Rend. Lincei Mat. Appl. 36 (2025), 89–137
DOI 10.4171/RLM/1066

Probability Theory. – *On mean field monotonicity conditions from a control theoretical perspective*, by Alain Bensoussan, Ziyu Huang, Shanjian Tang and Sheung Chi Phillip Yam, accepted on 24 July 2025.

We dedicate this paper to the memory of Professor Giuseppe Da Prato.

Abstract. – In this article, from the viewpoint of the control theory, we discuss the relationships among the commonly used monotonicity conditions that ensure the well-posedness of the solutions arising from problems of mean field games (MFGs) and mean field type control (MFTC). We first introduce the well-posedness of general forward-backward stochastic differential equations (FBSDEs) defined on some suitably chosen Hilbert spaces under the β-monotonicity. We then propose a monotonicity condition for the MFG, namely, partitioning the running cost functional into two parts, so that both parts still depend on the control and the state distribution, yet one satisfies a strong convexity and a small mean field effect condition, while the other has a newly introduced displacement quasi-monotonicity. To the best of our knowledge, the latter quasi type condition has not yet been discussed in the contemporary literature, and it can be considered as a bit more general monotonicity condition than those commonly used. Besides, for the MFG, we show that convexity and small mean field effect condition for the first part of running cost functional and the quasi-monotonicity condition for the second part together imply the β-monotonicity and thus the well-posedness for the associated FBSDEs. For the MFTC problem, we show that the β-monotonicity for the corresponding FBSDEs is simply the convexity assumption on the cost functional. Finally, we consider a more general setting where the drift functional is allowed to be non-linear for both MFG and MFTC problems.

Keywords. – mean field games, mean field type control problems, forward-backward stochastic differential equations, monotonicity conditions, β-monotonicity, small mean field effect, displacement quasi-monotonicity, generic drift functions.

Mathematics Subject Classification 2020. – 91A16 (primary); 60H30, 60H10, 93E20, 35R15 (secondary).

1. Introduction

Mean field games (MFGs) and mean field type control (MFTC) problems have been widely studied in recent years. For each of them, the underlying controlled dynamical system involves the probability distribution of the state, in addition to the state and the control. A MFG is essentially a fixed point problem which was first proposed by Lasry

and Lions in a series of papers [18, 37–39] and also independently by Huang, Caines and Malhamé [32, 33]. In a MFG, the coefficients of the controlled system are affected by the equilibrium probability distribution of the state of the overall population. In contrast, a MFTC problem is a McKean–Vlasov control problem, the system of which depends on the distribution of the current controlled state. There are numerous works in various settings in this area. For PDE approaches to the forward-backward system for MFGs, we refer to Bensoussan–Frehse–Yam [4], Gomes–Pimentel–Voskanyan [29], Graber–Mészáros [30], Huang–Tang [35] and Porretta [45]. For the master equation analytical methods to MFGs, we refer to Cardaliaguet–Cirant–Porretta [19], Cardaliaguet–Delarue–Lasry–Lions [20], Gangbo–Mészáros–Mou–Zhang [28], Mészáros–Mou [40] and Mou–Zhang [41]. For probabilistic approaches to MFGs, we refer to Ahuja–Ren–Yang [2], Bensoussan–Tai–Wong–Yam [11], Bensoussan–Wong–Yam–Yuan [14], Buckdahn–Li–Peng– Rainer [17], Carmona–Delarue [23, 24], Chassagneux–Crisan–Delarue [25] and Huang–Tang [34]. For probabilistic approaches to MFTC problems, we refer to Buckdahn–Li–Peng–Rainer [17], Cardaliaguet–Delarue–Lasry–Lions [20], Carmona–Delarue [22–24] and Chassagneux–Crisan–Delarue [25]. For the dynamic programming principle and HJB equation of McKean–Vlasov control problem, we refer to Djete–Possamai–Tan [26] and Pham–Wei [44]. For the lifting method and a Hilbert spaces approach for the MFTC problem, we refer to Bensoussan–Graber–Yam [5], Bensoussan–Huang–Yam [6, 9, 10], Bensoussan–Tai–Yam [12] and Bensoussan–Yam [15]. The MFTC problem is also studied together with the so-called "potential mean field game", which is first observed by Lasry and Lions in [39] (although under a different name), and then subsequently well studied in the literature; see [16, 21, 42] as examples.

In this article, we adopt a stochastic control method to discuss monotonicity conditions to solve MFG and MFTC problems with control-dependent diffusion functions which can be degenerate. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t, 0 \le t \le T\}, \mathbb{P})$ be a complete filtered probability space (with the filtration being augmented by all the $\mathbb{P}$-null sets) on which an n-dimensional Brownian motion $\{B_t, 0 \le t \le T\}$ is defined and is $\mathcal{F}_t$-adapted. We denote by $\mathcal{P}_2(\mathbb{R}^n)$ the space of all probability measures with finite second-order moments on $\mathbb{R}^n$, equipped with the 2-Wasserstein metric W_2. Given the functional coefficients:

$$b : [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}^n, \quad \sigma : [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}^{n \times n},$$
$$f : [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}, \quad g : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R},$$

for $(t, \mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^n)$, we choose a random vector $\xi \in L^2_{\mathcal{F}_t}$ independent of the Brownian motion $\mathcal{W}^t_s := \{B_s - B_t,\ t \le s \le T\}$ such that $\mathcal{L}(\xi) = \mu$; then, the limiting problem of the MFG can be formulated in the following probabilistic way, which is a fixed point problem. For a given distribution flow $m = \{m_s \in \mathcal{P}_2(\mathbb{R}^n) : t \le s \le T\}$,

consider the following stochastic optimal control problem:

$$
\begin{aligned}
v_\cdot^m &\in \operatorname*{argmin}_{v_\cdot \in \mathcal{L}^2_{\mathcal{F}}(t,T)} J(v_\cdot; m_s, t \le s \le T) \\
&:= \mathbb{E}\bigg[\int_t^T f(s, X_s^{m,v}, m_s, v_s)ds + g(X_T^{m,v}, m_T)\bigg], \qquad (1.1)\\
X_s^{m,v} &= \xi + \int_t^s b(r, X_r^{m,v}, m_r, v_r)dr + \int_t^s \sigma(r, X_r^{m,v}, m_r, v_r)dB_r, \quad s \in [t,T].
\end{aligned}
$$

If the following consistence condition holds:

$$
m_s = \mathcal{L}(X_s^{m,v^m}), \quad s \in [t,T] \qquad (1.2)
$$

for some $\hat{m}$, then the distribution flow $\hat{m}$ is called an equilibrium for the MFG, and the corresponding control $\hat{v} := v^{\hat{m}}$ is called the solution of MFG (1.1)–(1.2). From the stochastic maximum principle [8,23,24,34,46], the solution of MFG (1.1)–(1.2) (if it exists) gives a solution to the following system of forward-backward stochastic differential equations (FBSDEs) for processes $(X, P, Q) \in (\mathcal{S}^2_{\mathcal{F}} \times \mathcal{S}^2_{\mathcal{F}} \times \mathcal{L}^2_{\mathcal{F}})(0,T)$:

$$
\begin{cases}
X_s = \xi + \displaystyle\int_t^s D_p H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big)dr \\
\qquad\quad + \displaystyle\int_t^s D_q H\big(X_r, \mathcal{L}(X_r), P_r, Q_r\big)dB_r, \\
P_s = D_x g\big(X_T, \mathcal{L}(X_T)\big) + \displaystyle\int_s^T D_x H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big)dr \\
\qquad\quad - \displaystyle\int_s^T Q_r dB_r, \quad s \in [t,T],
\end{cases} \qquad (1.3)
$$

where X is the state process corresponding to the solution of the MFG, and (P, Q) is the related adjoint process. Here, the Hamiltonian

$$
H : [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^{n\times n} \to \mathbb{R}
$$

is defined as

$$
\begin{aligned}
H(s,x,m,p,q) &:= \inf_{v\in\mathbb{R}^d} L(s,x,m,v,p,q), \\
L(s,x,m,v,p,q) &:= p^\top b(s,x,m,v) + \sum_{j=1}^n (q^j)^\top \sigma^j(s,x,m,v) + f(s,x,m,v).
\end{aligned}
$$

Reversely, the well-posedness of FBSDEs (1.3) also gives a solution to MFG (1.1)–(1.2). Another main concern of this article is to study the following MFTC problem in

a more probabilistic favor:

$$\inf_{v.\in\mathcal{L}^2_{\mathcal{F}}(t,T)} J(v.) := \left[\mathbb{E}\int_t^T f\big(s, X_s^v, \mathcal{L}(X_s^v), v_s\big)ds + g\big(X_T^v, \mathcal{L}(X_T^v)\big)\right],$$

$$\begin{aligned} X_s^v &= \xi + \int_t^s b\big(r, X_r^v, \mathcal{L}(X_r^v), v_r\big)dr \\ &\quad + \int_t^s \sigma\big(r, X_r^v, \mathcal{L}(X_r^v), v_r\big)dB_r, \quad s \in [t,T]. \end{aligned} \tag{1.4}$$

From the stochastic maximum principle for McKean–Vlasov type control problems [10,22], the optimal control of the MFTC problem (1.4) is associated with the following FBSDEs:

$$\begin{cases} X_s = \xi + \displaystyle\int_t^s D_p H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big)dr \\ \qquad + \displaystyle\int_t^s D_q H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big)dB_r, \\ P_s = D_x g\big(X_T, \mathcal{L}(X_T)\big) + \widetilde{\mathbb{E}}\Big[D_y \dfrac{dg}{dv}\big(\widetilde{X_T}, \mathcal{L}(X_T)\big)(X_T)\Big] \\ \qquad + \displaystyle\int_s^T \Big\{D_x H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big) \\ \qquad\qquad + \widetilde{\mathbb{E}}\Big[D_y \dfrac{dH}{dv}\big(r, \widetilde{X_r}, \mathcal{L}(X_r), \widetilde{P_r}, \widetilde{Q_r}\big)(X_r)\Big]\Big\}dr \\ \qquad - \displaystyle\int_s^T Q_r dB_r, \quad s \in [t,T], \end{cases} \tag{1.5}$$

where the processes $\widetilde{X_s}$, $\widetilde{P_s}$ and $\widetilde{Q_s}$ are respectively independent copies of X_s, P_s and Q_s, and $\frac{dH}{dv}$ is the linear derivative of H with respect to the measure variable (see Section 1.1 below). The backward equation in FBSDEs (1.5) is different from that of FBSDEs (1.3) by including the terms involving $D_y \frac{dH}{dv}$ and $D_y \frac{dg}{dv}$; this is because the state process of the MFTC problem (1.4) depends simultaneously on the distribution of the current controlled state, while the state process of MFG (1.1)–(1.2) depends on the equilibrium distribution from the population. The main result of this work is to give the well-posedness of FBSDEs (1.3) under different kinds of monotonicity conditions on f and also give the well-posedness of FBSDEs (1.5) under a convexity assumption on f in $(x,v) \in \mathbb{R}^n \times \mathbb{R}^d$ and $m \in \mathcal{P}_2(\mathbb{R}^n)$. Our approach is more aligned with the traditional stochastic control method, which is significantly different from the recently advocated analytical approach to MFGs via master equation or PDE-based HJB-FP equations in the existing literature. The control theoretic perspective allows us to cope with the state and control linear diffusion σ, which is unbounded and possibly

degenerate. We allow our running cost functional f to be non-separable in v and m with a quadratic growth, and to satisfy a strong convexity or even a more general displacement quasi-monotonicity. Moreover, we also solve both MFG (1.1)–(1.2) and the MFTC problem (1.4) with generic drift functionals b. We also require less regularity of the cost functions than the existing literature to solve the original mean field problems.

To study the well-posedness of FBSDEs (1.3) and (1.5), we first establish the solvability of a general system of FBSDEs defined on Hilbert spaces (see FBSDEs (2.1)) under our β-monotonicity (see Condition 2.1) and then apply this general well-posedness result for FBSDEs (1.3) and (1.5), both of which can be viewed as particular cases of FBSDEs (2.1). The monotonicity condition for fully coupled FBSDEs was first introduced by Hu–Peng [31] and Peng–Wu [43] together with the continuation method. This monotonicity condition was then used as a convexity condition in [22] for FBSDEs arising from the MFTC problem, and it was also used as a weak monotonicity condition (displacement monotonicity condition) in [2, 34] for FBSDEs arising from MFG with a common noise. Compared with the usual monotonicity condition for FBSDEs, our β-monotonicity allows us to include more general situations by choosing a suitable candidate of β; for instance, it can be used to establish the well-posedness of the Jacobian and Hessian flows for FBSDEs (1.3) and (1.5) by choosing suitable maps of β; see our previous work [6, 8]. In this work, we further extend our previous results in [6, 8] by relaxing assumptions on the coefficients and show that these new conditions can imply the β-monotonicity and therefore ensure the well-posedness of the FBSDEs for MFGs and MFTC problems. The main results are Theorems 3.3–3.5 for FBSDEs (1.3) (under Conditions 3.1–3.3, respectively), and Theorem 4.2 for FBSDEs (1.5). They are further extended in Section 5 to a more general drift and a varying diffusion.

For the FBSDEs (1.3) arising from MFG (1.1)–(1.2), we propose a new monotonicity condition on f (see Condition 3.3): the functional f is assumed to be divided into two parts, both dependent on (s, x, m, v), with one part satisfying a strong convexity condition in (x, v), and the other part satisfying the newly proposed displacement quasi-monotonicity. To the best of our knowledge, this condition is brand-new in mean field theory, and it can be considered a more general monotonicity condition; for example, it can include as interesting special cases the one with separability and displacement monotonicity proposed in [2] and also the "strong convexity and small mean field effect" condition proposed in our previous work [8], and it also overlaps with monotonicity conditions proposed in [28, 30, 41] from an analytical viewpoint. We show that the convexity and the quasi-monotonicity conditions in our Condition 3.3 can imply the corresponding β-monotonicity for FBSDEs (1.3) and therefore ensure the well-posedness of the FBSDEs. As two particular cases of our Condition 3.3, the classical displacement monotonicity and the small mean field effect can both be viewed as a condition to ensure the β-monotonicity. For the FBSDEs (1.5) arising

from the MFTC problem (1.4), we need the cost functional f to be jointly convex in (x, v) and also convex in m; see Assumption (B3) in Section 4. We show that the corresponding β-monotonicity for FBSDEs (1.5) is exactly this notion of convexity assumption. Moreover, for both FBSDEs (1.3) and (1.5), we also study the case when the drift functional b can be non-linear in x, v and m; see Assumptions (A1′) and (B1′) in Section 5. We can prove that, in these cases, the coefficients satisfy the β-monotonicity; therefore, our well-posedness result for general FBSDEs defined on Hilbert spaces under β-monotonicity is feasible.

MFG (1.1)–(1.2) is usually associated with a mean field master equation, and the MFTC problem (1.4) is usually associated with a Bellman equation. In [6], we establish the classical solutions of the Bellman equation corresponding to the MFTC problem; and in [8], we also establish the classical solution of the MFG master equation. We can see that, to obtain a classical solution of a master equation or HJB equation requires more restrictive assumptions and higher regularity on coefficients than just to obtain an equilibrium solution of MFG or an optimal control for the MFTC problem. Therefore, we study MFG (1.1)–(1.2) and the MFTC problem (1.4) by a stochastic control method, which can include more cases. To make the control perspective more clear, we prefer to give conditions directly on coefficients b, σ, f and g, rather than give conditions on the Hamiltonian functional H or other feedback maps of m. We also refer to other discussions on monotonicity conditions for MFG from an analytical perspective; see [28, 30, 41].

The rest of the paper is organized as follows. In Section 2, we establish the well-posedness of general FBSDEs defined on Hilbert spaces under the β-monotonicity. In Section 3, we give a sufficient condition of maximum principle to hold for MFG (1.1)–(1.2) and propose different monotonicity conditions on f to ensure the well-posedness of FBSDEs (1.3). Section 4 gives a sufficient condition of maximum principle for the MFTC problem (1.4) and shows that the convexity of f can ensure the well-posedness of FBSDEs (1.5). In Section 5, we study the case when the drift functional b is non-linear and generic for both MFG (1.1)–(1.2) and the MFTC problem (1.4).

1.1. Notations

For any $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$, we denote by $\mathcal{L}(X)$ its law and by $\|X\|_2$ the L^2-norm. For every $t \in [0, T]$, we denote by $L^2_{\mathcal{F}_t}$ the set of all $\mathcal{F}_t$-measurable square-integrable $\mathbb{R}^n$-valued random vectors and denote by $\mathcal{L}^2_{\mathcal{F}}(0, T)$ the set of all $\mathcal{F}_t$-progressively-measurable $\mathbb{R}^n$-valued processes $\alpha_\cdot = \{\alpha_t,\ 0 \le t \le T\}$ such that $\mathbb{E}[\int_0^T |\alpha_t|^2 dt] < +\infty$. We denote by $\mathcal{S}^2_{\mathcal{F}}(0, T)$ the set of all $\mathcal{F}_t$-progressively-measurable $\mathbb{R}^n$-valued processes $\alpha_\cdot = \{\alpha_t,\ 0 \le t \le T\}$ such that $\mathbb{E}[\sup_{0\le t\le T} |\alpha_t|^2] < +\infty$. We denote by $\mathcal{P}_2(\mathbb{R}^n)$ the space of all probability measures of finite second-order moments on $\mathbb{R}^n$, equipped with

the 2-Wasserstein metric: $W_2(m,m') := \inf_{\pi\in\Pi(m,m')} \sqrt{\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-x'|^2\pi(dx,dx')}$, where $\Pi(m,m')$ is the set of joint probability measures with respective marginals m and m'. We denote by δ_0 the point mass distribution of the random variable ξ such that $\mathbb{P}(\xi=\mathbf{0})=1$. Also see [3] for more results on Wasserstein metric space.

The linear derivative of a functional $k(\cdot):\mathcal{P}_2(\mathbb{R}^n)\to\mathbb{R}$ at $m\in\mathcal{P}_2(\mathbb{R}^n)$ is another functional $\mathcal{P}_2(\mathbb{R}^n)\times\mathbb{R}^n\ni(m,y)\mapsto\frac{dk}{dv}(m)(y)$, being jointly continuous and satisfying $\int_{\mathbb{R}^n}|\frac{dk}{dv}(m)(y)|^2dm(y)\le c(m)$ for some positive constant $c(m)$ which is bounded on any bounded subsets of $\mathcal{P}_2(\mathbb{R}^n)$, such that

$$\lim_{\varepsilon\to0}\frac{k\big((1-\varepsilon)m+\varepsilon m'\big)-k(m)}{\varepsilon}=\int_{\mathbb{R}}^{n}\frac{dk}{dv}(m)(y)\big(dm'(y)-dm(y)\big),\quad\forall m'\in\mathcal{P}_2(\mathbb{R}^n);$$

we refer the reader to [15,23,24] for more details about the notion of linear derivatives. In particular, the linear derivatives in $\mathcal{P}_2(\mathbb{R}^n)$ are connected to the Gâteaux derivatives in $L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$ in the following way. For a linearly differentiable functional $k:\mathcal{P}_2(\mathbb{R}^n)\to\mathbb{R}$ such that the derivative $D_y\frac{dk}{dv}(\mu)(y)$ is jointly continuous in (μ,y) and $D_y\frac{dk}{dv}(\mu)(y)\le c(\mu)(1+|y|)$ for $(\mu,y)\in\mathcal{P}_2(\mathbb{R}^n)\times\mathbb{R}^n$, the functional $K(X):=k(\mathcal{L}(X))$, $X\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$ has the following Gâteaux derivative:

$$D_XK(X)(\omega)=D_y\frac{dk}{dv}\big(\mathcal{L}(X)\big)\big(X(\omega)\big).\tag{1.6}$$

Furthermore, if k is twice linearly differentiable, then the functional K is also twice Gâteaux differentiable, and the Gâteaux derivative at X along a direction $Z\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$ is

$$D_X^2K(X)(Z)=\Big(D_y^2\frac{dk}{dv}\big(\mathcal{L}(X)\big)(X)\Big)^\top Z+\widetilde{\mathbb{E}}\Big[\Big(D_{y'}D_y\frac{d^2k}{dv}\big(\mathcal{L}(X)\big)(X,\widetilde{X})\Big)^\top\widetilde{Z}\Big].$$

Here and in the following of the paper, for any random variable ξ, we write $\widetilde{\xi}$ for its independent copy, and $\widetilde{\mathbb{E}}[\widetilde{\xi}]$ for the corresponding expectation.

For convenience, in this article, we write $f|_a^b:=f(b)-f(a)$ for the difference of a functional f between two points b and a.

2. Monotonicity condition for FBSDEs on Hilbert spaces

We here give the monotonicity condition to ensure the well-posedness of general FBSDEs defined on Hilbert spaces, which will be used to give the respective monotonicity conditions so as to ensure the well-posedness of FBSDEs (1.3) and (1.5) in

the following sections. We consider the following FBSDEs defined on Hilbert spaces of $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$: for an initial $(t, \xi) \in [0, T] \times L^2_{\mathcal{F}_t}$,

$$\begin{cases} X_s = \xi + \int_t^s \mathbf{B}(r, X_r, P_r, Q_r)dr + \int_t^s \mathbf{A}(r, X_r, P_r, Q_r)dB_r, \\ P_s = \mathbf{G}(X_T) - \int_s^T \mathbf{F}(r, X_r, P_r, Q_r)dr - \int_s^T Q_r dB_r, \quad s \in [t, T], \end{cases} \tag{2.1}$$

where

$$\begin{aligned} &\mathbf{B}, \mathbf{F} : [0, T] \times (L^2 \times L^2 \times (L^2)^n)(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n) \to L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n), \\ &\mathbf{A} : [0, T] \times (L^2 \times L^2 \times (L^2)^n)(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n) \to L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{n\times n}), \\ &\mathbf{G} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n) \to L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n). \end{aligned}$$

The system of FBSDEs (2.1) can be viewed as a lifted version of FBSDEs (1.3) and (1.5) via (3.6) and (4.2) respectively, under assumptions (A)'s and (B)'s to be stated in the following sections. The well-posedness of the generic FBSDEs (2.1) will be established by requiring the following condition.

Condition 2.1. There exists a map

$$\begin{aligned} \beta : [0, T] \times (L^2 \times L^2 \times (L^2)^n)(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n) \ni (s, X, P, Q) \\ \mapsto \beta(s, X, P, Q) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \end{aligned}$$

and constants $\Lambda_\beta > 0$, $\Gamma_\beta \geq 0$ and $K_\beta > 0$, such that for any $X, X', P, P' \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ and $Q, Q' \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{n\times n})$, we have the following:

(i) (β-*Monotonicity*)

(a) The maps $\mathbf{B}$, $\mathbf{A}$ and $\mathbf{F}$ satisfy

$$\begin{aligned} &\mathbb{E}\Big[\big(\mathbf{F}(s, X', P', Q') - \mathbf{F}(s, X, P, Q)\big)^\top (X' - X) \\ &\quad + \big(\mathbf{B}(s, X', P', Q') - \mathbf{B}(s, X, P, Q)\big)^\top (P' - P) \\ &\quad + \sum_{j=1}^n \big(\mathbf{A}^j(s, X', P', Q') - \mathbf{A}^j(s, X, P, Q)\big)^\top (Q'^j - Q^j)\Big] \\ &\leq -\Lambda_\beta \mathbb{E}\big[|\beta(s, X', P', Q') - \beta(s, X, P, Q)|^2\big] \\ &\quad + \Gamma_\beta \big(\|X' - X\|_2^2 + \|P' - P\|_2^2 + \|Q' - Q\|_2^2\big). \end{aligned} \tag{2.2}$$

(b) The map $\mathbf{G}$ satisfies

$$\mathbb{E}\big[\big(\mathbf{G}(X') - \mathbf{G}(X)\big)^\top (X' - X)\big] \geq 0.$$

(ii) (β-*Lipschitz*) The following continuity conditions hold:

$$\begin{aligned}
&\big\|\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big\|_2^2+\big\|\mathbf{A}(s,X',P',Q')-\mathbf{A}(s,X,P,Q)\big\|_2^2\\
&\qquad\le K_\beta\big(\|X'-X\|_2^2+\mathbb{E}\big[\big|\beta(s,X',P',Q')-\beta(s,X,P,Q)\big|^2\big]\big);\\
&\big\|\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big\|^2+\big\|\mathbf{G}(X')-\mathbf{G}(X)\big\|_2^2\\
&\qquad\le K_\beta\big(\|X'-X\|_2^2+\|P'-P\|_2^2+\|Q'-Q\|_2^2\\
&\qquad\qquad+\mathbb{E}\big[\big|\beta(s,X',P',Q')-\beta(s,X,P,Q)\big|^2\big]\big).
\end{aligned}\tag{2.3}$$

We now state the well-posedness of FBSDEs (2.1) when Condition 2.1 holds under suitable relations of the involved constants $(\Lambda_\beta, K_\beta, \Gamma_\beta)$.

LEMMA 2.1. *Under Condition 2.1, when*

$$\Lambda_\beta > 4T(T+1)K_\beta^2\exp\big(2T(T+1)K_\beta^2\big)\big[1+T(T+1)K_\beta^2\exp\big(2T(T+1)K_\beta^2\big)\big]\Gamma_\beta,\tag{2.4}$$

there is a unique adapted solution $(X,P,Q)\in\mathcal{S}_{\mathcal{F}}^2(t,T)\times\mathcal{S}_{\mathcal{F}}^2(t,T)\times(\mathcal{L}_{\mathcal{F}}^2(t,T))^n$ *of the FBSDEs* (2.1). *Furthermore, if the parameter* $\Gamma_\beta=0$ *in Condition* 2.1 (i) (a), *then for any* $\Lambda_\beta>0$, *the same assertion is still true.*

REMARK 2.1. The exponential term in (2.4) comes from the use of Grönwall's inequality. Actually, as a particular case, when the right-hand side of (2.2) is simply

$$-\Lambda_\beta\mathbb{E}\big[\big|\beta(s,X',P',Q')-\beta(s,X,P,Q)\big|^2\big]+\Gamma_\beta\|X'-X\|_2^2,\tag{2.5}$$

and the right-hand side of (2.3) is just written as

$$K_\beta\mathbb{E}\big[\big|\beta(s,X',P',Q')-\beta(s,X,P,Q)\big|^2\big],\tag{2.6}$$

then the condition (2.4) can reduce to $\Lambda_\beta>2T(T+1)K_\beta^2\Gamma_\beta$; for instance, for the FBSDEs (1.3) for the MFG (1.1)–(1.2) in connection with the setting (3.6), if both the coefficient functions b and σ do not functionally depend on x and m, then the right-hand side of (2.2) becomes (2.5) and the right-hand side of (2.3) reduces to (2.6).

The proof of the preceding well-posedness result appeals to the method of continuation in coefficients proposed in [31], is similar to that of [43, Theorem 2.3], [2, Theorem 1], [6, Lemma 4.1] and [8, Lemma 2.5] and is thus omitted here. Condition 2.1 (i) is actually the 'usual monotonicity condition' for FBSDEs, which can be dated back to Hu–Peng [31] and Peng–Wu [43] for fully coupled FBSDEs in Euclidean spaces. It is also referred to as the weak monotonicity condition (displacement monotonicity condition) in Ahuja et al. [2] for FBSDEs in Hilbert spaces, which can be applied to the FBSDEs arising from MFGs with a common noise. Here, our β-monotonicity in

Condition 2.1 allows us to extend our study to more general situations. It is also used in our previous works [6,8] to establish the well-posedness of the respective Jacobian and Hessian flows for FBSDEs (1.3) and (1.5) by choosing suitable maps β. In this paper, we further extend our previous works [6,8] by relaxing assumptions on the running cost functional f (see Section 3), and we can also include the generic drift functional cases (see Section 5). Lemma 2.1 with our β-monotonicity turns out to apply to all these cases.

3. Monotonicity conditions for MFG

We now study the solvability of MFG (1.1)–(1.2). We first give a sufficient condition of maximum principle to hold for MFG (1.1)–(1.2) and then introduce different monotonicity conditions for FBSDEs (1.3), and we will give in the last section with the help of Lemma 2.1 the well-posedness result under these monotonicity conditions.

3.1. *Maximum principle*

We need the following assumptions on coefficients.

(A1) The functions b and σ are linear in (x, v). That is,

$$\begin{aligned} b(s,x,m,v) &= b_0(s,m) + b_1(s)x + b_2(s)v, \\ \sigma(s,x,m,v) &= \sigma_0(s,m) + \sigma_1(s)x + \sigma_2(s)v. \end{aligned}$$

Here, the functions b_0 and σ_0 are L-Lipschitz continuous in $m \in \mathcal{P}_2(\mathbb{R}^n)$, and the norms of matrices $b_1(s), \sigma_1(s), b_2(s), \sigma_2(s)$ are bounded by L.

(A2) The cost functions f and g have a quadratic growth and satisfy for $(s,x,m,v) \in [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^d$,

$$\begin{aligned} \big|f(s,x,m,v)\big| &\le L\big(1 + |x|^2 + W_2^2(m,\delta_0) + |v|^2\big), \\ \big|g(x,m)\big| &\le L\big(1 + |x|^2 + W_2^2(m,\delta_0)\big). \end{aligned} \tag{3.1}$$

The derivatives $D_x f$, $D_v f$, $D_x g$ exist, and they are continuous in all their arguments, such that

$$\begin{aligned} &\big|(D_x f, D_v f)(s,x',m',v') - (D_x f, D_v f)(s,x,m,v)\big| \\ &\quad \le L\big(|x'-x| + |v'-v| + W_2(m,m')\big), \\ &\big|D_x g(x',m') - D_x g(x,m)\big| \le L\big(|x'-x| + W_2(m,m')\big). \end{aligned}$$

(A3) (i) The terminal cost function g is convex in x, and there exists $\lambda > 0$ such that for any $s \in [0,T]$ and $(x,m,v,v') \in \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^d \times \mathbb{R}^d$,

$$f(s,x,m,v') - f(s,x,m,v) \ge \big(D_v f(s,x,m,v)\big)^\top (v'-v) + \lambda |v'-v|^2;$$

(ii) moreover, the running cost function f is jointly convex in x and v, that is,

$$f(s,x',v',m)-f(s,x,v,m)\geq\left[\begin{pmatrix}D_xf\\D_vf\end{pmatrix}(s,x,v,m)\right]^\top\begin{pmatrix}x'-x\\v'-v\end{pmatrix}+\lambda|v'-v|^2.$$

Under above assumptions, we define the optimal control $\hat{v}$ as a map $[0,T]\times\mathbb{R}^n\times\mathcal{P}_2(\mathbb{R}^n)\times\mathbb{R}^n\times\mathbb{R}^{n\times n}\ni(s,x,m,p,q)\mapsto\hat{v}(s,x,m,p,q)\in\mathbb{R}^d$ so that

$$D_vL\big(s,x,m,\hat{v}(s,x,m,p,q),p,q\big)=0. \tag{3.2}$$

We now give the sufficiency for the maximum principle to hold for MFG (1.1)–(1.2) under above assumptions.

THEOREM 3.1. *Under Assumptions (A1), (A2) and (A3) (i), suppose that FBSDEs* (1.3) *have a unique solution*

$$(X,P,Q)\in\mathcal{S}^2_{\mathcal{F}}(t,T)\times\mathcal{S}^2_{\mathcal{F}}(t,T)\times\big(\mathcal{L}^2_{\mathcal{F}}(t,T)\big)^n.$$

Then, there is a constant $c(L,T)>0$, such that for $\lambda>c(L,T)$, MFG (1.1)–(1.2) *has a unique solution*

$$v_s:=\hat{v}\big(s,X_s,\mathcal{L}(X_s),P_s,Q_s\big),\quad s\in[t,T]. \tag{3.3}$$

Furthermore, if (A3) (ii) *is satisfied, then, for any $\lambda>0$, the same assertion still holds.*

PROOF. The first statement is proven in [8, Lemma 2.2]. Here, we only prove the second statement. We denote $m_s:=\mathcal{L}(X_s)$ for $s\in[t,T]$. For any control $u\in\mathcal{L}^2_{\mathcal{F}}(t,T)$, we denote by X^u the corresponding state:

$$\begin{aligned}X^u_s=\xi&+\int_t^s\big[b_0(r,m_r)+b_1(r)X^u_r+b_2(r)u_r\big]dr\\&+\int_t^s\big[\sigma_0(r,m_r)+\sigma_1(r)X^u_r+\sigma_2(r)u_r\big]dB_r.\end{aligned}$$

From Assumptions (A2) and (A3) (ii), we have

$$\begin{aligned}&J(u;m_s,\ t\leq s\leq T)-J(v;m_s,\ t\leq s\leq T)\\&=\mathbb{E}\bigg[\int_t^T\big(f(s,X^u_s,m_s,u_s)-f(s,X_s,m_s,v_s)\big)ds+g(X^u_T,m_T)-g(X_T,m_T)\bigg]\\&\geq\mathbb{E}\bigg[\int_t^T\Big(\lambda|u_s-v_s|^2+\big(D_vf(s,X_s,m_s,v_s)\big)^\top(u_s-v_s)\\&\quad+\big(D_xf(s,X_s,m_s,v_s)\big)^\top(X^u_s-X_s)\Big)ds+\big(D_xg(X_T,m_T)\big)^\top(X^u_T-X_T)\bigg].\end{aligned}\tag{3.4}$$

From Assumption (A1) (and the fact that $b_0(s, m_s)$ and $\sigma_0(s, m_s)$ are the same in the respective equations for X_s and X_s^u), we know that

$$X_s^u - X_s = \int_t^s \big[b_1(r)(X_r^u - X_r) + b_2(r)(u_r - v_r)\big]dr + \int_t^s \big[\sigma_1(r)(X_r^u - X_r) + \sigma_2(r)(u_r - v_r)\big]dB_r;$$

then, from Itô's formula, we have for $s \in (t, T)$,

$$\frac{d}{ds}\mathbb{E}\big[P_s^\top (X_s^u - X_s)\big] = \Big(P_s^\top b_2(s) + \sum_{j=1}^n (Q_s^j)^\top \sigma_2^j(s)\Big)(u_s - v_s) - \big(D_x f(s, X_s, m_s, v_s)\big)^\top (X_s^u - X_s).$$

From the optimality condition (3.3), we have

$$P_s^\top b_2(s) + \sum_{j=1}^n (Q_s^j)^\top \sigma_2^j(s) + \big(D_v f(s, X_s, m_s, v_s)\big)^\top = 0, \quad s \in [t, T].$$

Therefore, we deduce that

$$\mathbb{E}\big[\big(D_x g(X_T, m_T)\big)^\top (X_T^u - X_T)\big] = -\mathbb{E}\bigg[\int_t^T \Big(\big(D_v f(s, X_s, m_s, v_s)\big)^\top (u_s - v_s) + \big(D_x f(s, X_s, m_s, v_s)\big)^\top (X_s^u - X_s)\Big)ds\bigg]. \tag{3.5}$$

Substituting (3.5) into (3.4), we have

$$J(u; m_s,\ t \le s \le T) - J(v; m_s,\ t \le s \le T) \ge \lambda \mathbb{E}\bigg[\int_t^T |u_s - v_s|^2 ds\bigg].$$

Then, we know that the stochastic control problem (1.1) with the fixed distribution flow $\{m_s,\ t \le s \le T\}$ has a unique optimal control as in (3.3). Therefore, (3.3) gives a solution for MFG (1.1)–(1.2). For the uniqueness result, suppose that $v' \in \mathcal{L}_{\mathcal{F}}^2(0, T)$ is another solution for MFG (1.1)–(1.2), and we denote by X' its related controlled process and $m'_s := \mathcal{L}(X'_s)$ for $s \in [t, T]$. For the stochastic control problem (1.1) with the fixed distribution flow $\{m'_s,\ t \le s \le T\}$, we know that v' is an optimal control, and we denote by (P', Q') the corresponding adjoint process. Then, from the necessary condition of the maximum principle (see [46] for instance), we know that (X', P', Q') is also a solution for FBSDEs (1.3). From the uniqueness of the solution for FBSDEs (1.3), we know that $(X', P', Q') = (X, P, Q)$, and then, $v' = v$ in $\mathcal{L}_{\mathcal{F}}^2(0, T)$. ∎

In Theorem 3.1, $c(L,T) = 4L^2T(T+1)\exp(4L^2T(T+1))$, again the exponential term comes from the use of Grönwall's inequality. As a particular case, when both b and σ do not functionally depend on x, then one can take $c(L,T) = 2L^2T(T+1)$.

3.2. *Well-posedness of FBSDEs* (1.3)

We shall use Lemma 2.1 to give the well-posedness of our FBSDEs (1.3), which can be viewed as a subcase of FBSDEs (2.1) by setting

$$\begin{aligned}
\mathbf{B}(s,X,P,Q)(\omega) &:= D_p H\big(s, X(\omega), \mathcal{L}(X), P(\omega), Q(\omega)\big),\\
\mathbf{A}(s,X,P,Q)(\omega) &:= D_q H\big(s, X(\omega), \mathcal{L}(X), P(\omega), Q(\omega)\big),\\
\mathbf{F}(s,X,P,Q)(\omega) &:= -D_x H\big(s, X(\omega), \mathcal{L}(X), P(\omega), Q(\omega)\big),\\
\mathbf{G}(X)(\omega) &:= D_x g\big(X(\omega), \mathcal{L}(X)\big),\\
X, P \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n), &\quad Q \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{n\times n}).
\end{aligned} \tag{3.6}$$

From Assumptions (A1) and (A2), it is easy to check that the β-Lipschitz-continuities specified in Condition 2.1 (ii) are satisfied with the constant $K_\beta = C(L)$ and the following choice of β:

$$\beta(s,X,P,Q)(\omega) := \hat{v}\big(s, X(\omega), \mathcal{L}(X), P(\omega), Q(\omega)\big), \tag{3.7}$$

where the optimal control $\hat{v}$ is defined in (3.2). In view of the map G defined in (3.6), Condition 2.1 (i) motivates us to introduce the following displacement quasi-monotonicity condition.

Definition 3.1 (Displacement quasi-monotonicity condition). For a functional $g : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$ satisfying Assumption (A2), we say that g satisfies the displacement quasi-monotonicity condition with a parameter $\lambda \in \mathbb{R}$, if for any square-integrable random variables ξ and ξ' on the same probability space,

$$\mathbb{E}\big[\big(D_x g\big(\xi', \mathcal{L}(\xi')\big) - D_x g\big(\xi, \mathcal{L}(\xi)\big)\big)^\top (\xi' - \xi)\big] \geq -\lambda \|\xi' - \xi\|_2^2. \tag{3.8}$$

Remark 3.1. Here, λ can be an arbitrary real number. As one particular representative example of Definition 3.1, when $\lambda = 0$ in (3.8), it is reduced to the prevalent displacement monotonicity condition:

$$\mathbb{E}\big[\big(D_x g\big(\xi', \mathcal{L}(\xi')\big) - D_x g\big(\xi, \mathcal{L}(\xi)\big)\big)^\top (\xi' - \xi)\big] \geq 0. \tag{3.9}$$

This displacement monotonicity condition was first introduced in [1] under a different name "weak monotonicity condition", and then it was also used in [2, 34] for a probabilistic approach to the solvability of MFG with a common noise. It is also used in [27, 28] for an analytical method for the well-posedness of MFG master equations.

In view of the map G defined in (3.6), we know that when g satisfies the displacement monotonicity condition (3.9), we have

$$\begin{aligned}(3.10)\qquad &\mathbb{E}\big[\big(\mathbf{G}(X')-\mathbf{G}(X)\big)^{\top}(X'-X)\big]\\ &\quad=\mathbb{E}\big[\big(D_x g\big(X',\mathscr{L}(X')\big)-D_x g\big(X,\mathscr{L}(X)\big)\big)^{\top}(X'-X)\big]\geq 0,\end{aligned}$$

and therefore, Condition 2.1 (i) (b) is satisfied. In the rest of this section, we focus on conditions on f to ensure the β-monotonicity in Condition 2.1 (i) (a) with the choice of β in (3.7). Then, the well-posedness of FBSDEs (1.3) can be deduced by Lemma 2.1. The following result shows that the convexity of f in v in Assumption (A3) (i) can guarantee Condition 2.1 (i) (a) and then gives the local solvability of FBSDEs (1.3).

THEOREM 3.2. *Under Assumptions* (A1), (A2) *and* (A3) (i), *coefficients in* (3.6) *satisfy Condition* 2.1 *with the choice of* β *in* (3.7). *As a consequence, there is a constant* $c(L,T)>0$, *such that for* $\lambda>c(L,T)$, *there is a unique adapted solution of the FBSDEs* (1.3).

PROOF. From a similar approach as [8, Lemma 2.3], we can deduce that Condition 2.1 (i) (a) is satisfied with $\Lambda_\beta=\lambda$, $\Gamma_\beta=\frac{9L^2}{4\lambda}+4L$, and $K_\beta=4L$. Therefore, from (2.4), we obtain the desired result. Here, $c(L,T)=A+\sqrt{A^2+4A}$, where $A:=64T(T+1)L^3\exp(32T(T+1)L^2)[1+16T(T+1)L^2\exp(32T(T+1)L^2)]$. The exponential term comes from the use of Grönwall's inequality. As a particular case, when both b and σ do not functionally depend on x and m, we can take $A=2T(T+1)L^3$. This is similar to [11, condition (3.11)]. ∎

Theorem 3.2 only gives the solvability of FBSDEs (1.3) when T is small enough. From now on, we aim to obtain the global well-posedness of FBSDEs (1.3), which requires Condition 2.1 (i) (a) to be satisfied with the parameter $\Gamma_\beta=0$ for our settings in (3.6) and (3.7). From Assumption (A1), when the coefficients b_0 and σ_0 do not functionally depend on m, for $s\in[t,T]$, $X,X',P,P'\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$ and $Q,Q'\in(L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n))^n$, by denoting $\hat{v}:=\hat{v}(s,X,\mathscr{L}(X),P,Q)$ and $\hat{v}':=\hat{v}(s,X',\mathscr{L}(X'),P',Q')$, we can compute that

$$\begin{aligned}(3.11)\quad &\mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^{\top}(X'-X)\\ &\qquad+\big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^{\top}(P'-P)\\ &\qquad+\sum_{j=1}^{n}\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^{\top}(Q'^j-Q^j)\Big]\\ &=-\mathbb{E}\left\{\left[\begin{pmatrix}D_x f\\ D_v f\end{pmatrix}(s,\cdot)\,\middle|\,\begin{matrix}(X',\mathscr{L}(X'),\hat{v}(s,X',\mathscr{L}(X'),P',Q'))\\ (X,\mathscr{L}(X),\hat{v}(s,X,\mathscr{L}(X),P,Q))\end{matrix}\right]^{\top}\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix}\right\},\end{aligned}$$

with which we shall introduce different monotonicity assumptions on f so as to ensure the right side of (3.11) to be smaller than

$$\mathbb{E}\big[-\Lambda_\beta\big|\hat{v}(s,X',\mathcal{L}(X'),P',Q')-\hat{v}(s,X,\mathcal{L}(X),P,Q)\big|^2\big] \quad \text{for some } \Lambda_\beta>0.$$

Based on the displacement quasi-monotonicity condition in Definition 3.1, we first propose the following weak monotonicity condition on f for the separable case first, and then in Condition 3.3, we shall introduce the one for the non-separable case.

CONDITION 3.1 (Displacement quasi-monotonicity for separable cases). The functional f is separable in m and v: $f(s,x,m,v)=f_0(s,x,m)+f_1(s,x,v)$, where the functional f_1 satisfies the following strong convexity: there exists $\lambda_v>0$ and $\lambda_x\geq 0$, such that

$$\begin{aligned}(3.12)\qquad & f_1(s,x',v')-f_1(s,x,v)\\ &\geq\left[\begin{pmatrix}D_x f_1\\ D_v f_1\end{pmatrix}(s,x,v)\right]^\top\begin{pmatrix}x'-x\\ v'-v\end{pmatrix}+\lambda_x|x'-x|^2+\lambda_v|v'-v|^2;\end{aligned}$$

the functional f_0 satisfies the following displacement quasi-monotonicity: there exists $\lambda_m\leq 2\lambda_x$, such that

$$\begin{aligned}(3.13)\qquad \mathbb{E}\big[\big(D_x f_0\big(s,\xi',\mathcal{L}(\xi')\big)-D_x f_0\big(s,\xi,\mathcal{L}(\xi)\big)\big)^\top(\xi'-\xi)\big]\geq -\lambda_m\|\xi'-\xi\|_2^2,\\ \forall \xi,\xi'\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n).\end{aligned}$$

REMARK 3.2. Condition 3.1 allows λ_m to be positive; that is, the running cost functional f_0 can be displacement quasi-monotonic. Particularly, when $\lambda_m=0$, (3.13) is reduced to the classical displacement monotonicity condition as in Remark 3.1. In this case, we can allow $\lambda_x=0$, and then, Condition 3.1 coincides with the assumption on f in [2].

We next show that Condition 3.1 can ensure our β-monotonicity in Condition 2.1 with $\Gamma_\beta=0$.

THEOREM 3.3. *Under Assumptions* (A1)–(A3)*, suppose that the functions b_0 and σ_0 do not functionally depend on m, the terminal cost functional g satisfies the displacement monotonicity condition* (3.9) *and the running cost functional f satisfies Condition* 3.1*. Then, coefficients in* (3.6) *satisfy Condition* 2.1 (i) (a) *with the choice of β in* (3.7) *and the parameter $\Gamma_\beta=0$. As a consequence, FBSDEs* (1.3) *have a unique global solution.*

PROOF. For the sake of notational convenience, in this proof, we denote

$$\hat{v}:=\hat{v}(s,X,\mathcal{L}(X),P,Q)\quad\text{and}\quad \hat{v}':=\hat{v}(s,X',\mathcal{L}(X'),P',Q').$$

From (3.11) and the separable assumption in Condition 3.1, we have

$$\begin{aligned}
&\mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^\top(X'-X)\\
&\quad+\big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^\top(P'-P)\\
&\quad+\sum_{j=1}^n\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^\top(Q'^j-Q^j)\Big]\\
&=-\mathbb{E}\left\{\left[\begin{pmatrix}D_xf_1\\D_vf_1\end{pmatrix}(s,\cdot)\,\middle|\,\begin{matrix}(X',\hat v')\\(X,\hat v)\end{matrix}\right]^\top\begin{pmatrix}X'-X\\\hat v'-\hat v\end{pmatrix}\right.\\
&\qquad\left.+\left(D_xf_0(s,\cdot)\,\middle|\,\begin{matrix}(X',\mathcal{L}(X'))\\(X,\mathcal{L}(X))\end{matrix}\right)^\top(X'-X)\right\}.
\end{aligned}$$

From the displacement quasi-monotonicity (3.13), we know that

$$-\mathbb{E}\big[\big(D_xf_0\big(s,X',\mathcal{L}(X')\big)-D_xf_0\big(s,X,\mathcal{L}(X)\big)\big)^\top(X'-X)\big]\le\lambda_m\|\xi'-\xi\|_2^2,$$

and from the convexity (3.12), we have

$$\begin{aligned}
&-\mathbb{E}\big[\big(D_vf_1(s,X',\hat v')-D_vf_1(s,X,\hat v)\big)^\top(\hat v'-\hat v)\\
&\qquad+\big(D_xf_1(s,X',\hat v')-D_xf_1(s,X,\hat v)\big)^\top(X'-X)\big]\\
&\quad\le-2\lambda_x\|\xi'-\xi\|_2^2-2\lambda_v\|\hat v'-\hat v\|_2^2.
\end{aligned}$$

Combining the last two inequalities, we can see that

$$\begin{aligned}
&-\mathbb{E}\left\{\left[\begin{pmatrix}D_xf_1\\D_vf_1\end{pmatrix}(s,\cdot)\,\middle|\,\begin{matrix}(X',\hat v')\\(X,\hat v)\end{matrix}\right]^\top\begin{pmatrix}X'-X\\\hat v'-\hat v\end{pmatrix}+\left(D_xf_0(s,\cdot)\,\middle|\,\begin{matrix}(X',\mathcal{L}(X'))\\(X,\mathcal{L}(X))\end{matrix}\right)^\top(X'-X)\right\}\\
&\quad\le(\lambda_m-2\lambda_x)\|\xi'-\xi\|_2^2-2\lambda_v\|\hat v'-\hat v\|_2^2,
\end{aligned}$$

and since $\lambda_m\le2\lambda_x$, we know that Condition 2.1 (i) (a) is valid with the choice of β in (3.7) for $\Lambda_\beta=2\lambda_v$ and $\Gamma_\beta=0$. ∎

We now consider the case when f is not separable in x and v. We propose the following strong convexity and small mean field effect condition, which is also used in our previous work [8].

CONDITION 3.2 (Strong convexity and small mean field effect). There exist non-negative constants $L_x, L_v\le L$, such that

$$\begin{aligned}
&\big|D_xf(s,x',m',v')-D_xf(s,x,m,v)\big|\le L\big(|x'-x|+|v'-v|\big)+L_xW_2(m,m'),\\
&\big|D_vf(s,x',m',v')-D_vf(s,x,m,v)\big|\le L\big(|x'-x|+|v'-v|\big)+L_vW_2(m,m');
\end{aligned}\tag{3.14}$$

and there exist $\lambda_v > 0$ and $\lambda_x \geq \frac{L_v^2}{8\lambda_v} + \frac{L_x}{2}$ such that

$$
\begin{aligned}
(3.15)\quad & f(s,x',m,v') - f(s,x,m,v) \\
& \geq \left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix} (s,x,m,v) \right]^\top \begin{pmatrix} x'-x \\ v'-v \end{pmatrix} + \lambda_x |x'-x|^2 + \lambda_v |v'-v|^2.
\end{aligned}
$$

Here, in the relation $\lambda_x \geq \frac{L_v^2}{8\lambda_v} + \frac{L_x}{2}$, the parameter $\frac{1}{8}$ is not the optimal, but we do not drill down into the details. We first give the relation of Condition 3.2 with the displacement monotonicity condition when f is independent of v.

REMARK 3.3. For a functional f satisfying Condition 3.2 and independent of v, the continuity (3.14) reduces to

$$
(3.16)\qquad \left| D_x f(s,x',m') - D_x f(s,x,m) \right| \leq L|x'-x| + L_x W_2(m,m'),
$$

and the convexity (3.15) reduces to

$$
(3.17)\qquad f(s,x',m) - f(s,x,m) \geq \big(D_x f(s,x,m)\big)^\top (x'-x) + \lambda_x |x'-x|^2,
$$

with the constants satisfying $\lambda_x \geq \frac{L_x}{2}$. For any square-integrable random variables ξ and ξ' on the same probability space, we know that

$$
\begin{aligned}
& \mathbb{E}\big[\big(D_x f\big(s,\xi',\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi)\big)\big)^\top (\xi'-\xi)\big] \\
& \quad = \mathbb{E}\big[\big(D_x f\big(s,\xi',\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi')\big)\big)^\top (\xi'-\xi)\big] \\
& \qquad + \mathbb{E}\big[\big(D_x f\big(s,\xi,\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi)\big)\big)^\top (\xi'-\xi)\big].
\end{aligned}
$$

From the convexity (3.17), we know that

$$
\mathbb{E}\big[\big(D_x f\big(s,\xi',\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi')\big)\big)^\top (\xi'-\xi)\big] \geq 2\lambda_x \|\xi'-\xi\|_2^2,
$$

and from (3.16), we can compute that

$$
\mathbb{E}\big[\big(D_x f\big(s,\xi,\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi)\big)\big)^\top (\xi'-\xi)\big] \geq -L_x \|\xi'-\xi\|_2^2.
$$

Combining the last two inequalities, we have

$$
\mathbb{E}\big[\big(D_x f\big(s,\xi',\mathcal{L}(\xi')\big) - D_x f\big(s,\xi,\mathcal{L}(\xi)\big)\big)^\top (\xi'-\xi)\big] \geq (2\lambda_x - L_x)\|\xi'-\xi\|_2^2;
$$

since $\lambda_x \geq \frac{L_x}{2}$, we know that f satisfies the displacement monotonicity condition.

However, when f is dependent on v, Condition 3.2 cannot be included by Condition 3.1. For the very special linear-quadratic case $f(x,m,v) := |x|^2 + |v|^2 + v^\top \int_{\mathbb{R}^n} y m(dy)$, it can be easy to check that f satisfies Condition 3.2, yet f is not separable in m and v.

We now show that Condition 3.2 can also ensure our β-monotonicity in Condition 2.1 with $\Gamma_\beta = 0$.

THEOREM 3.4. *Under Assumptions* (A1)–(A3)*, suppose that the functions b_0 and σ_0 do not functionally depend on m, the terminal cost functional g satisfies the displacement monotonicity condition* (3.9) *and the running cost functional f satisfies Condition* 3.2*. Then, coefficients in* (3.6) *satisfy Condition* 2.1 (i) (a) *with the choice of β in* (3.7) *and the parameter $\Gamma_\beta = 0$. As a consequence, FBSDEs* (1.3) *have a unique global solution.*

PROOF. As in the proof of Theorem 3.3, we adopt the use the notations $\hat{v} := \hat{v}(s, X, \mathcal{L}(X), P, Q)$ and $\hat{v}' := \hat{v}(s, X', \mathcal{L}(X'), P', Q')$. Since f satisfies (3.14) and (3.15), we can deduce that

$$
\begin{aligned}
(3.18)\quad & -\mathbb{E}\left\{\left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s,\cdot)\,\middle|\,\begin{matrix}(X',\mathcal{L}(X'),\hat{v}')\\(X,\mathcal{L}(X),\hat{v})\end{matrix}\right]^\top \begin{pmatrix} X'-X \\ \hat{v}'-\hat{v}\end{pmatrix}\right\} \\
&\le -\mathbb{E}\left\{\left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s,\cdot,\mathcal{L}(X'),\cdot)\,\middle|\,\begin{matrix}(X',\hat{v}')\\(X,\hat{v})\end{matrix}\right]^\top \begin{pmatrix} X'-X \\ \hat{v}'-\hat{v}\end{pmatrix}\right\} \\
&\quad + \mathbb{E}\left\{\left|\left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s,X,\cdot,\hat{v}')\,\middle|\,\begin{matrix}\mathcal{L}(X')\\\mathcal{L}(X)\end{matrix}\right]^\top \begin{pmatrix} X'-X \\ \hat{v}'-\hat{v}\end{pmatrix}\right|\right\} \\
&\le -2\mathbb{E}\big[\lambda_v|\hat{v}'-\hat{v}|^2+\lambda_x|X'-X|^2\big]+L_v\|X'-X\|_2\cdot\|\hat{v}'-\hat{v}\|_2+L_x\|X'-X\|_2^2.
\end{aligned}
$$

Then, we see from (3.11) and Young's inequality that

$$
\begin{aligned}
&\mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^\top(X'-X) \\
&\quad+\big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^\top(P'-P) \\
&\quad+\sum_{j=1}^n\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^\top(Q'^j-Q^j)\Big] \\
&\le -2\mathbb{E}\big[\lambda_v|\hat{v}'-\hat{v}|^2+\lambda_x|X'-X|^2\big]+L_v\|X'-X\|_2\cdot\|\hat{v}'-\hat{v}\|_2+L_x\|X'-X\|_2^2 \\
&\le -\lambda_v\|\hat{v}'-\hat{v}\|_2^2-\left(2\lambda_x-L_x-\frac{L_v^2}{4\lambda_v}\right)\|X'-X\|_2^2 \\
&\le -\lambda_v\|\hat{v}'-\hat{v}\|_2^2.
\end{aligned}
$$

In the last inequality, we have used the inequality in Condition 3.2: $\lambda_x \ge \frac{L_x}{2}+\frac{L_v^2}{8\lambda_v}$. Therefore, we know that Condition 2.1 (i) (a) is valid with the choice of β in (3.7) by setting $\Lambda_\beta = \lambda_v$ and $\Gamma_\beta = 0$. ∎

REMARK 3.4. Although we here assume that b_0 and σ_0 do not functionally depend on m, since the functional f here can be non-separable, therefore, the optimal control $\hat{v}$ here can depend on m, and hence, the coefficients

$$D_p H(s,x,m,p,q) = b\big(s,x,m,\hat{v}(s,x,m,p,q)\big),$$
$$D_q H(s,x,m,p,q) = \sigma\big(s,x,m,\hat{v}(s,x,m,p,q)\big)$$

are dependent on m. Therefore, the coefficients of the SDE in the system (1.3) are distribution dependent. Moreover, since we allow the diffusion term σ to depend on v, the optimal control $\hat{v}$ can depend on q, which also extends the results in the existing literature via a PDE approach.

In Condition 3.2, the functional f does not need to be separable or displacement monotonic. From the relation $\lambda_x \geq \frac{L_v^2}{8\lambda_v} + \frac{L_x}{2}$, we can see that the dependence of $D_x f$ and $D_v f$ on m should be smaller than the convexity of f in x. Now, we further extend Conditions 3.1 and 3.2 and propose the following monotonicity assumption for f. The functional f is assumed to be divided into two parts (both dependent on (s,x,m,v)), with one part satisfying a strong convexity and small mean field effect condition and the other part satisfying a displacement quasi-monotonicity.

CONDITION 3.3. The running cost functional f is in the form:

$$f(s,x,m,v) = f_0(s,x,m,v) + f_1(s,x,m,v). \tag{3.19}$$

Here, the functional f_1 satisfies the continuity (3.14) with non-negative constants L_x and L_v, and it also satisfies the convexity (3.15) with constants $\lambda_x \geq 0$ and $\lambda_v > 0$. The functional f_0 is convex in v and satisfies the following displacement quasi-monotonicity with a constant $\lambda_m \in \mathbb{R}$: for any $(s,V) \in [0,T] \times L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$,

$$\mathbb{E}\big[\big(D_x f_0\big(s,\xi',\mathcal{L}(\xi'),V\big) - D_x f_0\big(s,\xi,\mathcal{L}(\xi),V\big)\big)^{\top}(\xi'-\xi)\big] \geq -\lambda_m \|\xi'-\xi\|_2^2, \tag{3.20}$$
$$\forall \xi,\xi' \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n);$$

and there exists non-negative constant l_x, such that

$$\begin{aligned} &\big|D_x f_0(s,x,m,v') - D_x f_0(s,x,m,v)\big| \leq l_x|v'-v|, \\ &\big|D_v f_0(s,x',m',v) - D_v f_0(s,x,m,v)\big| \leq l_x\big(|x'-x| + W_2(m,m')\big). \end{aligned} \tag{3.21}$$

These parameters satisfy the following inequality condition:

$$2\lambda_x - \lambda_m \geq L_x + \frac{(L_v + 3l_x)^2}{4\lambda_v}. \tag{3.22}$$

In Condition 3.3, we see that the dependence of $D_x f_0$ in m can be large, and the f_1 part does not need to be monotonic. From (3.22), we can see that the strong convexity of f_1 in x allows f_0 to be quasi-monotonic in m. Condition 3.3 includes Condition 3.1 as

a special case. Actually, when f_0 is independent of v, then the parameters $l_x = 0$, and the monotonicity condition (3.20) for f_0 is reduced to (3.13). Furthermore, if f_1 does not functionally depend on m, then the parameters $L_x = L_v = 0$, and the convexity condition (3.15) for f_1 is reduced to (3.12). Therefore, in this case, Condition 3.3 is reduced to Condition 3.1. We also elaborate more on the condition (3.20). When the derivative $D_x f_0$ is differentiable in x and is linearly differentiable in m (although we do not need such regularity in this article), in view of (1.6), the inequality (3.20) is equivalent to the following one: for any $(s, V) \in [0, T] \times L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ and $\xi, \eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$,

$$\mathbb{E}\Big\{\eta^\top\Big[D_x^2 f_0\big(s,\xi,\mathcal{L}(\xi),V\big)\eta+\widetilde{\mathbb{E}}\Big(D_x D_y \frac{df_0}{dv}\big(s,\xi,\mathcal{L}(\xi),V\big)(\widetilde{\xi})\widetilde{\eta}\Big)+\lambda_m\eta\Big]\Big\}\geq 0. \tag{3.23}$$

In particular, when the derivatives $D_x^2 f_0$ and $D_x D_y \frac{df_0}{dv}$ do not functionally depend on v, then (3.23) is reduced to the displacement quasi-monotonicity condition as in (3.13). We also refer to [28, Definition 3.4] for a similar assumption as (3.23) on the Hamiltonian functional H.

We next show that Condition 3.3 yields our β-monotonicity in Condition 2.1 with $\Gamma_\beta = 0$, and then FBSDEs (1.3) are well posed.

THEOREM 3.5. *Under Assumptions* (A1)–(A3)*, suppose that the functions b_0 and σ_0 do not functionally depend on m, the terminal cost functional g satisfies the displacement monotonicity condition* (3.9) *and the running cost functional f satisfies Condition* 3.3*. Then, coefficients in* (3.6) *satisfy Condition* 2.1 (i) (a) *with the choice of β in* (3.7) *and the parameter $\Gamma_\beta = 0$. As a consequence, FBSDEs* (1.3) *have a unique global solution.*

PROOF. We still use the notations $\hat{v} := \hat{v}(s, X, \mathcal{L}(X), P, Q)$ and $\hat{v}' := \hat{v}(s, X', \mathcal{L}(X'), P', Q')$. From (3.11) and (3.19), we know that

$$\begin{aligned}
&\mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^\top(X'-X)\\
&\quad+\big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^\top(P'-P)\\
&\quad+\sum_{j=1}^n\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^\top(Q'^j-Q^j)\Big]\\
&=-\mathbb{E}\Bigg\{\Bigg[\begin{pmatrix}D_x f_1\\ D_v f_1\end{pmatrix}(s,\cdot)\,\Bigg|\,\begin{matrix}(X',\mathcal{L}(X'),\hat{v}')\\ (X,\mathcal{L}(X),\hat{v})\end{matrix}\Bigg]^\top\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix}\\
&\qquad+\Bigg[\begin{pmatrix}D_x f_0\\ D_v f_0\end{pmatrix}(s,\cdot)\,\Bigg|\,\begin{matrix}(X',\mathcal{L}(X'),\hat{v}')\\ (X,\mathcal{L}(X),\hat{v})\end{matrix}\Bigg]^\top\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix}\Bigg\}.
\end{aligned} \tag{3.24}$$

Since f_1 satisfies (3.14) and (3.15), from (3.18), we then obtain

$$\begin{aligned}(3.25)\quad &-\mathbb{E}\left\{\left[\begin{pmatrix}D_x f_1\\ D_v f_1\end{pmatrix}(s,\cdot)\,\middle|\,\begin{pmatrix}X',\mathcal{L}(X'),\hat{v}'\\ X,\mathcal{L}(X),\hat{v}\end{pmatrix}\right]^{\top}\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix}\right\}\\ &\le -2\mathbb{E}\big[\lambda_v|\hat{v}'-\hat{v}|^2+\lambda_x|X'-X|^2\big]+L_v\|X'-X\|_2\|\hat{v}'-\hat{v}\|_2+L_x\|X'-X\|_2^2.\end{aligned}$$

Since f_0 satisfies the displacement quasi-monotonicity condition (3.20), we have

$$(3.26)\quad -\mathbb{E}\big[\big(D_x f_0\big(s,X',\mathcal{L}(X'),\hat{v}\big)-D_x f_0\big(s,X,\mathcal{L}(X),\hat{v}\big)\big)^{\top}(X'-X)\big]\le\lambda_m\|X'-X\|_2^2.$$

As f_0 is convex in v, we know that

$$(3.27)\qquad -\mathbb{E}\big[\big(D_v f_0\big(s,X,\mathcal{L}(X),\hat{v}'\big)-D_v f_0\big(s,X,\mathcal{L}(X),\hat{v}\big)\big)^{\top}(\hat{v}'-\hat{v})\big]\le 0.$$

Combining (3.21) with an application of Cauchy–Schwarz inequality, we also have

$$\begin{aligned}(3.28)\quad &-\mathbb{E}\left\{\left(D_v f_0(s,\cdot,\hat{v}')\,\middle|\,\begin{pmatrix}X',\mathcal{L}(X'))\\ (X,\mathcal{L}(X)\end{pmatrix}\right)^{\top}(\hat{v}'-\hat{v})\right\}\le 2l_x\|X'-X\|_2\|\hat{v}'-\hat{v}\|_2,\\ &-\mathbb{E}\left\{\left(D_x f_0\big(s,X',\mathcal{L}(X'),\cdot\big)\,\middle|\,\begin{matrix}\hat{v}'\\ \hat{v}\end{matrix}\right)^{\top}(X'-X)\right\}\le l_x\|X'-X\|_2\|\hat{v}'-\hat{v}\|_2.\end{aligned}$$

From (3.26)–(3.28), we deduce that

$$\begin{aligned}(3.29)\quad &-\mathbb{E}\left\{\left[\begin{pmatrix}D_x f_0\\ D_v f_0\end{pmatrix}(s,\cdot)\,\middle|\,\begin{pmatrix}X',\mathcal{L}(X'),\hat{v}'\\ X,\mathcal{L}(X),\hat{v}\end{pmatrix}\right]^{\top}\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix}\right\}\\ &\le\lambda_m\|X'-X\|_2^2+3l_x\|X'-X\|_2\|\hat{v}'-\hat{v}\|_2.\end{aligned}$$

Combining (3.24), (3.25) and (3.29), from Young's inequality, we have

(3.30)

$$\begin{aligned}&\mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^{\top}(X'-X)\\ &\quad+\big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^{\top}(P'-P)\\ &\quad+\sum_{j=1}^{n}\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^{\top}(Q'^j-Q^j)\Big]\\ &\le-2\lambda_v\|\hat{v}'-\hat{v}\|_2^2-(2\lambda_x-\lambda_m-L_x)\|X'-X\|_2^2+(L_v+3l_x)\|X'-X\|_2\|\hat{v}'-\hat{v}\|_2\\ &\le-\lambda_v\|\hat{v}'-\hat{v}\|_2^2-\left(2\lambda_x-\lambda_m-L_x-\frac{(L_v+3l_x)^2}{4\lambda_v}\right)\|X'-X\|_2^2\\ &\le-\lambda_v\|\hat{v}'-\hat{v}\|_2^2,\end{aligned}$$

where the last inequality follows in light of $2\lambda_x - \lambda_m \geq L_x + \frac{(L_v+3l_x)^2}{4\lambda_v}$. Therefore, Condition 2.1 (i) (a) is valid with the choice of β in (3.7) for $\Lambda_\beta = \lambda_v$ and $\Gamma_\beta = 0$. ∎

In the proof of Theorem 3.5, we have shown that Condition 3.3 for the running cost functional f ensures that the β-monotonicity in Condition 2.1 holds with the parameter $\Gamma_\beta = 0$. As two particular cases of Condition 3.3, the classical displacement monotonicity [1, 2, 27, 28] and the small mean field effect [8] can both be viewed as a condition to guarantee this β-monotonicity. We also refer to [14, Remark 7.2] and [11, Section 2.1] for more detailed discussions on the relationship between the small mean field effect and displacement monotonicity condition. As a direct consequence of Theorems 3.1, 3.2 and 3.5, we obtain the following solvability of MFG (1.1)–(1.2).

COROLLARY 3.6. *Under Assumptions* (A1)–(A3), (3.3) *gives a local solution of MFG* (1.1)–(1.2). *Furthermore, under assumptions in Theorem* 3.5, (3.3) *is the unique global solution.*

So far, we focus on the convexity setting on the terminal functional g; actually, our stochastic control method can be applied to some more general cases beyond usual convex settings. To this point, we first introduce the recent work of [41], in which the authors use an analytical method to study the MFG master equation with σ being a constant and in the absence of the displacement monotonicity condition; particularly, they assume certain non-convexity of the terminal cost g. More precisely, g is assumed to satisfy an anti-monotonicity condition, while the Hamiltonian H takes the form $H(s, m, p) = \langle A_0 x, p\rangle + H_0(x, m, p)$, where H_0 satisfies appropriate regularity conditions; see [41, Assumption 6.1] for details.

We now give a viewpoint of the above conditions from the control theoretical perspective. In our previous work [8], we can see that with a stochastic control approach, the convexity of g is also not necessary for the local solvability of the FBSDEs associated with the MFG. Actually, even for the global solvability, the convexity and the displacement monotonicity condition of g are not necessary either; indeed, we may use the following condition on g: there exists some $l_g > 0$, such that for any ξ and ξ' on the same probability space,

$$\left|\mathbb{E}\left[\left(D_x g\left(\xi', \mathcal{L}(\xi')\right) - D_x g\left(\xi, \mathcal{L}(\xi)\right)\right)^\top (\xi' - \xi)\right]\right| \leq l_g \|\xi' - \xi\|_2^2, \tag{3.31}$$

which resembles the anti-monotonicity condition [41, (1.5)]. Here, we assume that the running cost functional f satisfies Condition 3.3, which allows the Hamiltonian H to be a general one; and due to the page limit, we only give a sketch of motivation for a simple case with σ being constant. In view of the continuation method, we should study the continuity of the solution of the following FBSDEs in initial $\xi \in \mathcal{L}^2_{\mathcal{F}_t}$, with

an input $(I^b, I^f, I^g) \in \mathcal{L}^2_{\mathcal{F}}(0,T) \times \mathcal{L}^2_{\mathcal{F}}(0,T) \times \mathcal{L}^2_{\mathcal{F}_T}$:

$$\begin{cases} X_s = \xi + \int_t^s \big[\gamma D_p H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big) + I^b_r\big]dr + \int_t^s \sigma dB_r, \\ P_s = \gamma D_x g\big(X_T, \mathcal{L}(X_T)\big) + I^g \\ \qquad + \int_s^T \big[\gamma D_x H\big(r, X_r, \mathcal{L}(X_r), P_r, Q_r\big) + I^f_r\big]dr - \int_s^T Q_r dB_r, \end{cases} \tag{3.32}$$

where $\gamma \in [0,1]$. For any two initials ξ^1 and ξ^2 and any two inputs $((I^b)^1, (I^f)^1, (I^g)^1)$ and $((I^b)^2, (I^f)^2, (I^g)^2)$, we denote by (X^1, P^1, Q^1, v^1) and (X^2, P^2, Q^2, v^2) the corresponding solutions of FBSDEs (3.32) and denote $\Delta X := X' - X$ (we also adopt similar notations for all other differences). Then, from (3.30), we know that

$$\begin{aligned} &\mathbb{E}\big[\gamma\big[D_x g\big(X^2_T, \mathcal{L}(X^2_T)\big) - D_x g\big(X^1_T, \mathcal{L}(X^1_T)\big)\big]^\top \Delta X_T + (\Delta I^g)^\top \Delta X_T - (\Delta P_t)^\top \Delta\xi\big] \\ &\quad \le \mathbb{E}\int_t^T \gamma\Big[-\lambda_v|\Delta v_s|^2 - \Big(2\lambda_x - \lambda_m - L_x - \frac{(L_v + 3l_x)^2}{4\lambda_v}\Big)|\Delta X_s|^2\Big]ds \\ &\qquad + \mathbb{E}\int_t^T \big(|\Delta P_s| \cdot |\Delta I^b_s| + |\Delta X_s| \cdot |\Delta I^f_s|\big)ds. \end{aligned}$$

Then, from (3.31), we know that

$$\begin{aligned} &\gamma\mathbb{E}\int_t^T \Big[\lambda_v|\Delta v_s|^2 + \Big(2\lambda_x - \lambda_m - L_x - \frac{(L_v + 3l_x)^2}{4\lambda_v}\Big)|\Delta X_s|^2\Big]ds \\ &\quad \le \mathbb{E}\Big[\gamma l_g|\Delta X_T|^2 + |\Delta P_t| \cdot |\Delta\xi| + |\Delta X_T| \cdot |\Delta I^g| \\ &\qquad + \int_t^T \big(|\Delta P_s| \cdot |\Delta I^b_s| + |\Delta X_s| \cdot |\Delta I^f_s|\big)ds\Big]. \end{aligned} \tag{3.33}$$

By applying Itô's formula on $|\Delta X_s|^2$ and then using Young's inequality, we can compute that

$$\begin{aligned} &\mathbb{E}\big[|\Delta X_T|^2\big] \\ &= \mathbb{E}\Big[|\Delta\xi|^2 + 2\gamma\int_t^T \big[(\Delta X_s)^\top b_1(s)\Delta X_s + (\Delta X_s)^\top b_2(s)\Delta v_s\big]ds \\ &\qquad\qquad + 2\int_t^T |\Delta I^b_s| \cdot |\Delta X_s|ds\Big] \\ &\le \mathbb{E}\big[|\Delta\xi|^2 + 2\gamma L\int_t^T \big(|\Delta X_s|^2 + |\Delta X_s| \cdot |\Delta v_s|\big)ds + 2\int_t^T |\Delta I^b_s| \cdot |\Delta X_s|ds\big] \\ &\le \mathbb{E}\Big[|\Delta\xi|^2 + \gamma A\int_t^T \Big[\lambda_v|\Delta v_s|^2 + \Big(2\lambda_x - \lambda_m - L_x - \frac{(L_v + 3l_x)^2}{4\lambda_v}\Big)|\Delta X_s|^2\Big]ds\Big] \end{aligned}$$

where

$$A := \frac{L\lambda_v + L\sqrt{\lambda_v^2 + \lambda_v\big(2\lambda_x - \lambda_m - L_x - \frac{(L_v+3l_x)^2}{4\lambda_v}\big)}}{\lambda_v\big(2\lambda_x - \lambda_m - L_x - \frac{(L_v+3l_x)^2}{4\lambda_v}\big)}.$$

Then, by substituting (3.33) into the above inequality, we know that

$$\begin{aligned}(1-\gamma A l_g)\mathbb{E}\big[|\Delta X_T|^2\big] &\le (A+2)\mathbb{E}\bigg[|\Delta\xi|^2 + |\Delta P_t|\cdot|\Delta\xi| + |\Delta X_T|\cdot|I^g| \\ &\quad + \int_t^T \big(|\Delta P_s|\cdot|\Delta I_s^b| + |\Delta X_s|\cdot|\Delta I_s^f| + |\Delta X_s|\cdot|\Delta I_s^b|\big)ds\bigg].\end{aligned}$$

If

$$Al_g < 1, \tag{3.34}$$

by noting that $\gamma \in [0,1]$, we have

$$\begin{aligned}\mathbb{E}\big[|\Delta X_T|^2\big] &\le C\mathbb{E}\bigg[|\Delta\xi|^2 + |\Delta P_t|\cdot|\Delta\xi| + |\Delta X_T|\cdot|I^g| \\ &\quad + \int_t^T \big(|\Delta P_s|\cdot|\Delta I_s^b| + |\Delta X_s|\cdot|\Delta I_s^f|\big)ds\bigg],\end{aligned}$$

where C is a constant depending only on $(L, \lambda_x, \lambda_v, \lambda_m, L_x, L_v, l_x, l_g)$ and it is independent of γ. Substituting the last estimate back into (3.33), we know that for any $\varepsilon > 0$,

$$\begin{aligned}&\gamma\mathbb{E}\int_t^T |\Delta v_s|^2 ds \\ &\quad\le C\mathbb{E}\bigg[|\Delta\xi|^2 + |\Delta P_t|\cdot|\Delta\xi| + |\Delta X_T|\cdot|\Delta I^g| \\ &\qquad + \int_t^T \big(|\Delta P_s|\cdot|\Delta I_s^b| + |\Delta X_s|\cdot|\Delta I_s^f|\big)ds\bigg] \\ &\quad\le \varepsilon\mathbb{E}\Big[\sup_{t\le s\le T}|\Delta X_s|^2 + \sup_{t\le s\le T}|\Delta P_s|^2\Big] \\ &\qquad + C\bigg(1+\frac{1}{\varepsilon}\bigg)\mathbb{E}\bigg[|\Delta\xi|^2 + |\Delta I^g|^2 + \int_t^T \big(|\Delta I_s^b|^2 + |\Delta I_s^f|^2\big)ds\bigg].\end{aligned} \tag{3.35}$$

With standard estimates for SDEs, we know that

$$\mathbb{E}\Big[\sup_{t\le s\le T}|\Delta X_s|^2\Big] \le C(L,T)\mathbb{E}\bigg[|\Delta\xi|^2 + \int_t^T \big(\gamma|\Delta v_s|^2 + |\Delta I_s^b|^2\big)ds\bigg] \tag{3.36}$$

and then, with (3.36) and standard estimates for BSDEs,

$$\mathbb{E}\Big[\sup_{t\le s\le T}|\Delta P_s|^2+\int_t^T|\Delta Q_s|^2ds\Big] \le C(L,T)\mathbb{E}\Big[|\Delta I^g|^2+\int_t^T\big(\gamma|\Delta v_s|^2+|\Delta I_s^f|^2\big)ds\Big]. \tag{3.37}$$

Substituting (3.36) and (3.37) into (3.35), we can see that by choosing ε small enough,

$$\gamma\mathbb{E}\int_t^T|\Delta v_s|^2ds\le C\mathbb{E}\Big[|\Delta\xi|^2+|\Delta I^g|^2+\int_t^T\big(|\Delta I_s^b|^2+|\Delta I_s^f|^2\big)ds\Big]. \tag{3.38}$$

From (3.36)–(3.38), we obtain that

$$\mathbb{E}\Big[\sup_{t\le s\le T}|\Delta X_s|^2+\sup_{t\le s\le T}|\Delta P_s|^2+\int_t^T|\Delta Q_s|^2ds\Big] \le C\mathbb{E}\Big[|\Delta\xi|^2+|\Delta I^g|^2+\int_t^T\big(|\Delta I_s^b|^2+|\Delta I_s^f|^2\big)ds\Big], \tag{3.39}$$

where the constant C is independent of γ. In view of the estimate (3.39), we can recursively enlarge the parameter γ in FBSDEs (3.32) by using the Banach fixed point theorem; since C is independent of γ, we can obtain the global solvability. This approach is typical in the method of continuation in coefficients proposed in [31]. The above sketchy argument will work as long as the condition (3.34) holds; this in turn allows a flexibility to free g from being convex, indeed; we also refer to [41, Theorem 7.1] for similar conditions. We shall provide more details in our future work.

4. Solvability of the MFTC problem

We consider the solvability of the MFTC problem (1.4). As in Section 3, we first give a sufficient condition of a maximum principle for the MFTC problem (1.4) and then show that the convexity assumption on f is actually the β-monotonicity condition (Condition 2.1 (i)) for the well-posedness of FBSDEs (1.5) in view of the β-monotonicity in Condition 2.1.

4.1. Maximum principle

We take the following assumptions on coefficients.

(B1) The functions b and σ are linear. That is,

$$b(s,x,m,v)=b_0(s)+b_1(s)x+b_2(s)v+b_3(s)\int_{\mathbb{R}^n}y\,m(dy),$$

$$\sigma(s,x,m,v)=\sigma_0(s)+\sigma_1(s)x+\sigma_2(s)v+\sigma_3(s)\int_{\mathbb{R}^n}y\,m(dy),$$

with all the norms of $b_0(s)$, $b_1(s)$, $b_2(s)$, $b_3(s)$, $\sigma_0(s)$, $\sigma_1(s)$, $\sigma_2(s)$, $\sigma_3(s)$ being bounded by L.

(B2) The functions f and g have a quadratic growth (3.1). The derivatives $D_x f$, $D_v f$, $D_y \frac{df}{dv}$, $D_x g$ and $D_y \frac{dg}{dv}$ exist, and they are continuous in all of their own arguments and satisfy

$$\begin{aligned}\big|(D_x, D_v) f(s,x',m',v') - (D_x, D_v) f(s,x,m,v)\big| &\le L\big(|x'-x| + |v'-v| + W_2(m,m')\big),\\ \big|D_x g(x',m') - D_x g(x',m')\big| &\le L\big(|x'-x| + W_2(m,m')\big),\end{aligned}$$

and for any square-integrable random variables ξ and ξ' on the same probability space,

$$\begin{aligned}&\mathbb{E}\left[\left|D_y \frac{df}{dv}\big(s,x',\mathcal{L}(\xi'),v'\big)(\xi') - D_y \frac{df}{dv}\big(s,x',\mathcal{L}(\xi),v'\big)(\xi)\right|^2\right]\\ &\qquad\le L\big(|x'-x|^2 + |v'-v|^2 + \|\xi'-\xi\|_2^2\big),\\ &\mathbb{E}\left[\left|D_y \frac{dg}{dv}\big(x',\mathcal{L}(\xi')\big)(\xi') - D_y \frac{dg}{dv}\big(x',\mathcal{L}(\xi)\big)(\xi)\right|^2\right] \le L\big(|x'-x|^2 + \|\xi'-\xi\|_2^2\big).\end{aligned}$$

(B3) The terminal cost functional g is convex, that is,

$$\begin{aligned}&g\big(x',\mathcal{L}(\xi')\big) - g\big(x,\mathcal{L}(\xi)\big)\\ &\qquad\ge \big(D_x g\big(x,\mathcal{L}(\xi)\big)\big)^\top (x'-x) + \mathbb{E}\left[\left(D_y \frac{dg}{dv}\big(x,\mathcal{L}(\xi)\big)(\xi)\right)^\top (\xi'-\xi)\right].\end{aligned} \tag{4.1}$$

And there exists $\lambda > 0$ such that for any square-integrable random variables ξ and ξ' on the same probability space,

$$\begin{aligned}&f\big(s,x',\mathcal{L}(\xi'),v'\big) - f\big(s,x,\mathcal{L}(\xi),v\big)\\ &\quad\ge \left[\begin{pmatrix} D_x f\\ D_v f\end{pmatrix}\big(s,x,\mathcal{L}(\xi),v\big)\right]^\top \begin{pmatrix} x'-x\\ v'-v\end{pmatrix}\\ &\qquad + \lambda|v'-v|^2 + \mathbb{E}\left[\left(D_y \frac{df}{dv}\big(s,x,\mathcal{L}(\xi),v\big)(\xi)\right)^\top (\xi'-\xi)\right].\end{aligned}$$

Under above assumptions, we have the following sufficient maximum principle for MFTC (1.4).

THEOREM 4.1. *Under Assumptions* (B1)–(B3)*, suppose that FBSDEs* (1.5) *have a uniqueness solution* $(X,P,Q) \in \mathcal{S}_{\mathcal{F}}^2(t,T) \times \mathcal{L}_{\mathcal{F}}^2(t,T) \times (\mathcal{L}_{\mathcal{F}}^2(t,T))^n$*. Then, the MFTC problem* (1.4) *has a unique optimal control* $v_s := \hat{v}(s, X_s, \mathcal{L}(X_s), P_s, Q_s)$, $s \in [t,T]$*, where the map* $\hat{v}$ *is defined in* (3.2)*.*

The proof of Theorem 4.1 is similar to that in [10] and is thus omitted here.

4.2. *Well-posedness of the forward-backward system*

We shall use Lemma 2.1 to give the well-posedness of our FBSDEs (1.5), which can be viewed as a subcase of FBSDEs (2.1) by setting

$$
\begin{aligned}
\mathbf{B}(s,X,P,Q) &:= D_p H\big(s,X,\mathcal{L}(X),P,Q\big),\\
\mathbf{A}(s,X,P,Q) &:= D_q H\big(s,X,\mathcal{L}(X),P,Q\big),\\
\mathbf{F}(s,X,P,Q) &:= -D_x H\big(s,X,\mathcal{L}(X),P,Q\big)\\
&\quad -\widetilde{\mathbb{E}}\Big[D_y \frac{dH}{dv}\big(s,\widetilde{X},\mathcal{L}(X),\widetilde{P},\widetilde{Q}\big)(X)\Big],\\
\mathbf{G}(X) &:= D_x g\big(X,\mathcal{L}(X)\big)+\widetilde{\mathbb{E}}\Big[D_y \frac{dg}{dv}\big(\widetilde{X},\mathcal{L}(X)\big)(X)\Big],
\end{aligned}
\tag{4.2}
$$

for $X, P \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$ and $Q \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{n\times n})$, where $\widetilde{X}$, $\widetilde{P}$ and $\widetilde{Q}$ are respectively independent copies of X, P and Q. We now show that Assumptions (B1)–(B3) can ensure the β-monotonicity in Condition 2.1 and then imply with the global well-posedness of FBSDEs (1.5). The choices of $\mathbf{F}$ and $\mathbf{G}$ in (4.2) are different from that in (3.6) due to the arithmetic difference between the backward equation in FBSDEs (1.5) and in FBSDEs (1.3) respectively, which is because the state process of the MFTC problem (1.4) depends simultaneously on the distribution of the current controlled state, while the state process of MFG (1.1)–(1.2) depends on the equilibrium distribution from the population.

THEOREM 4.2. *Under Assumptions* (B1)–(B3)*, coefficients in* (4.2) *satisfy Condition* 2.1 *with the map*

$$\beta(s,X,P,Q) := \hat{v}\big(s,X,\mathcal{L}(X),P,Q\big)$$

and the parameter $\Gamma_\beta = 0$*. As a consequence, FBSDEs* (1.5) *have a unique adapted solution.*

PROOF. Here, we check Condition 2.1 (i). From (4.2), Assumption (B1) and the definition of $\hat{v}(\cdot)$ in (3.2), we can write

$$
\begin{aligned}
\mathbf{B}(s,X,P,Q) &= b_0(s)+b_1(s)X+b_2(s)\hat{v}+b_3(s)\mathbb{E}[X],\\
\mathbf{A}^j(s,X,P,Q) &= \sigma_0^j(s)+\sigma_1^j(s)X+\sigma_2^j(s)\hat{v}+\sigma_3^j(s)\mathbb{E}[X],\\
\mathbf{F}(s,X,P,Q) &= -b_1(s)^\top P-\sum_{j=1}^n \sigma_1^j(s)^\top Q^j - D_x f\big(s,X,\mathcal{L}(X),\hat{v}\big)\\
&\quad - b_3(s)^\top\mathbb{E}[P]-\sum_{j=1}^n \sigma_3^j(s)^\top\mathbb{E}[Q^j]-\widetilde{\mathbb{E}}\Big[D_y\frac{df}{dv}\big(s,\widetilde{X},\mathcal{L}(X),\tilde{v}\big)(X)\Big],
\end{aligned}
$$

where $\hat{v} := \hat{v}(s, X, \mathcal{L}(X), P, Q)$ and $\tilde{\hat{v}}$ is an independent copy of $\hat{v}$; the map $\hat{v}(\cdot)$ is defined in (3.2). Then, we can compute that

$$
\begin{aligned}
(4.3)\quad & \mathbb{E}\Big[\big(\mathbf{F}(s,X',P',Q')-\mathbf{F}(s,X,P,Q)\big)^{\top}(X'-X) \\
&\quad + \big(\mathbf{B}(s,X',P',Q')-\mathbf{B}(s,X,P,Q)\big)^{\top}(P'-P) \\
&\quad + \sum_{j=1}^{n}\big(\mathbf{A}^j(s,X',P',Q')-\mathbf{A}^j(s,X,P,Q)\big)^{\top}(Q'^j-Q^j)\Big] \\
&= -\mathbb{E}\Bigg\{\Bigg[\begin{pmatrix} D_x f \\ D_v f\end{pmatrix}(s,\cdot)\,\Bigg|\,\begin{matrix}(X',\mathcal{L}(X'),\hat{v}')\\(X,\mathcal{L}(X),\hat{v})\end{matrix}\Bigg]^{\top}\begin{pmatrix}X'-X\\ \hat{v}'-\hat{v}\end{pmatrix} \\
&\quad + \widetilde{\mathbb{E}}\Big[D_y\frac{df}{dv}\big(s,\widetilde{X'},\mathcal{L}(X'),\widetilde{\hat{v}'}\big)(X')-D_y\frac{df}{dv}\big(s,\tilde{X},\mathcal{L}(X),\tilde{\hat{v}}\big)(X)\Big]^{\top}(X'-X)\Bigg\}.
\end{aligned}
$$

By Fubini's lemma, we know that

$$
\begin{aligned}
&\mathbb{E}\Big\{\widetilde{\mathbb{E}}\Big[D_y\frac{df}{dv}\big(s,\widetilde{X'},\mathcal{L}(X'),\widetilde{\hat{v}'}\big)(X')-D_y\frac{df}{dv}\big(s,\tilde{X},\mathcal{L}(X),\tilde{\hat{v}}\big)(X)\Big]^{\top}(X'-X)\Big\} \\
&= \mathbb{E}\Big\{\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{df}{dv}\big(s,X',\mathcal{L}(X'),\hat{v}'\big)(\widetilde{X'})-D_y\frac{df}{dv}\big(s,X,\mathcal{L}(X),\hat{v}\big)(\tilde{X})\Big)^{\top}(\widetilde{X'}-\tilde{X})\Big]\Big\};
\end{aligned}
$$

then, from (4.3) and the convexity of f in Assumption (B3), we know that Condition 2.1 (i) (a) is valid with the map

$$\beta(s,X,P,Q) := \hat{v}\big(s,X,\mathcal{L}(X),P,Q\big)$$

by setting $\Lambda_\beta = 2\lambda$ and $\Gamma_\beta = 0$. In a similar way, we can show that Condition 2.1 (i) (b) is also satisfied. ∎

5. The case for generic drift

So far we mainly focused on the case when b is linear in x and v. We next proceed on the more general setting such that b can be non-linear in x, v and m. We study both MFG (1.3) and the MFTC problem (1.4) with a generic b.

5.1. *Mean field games with generic drifts*

In this subsection, we shall give the solvability of MFG (1.1)–(1.2) without Assumption (A1); instead, we adopt the following assumptions on b and σ.

(A1′) The coefficient b grows linearly and is L-Lipschitz continuous in (x, v) and l_m-Lipschitz continuous in m; it is continuously differentiable in x and v, with the derivatives $D_x b$ and $D_v b$ being bounded by L. Moreover, for any $s \in [0, T]$, $x, x' \in \mathbb{R}^n$, $v, v' \in \mathbb{R}^d$ and $m, m' \in \mathcal{P}_2(\mathbb{R}^n)$,

$$\begin{aligned} (5.1)\quad &\big|(D_x b, D_v b)(s, x', m', v') - (D_x b, D_v b)(s, x, m, v)\big| \\ &\qquad \le \frac{L_b^x |x' - x| + L_b^v |v' - v| + L_b^m W_2(m, m')}{1 + |x| \vee |x'| + |v| \vee |v'| + W_2(m, \delta_0) \vee W_2(m', \delta_0)}; \end{aligned}$$

and there exists $\lambda_b > 0$ such that

$$(5.2)\qquad (D_v b)(D_v b)^\top(s, x, m, v) \ge \lambda_b I_n, \quad \forall (s, x, v) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^d.$$

The coefficient σ is linear in x as $\sigma(s, x) = \sigma_0(s) + \sigma_1(s)x$, with the norms of matrices σ_0 and σ_1 being bounded by L.

In the condition (5.1), if the coefficient b demanded is linear in x and v, then

$$L_b^x = L_b^v = L_b^m = 0.$$

Therefore, (A1′) extends the linear assumption on b in Assumption (A1). Our Assumption (A1′) is also used in [13, 14] for the solvability of first-order MFGs and MFTC problems with a generic b. We also refer to [28] for similar conditions to obtain the solvability of mean field game master equation with an analytical approach, but on the Hamiltonian functional H, not the individual assumptions on the drift functional b and on the cost functionals. In view of Assumption (A1′), the system of FBSDEs associated with the MFG is as follows:

$$(5.3)\quad \begin{cases} X_s = \xi + \displaystyle\int_t^s b\big(r, X_r, \mathcal{L}(X_r), v_r\big)dr + \int_t^s \big[\sigma_0(r) + \sigma_1(r)X_r\big]dB_r, \\ P_s = -\displaystyle\int_s^T Q_r dB_r + D_x g\big(X_T, \mathcal{L}(X_T)\big) \\ \qquad + \displaystyle\int_s^T \Big[D_x b\big(r, X_r, \mathcal{L}(X_r), v_r\big)^\top P_r \\ \qquad\qquad + \displaystyle\sum_{j=1}^n \big(\sigma_1^j(r)\big)^\top Q_r^j + D_x f\big(r, X_r, \mathcal{L}(X_r), v_r\big)\Big]dr, \end{cases}$$

subject to

$$(5.4)\qquad D_v b\big(s, X_s, \mathcal{L}(X_s), v_s\big)^\top P_s + D_v f\big(s, X_s, \mathcal{L}(X_s), v_s\big) = 0, \quad s \in [t, T].$$

We first give the corresponding maximum principle to hold for the MFG with Assumption (A1′).

THEOREM 5.1. *Under Assumptions* (A1′), (A2) *and* (A3), *suppose that* f *satisfies the convexity* (3.15) *with constants*

$$\lambda_x \geq \frac{L^2 L_b^x}{\lambda_b}, \quad \lambda_v > \frac{L^2 L_b^v}{\lambda_b}. \tag{5.5}$$

If FBSDEs (5.3)–(5.4) *have a unique solution*

$$(\bar{X}, \bar{P}, \bar{Q}, \bar{v}) \in \mathcal{S}^2_{\mathcal{F}}(t,T) \times \mathcal{S}^2_{\mathcal{F}}(t,T) \times \big(\mathcal{L}^2_{\mathcal{F}}(t,T)\big)^n \times \mathcal{L}^2_{\mathcal{F}}(t,T),$$

then $\bar{v}$ *is the unique solution of MFG* (1.1)–(1.2).

PROOF. We now prove the coercive property for J. For any control $v \in \mathcal{L}^2_{\mathcal{F}}(t,T)$, we denote by X^v the corresponding state. From Assumptions (A2) and the convexity of f and g, we have

$$\begin{aligned} &J\big(v; \mathcal{L}(\bar{X}_s), t \leq s \leq T\big) - J\big(\bar{v}; \mathcal{L}(\bar{X}_s), t \leq s \leq T\big) \\ &\geq \mathbb{E}\Bigg\{\int_t^T \Bigg[\Bigg[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\Bigg]^\top \begin{pmatrix} X_s^v - \bar{X}_s \\ v_s - \bar{v}_s \end{pmatrix} \\ &\qquad + \lambda_x |X_s^v - \bar{X}_s|^2 + \lambda_v |v_s - \bar{v}_s|^2\Bigg] ds + D_x g(\bar{X}_T)^\top (X_T^v - \bar{X}_T)\Bigg\}. \end{aligned} \tag{5.6}$$

By applying Itô's formula on $\bar{P}_s^\top (X_s^v - \bar{X}_s)$ and taking expectation, we have

$$\begin{aligned} &\mathbb{E}\big[D_x g(\bar{X}_T)^\top (X_T^v - \bar{X}_T)\big] \\ &= \mathbb{E}\Bigg[\int_t^T \bar{P}_s^\top \big[b\big(s, X_s^v, \mathcal{L}(\bar{X}_s), v_s\big) - b\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big] + \sum_{j=1}^n (\bar{Q}_s^j)^\top \sigma_1^j(s)(X_s^v - \bar{X}_s) \\ &\quad - \Bigg[D_x b\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)^\top \bar{P}_s + \sum_{j=1}^n \big(\sigma_1^j(s)\big)^\top \bar{Q}_s^j \\ &\qquad + D_x f\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\Bigg]^\top (X_s^v - \bar{X}_s) ds\Bigg] \\ &= \mathbb{E}\Bigg[\int_t^T \big[b\big(s, X_s^v, \mathcal{L}(\bar{X}_s), v_s\big) - b\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big) \\ &\quad - D_x b(s, \bar{X}_s, \bar{v}_s)(X_s^v - \bar{X}_s)\big]^\top \bar{P}_s - D_x f\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)^\top (X_s^v - \bar{X}_s) ds\Bigg]. \end{aligned}$$

From the first-order condition (5.4), we know that

$$D_v b\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)^\top \bar{P}_s + D_v f\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big) = 0, \tag{5.7}$$

and therefore,

$$
\begin{aligned}
(5.8)\qquad & \mathbb{E}\big[D_x g(\bar{X}_T)^\top (X_T^v - \bar{X}_T)\big] \\
&= \mathbb{E}\Bigg[\int_t^T \Bigg[b\big(s,\cdot,\mathcal{L}(\bar{X}_s),\cdot\big)\,\Big|\,\begin{matrix}(X_s^v, v_s)\\ (\bar{X}_s,\bar{v}_s)\end{matrix} \\
&\qquad - \big[(D_x b, D_v b)\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\big]\begin{pmatrix} X_s^v - \bar{X}_s \\ v_s - \bar{v}_s\end{pmatrix}\Bigg]^\top \bar{P}_s \\
&\qquad - \Bigg[\begin{pmatrix} D_x f \\ D_v f\end{pmatrix}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\Bigg]^\top \begin{pmatrix} X_s^v - \bar{X}_s \\ v_s - \bar{v}_s\end{pmatrix} ds\Bigg].
\end{aligned}
$$

In view of (5.7), the condition (5.2) and the *Schur complement*, we can compute that

$$
\bar{P}_s = -\big[\big((D_v b)^\top (D_v b)\big)^{-1}(D_v b)^\top\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\big](D_v f)\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)
$$

(also see the following Remark 5.1), and then, from (A2), we know that

$$
(5.9)\qquad |\bar{P}_s| \le \frac{L^2}{\lambda_b}\big[1 + |\bar{X}_s| + W_2\big(\mathcal{L}(\bar{X}_s),\delta_0\big) + |\bar{v}_s|\big].
$$

Again from (5.7), we know that

$$
\begin{aligned}
&\big[D_v f\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\cdot\big)\big|_0^{\bar{v}_s}\big]^\top \bar{v}_s \\
&\qquad = -\big[D_v b\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)^\top \bar{P}_s + D_v f\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),0\big)\big]^\top \bar{v}_s,
\end{aligned}
$$

and from the convexity assumption (3.15), we have

$$
\begin{aligned}
2\lambda_v|\bar{v}_s| &\le \big|D_v b\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\big|\cdot|\bar{P}_s| + \big|D_v f\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),0\big)\big| \\
&\le L\big(1 + |\bar{X}_s| + W_2\big(\mathcal{L}(\bar{X}_s),\delta_0\big) + |\bar{P}_s|\big).
\end{aligned}
$$

Therefore, we deduce from the condition (5.1) that

$$
\begin{aligned}
(5.10)\qquad & \Bigg|\mathbb{E}\Bigg[\int_t^T \Bigg[b\big(s,\cdot,\mathcal{L}(\bar{X}_s),\cdot\big)\,\Big|\,\begin{matrix}(X_s^v, v_s)\\ (\bar{X}_s,\bar{v}_s)\end{matrix} \\
&\qquad - \big[(D_x b, D_v b)\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\big]\begin{pmatrix} X_s^v - \bar{X}_s \\ v_s - \bar{v}_s\end{pmatrix}\Bigg]^\top \bar{P}_s\Bigg]\Bigg| \\
&\le \frac{L^2}{\lambda_b}\mathbb{E}\Bigg[\int_t^T \big(L_b^x|X_s^v - \bar{X}_s|^2 + L_b^v|v_s - \bar{v}_s|^2\big)\,ds\Bigg].
\end{aligned}
$$

Substituting (5.10) into (5.8), we know that

$$\mathbb{E}\left\{\int_t^T\left[\begin{pmatrix}D_x f\\ D_v f\end{pmatrix}(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s)\right]^\top\begin{pmatrix}X_s^v-\bar{X}_s\\ v_s-\bar{v}_s\end{pmatrix}ds+D_x g(\bar{X}_T)^\top(X_T^v-\bar{X}_T)\right\}$$
$$\geq -\frac{L^2}{\lambda_b}\mathbb{E}\left[\int_t^T\left(L_b^x|X_s^v-\bar{X}_s|^2+L_b^v|v_s-\bar{v}_s|^2\right)ds\right].$$

Combining the last inequality and (5.6), from (5.5), we know that

$$J(v)-J(\bar{v})\geq\left(\lambda_v-\frac{L^2L_b^v}{\lambda_b}\right)\mathbb{E}\left[\int_t^T|v_s-\bar{v}_s|^2ds\right],$$

from which we obtain the claimed coerciveness, which shows that $\bar{v}$ is a solution for MFG (1.1)–(1.2). And the proof for the uniqueness result is similar to that for Theorem 3.1. ∎

REMARK 5.1. In the proof of Theorem 5.1, the condition (5.9) is also known as a cone property (cone condition) which was first proposed in [13, 14] for the study on the first-order mean field theory. We would like to emphasize that the cone condition is satisfied under Assumptions (A1′), (A2) and (A3), which does not require additional assumption. We now show that Assumption (A1′) and Condition 3.3 can ensure the β-monotonicity corresponding to FBSDEs (5.3) and then guarantee the global well-posedness of FBSDEs (1.3).

THEOREM 5.2. *Under Assumptions (A1′), (A2) and (A3), suppose that the terminal cost functional g satisfies the displacement monotonicity condition (3.9), and the running cost functional f satisfies Condition 3.3. Then, when*

$$\lambda_v-2kL_b^v,\ 2\lambda_x-\lambda_m\geq L_x+\frac{2L^2l_m}{\lambda_b}+\frac{3L^3(L_b^x+L_b^m)l_m}{\lambda_b^2}+k(2L_b^x+L_b^m)$$
$$+\frac{1}{4\lambda_v}\left(L_v+3l_x+\frac{L^2l_m}{\lambda_b}+\frac{3L^3L_b^vl_m}{\lambda_b^2}+2kL_b^x\right)^2,\tag{5.11}$$

where

$$k:=\frac{L^2}{\lambda_b},\tag{5.12}$$

FBSDEs (1.3) have a unique global solution.

PROOF. We shall show that the coefficients of FBSDEs (5.3) satisfy Condition 2.1. For $s\in[t,T]$, $X,X',P,P'\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$, $Q,Q'\in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{n\times n})$ and

$V, V' \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ respectively satisfying

$$\begin{aligned} D_v b\big(s, X, \mathcal{L}(X), V\big)^\top P + D_v f\big(s, X, \mathcal{L}(X), V\big) = 0, \\ D_v b\big(s, X', \mathcal{L}(X'), V'\big)^\top P + D_v f\big(s, X', \mathcal{L}(X'), V'\big) = 0, \end{aligned} \tag{5.13}$$

we can compute that

(5.14)

$$\begin{aligned} &\mathbb{E}\Big\{\big(b\big(s, X', \mathcal{L}(X'), V'\big) - b\big(s, X, \mathcal{L}(X), V\big)\big)^\top (P' - P) \\ &\quad + \sum_{j=1}^n \big(\sigma_1^j(s)(X' - X)\big)^\top (Q'^j - Q^j) \\ &\quad - \Big[D_x b\big(s, X', \mathcal{L}(X'), V'\big)^\top P' + \sum_{j=1}^n \big(\sigma_1^j(s)\big)^\top Q'^j \\ &\qquad + D_x f\big(s, X', \mathcal{L}(X'), V'\big) - D_x b\big(s, X, \mathcal{L}(X), V\big)^\top P \\ &\qquad + \sum_{j=1}^n \big(\sigma_1^j(s)\big)^\top Q^j + D_x f\big(s, X, \mathcal{L}(X), V\big)\Big]^\top (X' - X)\Big\} \\ &= \mathbb{E}\Big\{ - \big[b\big(s, X, \mathcal{L}(X), V\big) - b\big(s, X', \mathcal{L}(X'), V'\big) \\ &\qquad - D_x b\big(s, X', \mathcal{L}(X'), V'\big)(X - X')\big]^\top (P' - P) \\ &\qquad - \big[\big(D_x b\big(s, X', \mathcal{L}(X'), V'\big) - D_x b\big(s, X, \mathcal{L}(X), V\big)\big)(X' - X)\big]^\top P \\ &\qquad - \big[D_x f\big(s, X', \mathcal{L}(X'), V'\big) - D_x f\big(s, X, \mathcal{L}(X), V\big)\big]^\top (X' - X)\Big\} \\ &= \mathbb{E}\Big\{\big[b\big(s, X, \mathcal{L}(X'), V\big) - b\big(s, X, \mathcal{L}(X), V\big)\big]^\top (P' - P) \\ &\qquad + \big[b\big(s, X', \mathcal{L}(X), V'\big) + b\big(s, X, \mathcal{L}(X'), V\big) \\ &\qquad - b\big(s, X, \mathcal{L}(X), V\big) - b\big(s, X', \mathcal{L}(X'), V'\big)\big]^\top P \\ &\qquad - \left[b\big(s, \cdot, \mathcal{L}(X'), \cdot\big) \,\middle|\, \begin{matrix}(X, V) \\ (X', V')\end{matrix} - \big[(D_x b, D_v b)\big(s, X', \mathcal{L}(X'), V'\big)\big] \begin{pmatrix} X - X' \\ V - V' \end{pmatrix} \right]^\top P' \\ &\qquad - \left[b\big(s, \cdot, \mathcal{L}(X), \cdot\big) \,\middle|\, \begin{matrix}(X', V') \\ (X, V)\end{matrix} - \big[(D_x b, D_v b)\big(s, X, \mathcal{L}(X), V\big)\big] \begin{pmatrix} X' - X \\ V' - V \end{pmatrix} \right]^\top P \\ &\qquad - \left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s, \cdot) \,\middle|\, \begin{matrix}\big(X', \mathcal{L}(X'), V'\big) \\ \big(X, \mathcal{L}(X), V\big)\end{matrix} \right]^\top \begin{pmatrix} X' - X \\ V' - V \end{pmatrix} \Big\}. \end{aligned}$$

From (5.13) and Assumptions (A1′) and (A2), we can compute that

$$
\begin{aligned}
&|P' - P| \\
&= \left| \left[\left((D_v b)^\top (D_v b) \right)^{-1} (D_v b)^\top \big(s, X', \mathcal{L}(X'), V'\big) \right] (D_v f)(s, \cdot) \;\middle|\; \begin{matrix} (X', \mathcal{L}(X'), V') \\ (X, \mathcal{L}(X), V) \end{matrix} \right| \\
&\quad + \left\{ \left[\left((D_v b)^\top (D_v b) \right)^{-1} (D_v b)^\top \right](s, \cdot) \;\middle|\; \begin{matrix} (X', \mathcal{L}(X'), V') \\ (X, \mathcal{L}(X), V) \end{matrix} \right\} (D_v f)\big(s, X, \mathcal{L}(X), V\big) \Big| \\
&\le \frac{L}{\lambda_b} \left| (D_v f)(s, \cdot) \;\middle|\; \begin{matrix} (X', \mathcal{L}(X'), V') \\ (X, \mathcal{L}(X), V) \end{matrix} \right| \\
&\quad + \frac{3L^2}{\lambda_b^2} \left| (D_v b)(s, \cdot) \;\middle|\; \begin{matrix} (X', \mathcal{L}(X'), V') \\ (X, \mathcal{L}(X), V) \end{matrix} \right| \cdot \left| (D_v f)\big(s, X, \mathcal{L}(X), V\big) \right| \\
&\le \left(\frac{L^2}{\lambda_b} + \frac{3L^3 L_b^x}{\lambda_b^2} \right) |X' - X| + \left(\frac{L^2}{\lambda_b} + \frac{3L^3 L_b^v}{\lambda_b^2} \right) |V' - V| \\
&\quad + \left(\frac{L^2}{\lambda_b} + \frac{3L^3 L_b^m}{\lambda_b^2} \right) W_2\big(\mathcal{L}(X), \mathcal{L}(X')\big),
\end{aligned}
$$

and therefore,

$$
\begin{aligned}
&\mathbb{E}\left\{ \left[b\big(s, X, \mathcal{L}(X'), V\big) - b\big(s, X, \mathcal{L}(X), V\big) \right]^\top (P' - P) \right\} \\
&\qquad \le \left(\frac{2L^2 l_m}{\lambda_b} + \frac{3L^3 (L_b^x + L_b^m) l_m}{\lambda_b^2} \right) \|X' - X\|_2^2 \\
&\qquad\quad + \left(\frac{L^2 l_m}{\lambda_b} + \frac{3L^3 L_b^v l_m}{\lambda_b^2} \right) \|V' - V\|_2 \cdot \|X' - X\|_2.
\end{aligned} \tag{5.15}
$$

As in (5.9), from (5.13), the condition (5.2) and Assumption (A2), we also have the cone properties

$$
\begin{aligned}
|P| &\le k\big(1 + |X| + W_2\big(\mathcal{L}(X), \delta_0\big) + |V|\big), \\
|P'| &\le k\big(1 + |X'| + W_2\big(\mathcal{L}(X'), \delta_0\big) + |V'|\big),
\end{aligned}
$$

where constant k is defined in (5.12); therefore, from the condition (5.1), we know that

$$
\begin{aligned}
&\mathbb{E}\Bigg\{ \Big[b\big(s, X', \mathcal{L}(X), V'\big) + b\big(s, X, \mathcal{L}(X'), V\big) \\
&\quad - b\big(s, X, \mathcal{L}(X), V\big) - b\big(s, X', \mathcal{L}(X'), V'\big) \Big]^\top P \\
&\quad - \left[b\big(s, \cdot, \mathcal{L}(X'), \cdot\big) \;\middle|\; \begin{matrix} (X, V) \\ (X', V') \end{matrix} - \big[(D_x b, D_v b)\big(s, X', \mathcal{L}(X'), V'\big) \big] \begin{pmatrix} X - X' \\ V - V' \end{pmatrix} \right]^\top P'
\end{aligned} \tag{5.16}
$$

$$
\begin{aligned}
&-\left[b\big(s,\cdot,\mathcal{L}(X),\cdot\big)\,\middle|\,\begin{matrix}(X',V')\\(X,V)\end{matrix}-\big[(D_x b, D_v b)\big(s,X,\mathcal{L}(X),V\big)\big]\begin{pmatrix}X'-X\\V'-V\end{pmatrix}\right]^{\top}P\Big\}\\
&\le kL_b^m\big(\|X'-X\|_2^2+\|V'-V\|_2\cdot\|X'-X\|_2\big)+2kL_b^x\|X'-X\|_2^2+2kL_b^v\|V'-V\|_2^2.
\end{aligned}
$$

Since f satisfies Condition 3.3, from the proof of Theorem 3.5 (also referring to (3.25), (3.29) and (3.30)), we know that

$$
\begin{aligned}
(5.17)\qquad &\mathbb{E}\Big\{-\left[\begin{pmatrix}D_x f\\D_v f\end{pmatrix}(s,\cdot)\,\middle|\,\begin{matrix}(X',\mathcal{L}(X'),V')\\(X,\mathcal{L}(X),V)\end{matrix}\right]^{\top}\begin{pmatrix}X'-X\\V'-V\end{pmatrix}\Big\}\\
&\qquad\le -2\lambda_v\|\hat{v}'-\hat{v}\|_2^2-(2\lambda_x-\lambda_m-L_x)\|X'-X\|_2^2\\
&\qquad\quad+(L_v+3l_x)\|X'-X\|_2\cdot\|V'-V\|_2.
\end{aligned}
$$

Substituting (5.15)–(5.17) back into (5.14), from Young's inequality and condition (5.2), we know that

$$
\begin{aligned}
&\text{the left-hand side of (5.14)}\\
&\le -(2\lambda_v-2kL_b^v)\|V'-V\|_2^2\\
&\quad+\left(L_v+3l_x+\frac{L^2l_m}{\lambda_b}+\frac{3L^3L_b^vl_m}{\lambda_b^2}+2kL_b^x\right)\|X'-X\|_2\cdot\|V'-V\|_2\\
&\quad-\left(2\lambda_x-\lambda_m-L_x-\frac{2L^2l_m}{\lambda_b}-\frac{3L^3(L_b^x+L_b^m)l_m}{\lambda_b^2}-k(2L_b^x+L_b^m)\right)\|X'-X\|_2^2\\
&\le -(\lambda_v-2kL_b^v)\|V'-V\|_2^2,
\end{aligned}
$$

from which we know that Condition 2.1 (i) (a) is valid with $\Lambda_\beta=\lambda_v-2kL_b^v$ and $\Gamma_\beta=0$. Since g satisfies the displacement monotonicity condition, from (3.10), we know that Condition 2.1 (i) (b) is satisfied. As a consequence of Lemma 2.1, we obtain the well-posedness of FBSDEs (5.3). ∎

REMARK 5.2. In the relation (5.11), the parameters are not the optimal, but we do not drill down into the details in this article. The condition (5.11) in Theorem 5.2 seems to be restrictive; however, when the coefficient b is linear in x and v and does not functionally depend on m, then $L_b^x=L_b^v=L_b^m=l_m=0$, and the conditions above are reduced to that in Condition 3.3.

As a direct consequence of Theorems 5.1 and 5.2, we obtain the following solvability of the MFG with the generic b.

COROLLARY 5.3. *Under assumptions in Theorem 5.2, MFG (1.1)–(1.2) has a unique global solution.*

The further Jacobian (and Hessian) flows for the FBSDEs and classical solutions to the associated master equations for MFGs under non-linear settings are studied in our new article [7].

5.2. *Relation with the mean field game HJB-FP system*

MFG (1.1)–(1.2) is associated with the following HJB-FP system:

$$
\begin{cases}
\partial_s v^{t,\mu}(s,x) \\
\quad + H\Big(s,x,m^{t,\mu}(s), D_x v^{t,\mu}(s,x), \dfrac{1}{2} D_x^2 v^{t,\mu}(s,x)\sigma\big(s,x,m^{t,\mu}(s)\big)\Big) = 0, \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad s \in [t,T), \\
\partial_s m^{t,\mu} + \operatorname{div}\big[D_p H\big(s,x,m^{t,\mu}(s), D_x v^{t,\mu}(s,x)\big) m^{t,\mu}\big] \\
\qquad\qquad - \displaystyle\sum_{i,j=1}^{n} \partial_{x_i}\partial_{x_j}\big[a_{ij}\big(s,x,m^{t,\mu}(s)\big) m^{t,\mu}\big] = 0, \quad s \in (t,T], \\
m^{t,\mu}(t) = \mu, \quad v^{t,\mu}(T,x) = g\big(x, m^{t,\mu}(T)\big),
\end{cases}
\tag{5.18}
$$

where $a_{ij}(s,x,m) := \frac{1}{2}(\sigma\sigma^\top)_{ij}(s,x,m)$. This HJB-FP system is related to our FBSDEs (1.3) in the following manner: suppose that $(X^{t,\xi}, P^{t,\xi}, Q^{t,\xi})$ is an adapted solution of FBSDEs (1.3) with initial condition $X_t^{t,\xi} = \xi \sim \mu$; then,

$$
m^{t,\mu}(s) = \mathcal{L}(X_s^{t,\xi}), \quad D_x v^{t,\mu}(s, X_s^{t,\xi}) = P_s^{t,\xi}, \quad s \in [t,T]. \tag{5.19}
$$

This relation (5.19) also gives a decoupling field of FBSDEs (1.3). The HJB-FP system (5.18) is widely studied in the contemporary literature; see [18, 20, 29, 37–39] for instance. In such works, σ is usually assumed to be constant or non-degenerate, and system (5.18) can be studied via the solution a parabolic equation. For the existence result for system (5.18) and also the solution regularity, the regularity of the functional Hamiltonian $H(s,x,m,p,q)$ is directly assumed. For the uniqueness result for system (5.18), some monotonicity conditions are required for H and g. The Lasry–Lions monotonicity condition:

$$
\int_{\mathbb{R}^n} \big[g(x,m') - g(x,m)\big](m' - m)(dx) \geq 0,
$$

which is introduced by Lasry and Lions [37, 38] and then used in the literature [20, 29], is now well known. Its relation with the displacement monotonicity condition is referred to in [1, 11, 28, 30].

In our work, to see the matter from the control perspective more clearly, we give conditions directly on coefficients b, σ and f, rather than on the induced functional

Hamiltonian H as in the above mentioned literature studying the HJB-FP system (5.18) via analytical methods. However, we can show that our assumptions on b, σ, f for the control method are related to those on H for the analytical method. For example, we show that the condition $\lambda_v > \frac{2L^2L_b^v}{\lambda_b}$ in Theorem 5.2 can guarantee the concavity for the Hamiltonian functional H in p; this concavity assumption is widely used in the existing literature, such as [18,20,28,29,36]. For convenience, we here prove the particular case when the dimension $n = d = 1$ and the coefficients b and f are twice differentiable in v although we do not need coefficients to be twice differentiable just for the sake of the well-posedness of the MFGs and MFTC problem. We have

$$\begin{aligned} D_p^2 H(s,x,m,p,q) &= D_p\big[b\big(s,x,m,\hat{v}(s,x,m,p)\big)\big] \\ &= D_v b\big(s,x,m,\hat{v}(s,x,m,p)\big) D_p \hat{v}(s,x,m,p). \end{aligned} \tag{5.20}$$

From the first-order optimality condition, we know that

$$D_v b\big(s,x,m,\hat{v}(s,x,m,p)\big)\, p + D_v f\big(s,x,m,\hat{v}(s,x,m,p)\big) = 0,$$

and by differentiating p in this last equation, we have

$$\begin{aligned} \big[D_v^2 b\big(s,x,m,\hat{v}(s,x,m,p)\big)\, p + D_v^2 f\big(s,x,m,\hat{v}(s,x,m,p)\big)\big] D_p \hat{v}(s,x,m,p) \\ = -D_v b\big(s,x,m,\hat{v}(s,x,m,p)\big). \end{aligned}$$

Then, in view of the condition (5.2), we know that

$$D_p \hat{v}(s,x,m,p) = \frac{|D_v b|^2}{(D_v^2 b)(D_v f) - (D_v^2 f)(D_v b)}\big(s,x,m,\hat{v}(s,x,m,p)\big);$$

we shall explain its well-posedness without exploding to infinity in the following. Substituting this last expression into (5.20), we can compute that

$$D_p^2 H(s,x,m,p) = \frac{|D_v b|^2}{\frac{(D_v b)(D_v^2 b)(D_v f)}{|D_v b|^2} - D_v^2 f}\big(s,x,m,\hat{v}(s,x,m,p)\big).$$

From Assumption (A1′), we know that

$$\left|\frac{(D_v b)(D_v^2 b)(D_v f)}{|D_v b|^2}\right| \le \frac{L^2 L_b^v}{\lambda_b},$$

and together with the convexity of f in v, the inequality $\lambda_v > \frac{2L^2L_b^v}{\lambda_b}$ yields that $D_p^2 H(s,x,m,p,q) < 0$, which implies that H is concave in p.

In the HJB-FP system (5.18), if we define

$$U(t,x,\mu) := v^{t,\mu}(t,x), \quad (t,x,\mu) \in [0,T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n),$$

then, from system (5.18), we can immediately see that it satisfies the following MFG

master equation:

$$
\begin{cases}
\partial_t U(t,x,\mu) + H\Big(t,x,\mu,D_x U(t,x,\mu),\dfrac{1}{2}D_x^2 U(t,x,\mu)\sigma(t,x,\mu)\Big) \\
+\displaystyle\int_{\mathbb{R}}^{n}\Big[D_p H\big(t,y,\mu,D_x U(t,y,\mu)\big)^{\top} D_y \frac{dU}{dv}(t,x,\mu)(y) \\
+\dfrac{1}{2}\operatorname{Tr}\Big((\sigma\sigma^{\top})(t,y,\mu)D_y^2\frac{dU}{dv}(t,x,\mu)(y)\Big)\Big]d\mu(y) = 0, \quad t\in[0,T), \\
U(T,x,\mu) = g(x,\mu), \quad \forall (x,\mu)\in\mathbb{R}^n\times\mathcal{P}_2(\mathbb{R}^n).
\end{cases}
\tag{5.21}
$$

The solution of the master equation (5.21) is a decoupling field of the mean field game HJB-FP system (5.18):

$$
v^{t,\mu}(s,x) := U\big(s,x,m^{t,\mu}(s)\big), \quad (s,x)\in[t,T]\times\mathbb{R}^n;
$$

we also refer to [8, 11, 36] for more discussions on the relations between the mean field game master equation and the HJB-FP system. The master equation (5.21) is widely studied in the existing literature; see [28,41] for an analytical method for well-posedness when σ is constant, and also see our previous work [8] for a probabilistic approach when b is linear. Note that in [40], an analytical approach is used to study the HJB-FP equations for MFGs under displacement monotonicity conditions on the cost function, assuming the special drift $b \equiv v$ and the constant diffusion σ. Our coefficients b and σ are fairly general. From [8, 11], we see that the classical solution of the master equation (5.21) or the HJB-FP equations (5.18) requires more restrictive assumptions and higher regularity on coefficients than just the solution of the MFG. Therefore, the stochastic control method in the paper allows us to include more cases and also to impose fewer regularity conditions on the coefficients, so as to guarantee the solvability of the original mean field problem.

5.3. Mean field type control problems with generic drifts

We next study the solvability of the MFTC problem (1.4) without Assumption (B1); instead, we take the following assumptions on b and σ:

(B1′) The coefficient b grows linearly and is L-Lipschitz continuous in (x,v,m); it is continuously differentiable in x, v and m, with the derivatives $D_x b$, $D_x b$ and $D_y \frac{db}{dv}$ being bounded by L, and they are L-Lipschitz continuous in (x,v,m). Moreover, for any $s\in[0,T]$, $x,x'\in\mathbb{R}^n$, $v,v'\in\mathbb{R}^d$ and square-integrable random variables ξ and ξ' on the same probability space,

$$
\begin{aligned}
&\Bigg| b(s,\cdot)\,\Bigg|\,\begin{matrix}(x',\mathcal{L}(\xi'),v')\\(x,\mathcal{L}(\xi),v)\end{matrix} - \big[(D_x,D_v)b\big(s,x,\mathcal{L}(\xi),v\big)\big]\begin{pmatrix}x'-x\\v'-v\end{pmatrix} \\
&- \mathbb{E}\Big[D_y\frac{db}{dv}\big(s,x,\mathcal{L}(\xi),v\big)(\xi)(\xi'-\xi)\Big]\Bigg| \le \frac{L_b^x|x'-x|^2 + L_b^v|v'-v|^2 + L_b^m\|\xi'-\xi\|_2^2}{1+|x|\vee|x'|+|v|\vee|v'|+\|\xi\|_2\vee\|\xi'\|_2},
\end{aligned}
$$

and there exists $\lambda_b > 0$, such that

$$(D_v b)(D_v b)^\top(s,x,m,v) \geq \lambda_b I_n, \quad \forall (s,x,v) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^d.$$

The coefficient σ is linear in (x,v) as $\sigma(s,x,m) = \sigma_0(s) + \sigma_1(s)x + \sigma_3(s)\int_{\mathbb{R}^n} y m(dy)$, with the norms of matrices σ_0, σ_1 and σ_3 being bounded by L.

REMARK 5.3. Assumption (B1′) is different from Assumption (A1′). This is because the state process of the MFTC problem (1.4) depends simultaneously on the distribution of the current controlled state, while the state process of MFG (1.1)–(1.2) depends on the population equilibrium distribution. In Assumption (B1′), when the coefficient b is linear in (x,m,v), then $L_b^x = L_b^v = L_b^m = 0$. Therefore, (B1′) is a non-linear extension of the linear assumption on b in Assumption (B1).

We also need the following strong convexity on f:

(B3′) The functional g satisfies the convexity (4.1), and there exist $\lambda_v > 0$ and $\lambda_x, \lambda_m \geq 0$, such that for any square-integrable random variables ξ and ξ' on the same probability space,

(5.22)
$$\begin{aligned} &f\big(s,x',\mathcal{L}(\xi'),v'\big) - f\big(s,x,\mathcal{L}(\xi),v\big) \\ &\geq \left[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}\big(s,x,\mathcal{L}(\xi),v\big)\right]^\top \begin{pmatrix} x'-x \\ v'-v \end{pmatrix} + \mathbb{E}\Big[\Big(D_y \frac{df}{dv}\big(s,x,\mathcal{L}(\xi),v\big)(\xi)\Big)^\top(\xi'-\xi)\Big] \\ &\quad + \lambda_v|v'-v|^2 + \lambda_x|x'-x|^2 + \lambda_m\|\xi'-\xi\|_2^2. \end{aligned}$$

In view of Assumption (B1′), the system of FBSDEs associated with the MFTC problem (1.4) are as follows:

(5.23)
$$\begin{cases} X_s = \xi + \displaystyle\int_t^s b\big(r,X_r,\mathcal{L}(X_r),v_r\big)dr + \int_t^s \big\{\sigma_0(r) + \sigma_1(r)X_r + \sigma_3(r)\mathbb{E}[X_r]\big\}dB_r, \\ P_s = -\displaystyle\int_s^T Q_r dB_r + D_x g\big(X_T,\mathcal{L}(X_T)\big)D_x g\big(X_T,\mathcal{L}(X_T)\big) \\ \qquad + \widetilde{\mathbb{E}}\Big[D_y \frac{dg}{dv}\big(\widetilde{X_T},\mathcal{L}(X_T)\big)(X_T)\Big] \\ \qquad + \displaystyle\int_s^T \Big\{D_x b\big(r,X_r,\mathcal{L}(X_r),v_r\big)^\top P_r + \sum_{j=1}^n \big(\sigma_1^j(r)\big)^\top Q_r^j \\ \qquad + D_x f\big(r,X_r,\mathcal{L}(X_r),v_r\big) + \widetilde{\mathbb{E}}\Big[\widetilde{P_r}^\top\Big(D_y \frac{db}{dv}\big(r,\widetilde{X_r},\mathcal{L}(X_r),\widetilde{v_r}\big)(X_r)\Big)\Big] \\ \qquad + \displaystyle\sum_{j=1}^n \big(\sigma_3^j(r)\big)^\top \mathbb{E}[Q_r^j] + \widetilde{\mathbb{E}}\Big[D_y \frac{df}{dv}\big(r,\widetilde{X_r},\mathcal{L}(X_r),\widetilde{v_r}\big)(X_r)\Big]\Big\}dr, \quad s \in [t,T], \end{cases}$$

with the condition

$$D_v b\big(s, X_s, \mathcal{L}(X_s), v_s\big)^\top P_s + D_v f\big(s, X_s, \mathcal{L}(X_s), v_s\big) = 0, \quad s \in [t, T]. \tag{5.24}$$

We now give the corresponding maximum principle for MFTC (1.4) with a generic b.

THEOREM 5.4. *Under Assumptions* (B1′), (B2) *and* (B3′), *suppose that*

$$\lambda_x + \lambda_m \geq k(L_b^x + L_b^m), \quad \lambda_v > kL_b^v, \tag{5.25}$$

where the constant k is defined in (5.12). *If FBSDEs* (5.23)–(5.24) *have a solution* $(\bar{X}, \bar{P}, \bar{Q}, \bar{v}) \in \mathcal{S}^2_{\mathcal{F}}(t,T) \times \mathcal{S}^2_{\mathcal{F}}(t,T) \times (\mathcal{L}^2_{\mathcal{F}}(t,T))^n \times \mathcal{L}^2_{\mathcal{F}}(t,T)$, *then, $\bar{v}$ is the unique optimal control of the MFTC problem* (1.4).

PROOF. Similar to the proof of Theorem 5.1, we estimate the difference $J(v) - J(\bar{v})$ for any control $v \in \mathcal{L}^2_{\mathcal{F}}(t,T)$. From Assumption (B3′), denoting by X^v the state process corresponding to a control v, we have

$$\begin{aligned}
&J(v) - J(\bar{v}) \\
&\geq \mathbb{E}\bigg\{\int_t^T \bigg[\bigg[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\bigg]^\top \begin{pmatrix} X_s^v - \bar{X}_s \\ v_s - \bar{v}_s \end{pmatrix} + \lambda_v |v_s - \bar{v}_s|^2 + \lambda_x |X_s^v - \bar{X}_s|^2 \\
&\quad + \widetilde{\mathbb{E}}\bigg[\bigg(D_y \frac{df}{d\nu}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big(\widetilde{\bar{X}_s}\big)\bigg)^\top \big(\widetilde{X_s^v} - \widetilde{\bar{X}_s}\big)\bigg] + \lambda_m \|\widetilde{X_s^v} - \widetilde{\bar{X}_s}\|_2^2\bigg] ds \\
&\quad + D_x g(\bar{X}_T)^\top (X_T^v - \bar{X}_T) + \widetilde{\mathbb{E}}\bigg[\bigg(D_y \frac{dg}{d\nu}\big(\bar{X}_T, \mathcal{L}(\bar{X}_T)\big)\big(\widetilde{\bar{X}_T}\big)\bigg)^\top \big(\widetilde{X_T^v} - \widetilde{\bar{X}_T}\big)\bigg]\bigg\}.
\end{aligned} \tag{5.26}$$

By applying Itô's formula to the product $\bar{P}_s^\top (X_s^v - \bar{X}_s)$ and then taking expectation, we have

$$\begin{aligned}
&\mathbb{E}\big[D_x g(\bar{X}_T)^\top (X_T^v - \bar{X}_T)\big] + \mathbb{E}\widetilde{\mathbb{E}}\bigg[\bigg(D_y \frac{dg}{d\nu}\big(\widetilde{\bar{X}_T}, \mathcal{L}(\bar{X}_T)\big)(\bar{X}_T)\bigg)^\top (X_T^v - \bar{X}_T)\bigg] \\
&= \mathbb{E}\bigg\{\int_t^T \bar{P}_s^\top \bigg[b(s,\cdot) \,\bigg|\, \begin{matrix} (X_s^v, \mathcal{L}(X_s^v), v_s) \\ (\bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s) \end{matrix}\bigg] + \sum_{j=1}^n (\bar{Q}_s^j)^\top \sigma_3^j(s) \widetilde{\mathbb{E}}[\widetilde{X_s^v} - \widetilde{\bar{X}_s}] \\
&\quad - \big[D_x b\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)^\top \bar{P}_s + D_x f\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big]^\top (X_s^v - \bar{X}_s) \\
&\quad - \widetilde{\mathbb{E}}\bigg[\widetilde{\bar{P}_s}^\top \bigg(D_y \frac{db}{d\nu}\big(s, \widetilde{\bar{X}_s}, \mathcal{L}(\bar{X}_s), \widetilde{\bar{v}_s}\big)(\bar{X}_s)\bigg)(X_s^v - \bar{X}_s)\bigg] \\
&\quad - \sum_{j=1}^n \widetilde{\mathbb{E}}\big[\widetilde{\bar{Q}_s^j}^\top \sigma_3^j(s)(X_s^v - \bar{X}_s)\big] - \widetilde{\mathbb{E}}\bigg[\bigg(D_y \frac{df}{d\nu}\big(s, \widetilde{\bar{X}_s}, \mathcal{L}(\bar{X}_s), \widetilde{\bar{v}_s}\big)(\bar{X}_s)\bigg)^\top (X_s^v - \bar{X}_s)\bigg] ds\bigg\}.
\end{aligned}$$

By Fubini's lemma, we know that

(i) $$\begin{aligned}&\mathbb{E}\widetilde{\mathbb{E}}\big[\big(D_y\tfrac{dg}{dv}(\widetilde{\bar{X}_T},\mathcal{L}(\bar{X}_T))(\bar{X}_T)\big)^\top(X_T^v-\bar{X}_T)\big]\\&=\mathbb{E}\widetilde{\mathbb{E}}\big[\big(D_y\tfrac{dg}{dv}(\bar{X}_T,\mathcal{L}(\bar{X}_T))(\widetilde{\bar{X}_T})\big)^\top(\widetilde{X_T^v}-\widetilde{\bar{X}_T})\big],\end{aligned}$$

(ii) $$\begin{aligned}&\mathbb{E}\widetilde{\mathbb{E}}\big[\widetilde{\bar{P}_s}^\top\big(D_y\tfrac{db}{dv}(s,\widetilde{\bar{X}_s},\mathcal{L}(\bar{X}_s),\widetilde{\bar{v}_s})(\bar{X}_s)\big)(X_s^v-\bar{X}_s)\big]\\&=\mathbb{E}\widetilde{\mathbb{E}}\big[\bar{P}_s^\top\big(D_y\tfrac{db}{dv}(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s)(\widetilde{\bar{X}_s})\big)(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\big],\end{aligned}$$

(iii) $$\mathbb{E}\widetilde{\mathbb{E}}\big[\widetilde{\bar{Q}_s^j}^\top\sigma_3^j(s)(X_s^v-\bar{X}_s)\big]=\mathbb{E}\widetilde{\mathbb{E}}\big[\bar{Q}_s^{j\top}\sigma_3^j(s)(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\big],$$

(iv) $$\begin{aligned}&\mathbb{E}\widetilde{\mathbb{E}}\big[\big(D_y\tfrac{df}{dv}(s,\widetilde{\bar{X}_s},\mathcal{L}(\bar{X}_s),\widetilde{\bar{v}_s})(\bar{X}_s)\big)^\top(X_s^v-\bar{X}_s)\big]\\&=\mathbb{E}\widetilde{\mathbb{E}}\big[\big(D_y\tfrac{df}{dv}(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s)(\widetilde{\bar{X}_s})\big)^\top(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\big].\end{aligned}$$

Therefore, we know that

$$\begin{aligned}&\mathbb{E}\Big\{D_xg(\bar{X}_T)^\top(X_T^v-\bar{X}_T)+\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{dg}{dv}\big(\bar{X}_T,\mathcal{L}(\bar{X}_T)\big)(\widetilde{\bar{X}_T})\Big)^\top(\widetilde{X_T^v}-\widetilde{\bar{X}_T})\Big]\Big\}\\&\quad=\mathbb{E}\Big\{\int_t^T\bar{P}_s^\top\Big[b(s,\cdot)\,\Big|\,\begin{matrix}(X_s^v,\mathcal{L}(X_s^v),v_s)\\(\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s)\end{matrix}\Big)-D_xb\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)(X_s^v-\bar{X}_s)\\&\qquad-\widetilde{\mathbb{E}}\Big[\bar{P}_s^\top\Big(D_y\frac{db}{dv}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)(\widetilde{\bar{X}_s})\Big)(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\Big]\Big]\\&\qquad-D_xf\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)^\top(X_s^v-\bar{X}_s)\\&\qquad-\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{df}{dv}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)(\widetilde{\bar{X}_s})\Big)^\top(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\Big]ds\Big\}.\end{aligned}$$

From (5.24), we know that

(5.27)
$$\begin{aligned}&\mathbb{E}\Big\{D_xg(\bar{X}_T)^\top(X_T^v-\bar{X}_T)+\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{dg}{dv}\big(\bar{X}_T,\mathcal{L}(\bar{X}_T)\big)(\widetilde{\bar{X}_T})\Big)^\top(\widetilde{X_T^v}-\widetilde{\bar{X}_T})\Big]\\&\quad+\int_t^T\Big[\begin{pmatrix}D_xf\\D_vf\end{pmatrix}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\Big]^\top\begin{pmatrix}X_s^v-\bar{X}_s\\v_s-\bar{v}_s\end{pmatrix}\\&\quad+\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{df}{dv}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)(\widetilde{\bar{X}_s})\Big)^\top(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\Big]ds\Big\}\\&=\mathbb{E}\Big\{\int_t^T\bar{P}_s^\top\Big[b(s,\cdot)\,\Big|\,\begin{matrix}(X_s^v,\mathcal{L}(X_s^v),v_s)\\(\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s)\end{matrix}\Big)-\big[(D_xb,D_vb)\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)\big]\begin{pmatrix}X_s^v-\bar{X}_s\\v_s-\bar{v}_s\end{pmatrix}\\&\qquad-\widetilde{\mathbb{E}}\Big[\Big(D_y\frac{db}{dv}\big(s,\bar{X}_s,\mathcal{L}(\bar{X}_s),\bar{v}_s\big)(\widetilde{\bar{X}_s})\Big)(\widetilde{X_s^v}-\widetilde{\bar{X}_s})\Big]\Big]ds\Big\}.\end{aligned}$$

In view of (5.24) and Assumption (B1′), we can compute that

$$\bar{P}_s = -\big[\big((D_v b)(D_v b)^\top\big)^{-1}(D_v b)\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big](D_v f)\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big),$$

and then, as in (5.9), we have the following cone property:

$$|\bar{P}_s| \le k\big(1 + |\bar{X}_s| + W_2\big(\mathcal{L}(\bar{X}_s), \delta_0\big) + |\bar{v}_s|\big), \tag{5.28}$$

where the constant k is defined in (5.12). Therefore, we deduce from Assumption (B1′) that

(5.29)

$$\begin{aligned}
&\bigg|\mathbb{E}\bigg\{\int_t^T \bar{P}_s^\top\bigg[b(s,\cdot)\,\bigg|\,\begin{pmatrix}X_s^v, \mathcal{L}(X_s^v), v_s\\ \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\end{pmatrix} - \big[(D_x b, D_v b)\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big]\begin{pmatrix}X_s^v - \bar{X}_s\\ v_s - \bar{v}_s\end{pmatrix}\\
&\quad - \widetilde{\mathbb{E}}\bigg[\bar{P}_s^\top\bigg(D_y\frac{db}{dv}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big(\widetilde{\bar{X}_s}\big)\big)\big(\widetilde{X_s^v} - \widetilde{\bar{X}_s}\big)\bigg]\bigg]ds\bigg\}\bigg|\\
&\le k\mathbb{E}\bigg\{\int_t^T \big[L_b^x|X_s^v - \bar{X}_s|^2 + L_b^v|v_s - \bar{v}_s|^2 + L_b^m\|\widetilde{X_s^v} - \widetilde{\bar{X}_s}\|_2^2\big]ds\bigg\}.
\end{aligned}$$

Substituting (5.29) into (5.27), we know that

$$\begin{aligned}
&\mathbb{E}\bigg\{D_x g(\bar{X}_T)^\top(X_T^v - \bar{X}_T) + \widetilde{\mathbb{E}}\bigg[\bigg(D_y\frac{dg}{dv}\big(\bar{X}_T, \mathcal{L}(\bar{X}_T)\big)\big(\widetilde{\bar{X}_T}\big)\bigg)^\top\big(\widetilde{X_T^v} - \widetilde{\bar{X}_T}\big)\bigg]\\
&\quad + \int_t^T\bigg[\begin{pmatrix}D_x f\\ D_v f\end{pmatrix}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\bigg]^\top\begin{pmatrix}X_s^v - \bar{X}_s\\ v_s - \bar{v}_s\end{pmatrix}ds\\
&\quad + \int_t^T\widetilde{\mathbb{E}}\bigg[\bigg(D_y\frac{df}{dv}\big(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{v}_s\big)\big(\widetilde{\bar{X}_s}\big)\bigg)^\top\big(\widetilde{X_s^v} - \widetilde{\bar{X}_s}\big)\bigg]ds\bigg\}\\
&\ge -k\mathbb{E}\bigg[\int_t^T (L_b^x + L_b^m)|X_s^v - \bar{X}_s|^2 + L_b^v|v_s - \bar{v}_s|^2 ds\bigg].
\end{aligned}$$

Combining the last inequality with (5.26), we know that

$$\begin{aligned}
&J(v) - J(\bar{v})\\
&\quad \ge \mathbb{E}\bigg\{\int_t^T\big[(\lambda_x + \lambda_m - k(L_b^x + L_b^m))|X_s^v - \bar{X}_s|^2 + (\lambda_v - kL_b^v)|v_s - \bar{v}_s|^2\big]\,ds\bigg\}\\
&\quad \ge (\lambda_v - kL_b^v)\mathbb{E}\bigg[\int_t^T |v_s - \bar{v}_s|^2 ds\bigg],
\end{aligned}$$

and from (5.25), we obtain the desired result. ∎

We next show that Assumption (B1′) and the strong convexity condition in (B3′) yield that the coefficients of FBSDEs (5.23) satisfy the β-monotonicity, which consequently implies the global well-posedness of FBSDEs (5.23).

THEOREM 5.5. *Under Assumptions* (B1′), (B2) *and* (B3′), *suppose that conditions in* (5.25) *hold; then FBSDEs* (5.23) *have a unique global solution.*

PROOF. We shall show that the coefficients of FBSDEs (5.23) satisfy Condition 2.1. For $s \in [t,T]$, $X, X', P, P' \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^n)$, $Q, Q' \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{n\times n})$ and $V, V' \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$ satisfying

$$\begin{aligned} D_v b\big(s,X,\mathcal{L}(X),V\big)^\top P + D_v f\big(s,X,\mathcal{L}(X),V\big) = 0,\\ D_v b\big(s,X',\mathcal{L}(X'),V'\big)^\top P + D_v f\big(s,X',\mathcal{L}(X'),V'\big) = 0, \end{aligned} \tag{5.30}$$

and by Fubini's lemma, we can compute that

(5.31)

$$\begin{aligned} &\mathbb{E}\Big\{\big[b(s,X',\mathcal{L}(X'),V') - b\big(s,X,\mathcal{L}(X),V\big)\big]^\top (P'-P)\\ &\quad + \sum_{j=1}^n \big[\sigma_1^j(s)(X'-X) + \sigma_3^j(s)\mathbb{E}[X'-X]\big]^\top (Q'^j - Q^j)\\ &\quad - \Big\{ D_x b\big(s,X',\mathcal{L}(X'),V'\big)^\top P' - D_x b\big(s,X,\mathcal{L}(X),V\big)^\top P\\ &\qquad + \widetilde{\mathbb{E}}\Big[\widetilde{P'}^\top \Big(D_y \frac{db}{dv}\big(s,\widetilde{X'},\mathcal{L}(X'),\widetilde{V'}\big)(X')\Big)\Big]\\ &\qquad - \widetilde{\mathbb{E}}\Big[\widetilde{P}^\top \Big(D_y \frac{db}{dv}\big(s,\widetilde{X},\mathcal{L}(X),\widetilde{V}\big)(X)\Big)\Big]\\ &\qquad + \sum_{j=1}^n \big(\sigma_1^j(s)\big)^\top (Q'^j - Q^j) + \sum_{j=1}^n \big(\sigma_3^j(s)\big)^\top \mathbb{E}[Q'^j - Q^j]\\ &\qquad + D_x f\big(s,X',\mathcal{L}(X'),V'\big) - D_x f\big(s,X,\mathcal{L}(X),V\big)\\ &\qquad + \widetilde{\mathbb{E}}\Big[D_y \frac{df}{dv}\big(s,\widetilde{X'},\mathcal{L}(X'),\widetilde{V'}\big)(X')\Big]\\ &\qquad - \widetilde{\mathbb{E}}\Big[D_y \frac{df}{dv}\big(s,\widetilde{X},\mathcal{L}(X),\widetilde{V}\big)(X)\Big]\Big\}^\top (X'-X)\Big\}\\ &= \mathbb{E}\Big\{ -\Big[b(s,\cdot)\,\Big|\,\begin{matrix}(X,\mathcal{L}(X),V)\\(X',\mathcal{L}(X'),V')\end{matrix}\Big] - D_x b\big(s,X',\mathcal{L}(X'),V'\big)(X-X') \end{aligned}$$

$$
\begin{aligned}
&- \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s, \widetilde{X'}, \mathcal{L}(X'), \widetilde{V'}\big)(X') \Big](X - X')\Big]^\top P' \\
&- \Big[b(s,\cdot) \,\Big|\, \begin{matrix} (X, \mathcal{L}(X), V) \\ (X', \mathcal{L}(X'), V') \end{matrix} - D_x b\big(s, X, \mathcal{L}(X), V\big)(X' - X) \\
&- \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s, \widetilde{X}, \mathcal{L}(X), \widetilde{V}\big)(X) \Big](X' - X)\Big]^\top P \\
&- \big(D_x f\big(s, X', \mathcal{L}(X'), V'\big) - D_x f\big(s, X, \mathcal{L}(X), V\big)\big)^\top (X' - X) \\
&- \widetilde{\mathbb{E}}\Big[D_y \frac{df}{dv}\big(s, \widetilde{X'}, \mathcal{L}(X'), \widetilde{V'}\big)(X') - D_y \frac{df}{dv}\big(s, \widetilde{X}, \mathcal{L}(X), \widetilde{V}\big)(X) \Big]^\top (X' - X) \Big\} \\
= \mathbb{E}\Big\{ &- \Big[b(s,\cdot) \,\Big|\, \begin{matrix} (X, \mathcal{L}(X), V) \\ (X', \mathcal{L}(X'), V') \end{matrix} - \big[(D_x b, D_v b)\big(s, X', \mathcal{L}(X'), V'\big)\big] \begin{pmatrix} X - X' \\ V - V' \end{pmatrix} \\
&- \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s, X', \mathcal{L}(X'), V'\big)(\widetilde{X'})(\widetilde{X} - \widetilde{X'}) \Big]\Big]^\top P' \\
&- \Big[b(s,\cdot) \,\Big|\, \begin{matrix} (X, \mathcal{L}(X), V) \\ (X', \mathcal{L}(X'), V') \end{matrix} - \big[(D_x b, D_v b)\big(s, X, \mathcal{L}(X), V\big)\big] \begin{pmatrix} X' - X \\ V' - V \end{pmatrix} \\
&- \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s, X, \mathcal{L}(X), V\big)(\widetilde{X})(\widetilde{X'} - \widetilde{X}) \Big]\Big]^\top P \\
&- \Big[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s,\cdot) \,\Big|\, \begin{matrix} (X', \mathcal{L}(X'), V') \\ (X, \mathcal{L}(X), V) \end{matrix} \Big]^\top \begin{pmatrix} X' - X \\ V' - V \end{pmatrix} \\
&- \widetilde{\mathbb{E}}\Big[\Big(D_y \frac{df}{dv}\big(s, X', \mathcal{L}(X'), V'\big)(\widetilde{X'}) - D_y \frac{df}{dv}\big(s, X, \mathcal{L}(X), V\big)(\widetilde{X}) \Big)^\top (\widetilde{X'} - \widetilde{X}) \Big] \Big\}.
\end{aligned}
$$

Similar to (5.28), from (5.30) and Assumptions (B1′) and (B2), we have

$$|P'| \le k\big(1 + |X'| + W_2\big(\mathcal{L}(X'), \delta_0\big) + |V'|\big).$$

Therefore, from Assumption (B1′), we arrive with the following:

$$
\begin{aligned}
(5.32)\quad \mathbb{E}\Big\{ &- \Big[b(s,\cdot) \,\Big|\, \begin{matrix} (X, \mathcal{L}(X), V) \\ (X', \mathcal{L}(X'), V') \end{matrix} - \big[(D_x b, D_v b)\big(s, X', \mathcal{L}(X'), V'\big)\big] \begin{pmatrix} X - X' \\ V - V' \end{pmatrix} \\
&- \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s, \widetilde{X'}, \mathcal{L}(X'), \widetilde{V'}\big)(X') \Big](X - X')\Big]^\top P' \Big\} \\
&\le k\mathbb{E}\big[(L_b^x + L_b^m)|X' - X|^2 + L_b^v |V' - V|^2\big].
\end{aligned}
$$

In a similar fashion, we also have

$$
\begin{aligned}
(5.33)\quad \mathbb{E}\Bigg\{ & -\Big[b(s,\cdot) \,\Big|\, \begin{matrix} (X,\mathcal{L}(X),V) \\ (X',\mathcal{L}(X'),V') \end{matrix} \Big] - \big[(D_x b, D_v b)(s,X,\mathcal{L}(X),V)\big] \begin{pmatrix} X'-X \\ V'-V \end{pmatrix} \\
& - \widetilde{\mathbb{E}}\Big[D_y \frac{db}{dv}\big(s,\widetilde{X},\mathcal{L}(X),\widetilde{V}\big)(X) \Big](X'-X) \Big]^{\top} P \Bigg\} \\
& \le k\mathbb{E}\big[(L_b^x + L_b^m)|X'-X|^2 + L_b^v |V'-V|^2\big].
\end{aligned}
$$

Since f satisfies Assumption (B3′), we know that

$$
\begin{aligned}
(5.34)\quad \mathbb{E}\Bigg\{ & -\Bigg[\begin{pmatrix} D_x f \\ D_v f \end{pmatrix}(s,\cdot) \,\Bigg|\, \begin{matrix} (X',\mathcal{L}(X'),V') \\ (X,\mathcal{L}(X),V) \end{matrix} \Bigg]^{\top} \begin{pmatrix} X'-X \\ V'-V \end{pmatrix} \\
& - \widetilde{\mathbb{E}}\Big[\Big(D_y \frac{df}{dv}\big(s,X',\mathcal{L}(X'),V'\big)(\widetilde{X'}) - D_y \frac{df}{dv}\big(s,X,\mathcal{L}(X),V\big)(\widetilde{X}) \Big)^{\top} (\widetilde{X'}-\widetilde{X}) \Big] \Bigg\} \\
& \le -2\lambda_v \|V'-V\|_2^2 - 2(\lambda_x+\lambda_m)\|X'-X\|_2^2.
\end{aligned}
$$

Substituting (5.32)–(5.34) back into (5.31), we conclude with the following:

the left-hand side of (5.31)

$$
\le -2(\lambda_v - kL_b^v)\|V'-V\|_2^2 - 2\big[\lambda_x + \lambda_m - k(L_b^x + L_b^m)\big]\|X'-X\|_2^2,
$$

from which we know that Condition 2.1 (i) (a) is valid with $\Lambda_\beta = 2\lambda_v - 2kL_b^v$ and $\Gamma_\beta = 0$. In a similar way, we can show that Condition 2.1 (i) (b) is also satisfied. Therefore, as a consequence of Lemma 2.1, we obtain the well-posedness of FBSDEs (5.23). ∎

As an immediate consequence of Theorems 5.4 and 5.5, we obtain the following solvability of the MFTC problem (1.4) with a generic b.

COROLLARY 5.6. *Under assumptions in Theorem 5.5, the MFTC problem* (1.4) *has a unique optimal control.*

ACKNOWLEDGMENTS. – Phillip Yam thanks University of Texas at Dallas for the kind invitation to be a Visiting Professor in the Naveen Jindal School of Management during his sabbatical leave.

FUNDING. – Alain Bensoussan is supported by the National Science Foundation under grant NSF-DMS-2204795. Ziyu Huang acknowledges the financial supports as a postdoctoral fellow from Department of Statistics and Data Science of The Chinese University of Hong Kong. Shanjian Tang is supported by the National Natural Science Foundation of China under grant nos. 12031009 and 11631004. Phillip Yam acknowledges the financial supports from HKGRF-14301321 with the project title

"General Theory for Infinite Dimensional Stochastic Control: Mean Field and Some Classical Problems", and HKGRF-14300123 with the project title "Well-posedness of Some Poisson-driven Mean Field Learning Models and their Applications". The work described in this article was also supported by a grant from the Germany/Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the German Academic Exchange Service of Germany (Reference No. G-CUHK411/23).

References

[1] S. Ahuja, Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM J. Control Optim.* **54** (2016), no. 1, 30–48. Zbl 1327.93403 MR 3439756

[2] S. Ahuja – W. Ren – T.-W. Yang, Forward-backward stochastic differential equations with monotone functionals and mean field games with common noise. *Stochastic Process. Appl.* **129** (2019), no. 10, 3859–3892. Zbl 1479.60110 MR 3997664

[3] L. Ambrosio – N. Gigli – G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. Lectures Math. ETH, Birkhäuser, Basel, 2005. Zbl 1090.35002 MR 2129498

[4] A. Bensoussan – J. Frehse – P. Yam, *Mean field games and mean field type control theory*. SpringerBriefs Math., Springer, New York, 2013. Zbl 1287.93002 MR 3134900

[5] A. Bensoussan – P. J. Graber – S. C. P. Yam, Control on Hilbert spaces and application to some mean field type control problems. *Ann. Appl. Probab.* **34** (2024), no. 4, 4085–4136. Zbl 1554.35339 MR 4783038

[6] A. Bensoussan – Z. Huang – S. Tang – S. C. P. Yam, Degenerate mean field type control with linear and unbounded diffusion, and their associated equations. *SIAM J. Math. Anal.* **57** (2025), no. 5, 5494–5530. MR 4959955

[7] A. Bensoussan – Z. Huang – S. Tang – S. C. P. Yam, Second order fully nonlinear mean field games with degenerate diffusions. 2025, arXiv:2503.16869v1.

[8] A. Bensoussan – Z. Huang – S. Tang – S. C. P. Yam, A class of degenerate mean field games, associated FBSDEs and master equations. [v1] 2024, [v2] 2025, arXiv:2410.12404v2.

[9] A. Bensoussan – Z. Huang – S. C. P. Yam, Control theory on Wasserstein space: a new approach to optimality conditions. *Ann. Math. Sci. Appl.* **8** (2023), no. 3, 565–628. Zbl 1532.35460 MR 4667976

[10] A. Bensoussan – Z. Huang – S. C. P. Yam, Maximum principle for mean field type control problems with general volatility functions. *Int. Game Theory Rev.* **26** (2024), no. 2, article no. 2440003. Zbl 08050102 MR 4760578

[11] A. Bensoussan – H. M. Tai – T. K. Wong – S. C. P. Yam, A control theoretical approach to mean field games: Part I—Global equilibrium solution. *Appl. Math. Optim.* **92** (2025), no. 2, article no. 44. Zbl 08104602 MR 4965816

[12] A. Bensoussan – H. M. Tai – S. C. P. Yam, Mean field type control problems, some Hilbert-space-valued FBSDEs, and related equations. *ESAIM Control Optim. Calc. Var.* **31** (2025), article no. 33. Zbl 08027667 MR 4885886

[13] A. Bensoussan – T. K. Wong – S. C. P. Yam – H. Yuan, A theory of first order mean field type control problems and their equations. *J. Eur. Math. Soc.* (to appear); arXiv:2305.11848v2.

[14] A. Bensoussan – T. K. Wong – S. C. P. Yam – H. Yuan, Global well-posedness of first-order mean field games and master equations with nonlinear dynamics. 2023, arXiv:2311.11896v1.

[15] A. Bensoussan – S. C. P. Yam, Control problem on space of random variables and master equation. *ESAIM Control Optim. Calc. Var.* **25** (2019), article no. 10. Zbl 1450.35305 MR 3943358

[16] A. Briani – P. Cardaliaguet, Stable solutions in potential mean field game systems. *NoDEA Nonlinear Differential Equations Appl.* **25** (2018), no. 1, article no. 1. Zbl 1390.35365 MR 3735559

[17] R. Buckdahn – J. Li – S. Peng – C. Rainer, Mean-field stochastic differential equations and associated PDEs. *Ann. Probab.* **45** (2017), no. 2, 824–878. Zbl 1402.60070 MR 3630288

[18] P. Cardaliaguet, Notes on mean field games. 2010, from P.-L. Lions' lectures at Collège de France.

[19] P. Cardaliaguet – M. Cirant – A. Porretta, Splitting methods and short time existence for the master equations in mean field games. *J. Eur. Math. Soc. (JEMS)* **25** (2023), no. 5, 1823–1918. Zbl 1522.35511 MR 4592861

[20] P. Cardaliaguet – F. Delarue – J.-M. Lasry – P.-L. Lions, *The master equation and the convergence problem in mean field games*. Ann. of Math. Stud. 201, Princeton University Press, Princeton, NJ, 2019. Zbl 1430.91002 MR 3967062

[21] P. Cardaliaguet – P. J. Graber – A. Porretta – D. Tonon, Second order mean field games with degenerate diffusion and local coupling. *NoDEA Nonlinear Differential Equations Appl.* **22** (2015), no. 5, 1287–1317. Zbl 1344.49061 MR 3399179

[22] R. Carmona – F. Delarue, Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.* **43** (2015), no. 5, 2647–2700. Zbl 1322.93103 MR 3395471

[23] R. Carmona – F. Delarue, *Probabilistic theory of mean field games with applications. I. Mean field FBSDEs, control, and games*. Probab. Theory Stoch. Model. 83, Springer, Cham, 2018. Zbl 1422.91014 MR 3752669

[24] R. Carmona – F. Delarue, *Probabilistic theory of mean field games with applications. II. Mean field games with common noise and master equations*. Probab. Theory Stoch. Model. 84, Springer, Cham, 2018. Zbl 1422.91015 MR 3753660

[25] 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.* **280** (2022), no. 1379, v+123. Zbl 1520.91003 MR 4493576

[26] M. F. Djete – D. Possamaï – X. Tan, McKean-Vlasov optimal control: the dynamic programming principle. *Ann. Probab.* **50** (2022), no. 2, 791–833. Zbl 1491.49018 MR 4399164

[27] W. Gangbo – A. R. Mészáros, Global well-posedness of master equations for deterministic displacement convex potential mean field games. *Comm. Pure Appl. Math.* **75** (2022), no. 12, 2685–2801. Zbl 1510.35342 MR 4509653

[28] W. Gangbo – A. R. Mészáros – C. Mou – J. Zhang, Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity. *Ann. Probab.* **50** (2022), no. 6, 2178–2217. Zbl 1501.35403 MR 4499277

[29] D. A. Gomes – E. A. Pimentel – V. Voskanyan, *Regularity theory for mean-field game systems*. SpringerBriefs Math., Springer, Cham, 2016. Zbl 1391.91003 MR 3559742

[30] P. J. Graber – A. R. Mészáros, On monotonicity conditions for mean field games. *J. Funct. Anal.* **285** (2023), no. 9, article no. 110095. Zbl 1519.91034 MR 4620724

[31] Y. Hu – S. Peng, Solution of forward-backward stochastic differential equations. *Probab. Theory Related Fields* **103** (1995), no. 2, 273–283. Zbl 0831.60065 MR 1355060

[32] M. Huang – P. E. Caines – R. P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automat. Control* **52** (2007), no. 9, 1560–1571. Zbl 1366.91016 MR 2352434

[33] M. Huang – R. P. Malhamé – P. E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006), no. 3, 221–251. Zbl 1136.91349 MR 2346927

[34] Z. Huang – S. Tang, Mean field games with common noises and conditional distribution dependent FBSDEs. *Chinese Ann. Math. Ser. B* **43** (2022), no. 4, 523–548. Zbl 1500.91014 MR 4487652

[35] Z. Huang – S. Tang, State-density flows of non-degenerate density-dependent mean field SDEs and associated PDEs. *Trans. Amer. Math. Soc.* **376** (2023), no. 3, 1553–1599. Zbl 1532.60149 MR 4549685

[36] Z. Huang – S. Tang, Fully coupled nonlocal quasilinear forward-backward parabolic equations arising from mean field games. [v1] 2021, [v3] 2022, arXiv:2105.13213v3.

[37] J.-M. Lasry – P.-L. Lions, Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 9, 619–625. Zbl 1153.91009 MR 2269875

[38] J.-M. Lasry – P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 10, 679–684. Zbl 1153.91010 MR 2271747

[39] J.-M. Lasry – P.-L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007), no. 1, 229–260. Zbl 1156.91321 MR 2295621

[40] A. R. Mészáros – C. Mou, Mean field games systems under displacement monotonicity. *SIAM J. Math. Anal.* **56** (2024), no. 1, 529–553. Zbl 08097053 MR 4686386

[41] C. Mou – J. Zhang, Mean field game master equations with anti-monotonicity conditions. *J. Eur. Math. Soc. (JEMS)* **27** (2025), no. 11, 4469–4499. Zbl 08101463 MR 4956554

[42] C. Orrieri – A. Porretta – G. Savaré, A variational approach to the mean field planning problem. *J. Funct. Anal.* **277** (2019), no. 6, 1868–1957. Zbl 1423.49044 MR 3985520

[43] S. Peng – Z. Wu, Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.* **37** (1999), no. 3, 825–843. Zbl 0931.60048 MR 1675098

[44] H. Pham – X. Wei, Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.* **55** (2017), no. 2, 1069–1101. Zbl 1361.93069 MR 3631380

[45] A. Porretta, Weak solutions to Fokker-Planck equations and mean field games. *Arch. Ration. Mech. Anal.* **216** (2015), no. 1, 1–62. Zbl 1312.35168 MR 3305653

[46] J. Yong – X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*. Appl. Math. (N.Y.) 43, Springer, New York, 1999. Zbl 0943.93002 MR 1696772

Received 9 December 2024,
and in revised form 23 July 2025

Alain Bensoussan
International Center for Decision and Risk Analysis, Naveen Jindal School of Management, University of Texas at Dallas
Richardson, TX 75080, USA
axb046100@utdallas.edu

Ziyu Huang
Department of Statistics and Data Science, The Chinese University of Hong Kong
Shatin, N.T., Hong Kong SAR, P. R. China
zyhuang19@fudan.edu.cn

Shanjian Tang
Department of Finance and Control Sciences, School of Mathematical Sciences, Fudan University
200433 Shanghai, P. R. China
sjtang@fudan.edu.cn

Sheung Chi Phillip Yam
Department of Statistics and Data Science, The Chinese University of Hong Kong
Shatin, N.T., Hong Kong SAR, P. R. China
scpyam@sta.cuhk.edu.hk