Mean Field Games with Mean-Field-Dependent Volatility, and Associated Fully Coupled Nonlocal Quasilinear Forward-Backward Parabolic Equations

Ziyu Huang∗ Shanjian Tang†

Abstract

In this paper, we study mean field games with mean-field-dependent volatility, and associated fully coupled nonlocal quasilinear forward-backward PDEs (FBPDEs). We show the global in time existence of classical solutions of the FBPDEs in space $C^{1+\frac{1}{4},2+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$, and also the uniqueness under an additional monotonicity condition. We prove a verification theorem and use the solution of the PDEs to give an optimal strategy of the mean field game. Finally, we study the linear-quadratic problems to illustrate the role of our main results.

Keywords mean field games, forward-backward parabolic equations, quasilinear PDEs, mean field equations.

2000 MR Subject Classification 93E20, 60H30, 35K55, 35K10

1 Introduction

In this paper, we consider the following forward-backward parabolic equations

$$\begin{cases}
\frac{\partial u}{\partial t}(t, x) + \sum_{i,j=1}^{n} a_{ij}(t, x, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) \\
\quad + \mathcal{H}(t, x, m(t, \cdot), Du(t, x)) = 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
\frac{\partial m}{\partial t}(t, x) - \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}[a_{ij}(t, x, m(t, \cdot))m(t, x)] \\
\quad + \text{div}[\frac{\partial \mathcal{H}}{\partial p}(t, x, m(t, \cdot), Du(t, x))m(t, x)] = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n; \\
m(0, x) = m_0(x), \quad u(T, x) = g(x, m(T, \cdot)), \quad x \in \mathbb{R}^n,
\end{cases}$$

(1)

It is a system of fully coupled nonlocal quasilinear forward-backward partial differential equations (FBPDEs). One use of studying FBPDEs (1) is, by setting

$$\mathcal{H}(t, x, m, p) := H(t, x, m, \phi(t, x, m, p), p), \quad (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n,$$

1School of Mathematical Sciences, Fudan University, Shanghai 200433, China. E-mail: zyhuang19@fudan.edu.cn.

†Department of Finance and Control Sciences, School of Mathematical Sciences, Fudan University, Shanghai 200433, China. E-mail: sjtang@fudan.edu.cn. Research partially supported by National Natural Science Foundation of China (Grants No. 11631004 and No. 12031009).
where $H$ is the Hamiltonian defined as
\[
H(t, x, m, \alpha, p) := \langle p, b(t, x, m, \alpha) \rangle + f(t, x, m, \alpha),
\]
\[(t, x, m, \alpha, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \times \mathbb{R}^n,
\]
and $\phi$ is the feedback function defined as
\[
\phi(t, x, m, p) := \operatorname{argmin}_{\alpha \in \mathbb{R}^d} H(t, x, m, \alpha, p), \quad (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n,
\]
we can obtain a classical solution of mean field equations (MFEs), which is a system of forward-backward parabolic equations, where the backward one is a Hamilton-Jacobi-Bellman (HJB) equation and the forward one is a Fokker-Planck (FP) equation. A classical solution of such MFEs gives an optimal control (as a feedback form) of the following mean field game (MFG)
\[
\begin{align*}
\dot{\alpha} & \in \operatorname{argmin}_{\alpha} J(\alpha|\hat{m}) := \mathbb{E} \left[ \int_0^T f(t, X^\alpha_t, \hat{m}_t, \alpha_t)dt + g(X^\alpha_T, \hat{m}_T) \right]; \\
X^\alpha_t & = \xi_0 + \int_0^t b(s, X^\alpha_s, \hat{m}_s, \alpha_s)ds + \int_0^t \sigma(s, X^\alpha_s, \hat{m}_s)dW_s, \quad t \in [0, T]; \\
\hat{m}_t & = \mathcal{L}(X^\alpha_t), \quad t \in [0, T], \quad \xi_0 \sim \mu_0.
\end{align*}
\]
where $\{W_t, \ 0 \leq t \leq T\}$ is a Brownian motion in $\mathbb{R}^n$, $\alpha$ is an admissible control, taking values in $\mathbb{R}^d$, adapted to the filtration $\mathcal{F} = \{\mathcal{F}_t, \ 0 \leq t \leq T\}$ generated by $(W, \xi_0)$ and $\mathcal{L}(\cdot)$ is the law. MFEs were first proposed by Lasry and Lions [29, 30, 31] for solving MFGs. Since introduced, this topic has aroused great interest. See [4, 7, 8] for the solvability of MFEs for different sakes and in different settings. In our model, the volatility is mean-field dependent. The appearance of the distribution has increased the complexity of MFEs and makes it difficult to solve. And the existence and uniqueness of classical solutions of fully coupled nonlocal quasilinear FBPDEs with distribution-dependent quasilinear part is new, to the best of our knowledge.

The objective of our paper is two-fold. The first objective is to show the existence of classical solutions of FBPDEs [11] in space $C^{1+\frac{1}{2}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n) \times C^{1+\frac{1}{2}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n)$ under appropriate assumptions (see Theorem [22]) and also the uniqueness under additional monotonicity conditions (see Theorem [2.1]). The second objective is to show that for a solution $(u, m)$ of the MFEs, the feedback strategy $\hat{a}(t, x) := \phi(t, x, m(t, \cdot), Du(t, x))$ is optimal for the MFG with $m(t, \cdot)$ being the distribution of the state corresponding to the optimal strategy (see Theorem [3.5]). Moreover, our results can be applied to linear-quadratic cases (see Section [4]).

MFGs were proposed by Lasry and Lions in a serie of papers [29, 30, 31] and also independently by Huang, Caines and Malhamé [21], under the different name of Nash Certainty Equivalence. A MFG is a limiting model of a symmetric, non-cooperative stochastic differential game with a large number of players. To be specific, each player solves a stochastic control problem with the cost and the state dynamics depending not only on his own state and control but also on other players’ states. The interaction among the players can be weak in the sense that one player is influenced by the other players only through the empirical distribution. Mean field type control problems, which aim to search the optimal controls of dynamics driven by McKean-Vlasov stochastic differential equations (SDEs), bear lots of resemblance to MFGs. The former are optimal control problems driven by distribution-dependent SDEs, and the latter are to optimize first and search for fixed points afterwards, which is more difficult. See Carmona et al. [12] for a discussion on the similarities and differences between the two problems. We also refer to [34, 35, 36] for works for mean field type control problems.
MFGs and mean field control problems are closely associated with PDEs. In a series of lectures given at the Collège de France [7], Lions showed the existence and uniqueness results for the particular second order MFEs:

\[
\begin{align*}
\frac{\partial u}{\partial t}(t,x) &+ \Delta u(t,x) - \frac{1}{2}|Du(t,x)|^2 + F(x,m(t,\cdot)) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n; \\
\frac{\partial m}{\partial t}(t,x) &- \Delta m(t,x) - \text{div}(Du(t,x)m(t,x)) = 0, \quad (t,x) \in (0,T] \times \mathbb{R}^n; \\
m(0,x) &= m_0(x), \quad u(T,x) = G(x,m(T,\cdot)), \quad x \in \mathbb{R}^n,
\end{align*}
\]

and showed that for a classical solution \((u, m)\) of the MFEs [5], the feedback strategy \(\hat{\alpha}(t,x) := -D_xu(t,x)\) is optimal for the MFG:

\[
\begin{align*}
\hat{\alpha} &\in \arg\min_{\alpha} J(\alpha) := \mathbb{E}\left[\int_0^T \left(\frac{1}{2}\alpha_s^2 + F(X_s, \hat{m}_s)\right) dt + G(X_T, \hat{m}_T)\right]; \\
X_t^\alpha &= X_0 + \int_0^t \alpha_s ds + \sqrt{2}dW_s, \quad \hat{m}_t = \mathcal{L}(X_t^\hat{\alpha}), \quad t \in [0,T], \quad X_0 \sim m_0.
\end{align*}
\]

Bensoussan et al. [4] solve the FBPDEs for linear quadratic MFGs under the condition that the volatility \(\sigma\) is a constant. Pham and Wei [34, 35, 36] use dynamic programming for optimal control of stochastic McKean-Vlasov dynamics and prove a verification theorem in the McKean-Vlasov framework. They give explicit solutions to Bellman equation for the linear quadratic mean field control problem. In this paper, we generalize Lions’ existence and uniqueness results [7, Theorems 3.1 and 3.6] of equation (5) to the more general case [11]. We do not need the linear quadratic condition and allow the volatility \(\sigma\) to depend on the state and the distribution of the state. The proof of the existence result is based on the Schauder fixed point theorem and relies on some results of quasilinear equations for parabolic type [27]. We also generalize Lions’ result [7, Lemma 3.7, p.15] of MFGs to the more general case [11] and apply our result to linear-quadratic problems.

The FP equations for McKean-Vlasov SDEs were first introduced by McKean [32] and were later studied in a more general setting in [24, 37]. Barbu and Röckner [3] study the FP equations for the McKean-Vlasov SDEs for the case of Nemytskii-type coefficients. Tse [38] study the higher order regularity of nonlinear FP equations. In this paper, our existence result of FBPDEs (11) includes the existence result of the FP equation for McKean-Vlasov SDEs with mean-field-dependent volatility (see Remark 3.3).

A method to decouple FBPDEs (11) is to study the so-called ”master equation”, a kind of second-order PDE stated on the space of probability measures. The concept of master equation is introduced by P. L. Lions in his lectures at Collège de France. Bensoussan et al. [4] interpret the master equation for both MFGs and mean field type control problems. Following a probabilistic method, Buckdahn et al. [6] solve the linear Kolmogorov PDE under appropriate regularity conditions of the coefficients with respect to the distribution variable. Crisan and McMurray [17] generalize the PDE studied by Buckdahn et al. [6] to include the cases where the terminal condition is not differentiable under assumptions of higher regularity of coefficients. Chassagneux et al. [16] show that the master equations admit a classical solution on sufficiently small time intervals. Cardaliaguet et al. [8] study the well-posedness of master equation and give a convergence result when the diffusion \(\sigma\) is a constant. Studying the master equation is an elegant method to study the FBPDEs (11), however, it requires regularity conditions of the coefficients with respect to the distribution variable. The proof of existence and uniqueness of solutions of the master equation is complicated, which will be discussed in our future works.
Carmona et al. [10, 11, 13, 14, 15] discussed MFGs with a probabilistic approach. Carmona and Delarue [11] consider MFGs under the nondegenerate condition with a probabilistic approach. And following the stochastic maximum principle for optimal controls [20], Carmona and Delarue [10] transform MFGs into existence and uniqueness of solutions of forward-backward stochastic differential equations (FBSDEs) within a linear-convex framework. The volatility is assumed to be a constant and the cost functions are assumed to be convex. Moreover, a weak mean-reverting condition is required. In this paper, we use a PDE approach, and we do not need the linear condition, and allow the drift, volatility and running cost to depend upon the distribution of the state. However, we need the differentiability of coefficients with respect to $x$. That is because we use the analytical method to consider the classical solutions to the FBPDEs. These differentiability assumptions are expected to be relaxed to more general cases by using the appropriate approximation technique. Actually, the mean field FBSDEs and mean field FBPDEs are interchangeable. We can get a solution of FBSDEs from a solution of FBPDEs, however, obtaining a solution of FBPDEs from a solution of FBSDEs requires additional regularity conditions of the coefficients. And our work builds a bridge.

MFGs are sometimes approached by symmetric, non-cooperative stochastic differential games of interacting $N$ players. On results about construction of $\varepsilon$-Nash equilibria for $N$-player games, see [10, 14, 22, 25, 28, 33]. For results about MFGs and mean field control problems with common noises, we refer to [1, 2, 23]. For other works about MFGs, we refer to [5, 18, 19, 39].

Summarizing our contributions in this paper, we
– show the existence and uniqueness of solutions of FBPDEs (1) in the class $C^{1+\frac{1}{2}+\frac{3}{2}}([0, T] \times \mathbb{R}^n)$;
– show the solvability of MFEs corresponding to MFG (4) and prove a verification theorem;
– study the linear-quadratic problems to illustrate the role of our main results.

The paper is organized as follows. In Section 2 we give the existence and uniqueness results of the solution to FBPDEs (1) in Theorems 2.1 and 2.11. In Section 3 we show the solvability of MFEs corresponding to MFG (4) and prove a verification theorem to construct an optimal strategy of MFG (4). In Section 4 we study the linear-quadratic problems to illustrate the role of our main results.

1.1 Notations

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t, 0 \leq t \leq T\}, \mathbb{P})$ denote a complete filtered probability space augmented by all the $\mathbb{P}$-null sets on which an $n$-dimensional Brownian motion $\{W_t, 0 \leq t \leq T\}$ is defined. $\mathcal{L}(\cdot)$ is the law. Let $S^2_\mathcal{F}(0, T)$ denote the set of all $\mathcal{F}_t$-progressively-measurable $\mathbb{R}^n$-valued processes $\beta = \{\beta_t, 0 \leq t \leq T\}$ such that

$$\mathbb{E}[\sup_{0 \leq t \leq T} |\beta_t|^2] < +\infty.$$  

Let $\mathcal{P}(\mathbb{R}^n)$ denote the space of all Borel probability measures on $\mathbb{R}^n$, and $\mathcal{P}_1(\mathbb{R}^n)$ the space of all probability measures $m \in \mathcal{P}(\mathbb{R}^n)$ such that

$$|m|_{\mathcal{P}_1(\mathbb{R}^n)} := \int_{\mathbb{R}^n} |x| m(dx) < \infty.$$  

The Kantorovitch-Rubinstein distance is defined on $\mathcal{P}_1(\mathbb{R}^n)$ by

$$d_1(m_1, m_2) := \inf_{\gamma \in \Pi(m_1, m_2)} \int_{\mathbb{R}^{2n}} |x - y| d\gamma(x,y), \quad m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n),$$  

4
where \( \Gamma(m_1, m_2) \) denotes the collection of all probability measures on \( \mathbb{R}^{2n} \) with marginals \( m_1 \) and \( m_2 \). The space \( (\mathcal{P}_1(\mathbb{R}^n), d_1) \) is a complete separable metric space \[11\]. Let \( \mathcal{P}_2(\mathbb{R}^n) \) the space of all probability measures \( m \in \mathcal{P}(\mathbb{R}^n) \) such that

\[
|m|_{\mathcal{P}_2(\mathbb{R}^n)} := \left( \int_{\mathbb{R}^n} |x|^2 m(dx) \right)^{\frac{1}{2}} < \infty.
\]

Let \( O \) be a domain in \( \mathbb{R}^n \), i.e. an arbitrary open connected subset of \( \mathbb{R}^n \), and \( Q_T \) be the cylinder \( (0, T) \times O \). For \( l > 0 \), let \( C^{\frac{l}{2}, l}(\bar{Q}_T) \) denote the set of all functions on \( \bar{Q}_T \) having all the continuous derivatives \( D^l_tD^s_x \) with \( 2r + s < l \) and having a finite norm

\[
|u|^{(l)}_{Q_T} = \sum_{2r+s \leq [l]} \sup_{Q_T} |D^l_tD^s_x u| + \sum_{0 < l - 2r - s < 2} \langle D^l_tD^s_x u \rangle_{t, Q_T}^{(\frac{l-s-2}{2})} + \sum_{2r+s = [l]} \langle D^l_tD^s_x u \rangle_{t, Q_T}^{(l-[l])},
\]

where

\[
\langle u \rangle_{t, Q_T}^{(\alpha)} = \sup_{(t,x), (t',x') \in Q_T, |t - t'| \leq \rho_0} \frac{|u(t, x) - u(t', x)|}{|t - t'|^\alpha}, \quad 0 < \alpha < 1,
\]

\[
\langle u \rangle_{x, Q_T}^{(\alpha)} = \sup_{(t,x), (t,x') \in Q_T, |x - x'| \leq \rho_0} \frac{|u(t, x) - u(t, x')|}{|x - x'|^\alpha}, \quad 0 < \alpha < 1.
\]

\( C^{\frac{l}{2}, l}(\bar{Q}_T) \) is a Banach space \[27\]. We also denote by \( C^l(\bar{\Omega}) \) the space of all time-invariant fields \( u \in C^{\frac{l}{2}, l}(\bar{Q}_T) \).

## 2 Solvability of fully coupled quasilinear FBPDEs

In this section, we investigate the following second order FBPDEs:

\[
\begin{align*}
\frac{\partial u}{\partial t} (t, x) + \sum_{i,j=1}^{n} a_{ij}(t, x, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j} (t, x) \\
+ \mathcal{H}(t, x, m(t, \cdot), Du(t, x)) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n;
\end{align*}
\]

\[
\begin{align*}
\frac{\partial m}{\partial t} (t, x) - \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x, m(t, \cdot))m(t, x)] \\
+ \text{div} \left[ \frac{\partial \mathcal{H}}{\partial p} (t, x, m(t, \cdot), Du(t, x))m(t, x) \right] = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n;
\end{align*}
\]

\[
m(0, x) = m_0(x), \quad u(T, x) = g(x, m(T, \cdot)), \quad x \in \mathbb{R}^n.
\]

It is a system of fully coupled quasilinear forward-backward parabolic equations. Here,

\[
\sigma : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}^{n \times n}, \quad a_{ij} = \frac{1}{2} (\sigma \sigma^T)_{ij}, \quad 1 \leq i, j \leq n;
\]

\[
\mathcal{H} : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}; \quad g : \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}.
\]

Our aim in this section is to prove the existence and uniqueness of classical solutions of FBPDEs \[6\] under appropriate assumptions.
2.1 Existence

In this subsection, we give the existence result of FBPDEs \([6]\). First we state our main assumptions in this subsection. For notational convenience, we use the same constant \(L\) for all the conditions below.

\textbf{(A1)} (Assumptions on \(a\)) There exists \(0 < \gamma_1 \leq \gamma_2 < +\infty\), such that

\[
\gamma \xi^2 \leq \sum_{i,j=1}^{n} a_{ij}(t, x, m) \xi_i \xi_j \leq L \xi^2, \quad \forall \xi \in \mathbb{R}^n, \quad (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n).
\]

The function \(a(t, \cdot, m) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}\) is twice differentiable for all \((t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)\), with the function and derivatives being bounded by \(L\), Lipschitz continuous in \((x, m) \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)\) and \(\frac{1}{2}\)-Hölder continuous in \(t \in [0, T]\).

\textbf{(A2)} (Assumptions on \(g\)) The function \(g\) is Lipschitz continuous in \((x, m) \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)\). For all \(m \in \mathcal{P}_1(\mathbb{R}^n)\), \(g(\cdot, m) \in C^{2+\frac{1}{2}}(\mathbb{R}^n)\). And there exists \(0 < \beta < 1\) such that

\[
\sup_{m \in \mathcal{P}_1(\mathbb{R}^n)} |g(\cdot, m)|_{L^2}^{(1+\beta)} \leq L.
\]

\textbf{(A3)} (Assumptions on \(m_0\)) The distribution \(m_0 \in \mathcal{P}_2(\mathbb{R}^n)\) is absolutely continuous with respect to the Lebesgue measure, with a density (still denoted by \(m_0\)) in \(C^{2+\frac{1}{2}}(\mathbb{R}^n)\).

\textbf{(A4)} (Assumptions on \(\mathcal{H}\)) The function \(\mathcal{H}(t, x, m, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}\) is differentiable for all \((t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)\) and the derivative \(\frac{\partial \mathcal{H}}{\partial p}(t, \cdot, m, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n\) is differentiable for all \((t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)\). For all \((t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n\),

\[
|\mathcal{H}(t, x, m, p)| \leq L(1 + |p|^2).
\]

For \(N \in (0, +\infty)\) and \((t, x, m, p), (t', x', m', p') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times B(0, N)\), where \(B(0, N) := \{p \in \mathbb{R}^n : |p| \leq N\}\),

\[
|\mathcal{H}(t, x, m, p)| \leq L_N,
\]

\[
|\mathcal{H}(t, x, m, p)| \leq L_N(1 + |p|^2).
\]

for some constant \(L_N\) depending on \(N\).

We have the following main result in this section:

**Theorem 2.1.** Suppose that Assumptions (A1)-(A4) hold. Then, FBPDEs \([6]\) has at least one classical solution \((u, m) \in C^{1+\frac{1}{2}+\frac{1}{2}}([0, T] \times \mathbb{R}^n) \times C^{1+\frac{1}{2}+\frac{1}{2}}([0, T] \times \mathbb{R}^n)\).

**Remark 2.2.** Our Theorem 2.1 includes as a special case Lions’ existence result of MFEs \([5]\) \([4]\, Theorem 3.1, p.10\) with \(\sigma = \sqrt{2}\) and \(\mathcal{H}(x, m, p) = \frac{1}{2}|p|^2 + F(x, m)\). The proof of Theorem 2.1 relies on the use of Schauder fixed point theorem. We also refer to \([26]\) for alternative application of Schauder theorem for nonlinear Markov processes.
The proof of Theorem 2.1 is divided into several parts and relies on the use of Schauder fixed point theorem. As in the proof of Theorem 3.1 [7, Section 3.2, p.13], we define the set $D_{C_1} \subseteq C^0([0, T], \mathcal{P}_1(\mathbb{R}^n))$ as follows:

$$D_{C_1} := \{ \mu \in C^0([0, T], \mathcal{P}_1(\mathbb{R}^n)) : \sup_{s \neq t} \frac{d_1(\mu(s), \mu(t))}{|t - s|^{1/2}} \leq C_1, \sup_{t \in [0, T]} |\mu(t)|^2_{\mathcal{P}_2(\mathbb{R}^n)} \leq C_1 \}, \quad (7)$$

where $C_1 \in (0, +\infty)$ is waiting to be determined. $D_{C_1}$ is a convex closed subset of $C^0([0, T], \mathcal{P}_1(\mathbb{R}^n))$, and is actually compact, due to Arzelà–Ascoli Theorem and the fact that the set of probability measures $\mu$ for which $|\mu|^2_{\mathcal{P}_2(\mathbb{R}^n)} \leq C_1$ is compact in $\mathcal{P}_1(\mathbb{R}^n)$ (see [7, Lemma 5.7]). Now we define a map $\Phi : D_{C_1} \to D_{C_1}$. For any $\mu \in D_{C_1}$, let $u$ be the unique solution to the following PDE:

$$\begin{cases}
\frac{\partial u}{\partial t}(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) \\
\quad + \mathcal{H}(t, x, \mu(t, \cdot), Du(t, x)) = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n; \\
u(T, x) = g(x, \mu(T, \cdot)), \quad x \in \mathbb{R}^n.
\end{cases} \quad (8)$$

Then we set $\Phi(\mu) = m$ as the unique solution of the following Fokker-Planck equation:

$$\begin{cases}
\frac{\partial m}{\partial t}(t, x) - \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(t, x, \mu(t, \cdot))m(t, x)) \\
\quad + \text{div}[\mathcal{D}(t, x, \mu(t, \cdot), Du(t, x))m(t, x)] = 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
m(0, x) = m_0(x), \quad x \in \mathbb{R}^n.
\end{cases} \quad (9)$$

We first show that $\Phi$ is well-defined, i.e. equations (8) and (9) have a unique solution, and further that the solution of (9) belongs to $D_{C_1}$. We then show the continuity of $\Phi$, and Theorem 2.1 as a consequence of Schauder fixed point theorem.

2.1.1 Existence and uniqueness of the solution of PDE (8)

In this subsection, we show the existence and uniqueness of the solution of equation (8), and give an estimate of the gradient $Du$. We first give the existence and uniqueness results.

Lemma 2.3. Let Assumptions (A1), (A2) and (A4) be satisfied. Then, PDE (8) has a unique solution $u \in C^{1+\frac{1}{4}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n)$ such that $|u| \leq M$ where $M$ is a constant depending only on $(L, T)$.

Proof. The proof is based on the existence and uniqueness theorem of the Cauchy problem of quasi-linear parabolic PDEs [27, Theorem 8.1, p.495]. In view of this theorem and our Assumptions (A1), (A2) and (A4), to show the existence and uniqueness of solution of (8), we only need to check the following conditions:

(a) The functions $(a_{ij}, \frac{\partial a_{ij}}{\partial x_j})(t, x, \mu(t, \cdot))$ are Hölder continuous in $(t, x)$ with exponents ($\frac{1}{4}, \frac{1}{2}$), and for $N \in (0, +\infty)$, the function $K(t, x, p) := -\mathcal{H}(t, x, \mu(t, \cdot), p)$ is Hölder continuous in $(t, x, p)$ with exponents ($\frac{1}{4}, \frac{1}{2}, \frac{1}{2}$) for $(t, x, p) \in [0, T] \times \mathbb{R}^n \times B(0, N)$. 

7
(b) There exists a constant $C > 0$ such that for $(t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$,

$$\sum_{i,j=1}^n |(a_{ij}, \frac{\partial a_{ij}}{\partial x_i})(t, x, \mu(t, \cdot))|(1 + |p|)^2 + |K(t, x, p)| \leq C(1 + |p|)^2.$$  

(c) For $N \in (0, +\infty)$, the derivative $\frac{\partial K}{\partial p}$ is bounded in $[0, T] \times \mathbb{R}^n \times B(0, N)$.

We first check condition (a). For any $(t', x', t, x) \in [0, T] \times \mathbb{R}^n$ and $|p|, |p| \leq N$, from Assumption (A4), we have

$$|K(t', x', p) - K(t, x, p)| \leq C(N)(|t' - t| \frac{1}{2} + |x' - x| + d_1(\mu(t'), \mu(t)) + |p' - p|)$$

$$\leq C(N, C_1)(|t' - t| \frac{1}{2} + |x' - x| + |p' - p|).$$

Here, we use the fact that $d_1(\mu(t'), \mu(t)) \leq C_1 |t' - t| \frac{1}{2}$ since $\mu \in D_{C_1}$, and the notation $C(N, C_1)$ stands for a constant depending only on $(N, C_1)$. Then from the boundedness of $K$ when $|p| \leq N$, we know that $K$ is Hölder continuous in $(t, x, p)$ with exponents $(\frac{1}{2}, \frac{1}{2})$. The proof of the Hölder continuity of functions $(a_{ij}, \frac{\partial a_{ij}}{\partial x_i})(t, x, \mu(t, \cdot))$ is similar, which is omitted here. Condition (b) is a direct consequence of Assumptions (A1) and (A4), with the constant $C$ depending only on $L$. And since $\frac{\partial K}{\partial p}(t, x, p) = -\frac{\partial H(t, x, \mu(t, \cdot), p)$, condition (c) is a direct consequence of Assumption (A4). Up to now, we have already checked conditions (a)-(c). In view of [27, Theorem 8.1, p.495], PDE (8) has a unique solution $u \in C^{1+\frac{\gamma}{2}, \frac{\gamma}{2}}([0, T] \times \mathbb{R}^n)$. And in view of [27, Theorem 2.9, p.23], $|u| \leq M$ for all $(t, x) \in [0, T] \times \mathbb{R}^n$, where $M$ depends only on $(L, T)$. \hfill \Box

In the following steps of the proof of Theorem 2.1 (subsections 2.1.3 and 2.1.4), we will use the estimate of the gradient $Du$. Now we give the following estimate of the gradient $Du$.

**Lemma 2.4.** Let Assumptions (A1), (A2) and (A4) be satisfied and $u \in C^{1,2}([0, T] \times \mathbb{R}^n)$ is a solution of PDE (8) such that $|u| \leq M$ where $M$ is a constant. Then, there exists $\alpha > 0$ depending only on $(n, \gamma, L, \beta)$, such that $|Du|_{[0, T] \times \mathbb{R}^n}^{(\alpha)}$ is bounded by a constant depending only on $(n, M, \gamma, L, \beta)$.

**Proof.** The proof is based on [27, Theorem 3.1, p.437] for estimates of $Du$ on a bounded domains of $[0, T] \times \mathbb{R}^n$. We choose

$$O_1 \subset O_2 \subset \cdots \subset O_k \subset \cdots, \quad \bigcup_{k=1}^\infty O_k = \mathbb{R}^n,$$

and set $Q_T^k := [0, T] \times O_k$ and $S_T^k := [0, T] \times \partial O_k$. Then, for any $k \geq 1$, we have

$$u|_{Q_T^k} \in C^{1,2}(Q_T^k), \quad \max_{Q_T^k} |u| \leq \max_{[0, T] \times \mathbb{R}^n} |u| \leq M,$$

$$|g_x(\cdot, \mu(T))| \leq |g_x(\cdot, \mu(T))| \leq L.$$

In view of [27, Theorem 3.1, p.437] and our Assumptions (A1), (A2) and (A4), there exists a constant $\alpha > 0$ depending only on $(n, \gamma, L, \beta)$, such that $|Du|_{Q_T^k}^{(\alpha)}$ is bounded by a constant depending only on $(n, M, \gamma, L, \beta)$ for any $k \geq 1$, where $\bar{Q}_T^k \subset Q_T^k$ is separated from the boundary $S_T^k$ by 1. Since $\bigcup_{k=1}^\infty \bar{Q}_T^k = [0, T] \times \mathbb{R}^n$, and the constants above are independent of $k$, so we have $|Du|_{[0, T] \times \mathbb{R}^n}^{(\alpha)}$ is bounded by a constant depending only on $(n, M, \gamma, L, T, \beta)$. \hfill \Box

As a summary of Lemmas 2.3 and 2.4 we have the following result.
Proposition 2.5. Let Assumptions (A1), (A2) and (A4) be satisfied. For $\mu \in \mathcal{D}_{C_{1}}$, PDE (8) has a unique solution $u \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$. Moreover, there exists $\alpha > 0$ depending only on $(n, \gamma, L, \beta)$, such that $|Du|_{[0,T] \times \mathbb{R}^n}$ is bounded by a constant depending only on $(n, \gamma, L, T, \beta)$.

2.1.2 Existence and uniqueness of the solution of PDE (9)

In this subsection, we show the existence and uniqueness of the solution of equation (9).

Proposition 2.6. Let Assumptions (A1)-(A4) be satisfied. For $\mu \in \mathcal{D}_{C_{1}}$ and $u \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$ as the solution of PDE (9), PDE (9) has a unique solution $m \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$.

Proof. Note that PDE (9) also reads

\[
\begin{aligned}
\frac{\partial m}{\partial t}(t,x) - \sum_{i,j=1}^{n} a_{ij}(t,x,\mu(t,\cdot)) \frac{\partial^2 m}{\partial x_i \partial x_j}(t,x) \\
+ \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}(t,x,\mu(t,\cdot), Du(t,x)) - \sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j}(t,x,\mu(t,\cdot)) \frac{\partial m}{\partial x_i}(t,x) \\
+ \sum_{i=1}^{n} \left[ - \sum_{j=1}^{n} \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}(t,x,\mu(t,\cdot)) + \frac{\partial^2 H}{\partial x_i \partial p_i}(t,x,\mu(t,\cdot), Du(t,x)) \right] m(t,x) = 0, \quad (t,x) \in (0,T) \times \mathbb{R}^n; \\
m(0,x) = m_0(x), \quad x \in \mathbb{R}^n.
\end{aligned}
\]

The proof is based on the existence and uniqueness theorem of the Cauchy problem of linear parabolic PDEs [27, Theorem 5.1, p.320]. In view of this theorem and our Assumptions (A1)-(A4), we only need to show that the coefficients of PDE (9) belong to the class $C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$. From Assumption (A1) and the fact that $\mu \in \mathcal{D}_{C_{1}}$, we know that for $1 \leq i,j \leq n$,

\[
(a_{ij}, \frac{\partial a_{ij}}{\partial x_j}, \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j})(t,x,\mu(t,\cdot)) \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n).
\]

From Proposition 2.5, we know that

\[
Du \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n), \quad D^2 u \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n).
\]

So from (11), Assumption (A4) and the fact that $\mu \in \mathcal{D}_{C_{1}}$, we know that for $1 \leq i,j \leq n$,

\[
\begin{aligned}
\left( \frac{\partial H}{\partial p_i} \frac{\partial^2 H}{\partial x_i \partial p_i}, \frac{\partial^2 H}{\partial x_i \partial p_i}, \frac{\partial^2 H}{\partial p_j \partial p_i}(t,x,\mu(t,\cdot), Du(t,x)) \right) \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n).
\end{aligned}
\]

It remains to show $h_{ij} \in C^{1+\frac{1}{4},\frac{1}{2}+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$ for $1 \leq i,j \leq n$, where

\[
h_{ij}(t,x) := \frac{\partial^2 H}{\partial p_j \partial p_i}(t,x,\mu(t,\cdot), Du(t,x)) \frac{\partial (Du)^j}{\partial x_i}(t,x).
\]

Actually, from Assumption (A4) and the fact that $\mu \in \mathcal{D}_{C_{1}}$, we have

\[
|h_{ij}(t',x') - h_{ij}(t,x)| \leq C(1, |Du|^{1}_{[0,T] \times \mathbb{R}^n})(|t' - t|^{\frac{1}{2}+\frac{1}{2}} + |x' - x|)
\]
Then from (11) we know that \( h_{ij} \in C^{\frac{3}{2}}([0,T] \times \mathbb{R}^n) \) for \( 1 \leq i \leq n \). Up to now, we have shown that all the coefficients in PDE (10) belong to the class \( C^{\frac{3}{2}}([0,T] \times \mathbb{R}^n) \). From [27, Theorem 5.1, p.320], there is a unique solution in the class \( C^{1+\frac{1}{2},2+\frac{1}{2}}([0,T] \times \mathbb{R}^n) \) of PDE (10), equivalently, of PDE (9).

### 2.1.3 The solution of PDE (9) belongs to \( \mathcal{D}_{C_1} \)

In this subsection, we show that the solution \( m \) of equation (9) belongs to the set \( \mathcal{D}_{C_1} \). Consider the following SDE:

\[
\begin{align*}
\left\{ 
\begin{array}{l}
\mathrm{d}X_t = \frac{\partial \mathcal{H}}{\partial p}(t, X_t, \mu(t, \cdot), Du(t, X_t)) dt + \sigma(t, X_t, \mu(t, \cdot)) dW_t, \quad t \in (0, T]; \\
X_0 = \xi_0 \sim m_0.
\end{array}
\right.
\end{align*}
\]

(12)

From Assumptions (A1), (A3) and (A4), the fact that \( Du \in C^{\frac{3}{2},\frac{3}{2}}([0,T] \times \mathbb{R}^n) \) and standard arguments of SDE, we know that (12) has a unique solution \( X = \{ X_t, \; 0 \leq t \leq T \} \in \mathcal{S}^2 \). We set \( \tilde{m}(t, \cdot) = \mathcal{L}(X_t) \). The following definition is borrowed from [7, Definition 3.2, p.11].

**Definition 2.7 (Weak solution of (9)).** We say that \( m \) is a weak solution of PDE (9), if \( m \in L^1([0,T], \mathcal{P}_1(\mathbb{R}^n)) \) such that for any test function \( \varphi \in C_c^\infty([0,T] \times \mathbb{R}^n) \), we have

\[
\int_{\mathbb{R}^n} \varphi(0, x) m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot)) \partial_{x_i} \partial_{x_j} \varphi(t, x) + \sum_{i=1}^n \frac{\partial \mathcal{H}}{\partial p_i}(t, x, \mu(t, \cdot), Du(t, x)) \partial_{x_i} \varphi(t, x) \right] m(t, dx) dt = 0.
\]

The following lemma shows the uniqueness of the weak solutions of PDE (9), which is available in [7, p.12].

**Lemma 2.8.** Suppose that there exists \( l \in (0, 1) \) such that, the coefficients of the linear parabolic PDE (10) belong to the class \( C^{\frac{2}{1+l}}([0,T] \times \mathbb{R}^n) \) and the initial function belongs to \( C^l(\mathbb{R}^n) \). Then, PDE (9) has a unique weak solution.

Under Assumptions (A1)-(A4), from the proof of Proposition 2.6 we know that the coefficients of PDE (10) belong to the class \( C^{\frac{3}{2},\frac{3}{2}}([0,T] \times \mathbb{R}^n) \). So in view of Lemma 2.8, \( m \) is the unique weak solution of PDE (9). The next lemma shows that \( \tilde{m} \) is also a weak solution of PDE (9), which implies that

\[
m(t, \cdot) = \tilde{m}(t, \cdot) = \mathcal{L}(X_t), \quad t \in [0, T].
\]

**Lemma 2.9.** \( \tilde{m} \) defined above is a weak solution of PDE (9).

**Proof.** For a test function \( \varphi \in C_c^\infty([0,T] \times \mathbb{R}^n) \), by applying Itô’s formula, we have

\[
\varphi(t, X_t) = \varphi(0, \xi_0) + \int_0^t (\partial_t \varphi(s, X_s) + \frac{1}{2} \text{Tr} [\sigma \sigma^T(s, X_s, \mu(s, \cdot))] D^2 \varphi(s, X_s)) ds + \int_0^t \sigma(s, X_s, \mu(s, \cdot)) D \varphi(s, X_s) dW_s.
\]

10
So we have □
\[ \text{Taking the expectation on both sides and noting that } \varphi \in C_c^\infty([0, T] \times \mathbb{R}^n), \text{ we have} \]
\[ \mathbb{E}[\varphi(0, \xi_0)] + \mathbb{E}\left[\int_0^T (\partial_t \varphi(s, X_s) + \frac{1}{2} \text{Tr} [\sigma \sigma^T(s, X_s, \mu(s, \cdot)) D^2 \varphi(s, X_s)]) \right. \]
\[ + \left. \langle D\varphi(s, X_s), \frac{\partial H}{\partial p}(s, X_s, \mu(s, \cdot), D u(s, X_s)) \rangle ds \right] = 0. \]

By definition of \( \tilde{m} \), we know that \( \tilde{m} \) is a weak solution of PDE (9). □

Now we can prove that the solution of PDE (9) belongs to \( \mathcal{D}_{C_1} \).

**Proposition 2.10.** Let Assumptions (A1)-(A4) be satisfied. There is \( C_1 \in (0, +\infty) \) depending only on \( (n, \gamma, L, \beta, T, m_0) \), such that the solution \( m \) of PDE (9) belongs to \( \mathcal{D}_{C_1} \).

**Proof.** From the definition of \( d_1 \), we have for \( 0 \leq s < t \leq T \),
\[ d_1(m(s), m(t)) \leq \mathbb{E}|X_s - X_t| \]
\[ = \mathbb{E}\left[ \int_s^t \frac{\partial H}{\partial p}(\tau, X_\tau, \mu(\tau, \cdot), D u(\tau, X_\tau)) d\tau + \int_s^t \sigma(\tau, X_\tau, \mu(\tau, \cdot)) dW_\tau \right] \]
\[ \leq |t - s| \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} \left| \frac{\partial H}{\partial p}(t, x, \mu(t, \cdot), D u(t, x)) \right| + L|t - s|^{\frac{3}{2}}. \]

From Proposition 2.5, we have
\[ \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} |D u(t, x)| \leq C(n, \gamma, L, T, \beta). \]

Therefore, from Assumption (A4),
\[ \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} \left| \frac{\partial H}{\partial p}(t, x, \mu(t, \cdot), D u(t, x)) \right| \leq C(n, \gamma, L, T, \beta). \]

So we have
\[ \sup_{s \neq t} \frac{d_1(m(s), m(t))}{|t - s|^{\frac{3}{2}}} \leq C(n, \gamma, L, T, \beta). \]

Here, the notation \( C(n, \gamma, L, T, \beta) \) stands for a constant depending only on \( (n, \gamma, L, T, \beta) \). Similarly,
\[ \int_{\mathbb{R}^n} |x|^2 m(t, dx) = \mathbb{E}|X_t|^2 \leq C(L, T) \mathbb{E}|\xi_0|^2 + 1 + \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} \left[ \frac{\partial H}{\partial p}(t, x, \mu(t, \cdot), D u(t, x))^2 \right] \]
\[ \leq C(n, \gamma, L, T, \beta, m_0). \]

Therefore,
\[ \sup_{t \in [0, T]} \int_{\mathbb{R}^n} |x|^2 m(t, dx) \leq C(n, \gamma, L, T, \beta, m_0). \]

In summary, we have \( m \in \mathcal{D}_{C_1} \) with \( C_1 = C(n, \gamma, L, \beta, T, m_0) \) independent of \( \mu \). The proof is complete. □
2.1.4 Continuity of $\Phi$

Now it is clear that the mapping $\Phi : D_{C_1} \rightarrow D_{C_1}$ is well-defined. Next, we prove that it is continuous. For notational convenience, we set

$$\rho(\mu, \mu') := d_{C^0([0,T], \mathcal{P}(\mathbb{R}^n))}(\mu, \mu') = \sup_{0 \leq t \leq T} d_1(\mu(t), \mu'(t)), \quad \mu, \mu' \in D_{C_1}.$$ 

Let $\{\mu_k, \ k \geq 1\} \subset D_{C_1}$ converge to some $\mu \in D_{C_1}$ with respect to the metric $\rho$. Let $(u_k, m_k)$ and $(u, m)$ be the solutions of equations (8)-(9) corresponding to $\mu_k$ and $\mu$, respectively. From Assumptions (A1), (A2) and (A4), we know that the coefficients in (8) corresponding to $\mu_k$ uniformly converge to the coefficients in (8) corresponding to $\mu$. Then we get the local uniform convergence of $\{u_k, \ k \geq 1\}$ to $u$ by standard arguments. From Proposition 2.5, we know that the gradients $\{Du_k, \ k \geq 1\}$ are uniformly bounded and uniformly Hölder continuous and therefore locally uniformly converges to $Du$.

For any converging subsequence $\{m_{k_l}, \ l \geq 1\}$ of the relatively compact sequence $\{m_k, \ k \geq 1\}$ (since $D_{C_1}$ is compact), we assume that $\{m_{k_l}, \ l \geq 1\}$ converge to some $m \in D_{C_1}$. For any $\varphi \in C_c^\infty([0,T] \times \mathbb{R}^n)$, since $m_{k_l}$ is a weak solution of (9) corresponding to $(\mu_{k_l}, u_{k_l})$, we have

$$\int_{\mathbb{R}^n} \phi(0,x)m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[\partial_t \varphi(t,x) + \sum_{i,j=1}^n a_{ij}(t,x,\mu_{k_l}(t,\cdot))\partial_{x_i} \varphi(t,x) \right] m_{k_l}(t, dx)dt = 0. \tag{13}$$

From Assumption (A1) and the facts that $\varphi \in C_c^\infty([0,T] \times \mathbb{R}^n)$, we have

$$I_{k_l}^1 := \left| \int_0^T \int_{\mathbb{R}^n} \left[ \sum_{i,j=1}^n a_{ij}(t,x,\mu_{k_l}(t,\cdot))\partial_{x_i} \varphi(t,x) - \sum_{i,j=1}^n a_{ij}(t,x,\mu(t,\cdot))\partial_{x_i} \varphi(t,x) \right] m_{k_l}(t, dx)dt \right| \leq C(L)\rho(\mu_{k_l}, \mu) \int_0^T \int_{\mathbb{R}^n} \left( \sum_{i,j=1}^n |\partial_{x_i} \varphi(t,x)| \right) m_{k_l}(t, dx)dt. \tag{14}$$

From Proposition 2.5 we know that there is a constant $C(n, \gamma, L, T, \beta)$ depending only on $(n, \gamma, L, T, \beta)$, such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |Du(t,x)| \leq C(n, \gamma, L, T, \beta), \quad \sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |Du_{k_l}(t,x)| \leq C(n, \gamma, L, T, \beta), \quad l \geq 1.$$ 

Therefore, from Assumption (A4) and the fact that $\varphi \in C_c^\infty([0,T] \times \mathbb{R}^n)$

$$I_{k_l}^2 := \left| \int_0^T \int_{\mathbb{R}^n} \left[ \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t,x,\mu_{k_l}(t,\cdot), Du_{k_l}(t,x))\partial_{x_i} \varphi(t,x) \right. \right. \left. \left. - \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t,x,\mu(t,\cdot), Du(t,x))\partial_{x_i} \varphi(t,x) \right] m_{k_l}(t, dx)dt \right| \leq C(n, \gamma, L, T, \beta) \int_0^T \int_{\mathbb{R}^n} \left( \sum_{i=1}^n |\partial_{x_i} \varphi(t,x)| \right) m_{k_l}(t, dx)dt \tag{15} \cdot \left| \rho(\mu_{k_l}, \mu) + \sup_{(t,x) \in \text{supp}(\varphi)} |Du_{k_l}(t,x) - Du(t,x)| \right|.$$ 

12
From Kantorovich-Rubinstein Theorem [7, Theorem 5.5, p.36] and the fact that \( \{m_k, \ l \geq 1\} \) converge to \( \hat{m} \) in \( C^0([0,T], \mathcal{P}_1(\mathbb{R}^n)) \), we have for \( 1 \leq i, j \leq n \),

\[
\lim_{t \to +\infty} \int_0^T \int_{\mathbb{R}^n} |\partial_{x,ij} \varphi(t, x)|m_k(t, dx)dt = \int_0^T \int_{\mathbb{R}^n} |\partial_{x,ij} \varphi(t, x)|\hat{m}(t, dx)dt;
\]

\[
\lim_{t \to +\infty} \int_0^T \int_{\mathbb{R}^n} |\partial_{x,i} \varphi(t, x)|m_k(t, dx)dt = \int_0^T \int_{\mathbb{R}^n} |\partial_{x,i} \varphi(t, x)|\hat{m}(t, dx)dt.
\]

Since \( \{\mu_k, \ k \geq 1\} \) converge to \( \mu \) with respect to the norm \( \rho \) and \( \{Du_k, \ k \geq 1\} \) locally uniformly converges to \( Du \), we have

\[
\lim_{t \to +\infty} \rho(\mu_k, \mu) = 0;
\]

\[
\lim_{t \to +\infty} \sup_{(t,x) \in \text{supp} \varphi} |Du_k(t, x) - Du(t, x)| = 0.
\]

Plugging (16) and (17) into (14) and (15), we have

\[
\lim_{t \to +\infty} I_{k_i}^1 = \lim_{t \to +\infty} I_{k_i}^2 = 0.
\]

Plugging (18) into (13), we have

\[
\int_{\mathbb{R}^n} \phi(0, x)m_0(dx) + \lim_{t \to +\infty} \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot))\partial_{x,ij} \varphi(t, x) + \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t, x, \mu(t, \cdot), Du(t, x))\partial_{x,i} \varphi(t, x) \right]m_k(t, dx)dt = 0.
\]

Again from Kantorovich-Rubinstein Theorem, we have

\[
\lim_{t \to +\infty} \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot))\partial_{x,ij} \varphi(t, x) + \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t, x, \mu(t, \cdot), Du(t, x))\partial_{x,i} \varphi(t, x) \right]m_k(t, dx)dt
\]

\[
= \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot))\partial_{x,ij} \varphi(t, x) + \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t, x, \mu(t, \cdot), Du(t, x))\partial_{x,i} \varphi(t, x) \right]\hat{m}(t, dx)dt.
\]

Plugging (20) into (19), we have eventually

\[
\int_{\mathbb{R}^n} \phi(0, x)m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t))\partial_{x,ij} \varphi(t, x) + \sum_{i=1}^n \frac{\partial H}{\partial p_i}(t, x, \mu(t, \cdot), Du(t, x))\partial_{x,i} \varphi(t, x) \right]\hat{m}(t, dx)dt = 0,
\]
which means that $\hat{m}$ is a weak solution to (9) corresponding to $(\mu, u)$. Then, from Lemma 2.8 we know that $m = \hat{m}$. Up to now we can see that, any converging subsequence $\{m_k, l \geq 1\}$ of the relatively compact sequence $\{m_k, k \geq 1\}$ converge to $m$. So we know that $\{m_k, k \geq 1\}$ converge to $m$. Thus, $\Phi$ is continuous.

We conclude by Schauder fixed point theorem that the continuous map $\Phi$ has a fixed point in $\mathcal{C}_{C, t}$. This fixed point is a classical solution of FBPDEs (6). The proof of Theorem 2.1 is complete.

### 2.2 Uniqueness

In this subsection, we give the uniqueness result of FBPDEs (6) when the coefficient $\sigma$ is independent of $m$ and the function $H$ is of the form

$$H(t, x, m, p) = H^0(t, x, m) + H^1(t, x, p), \quad (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n,$$

with the function $H^1$ satisfying for any $(t, x, p_1, p_2) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$,

$$H^1(t, x, p_2) - H^1(t, x, p_1) \leq \left( \frac{\partial H^1}{\partial p}(t, x, p_1), p_2 - p_1 \right). \quad (21)$$

Moreover, let us assume that, besides Assumptions (A1)-(A4), the following monotonicity conditions hold:

$$\int [H^0(t, x, m_2) - H^0(t, x, m_1)](m_2 - m_1)(dx) > 0, \quad t \in [0, T], \ m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n), \ m_1 \neq m_2,$$

$$\int [g(x, m_2) - g(x, m_1)](m_2 - m_1)(dx) \geq 0, \quad m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n).$$

**Theorem 2.11.** Under the above conditions, FBPDEs (6) has no more than one classical solution.

**Proof.** Let $(u^1, m^1)$ and $(u^2, m^2)$ be two classical solutions of FBPDEs (6). We set $(\Delta u, \Delta m) := (u^2 - u^1, m^2 - m^1)$. Then

$$\frac{\partial \Delta u}{\partial t} + \sum_{i,j=1}^n a_{ij}(t, x) \frac{\partial^2 \Delta u}{\partial x_i \partial x_j} + \left[ H^1(t, x, Du^2) - H^1(t, x, Du^1) \right]$$

$$+ \left[ H^0(t, x, m^2(t, \cdot)) - H^0(t, x, m^1(t, \cdot)) \right] = 0, \quad (22)$$

$$\frac{\partial \Delta m}{\partial t} - \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x) \Delta m]$$

$$+ \text{div} \left[ \frac{\partial H^1}{\partial p}(t, x, Du^2) m^2 - \frac{\partial H^1}{\partial p}(t, x, Du^1) m^1 \right] = 0, \quad (23)$$

$$\Delta m(0, x) = 0, \quad \Delta u(T, x) = g(x, m^2(T, \cdot)) - g(x, m^1(T, \cdot)).$$

Let us use $\Delta u$ as a test function in equation (22), multiply equation (22) by $\Delta m$ and integrate over $[0, T] \times \mathbb{R}^n$, and add them together. We have

$$\int [g(x, m^2(T, \cdot)) - g(x, m^1(T, \cdot))] \Delta m(T, dx)$$

$$+ \int_0^T \int [H^0(t, x, m^2(t, \cdot)) - H^0(t, x, m^1(t, \cdot))] \Delta m(t, dx) dt$$

14
\[
\begin{align*}
&= \int_0^T [\mathcal{H}^1(t, x, Du^1(t, x)) - \mathcal{H}^1(t, x, Du^2(t, x))] \Delta m(t, dx) \\
& \quad + \langle D\Delta u(t, x), \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^2(t, x)) m^2(t, dx) - \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^1(t, x)) m^1(t, dx) \rangle dt.
\end{align*}
\]

From (21), we have
\[
\begin{align*}
[\mathcal{H}^1(t, x, Du^1) - \mathcal{H}^1(t, x, Du^2)] \Delta m + \langle D\Delta u, \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^2) m^2 - \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^1) m^1 \rangle \\
= m_2[\mathcal{H}^1(t, x, Du^1) - \mathcal{H}^1(t, x, Du^2) - \langle \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^2), Du^1 - Du^2 \rangle] \\
+ m_1[\mathcal{H}^1(t, x, Du^2) - \mathcal{H}^1(t, x, Du^1) - \langle \frac{\partial \mathcal{H}^1}{\partial p}(t, x, Du^1), Du^2 - Du^1 \rangle] \\
\leq 0.
\end{align*}
\]

So from the monotonicity conditions of \(\mathcal{H}^0\) and \(g\), we have \(\Delta m = 0\) and, therefore, \(\Delta u = 0\). \(\square\)

### 3 Application to mean field games

In this section, we investigate the optimal strategy of the MFG:

\[
\begin{align*}
\hat{\alpha} & \in \arg\min_{\alpha} J(\alpha | \hat{m}) := \mathbb{E} \left[ \int_0^T f(t, X_t^\alpha, \hat{m}_t, \alpha_t) dt + g(X_T^\alpha, \hat{m}_T) \right]; \\
X_t^\alpha & = \xi_0 + \int_0^t b(s, X_s^\alpha, \hat{m}_s, \alpha_s) ds + \int_0^t \sigma(s, X_s^\alpha, \hat{m}_s) dW_s, \quad t \in [0, T]; \\
\hat{m}_t & = \mathcal{L}(X_t^\hat{\alpha}), \quad t \in [0, T], \quad \xi_0 \sim m_0.
\end{align*}
\]

where \(\alpha\) is an admissible control, taking values in \(\mathbb{R}^d\), adapted to the filtration \(\mathcal{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}\). Here, functions \(\sigma\) and \(g\) are as defined in the last section, and

\[
\begin{align*}
& b : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}^n, \\
& f : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}.
\end{align*}
\]

We define the Hamiltonian \(H\) as

\[
H(t, x, m, \alpha, p) := \langle p, b(t, x, m, \alpha) \rangle + f(t, x, m, \alpha), \quad (t, x, m, \alpha, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \times \mathbb{R}^n,
\]

and the minimizing control function \(\phi\)

\[
\phi(t, x, m, p) := \arg\min_{\alpha \in \mathbb{R}^d} H(t, x, m, \alpha, p), \quad (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n,
\]

under appropriate conditions.
3.1 Solvability of MFEs

We first use the results in the last section to investigate the solvability of the following second order MFEs:

\[
\begin{aligned}
\frac{\partial u}{\partial t}(t, x) + H(t, x, m(t, \cdot), \phi(t, x, m(t, \cdot), Du(t, x)), Du(t, x)) \\
+ \sum_{i,j=1}^{n} a_{ij}(t, x, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) = 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n;
\end{aligned}
\]

\[
\begin{aligned}
\frac{\partial m}{\partial t}(t, x) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} [b_i(t, x, m(t, \cdot), \phi(t, x, m(t, \cdot), Du(t, x)))m(t, x)] \\
- \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x, m(t, \cdot))m(t, x)] = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n;
\end{aligned}
\]

\[
m(0, x) = m_0(x), \quad u(T, x) = g(x, m(T, \cdot)), \quad x \in \mathbb{R}^n.
\]

under Assumptions (A1)-(A3) and the following assumptions on functions \( b, f \) and \( \phi \).

\( (B1) \) (Assumptions on \( b \)) The function \( b(t, \cdot, m, \cdot) : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n \) is differentiable for all \( (t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n) \). For all \( (t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \),

\[
|b(t, x, m, \alpha)| \leq L(1 + |\alpha|).
\]

For \( N \in (0, +\infty) \) and \( (t, x, m, \alpha), (t', x', m', \alpha') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times B(0, N) \), where \( B(0, N) := \{ \alpha \in \mathbb{R}^d : |\alpha| \leq N \} \),

\[
|((\frac{\partial b}{\partial x'}, \frac{\partial b}{\partial \alpha}))(t, x, m, \alpha)| \leq L_N,
\]

\[
|(b, \frac{\partial b}{\partial x}, \frac{\partial b}{\partial \alpha})(t', x', m', \alpha') - (b, \frac{\partial b}{\partial x}, \frac{\partial b}{\partial \alpha})(t, x, m, \alpha)|
\]

\[
\leq L_N(|t' - t|^{\frac{1}{2}} + |x' - x| + d_1(m', m) + |\alpha' - \alpha|),
\]

for some constant \( L_N \) depending on \( N \).

\( (B2) \) (Assumptions on \( f \)) The function \( f(t, x, m, \cdot) : \mathbb{R}^d \to \mathbb{R} \) is differentiable for all \( (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \). For all \( (t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \),

\[
|f(t, x, m, \alpha)| \leq L(1 + |\alpha|^2).
\]

For \( N \in (0, +\infty) \) and \( (t, x, m, \alpha), (t', x', m', \alpha') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times B(0, N) \), where \( B(0, N) := \{ \alpha \in \mathbb{R}^d : |\alpha| \leq N \} \),

\[
|f(t', x', m', \alpha') - f(t, x, m, \alpha)| \leq L_N(|t' - t|^{\frac{1}{2}} + |x' - x| + d_1(m', m) + |\alpha' - \alpha|),
\]

for some constant \( L_N \) depending on \( N \).

\( (B3) \) (Assumptions on \( \phi \)) For each \( (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n \), there exits a unique vector \( \phi(t, x, m, p) \in \mathbb{R}^d \) satisfying (26). The function \( \phi(t, \cdot, m, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^d \) is differentiable for all \( (t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n) \). For all \( (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n \),

\[
|\phi(t, x, m, \alpha)| \leq L(1 + |p|).
\]
For \( N \in (0, +\infty) \) and \((t, x, m, p), (t', x', m', p') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times B(0, N)\), where \( B(0, N) := \{p \in \mathbb{R}^n : |p| \leq N\}\),

\[
|\left(\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial p}\right)(t, x, m, p)| \leq L_N,
\]

\[
|\phi, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial p}|(t', x', m', p') - (b, \frac{\partial b}{\partial x}, \frac{\partial b}{\partial m})(t, x, m, p)|
\leq L_N(|t' - t| + |x' - x| + d_1(m', m) + |p' - p|),
\]

for some constant \( L_N \) depending on \( N \).

Assumption (B3) is also used by Carmona and Delarue in [11] and by Nourian and Caines in [33]. Assumption (B3) is satisfied when function \( b \) is linear in \( \alpha \), function \( f \) is strongly convex in \( x, m \), and the derivatives \((\frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x \partial m}, \frac{\partial^2 f}{\partial m \partial x}, \frac{\partial^2 f}{\partial m^2})\) existing, being bounded for bounded \( \alpha \), and being Lipschitz continuous in \((x, m, \alpha)\) and \( \frac{1}{2}\)-Hölder continuous in \( t \in [0, T] \) for bounded \( \alpha \).

We give examples of functions \((\sigma, b, f)\) satisfying assumptions (A1) and (B1)-(B3).

**Example 3.1.** We define for \( 1 \leq i, j \leq n \) and \((x, m, \alpha) \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d\),

\[
b(x, m, \alpha) := B^0(x, m) + B^1 \alpha,
\]

\[
s_{ij}(x, m) := [2 + \cos(x_i - m_j)] \delta_{ij},
\]

\[
f(x, m, \alpha) := \frac{|\alpha|^2}{2} + \sum_{j=1}^n \sin(x_j - m_j),
\]

where \( B^1 \in \mathbb{R}^{n \times d}, \ m_i := \int_{\mathbb{R}^n} x_idm(dx) \) and \( \delta_{ij} = 1 \) when \( i = j \) and \( \delta_{ij} = 0 \) when \( i \neq j \). In view of Kantorovich-Rubinstein Theorem [7, Theorem 5.5], Assumptions (A1), (B1) and (B2) are satisfied. Moreover, \( \phi(p) = -B_1^1p \) for \( p \in \mathbb{R}^n \) and Assumption (B3) is satisfied. More generally, our assumptions are satisfied for the following class of functions:

\[
b(x, m, \alpha) := B(x, m) + B^1 \alpha, \quad f(x, m, \alpha) := \frac{|\alpha|^2}{2} + F(x, m),
\]

\[
(B, \sigma, F) : \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times n} \times \mathbb{R},
\]

where \( B^1 \in \mathbb{R}^{n \times d}, \sigma \) is non-degenerate and functions \((B, B^1, \sigma, \sigma_x, \sigma_{xx}, F)\) are bounded and Lipschitz continuous.

**Remark 3.1.** The uniform boundedness assumption in the state \( x \) of functions \((b, f, g, \phi)\) seems to be restrictive. However, these constraints can be relaxed to include the common linear-quadratic cases. We discuss these problems in detail in Section 4.

In view of FBPDEs [6], we set

\[
\mathcal{H}(t, x, m, p) := H(t, x, m, \phi(t, x, m, p), p), \quad (t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n.
\]

From Assumption (B3), we know that for \((t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n\),

\[
0 = \frac{\partial H}{\partial \alpha}(t, x, m, \alpha, p)|_{\alpha = \phi(t, x, m, p)}
\]

\[
= \frac{\partial b}{\partial \alpha}(t, x, m, \phi(t, x, m, p)p) + \frac{\partial f}{\partial \alpha}(t, x, m, \phi(t, x, m, p)).
\]
Proof. Since the coefficients \( (b, \sigma) \) are independent of \( m \) and the function \( f \) is of the form
\[
f(t, x, m, \alpha) = f^0(t, x, m) + f^1(t, x, \alpha), \quad (t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d.
\] (30)
Let us assume that, besides Assumptions (A1)-(A3) and (B1)-(B3), the following monotonicity conditions hold:
\[
\int [f^0(t, x, m_2) - f^0(t, x, m_1)](m_2 - m_1)(dx) > 0, \quad t \in [0, T], \ m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n), \ m_1 \neq m_2,
\]
\[
\int [g(x, m_2) - g(x, m_1)](m_2 - m_1)(dx) \geq 0, \quad m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n).
\]

**Theorem 3.4.** Under the above conditions, MFEs (27) has no more than one classical solution.

**Proof.** Since the coefficients \( (b, \sigma) \) are independent of \( m \) and the function \( f \) is of the form (30), the Hamiltonian \( H \) can be divided into two parts
\[
H(t, x, m, \alpha, p) = f^0(t, x, m) + H^1(t, x, \alpha, p)
\]
for \((t, x, m, \alpha, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \times \mathbb{R}^n\), where

\[H^1(t, x, \alpha, p) := (b(t, x, \alpha), p) + f^1(t, x, \alpha),\]

and the minimizing control function \(\phi\) satisfies

\[\phi(t, x, p) := \arg\min_{\alpha \in \mathbb{R}^d} H^1(t, x, \alpha, p), \quad (t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.\] (31)

We set for \((t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n\),

\[\mathcal{H}^0(t, x, m) := f^0(t, x, m), \quad \mathcal{H}^1(t, x, p) := H^1(t, x, \phi(t, x, p), p).\]

In view of (31), for \((t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n\)

\[\frac{\partial \mathcal{H}^1}{\partial p}(t, x, p) = b(t, x, \phi(t, x, p)).\]

Therefore, for any \((t, x, p_1, p_2) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n\), from (31), we have

\[\mathcal{H}^1(t, x, p_1) + (\frac{\partial \mathcal{H}^1}{\partial p}(t, x, p_1), p_2 - p_1)
= \langle b(t, x, \phi(t, x, p_1), p_1) + f^1(t, x, \phi(t, x, p_1)) + b(t, x, \phi(t, x, p_1)), p_2 - p_1 \rangle
= H^1(t, x, \phi(t, x, p_1), p_2)
\geq H^1(t, x, \phi(t, x, p_2), p_2) = \mathcal{H}^1(t, x, p_2).\]

Then, the uniqueness result of MFEs [27] is a consequence of Theorem 2.11. □

### 3.2 A verification theorem

In this subsection, we give a verification theorem for MFG [24] under the following assumptions.

(C1) For \(N \in (0, +\infty)\) and \((t, x, m), (t', x', m') \in [0, T] \times \mathbb{R}^n \times \{ m \in \mathcal{P}_1(\mathbb{R}^n) : |m|_{\mathcal{P}_1(\mathbb{R}^n)} \leq N\},\)

\[|a(t, x, m)| \leq L_N, \quad |a(t', x', m') - a(t, x, m)| \leq L_N||t' - t||^{\frac{1}{2}} + |x' - x| + d_1(m', m),\]

for some constant \(L_N\) depending on \(N\).

(C2) For \(N, M \in (0, +\infty)\) and \((t, x, m, \alpha), (t', x', m', \alpha') \in [0, T] \times \mathbb{R}^n \times \{ m \in \mathcal{P}_1(\mathbb{R}^n) : |m|_{\mathcal{P}_1(\mathbb{R}^n)} \leq N\} \times \{ \alpha \in \mathbb{R}^d : |\alpha| \leq M\},\)

\[|b(t, x, m, \alpha)| \leq L_{(N,M)}, \quad |b(t', x', m', \alpha') - b(t, x, m, \alpha)| \leq L_{(N,M)}||t' - t||^{\frac{1}{2}} + |x' - x| + d_1(m', m) + |\alpha' - \alpha|,\]

for some constant \(L_{(N,M)}\) depending on \((N, M)\).

(C3) The distribution \(m_0 \in \mathcal{P}_2(\mathbb{R}^n)\) is absolutely continuous with respect to the Lebesgue measure, with a density (still denoted by \(m_0\)) in \(C^{\frac{1}{2}}(\mathbb{R}^n)\).

(C4) For \((t, x, m, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n\), there exits a unique vector

\[\phi(t, x, m, p) := \arg\min_{\alpha \in \mathbb{R}^d} H(t, x, m, \alpha, p).\]
And for $N, M \in (0, +\infty)$ and $(t, x, m, p)$, $(t', x', m', p') \in [0, T] \times \mathbb{R}^n \times \{m \in \mathcal{P}_1(\mathbb{R}^n) : |m|_{\mathcal{P}_1(\mathbb{R}^n)} \leq N\} \times \{p \in \mathbb{R}^n : |p| \leq M\}$,

$$|\phi(t, x, m, p)| \leq L_{(N,M)},$$

$$|\phi(t', x', m', p') - \phi(t, x, m, p)| \leq L_{(N,M)}[|t' - t|^\frac{1}{2} + |x' - x| + d_1(m', m) + |p' - p|],$$

for some constant $L_{(N,M)}$ depending on $(N, M)$.

Suppose that $(u, m) \in C^{1+\frac{1}{4}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n) \times C^{\frac{1}{2}}([0, T], \mathcal{P}_1(\mathbb{R}^n))$ is a solution of the MFES (27), where

$$C^{\frac{1}{4}}([0, T], \mathcal{P}_1(\mathbb{R}^n)) := \{\mu \in C^0([0, T], \mathcal{P}_1(\mathbb{R}^n)) : \sup_{s \neq t} \frac{d_1(\mu(s), \mu(t))}{|t - s|^{\frac{1}{2}}} + \sup_{t \in [0, T]} |\mu(t)|_{\mathcal{P}_1(\mathbb{R}^n)} < +\infty\}.$$

The following theorem shows that the feedback strategy $\tilde{\alpha}(t, x) := \phi(t, x, m(t, \cdot), Du(t, x))$ is optimal for MFG (24).

**Theorem 3.5** (verification theorem). Let Assumptions (C1)-(C4) be satisfied and let $(u, m) \in C^{1+\frac{1}{4}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n) \times C^{\frac{1}{2}}([0, T], \mathcal{P}_1(\mathbb{R}^n))$ be a solution of FBPDEs (33). Then, $\tilde{\alpha} := \{\tilde{\alpha}(t, \tilde{X}_t), 0 \leq t \leq T\}$ is an optimal control for MFG (24), where $\tilde{X} = \{\tilde{X}_t, 0 \leq t \leq T\}$ is the solution of SDE

$$\tilde{X}_t = \xi_0 + \int_0^t b(s, \tilde{X}_s, m(s, \cdot), \tilde{\alpha}(s, \tilde{X}_s))ds + \int_0^t \sigma(s, \tilde{X}_s, m(s, \cdot))dW_s, \quad t \in [0, T].$$

**Proof.** Given $(u, m)$, similar as the proof of Lemma (24) we know that $\mathcal{L}(\tilde{X}_t)$ is a weak solution of the following PDE for $\tilde{m}$:

$$\begin{cases}
\frac{\partial \tilde{m}}{\partial t}(t, x) - \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x, m(t, \cdot))\tilde{m}(t, x)] + \sum_{i=1}^n \frac{\partial}{\partial x_i} [b_i(t, x, m(t, \cdot), \phi(t, x, m(t, \cdot), Du(t, x)))\tilde{m}(t, x)] = 0, & (t, x) \in (0, T) \times \mathbb{R}^n, \\
\tilde{m}(0, x) = m_0(x), & x \in \mathbb{R}^n.
\end{cases}$$

(33)

From the fact that $(u, m) \in C^{1+\frac{1}{4}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n) \times C^{\frac{1}{2}}([0, T], \mathcal{P}_1(\mathbb{R}^n))$ and Assumptions (C1)-(C4), it is easy to check that the coefficients of PDE (33) belong to the class $C^{\frac{1}{2}, \frac{1}{2}}([0, T] \times \mathbb{R}^n)$ and the initial function belongs to $C^{\frac{1}{2}}(\mathbb{R}^n)$. From Lemma (24) we know that $m$ is the unique weak solution to PDE (33). Therefore, $m(t, \cdot) = \mathcal{L}(\tilde{X}_t)$ for $t \in [0, T]$.

Let $\alpha$ be an adapted control and $X$ be the corresponding state:

$$\begin{cases}
dX_t = b(t, X_t, m(t, \cdot), \alpha_t)dt + \sigma(t, X_t, m(t, \cdot))dW_t, & t \in (0, T]; \\
X_0 = \xi_0.
\end{cases}$$

(34)

We have from Itô’s formula that

$$\mathbb{E}[u(T, X_t)] = \mathbb{E}[u(0, \xi_0)] + \int_0^T \left[ \frac{\partial u}{\partial t}(t, X_t) + \langle Du(t, X_t), b(t, X_t, m(t, \cdot), \alpha_t) \rangle \right.$$

$$+ \sum_{i,j=1}^n a_{ij}(t, X_t, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X_t) \left. \right] dt.$$
From the terminal condition of $u$ in (6) and the definition of the Hamiltonian, we have

$$J(\alpha|m) = E\left[ u(0, \xi_0) + \int_0^T \left( \frac{\partial u}{\partial t}(t, X_t) + \sum_{i,j=1}^n a_{ij}(t, X_t, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X_t) + H(t, X_t, m(t, \cdot), \alpha_t, Du(t, X_t)) \right) dt \right].$$

(35)

From Assumptions (C4), we know that

$$H(t, X_t, m(t, \cdot), \alpha_t, Du(t, X_t)) \geq H(t, X_t, m(t, \cdot), \phi(t, X_t, m(t, \cdot), Du(t, X_t)), Du(t, X_t)).$$

(36)

Plugging (36) into (35), since $u$ is a solution of (6), we have

$$J(\alpha|m) \geq E\left[ u(0, \xi_0) + \int_0^T \left( \frac{\partial u}{\partial t}(t, X_t) + \sum_{i,j=1}^n a_{ij}(t, X_t, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X_t) + H(t, X_t, m(t, \cdot), \phi(t, X_t, Du(t, X_t)), Du(t, X_t)) \right) dt \right].$$

$$= E[u(0, \xi_0)] = \int_{\mathbb{R}^n} u(0, x)m_0(dx).$$

This shows that $J(\alpha|m) \geq E[u(0, \xi_0)]$ for any adapted control $\alpha$. If we replace $\alpha$ with $\bar{\alpha}$ in the above computation, then the state process $X$ becomes $\bar{X}$ and all the above inequalities are equalities. So $J(\bar{\alpha}|m) = E[u(0, \xi_0)]$ and the proof is complete. \[\square\]

From Theorem 3.2, we know that under Assumptions (A1)-(A3) and (B1)-(B3), MFEs (27) has at least one classical solution $(u, m) \in C^{1+\frac{1}{2}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n) \times C^{1+\frac{1}{2}, 2+\frac{1}{4}}([0, T] \times \mathbb{R}^n)$. Moreover, from the proof of Theorem 2.1, we have $m \in D_{C_1} \subset C^2([0, T], \mathcal{P}_1(\mathbb{R}^n))$. Since Assumptions (A1)-(A3) and (B1)-(B3) can imply Assumptions (C1)-(C4), as a direct consequence of Theorem 3.5, we have the following result.

**Corollary 3.6.** Let Assumptions (A1)-(A3) and (B1)-(B3) be satisfied. Then, $\bar{\alpha} := \{\bar{\alpha}(t, X_t), 0 \leq t \leq T\}$ is an optimal control for MFG (24), where $(u, m)$ is a classical solution of MFEs (27) and $\bar{X} = \{\bar{X}_t, 0 \leq t \leq T\}$ is the solution of SDE (32).

**Remark 3.7.** Carmona and Delarue [17, Theorem 4.44, p.263] consider the MFG problem (24) under the nondegenerate condition with a probabilistic approach. A solution to the MFG problem is obtained by solving the appropriate FBSDE associated with the stochastic control problem. And the existence of solutions of McKean-Vlasov FBSDEs is proved under nondegenerate assumptions [17, Assumption, p.245] in [17, Theorem 4.29, p.246]. Following the stochastic maximum principle for optimality, Carmona and Delarue [17] transform the MFGs into solvability of distribution dependent FBSDEs. They prove the existence of the FBSDEs within a linear-convex framework where the volatility $\sigma$ is a constant. Moreover, a weak mean-reverting condition is needed.

In our work, we allow the volatility to depend upon the distribution of the state. However, we need the differentiability of coefficients in $x$, which seems to be restrictive. That is because we use the analytical method to consider the classical solutions to the FBPDEs. These differentiability assumptions are expected to be relaxed to more general cases by using some appropriate approximation techniques. But this will involve detailed analysis and go beyond the scope of this paper, so will not
be discussed here. Actually, the mean field FBSDEs and mean field FBPDEs are interchangeable. A classical solution of MFEs \([27]\) gives a decoupling function of associated mean field FBSDEs by setting \(Y_t = Du(t, X_t)\) and \(Z_t = \sigma(t, X_t, m(t, \cdot))D^2u(t, X_t)\), where \((Y, Z)\) are the associated adjoint processes. However, obtaining a solution of FBPDEs from a solution of FBSDEs requires additional regularity conditions of the coefficients. And our work builds a bridge.

4 Linear quadratic problems

In this section, we aim to relax the boundedness assumption in the state \(x\) of functions \((a, b, f, g, \phi)\) to include the linear-quadratic case. We denote by \(\mathcal{E}(b, \sigma, f, g, m_0)\) the MFEs with coefficients \((b, \sigma, f, g, m_0)\), and denote by \(\phi^{b,f}\) the feedback function \([20]\) corresponding to \((b, f)\). We denote by \(P(b, \sigma, f, g, \xi_0)\) the MFG \([24]\) corresponding to \((b, \sigma, f, g, \xi_0)\). We first give the following corollary of Theorems \([32, 33]\) which allows the drift \(b\) to be linear in the state \(x\).

**Corollary 4.1.** Suppose that functions \((\sigma, f, g)\) are independent of \(x\),

\[
 b(t, x, m, \alpha) = \lambda(t)x + b'(t, m, \alpha), \quad b' : [0, T] \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \to \mathbb{R}^n,
\]

with the function \(\lambda : [0, T] \to \mathbb{R}^{n \times n}\) being bounded and continuous and functions \((b', \sigma, f, g, m_0)\) satisfy Assumptions \((A1)-(A3)\) and \((B1)-(B3)\). Then, \(\mathcal{E}(b, \sigma, f, g, m_0)\) has at least one classical solution \((u, m) \in C^{1,2}([0, T] \times \mathbb{R}^n) \times C^{1,2}([0, T] \times \mathbb{R}^n)\), and the feedback \(\alpha(t, x) := \phi^{b,f}(t, Du(t, x))\) is an optimal control of \(P(b, \sigma, f, g, \xi_0)\).

**Proof.** We use the following transformation:

\[
 Y_t := e^{-\int_0^t \lambda(s)ds}X_t, \quad t \in [0, T]. \tag{37}
\]

Then, \(\mathcal{E}(b, \sigma, f, g, m_0)\) for \((u, m)\) is transformed into \(\mathcal{E}(\tilde{b}, \tilde{\sigma}, f, g, m_0)\) for \((\tilde{u}, \tilde{m})\), and \(P(b, \sigma, f, g, \xi_0)\) is transformed into MFG \(P(\tilde{b}, \tilde{\sigma}, f, g, \xi_0)\), where for \((t, m, \alpha, p) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^d \times \mathbb{R}^n\),

\[
 \tilde{b}(t, m, \alpha) = e^{-\int_0^t \lambda(s)ds}b'(t, m, \alpha), \quad \tilde{\sigma}(t, m) = e^{-\int_0^t \lambda(s)ds}\sigma(t, m),
\]

\[
 \phi^{\tilde{b},\tilde{f}}(t, m, p) = \phi^{b,f}(t, m, e^{-\int_0^t \lambda(s)ds}p); \tag{38}
\]

and for \((t, y) \in [0, T] \times \mathbb{R}^n\),

\[
 \tilde{u}(t, y) = u(t, e^{\int_0^t \lambda(s)ds}y), \quad \tilde{m}(t, y) = e^{\int_0^t \lambda(s)ds}m(t, e^{\int_0^t \lambda(s)ds}y). \tag{39}
\]

Since functions \((\sigma, b', \phi^{b,f})\) satisfy Assumptions \((A1), (B1)\) and \((B3)\) and function \(\lambda\) is bounded and continuous, it is easy to check that functions \((\tilde{\sigma}, \tilde{b}, \phi^{\tilde{b},\tilde{f}})\) satisfy Assumptions \((A1), (B1)\) and \((B3)\). From Theorem \([32]\) \(\mathcal{E}(\tilde{b}, \tilde{\sigma}, f, g, m_0)\) has at least one classical solution \((\tilde{u}, \tilde{m}) \in C^{1,2}([0, T] \times \mathbb{R}^n) \times C^{1,2}([0, T] \times \mathbb{R}^n)\). From Corollary \([32]\) the feedback \(\tilde{\alpha}(t, y) := \phi^{\tilde{b},\tilde{f}}(t, \tilde{m}(t, \cdot); Du(t, y))\) is an optimal control of \(P(\tilde{b}, \tilde{\sigma}, f, g, \xi_0)\). In view of \([38]-[39]\), \(\mathcal{E}(b, \sigma, f, g, m_0)\) has at least one classical solution \((u, m) \in C^{1,2}([0, T] \times \mathbb{R}^n) \times C^{1,2}([0, T] \times \mathbb{R}^n)\), and the feedback \(\alpha(t, x) := \phi^{b,f}(t, m(t, \cdot); Du(t, x))\) is an optimal control of \(P(b, \sigma, f, g, \xi_0)\). \qed

Now we consider the linear-quadratic case.
Corollary 4.2. Consider $P(b, \sigma, f, g, \xi_0)$ with coefficients

$$b(x, \alpha) = Ax + B\alpha, \quad \sigma : [0, T] \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}^{n \times n},$$

$$f(t, x, m, \alpha) = \frac{1}{2}x^*Qx + \frac{1}{2}m^*Ro + F(t, m), \quad g(x, m) = \frac{1}{2}x^*Mx + G(m),$$

where $A, B, Q, R, M$ are matrices of suitable sizes, $Q, R, M$ are symmetric and $R$ is positive definite. Suppose that $\sigma$ is non-degenerate, functions $(\sigma, F, G)$ are bounded, Lipschitz continuous in $m \in \mathcal{P}_1(\mathbb{R}^n)$ and $\frac{1}{2}$-Hölder continuous in $t \in [0, T]$, and the initial value $m_0$ satisfies Assumption (A3). Then, the optimal control is given by the feedback $\alpha(t, x) = -R^{-1}B^*P(t)x$, where $P(\cdot)$ is the solution of the standard Riccati equation (40).

Proof. We look for a solution of $E(b, \sigma, f, g, m_0)$ of the form $u(t, x) = \beta(t) + \frac{1}{2}x^*P(t)x$. Then, the feedback is of the form $\alpha(t, x) = -R^{-1}B^*P(t)x$, and $E(b, \sigma, f, g, m_0)$ is equivalent to the following equations:

$$\dot{P}(t) + A^*P(t) + P(t)A - 2P(t)BR^{-1}B^*P(t) + Q = 0, \quad P(T) = M,$$

$$\dot{\beta}(t) + \sum_{i,j=1}^n a_{ij}(t, m(t, \cdot))P_{ij}(t) + F(t, m(t, \cdot)) = 0, \quad \beta(T) = G(m(T, \cdot)),$$

$$\frac{\partial m}{\partial t}(t, x) + \text{div}[m(t, x)(A - BR^{-1}B^*P(t))x] - \sum_{i,j=1}^n a_{ij}(t, m(t, \cdot))\frac{\partial^2 m}{\partial x_i \partial x_j}(t, x) = 0, \quad m(0, x) = m_0(x).$$

Let $P(\cdot) \in C^1([0, T]; \mathbb{R}^{n \times n})$ be the solution of the standard Riccati equation (40). By using transformation (37) with $\lambda(t) := A - BR^{-1}B^*P(t)$, in view of Remark 3.3 we know that PDE (42) has a classical solution $m \in C^{1,2}([0, T] \times \mathbb{R}^n)$ and $m(t, \cdot) = \mathcal{L}(\hat{X}_t)$, where

$$\hat{X}_t = \xi_0 + \int_0^t (A - BR^{-1}B^*P(s))\hat{X}_s ds + \int_0^t \sigma(s, m(s, \cdot))dW_s, \quad t \in [0, T].$$

Then equation (41) has a classical solution $\beta \in C^1(0, T)$. Similar to the proof of Theorem 3.5 we know that the feedback $\alpha(t, x)$ is an optimal control of $P(b, \sigma, f, g, \xi_0)$. $\square$

Remark 4.3. Bensoussan et al. [4] solve the FBPDEs for linear quadratic mean field games under the condition that the volatility $\sigma$ is a constant. Our work cannot include Bensoussan’s results because our drift here is independent of the distribution variable $m$. However, our work also go beyond their framework by allowing the volatility $\sigma$ to depend on $m$. Our boundedness assumption in $m$ is expected to be relaxed by using some appropriate approximation techniques.

References

[1] Ahuja, S., Wellposedness of mean field games with common noise under a weak monotonicity condition, SIAM J. Control Optim., 54(1), 2016, 30-48.

[2] Ahuja, S., Ren, W. and Yang, T. W., Forward-backward stochastic differential equations with monotone functions and mean field games with common noise, Stoch. Proc. Appl., 129(10), 2019, 3859-3892.

[3] Barbu, V. and Röckner, M., From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE, Ann. Probab., 48(4), 2020, 1902-1920.
[4] Bensoussan, A., Frehse, J. and Yam, S. C. P., The Master equation in mean field theory, *Journal de Mathématiques Pures et Appliquées*, 103(6), 2015, 1441-1474.

[5] Bensoussan, A., Sung, K. C. J., Yam, S. C. P. and Yung, S. P., Linear-quadratic mean field games, *J. Optimiz. Theory App.*, 169(2), 2016, 496-529.

[6] Buckdahn, R., Li, J., Peng, S. and Rainer, C., Mean-field stochastic differential equations and associated pdes, *Ann. Probab.*, 45(2), 2017, 824-878.

[7] Cardaliaguet, P., Notes on mean field games, *From P.-L. Lions’ lectures at College de France*, 2010.

[8] Cardaliaguet, P., Delarue, F., Lasry, J.-M. and Lions, P.-L., The master equation and the convergence problem in mean field games, *Princeton University Press.*, 381, 2019, AMS-201.

[9] Carmona, R. and Delarue, F., Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics, *Ann. Probab.*, 43(5), 2015, 2647-2700.

[10] Carmona, R. and Delarue, F., Probabilistic analysis of mean-field games, *SIAM J. Control Optim.*, 51(4), 2013, 2705-2734.

[11] Carmona, R. and Delarue, F., Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games, Probability Theory and Stochastic Modelling, Springer International Publishing, 2018.

[12] Carmona, R., Delarue, F. and Lachapelle, A., Control of McKean–Vlasov dynamics versus mean field games, *Math. Financ. Econ.*, 7(2), 2013, 131–166.

[13] Carmona, R., Delarue, F. and Lack, D., Mean field games with common noise, *Ann. Probab.*, 44(6), 2016, 3740-3803.

[14] Carmona, R. and Lack, D., A probabilistic weak formulation of mean field games and applications, *Ann. Probab.*, 25(3), 2015, 1189–1231.

[15] Carmona, R. and Zhu, X., A probabilistic approach to mean field games with major and minor players, *Ann. Appl. Probab.*, 26(3), 2016, 1535-1580.

[16] Chassagneux, J. F., Crisan, D. and Delarue, F., A probabilistic approach to classical solutions of the master equation for large population equilibria, *Mathematics*, 2014.

[17] Crisan, D. and McMurray, E., Smoothing properties of McKean-Vlasov SDEs, *Probab. Theory Rel.*, 171(1), 2018, 97-148.

[18] Du, K. and Wu, Z., Linear quadratic mean-field-game of backward stochastic differential systems, *Math. Control Relat. F.*, 8(3&4), 2018, 653.

[19] Du, K. and Wu, Z., Linear-quadratic stackelberg game for mean-field backward stochastic differential system and application, *Math. Probl. Eng.*, 2019.

[20] Hu, Y. and Peng, S., Maximum principle for semilinear stochastic evolution control systems, *Stochastics*, 33(3), 1990, 159–180.

[21] Huang, M., Caines, P. E. and Malhamé, R. P., Large-population cost-coupled lqg problems with nonuniform agents: Individual-mass behavior and decentralized ε-nash equilibria, *IEEE T. Automat. Contr.*, 52(9), 2017, 1560–1571.

[22] Huang, M., Malhamé, R. P. and Caines, P. E., Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle, *Communications in Information and Systems*, 6(3), 2006, 221–252.

[23] Huang, Z. and Tang, S., Mean Field Games with Common Noises and Conditional Distribution Dependent FBSDEs, *arXiv preprint [arXiv:2105.11984]*, 2021.

[24] Jourdain, B., Méleard, S. and Woyczynski, W. A., Nonlinear SDEs driven by Lévy processes and related PDEs, *ALEA Lat. Am. J. Probab. Math. Stat.*, 4, 2008. 1-29.

[25] Kolokoltsov, V. N., Li, J. and Yang, W., Mean field games and nonlinear Markov processes, *Mathematics*, 2011.

[26] Kolokoltsov, V. N. and Yang, W., Existence of solutions to path-dependent kinetic equations and related forward - backward systems, *Open Journal of Optimization*, 2:2, 2013, 39-44.

[27] Ladyzenskaja, O. A., Solonnikov, V. A. and Ural’ceva, N. N., Linear and quasilinear equations for parabolic type, *Translations of Mathematical Monographs*, V. 23, American Mathematical Society, Providence, R.I. 1967.
[28] Lacker, D., A general characterization of the mean field limit for stochastic differential games, *Probab. Theory Rel.*, 165(3), 2016, 581-648.

[29] Lasry, J.-M. and Lions, P.-L., Jeux à champ moyen. i–le cas stationnaire, *C.R.A.S. Math.*, 343(9), 2006, 619–625.

[30] Lasry, J.-M. and Lions, P.-L., Jeux à champ moyen. ii–horizon fini et contrôle optimal, *C.R.A.S. Math.*, 343(10), 2006, 679–684.

[31] Lasry, J.-M. and Lions, P.-L., Mean field games, *Jpn. J. Math.*, 2(1), 2007, 229–260.

[32] McKean, H. P. Jr., A class of Markov processes associated with nonlinear parabolic equations, *Proc. Natl. Acad. Sci. USA*, 56(6), 1966, 1907-1911.

[33] Nourian, M. and Caines, P. E., $\epsilon$-nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents, *SIAM J. Control Optim.*, 51(4), 2012, 2090-2095.

[34] Pham, H., Linear quadratic optimal control of conditional McKean-Vlasov equation with random coefficients and applications, *Probability, Uncertainty and Quantitative Risk*, 1(1), 2016, 7.

[35] Pham, H. and Wei, X., Bellman equation and viscosity solutions for mean-field stochastic control problem, *Mathematics*, 47(1), 2015, 444-475.

[36] Pham, H. and Wei, X., Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics, *SIAM J. Control Optim.*, 55(2), 2017, 1069-1101.

[37] Sznitman, A.-S., Topics in propagation of chaos, Springer Berlin Heidelberg, 1991.

[38] Tse, A., Higher order regularity of nonlinear Fokker-Planck PDEs with respect to the measure component, *Journal de Mathématiques Pures et Appliquées* 150, 2021, 134-180.

[39] Yu, J. and Tang, S., Mean-field game with degenerate state- and distribution-dependent noise: a probabilistic approach, *Chinese Ann. Math.*, 2A(1), 2019, 1-21.