussan, P.J. Graber, S.C.P. Yam, Control on Hilbert spaces and application to some mean field type control problems, preprint, arXiv:2005.10770.
[4] C. Bertucci, Optimal stopping in mean field games, an obstacle problem approach, J. Math. Pures Appl. 120 (9) (2017) 165–194.
[5] C. Bertucci, Monotone solutions for mean field games master equations: finite state space and optimal stopping, J. Éc. Polytech. Math. 8 (2021) 1099–1132.
[6] C. Bertucci, Monotone solutions for mean field games master equations: continuous state space and common noise, Commun. Partial Differ. Equ. 48 (10-12) (2023) 1245–1285.
[7] C. Bertucci, A. Cecchin, Mean field games master equations: from discrete to continuous state space, preprint, arXiv:2207.03191.
[8] C. Bertucci, J.M. Lasry, P.L. Lions, Some remarks on mean field games, Commun. Partial Differ. Equ. 44 (3) (2019) 205–227.
[9] C. Bertucci, J.M. Lasry, P.L. Lions, Master equation for the finite state space planning problem, Arch. Ration. Mech. Anal. 242 (1) (2021) 327–342.
[10] C. Bertucci, J.M. Lasry, P.L. Lions, On Lipschitz solutions of mean field games master equations, preprint, arXiv:2302.05218.
[11] R. Buckdahn, J. Li, S. Peng, C. Rainer, Mean-field stochastic differential equations and associated PDEs, Ann. Probab. 45 (2017) 824–878.
[12] P.E. Caines, M. Huang, R.P. Malhamé, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, Commun. Inf. Syst. 6 (3) (2006) 221–251.
[13] P. Cardaliaguet, M. Cirant, A. Porretta, Splitting methods and short time existence for the master equations in mean field games, J. Eur. Math. Soc. 25 (5) (2023) 1823–1918.
[14] P. Cardaliaguet, A. Porretta, An introduction to mean field game theory, in: Lecture Notes in Mathematics, vol. 2281, 2020, pp. 1–158.
[15] P. Cardaliaguet, F. Delarue, J.M. Lasry, P.L. Lions, The Master Equation and the Convergence Problem in Mean Field Games, Annals of Mathematics Studies, vol. 201, Princeton University Press, Princeton, NJ, 2019, x+212 pp.
[16] P. Cardaliaguet, P. Souganidis, Monotone solutions of the master equation for mean field games with idiosyncratic noise, SIAM J. Math. Anal. 54 (4) (2022) 4198–4237.
[17] R. Carmona, F. Delarue, Probabilistic Theory of Mean Field Games with Applications I - Mean Field FBSDEs, Control, and Games, Probability Theory and Stochastic Modeling, vol. 83, Springer, Cham, 2018.
[18] R. Carmona, F. Delarue, Probabilistic Theory of Mean Field Games with Applications II - Mean Field Games with Common Noise and Master Equations, Probability Theory and Stochastic Modeling, vol. 84, Springer, Cham, 2018, xxv+697 pp.
[19] R. Carmona, F. Delarue, D. Lacker, Mean field games of timing and models for banks runs, Appl. Math. Optim. 76 (2017) 217–260.
[20] J.F. Chassagneux, D. Crisan, F. Delarue, A Probabilistic Approach to Classical Solutions of the Master Equation for Large Population Equilibria, Mem. Amer. Math. Soc., vol. 280(1379), 2022, v+123 pp.
[21] A. Cecchin, P. Dai Pra, M. Fischer, G. Pelino, On the convergence problem in mean field games: a two state model without uniqueness, SIAM J. Control Optim. 57 (4) (2019) 2443–2466.
[22] A. Cecchin, F. Delarue, Selection by vanishing common noise for potential finite state mean field games, Commun. Partial Differ. Equ. 47 (1) (2022) 89–168.
[23] A. Cecchin, F. Delarue, Weak solutions to the master equation of potential mean field games, preprint, arXiv:2204.04315.
[24] F. Delarue, R. Foguen Tchuendom, Selection of equilibria in a linear quadratic mean-field game, Stoch. Model. Appl. 130 (2) (2020) 1000–1040.
[25] J. Dianetti, Strong solutions to submodular mean field games with common noise and related McKean-Vlasov FBSDEs, preprint, arXiv:2212.12413.
[26] J. Dianetti, G. Ferrari, M. Fischer, M. Nendel, Submodular mean field games: existence and approximation of solutions, Ann. Appl. Probab. 31 (2021) 2538–2566.
[27] J. Dianetti, G. Ferrari, M. Fischer, M. Nendel, A unifying framework for submodular mean field games, Math. Oper. Res. 48 (3) (2023) 1679–1710.
[28] I. Ekren, N. Touzi, J. Zhang, Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I, Ann. Probab. 44 (2) (2016) 1212–1253.
[29] I. Ekren, N. Touzi, J. Zhang, Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part II, Ann. Probab. 44 (4) (2016) 2507–2553.

[30] W. Gangbo, A.R. Mészáros, Global well-posedness of master equations for deterministic displacement convex potential mean field games, Commun. Pure Appl. Math. 75 (12) (2022) 2685–2801.

[31] W. Gangbo, A.R. Mészáros, C. Mou, J. Zhang, Mean field games master equations with non-separable Hamiltonians and displacement monotonicity, Ann. Probab. 50 (6) (2022) 2178–2217.

[32] P.J. Graber, A.R. Mészáros, On mean field games and master equations through the lens of conservation laws, preprint, arXiv:2208.10360.

[33] P.J. Graber, A.R. Mészáros, On monotonicity conditions for mean field games, J. Funct. Anal. 285 (9) (2023), Paper No. 110095, 45 pp.

[34] S. Hamadene, M. Morlais, Viscosity solutions of systems of PDEs with interconnected obstacles and switching problem, Appl. Math. Optim. 67 (2) (2013) 163–196.

[35] Y. Hu, S. Peng, On the comparison theorem for multidimensional BSDEs, C. R. Math. Acad. Sci. Paris 343 (2006) 135–140.

[36] M. Iseri, J. Zhang, Set values for mean field games, preprint, arXiv:2107.01661.

[37] J.M. Lasry, P.L. Lions, Mean field games, Jpn. J. Math. 2 (2007) 229–260.

[38] P.L. Lions, Cours au Collège de France, http://www.college-de-france.

[39] P.L. Lions, B. Seeger, Linear and nonlinear transport equations with coordinate-wise increasing velocity fields, preprint, arXiv:2307.05819.

[40] P.L. Lions, P.E. Souganidis, Extended mean-field games, Atti Accad. Naz. Lincei, Rend. Lincei, Mat. Appl. 31 (3) (2020) 611–625.

[41] C. Mou, J. Zhang, Wellposedness of second order master equations for mean field games with nonsmooth data, Mem. Amer. Math. Soc. (2024), accepted in arXiv:1903.09907.

[42] C. Mou, J. Zhang, Mean field game master equations with anti-monotonicity conditions, J. Eur. Math. Soc. (2024), accepted in arXiv:2201.10762.

[43] S. Muñoz, Classical solutions to local first order extended mean field games, ESAIM Control Optim. Calc. Var. 29 (2023), Paper No. 14, 21 pp.

[44] J. Qiu, Viscosity solutions of stochastic Hamilton-Jacobi-Bellman equations, SIAM J. Control Optim. 56 (2018) 3708–3730.

[45] C. Wu, J. Zhang, Viscosity solutions to parabolic master equations and McKean-Vlasov SDEs with closed-loop controls, Ann. Appl. Probab. 30 (2020) 936–986.

[46] J. Zhang, Backward Stochastic Differential Equations – from Linear to Fully Nonlinear Theory, Springer, New York, 2017.

[47] J. Zhou, Viscosity solutions to second order path-dependent Hamilton-Jacobi-Bellman equations and applications, Ann. Appl. Probab. 33 (6B) (2023) 5564–5612.