Mean Field Games with Mean-Field-Dependent Volatility: A PDE Approach*

Ziyu HUANG\textsuperscript{1} \quad Shanjian TANG\textsuperscript{2}

Abstract

In this paper, we develop a PDE approach to consider the optimal strategy of mean field controlled stochastic system. Firstly, we discuss mean field SDEs and associated Fokker-Plank equations. Secondly, we consider a fully-coupled system of forward-backward PDEs. The backward one is the Hamilton-Jacobi-Bellman equation while the forward one is the Fokker-Planck equation. Our main result is to show the existence of classical solutions of the forward-backward PDEs in the class $H^{1+rac{1}{4},2+rac{1}{2}}([0,T] \times \mathbb{R}^n)$ by use of the Schauder fixed point theorem. Then, we use the solution to give the optimal strategy of the mean field stochastic control problem. Finally, we give an example to illustrate the role of our main result.

Keywords \ mean field games, PDE, mean field equations.

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

1 Introduction

Mean field games (MFGs) were proposed by Lasry and Lions in a series of papers \cite{22, 23, 24} and also independently by Huang, Caines and Malhamé \cite{17}, under the different name of Nash Certainty Equivalence. They are sometimes approached by symmetric, non-cooperative stochastic differential games of interacting $N$ 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. In view of the theory of McKean–Vlasov limits and propagation of chaos for uncontrolled weakly interacting particle systems \cite{29}, it is expected to have a convergence for $N$-player game Nash equilibria by assuming independence of the random noise in the players’ state processes and some symmetry conditions of the players. The literature in this area is huge. See \cite{6} for a summary of a series of Lions’ lectures given at the Collège de France. Carmona et al. discussed the MFG problem with a probabilistic approach. See \cite{7, 8, 9, 10, 11, 12}. On results about construction of $\epsilon$-Nash equilibria for $N$-player games, see \cite{8, 11, 18, 19, 21, 22, 27}. For results about mean field games and mean field control problems with common noises, we refer to \cite{11, 2, 26}. For some other works about mean field games, we refer to \cite{14, 15, 30}.

\textsuperscript{1}School of Mathematical Sciences, Fudan University, Shanghai 200433, China. E-mail: 19110180044@fudan.edu.cn
\textsuperscript{2}Department of Finance and Control Sciences, School of Mathematical Sciences, Fudan University, Shanghai 200433, China. E-mail: sjtang@fudan.edu.cn
\textsuperscript{*}This work was partially supported by National Key R&D Program of China under grant 2018YFA0703900 and by National Natural Science Foundation of China under grants 11631004 and 12031009.
In this paper, we consider the controlled mean field stochastic system

\[
\begin{aligned}
\frac{dX_t}{dt} &= b(t, X_t, \mathcal{L}(X_t^\alpha), \alpha_t)dt + \sigma(t, X_t, \mathcal{L}(X_t^\alpha))dW_t, \quad t \in (0, T]; \\
X_0 &= \xi_0 \sim m_0,
\end{aligned}
\]

where \(\{W_t, 0 \leq t \leq T\}\) is a Brownian motion in \(\mathbb{R}^n\), \(\alpha\) is an admissible control, taking values in \(\mathbb{R}^n\), adapted to the filtration \(\mathcal{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}\) generated by \((W, \xi_0)\), \(\hat{\alpha}\) is the optimal control and \(\mathcal{L}(\cdot)\) is the law. The goal is to minimize the cost functional

\[
J(\alpha|{(\mathcal{L}(X_t^\alpha))}_{0 \leq t \leq T}) = \mathbb{E}\left[ \int_0^T f(t, X_t, \mathcal{L}(X_t^\alpha), \alpha_t)dt + g(X_T, \mathcal{L}(X_T^\alpha)) \right].
\]

Carmona and Delarue [9] consider the MFG problem under the nondegenerate condition with a probabilistic approach. A solution to the MFG problem is obtained by solving the appropriate forward-backward stochastic differential equation (FBSDE) associated with the stochastic control problem. Following the stochastic maximum principle for optimality [16], Carmona and Delarue [8] transform the MFG problem into existence and uniqueness of solution of an FBSDE within a linear-convex framework. The drift \(b\) of the state is assumed to be linear and the volatility \(\sigma\) is assumed to be a constant. The terminal cost \(g\) is assumed to be convex in \(x\) and the running cost \(f\) has to satisfy a strong convexity assumption. Moreover, a weak mean-reverting condition is needed. Carmona and Delarue [7] also consider the controlled McKean-Vlasov dynamics. The solvability of the systems requires the drift and volatility to be linear, the differentiability of functions \(f\) and \(g\) of measures, and the convexity assumption of \(f\) and \(g\). In our work, we use a PDE approach, which does not require any convexity assumption usually imposed in Pontryagin principle. 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 make more differentiability assumptions on \(x\). That is because we use the analytical method to consider the classical solutions to the forward-backward PDE, rather than a probabilistic representation, so the differentiability of coefficients is needed. These differentiability assumptions are expected to be relaxed to more general cases by using the appropriate approximation technique.

The mean field control problems and mean field stochastic differential equations (SDEs) are closely associated with PDEs. There are three major issues of interest in MFGs and associated PDEs. (i) studying existence and uniqueness results of classical solutions of a coupled system of forward-backward PDEs by analytical methods, (ii) establishing connection of the optimal strategy of mean field stochastic control problem with the solutions of associated PDEs, (iii) investigating the mean field SDEs to give the solutions of associated PDEs by probabilistic methods. The main goal of this paper is to answer the questions in (i) and (ii) for the class of models we consider.

First we discuss issue (i). Lasry and Lions [22][23][24] are the first to reduce the MFGs to solving a fully-coupled system of forward-backward PDEs. The backward one is a Hamilton-Jacobi-Bellman (HJB) equation for the value function for each player while the forward one is the Fokker-Planck (FP) equation for the evolution of the player’s distribution. In a series of lectures given at the Collége de France [6], Lions showed the existence and uniqueness results for the particular second order mean field equations (MFEs):

\[
\begin{aligned}
\frac{\partial u}{\partial t}(t, x) + \Delta u(t, x) - \frac{1}{2} |Du(t, x)|^2 + F(x, m(t, \cdot)) &= 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n; \\
\frac{\partial m}{\partial t}(t, x) - \Delta m(t, x) - \text{div}(Du(t, x)m(t, x)) &= 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
m(0, x) &= m_0(x), \quad u(T, x) = G(x, m(T, \cdot)), \quad x \in \mathbb{R}^n.
\end{aligned}
\]
Pham and Wei [27, 28] use dynamic programming for optimal control and give explicit solutions to Bellman equation for the linear quadratic mean-field control problem, with applications to the mean-variance portfolio selection and a systemic risk model. Bensoussan et al. [3] solve the forward-backward partial PDEs for linear quadratic mean field games under the condition that the volatility $\sigma$ is a constant. In our work, we do not need the linear quadratic condition and allow the volatility $\sigma$ to depend on the state and the distribution of the state. In Section 3, we generalize Lions’ existence result of equation (3) [6, Theorem 3.1, pp.10] by considering the more general second order MFEs (4)-(6), the feedback strategy $\bar{\alpha}(t,x)$ is optimal for the optimal stochastic control problem

$$\min_{\alpha} J(\alpha) := \mathbb{E} \left[ \int_0^T \frac{1}{2} |\alpha_s|^2 + F(X_s, m_s) \right] dt + G(X_T, m_T);$$

where $\phi$ is the feedback strategy under suitable assumptions. We show the existence of solutions of MFEs (4)-(6) in the class $H^{1+\frac{1}{2}+\frac{n}{2}}([0,T] \times \mathbb{R}^n)$ (see Theorem 3.1). As is well known, equation (4) is the HJB equation for an optimal control problem and equation (5) is the FP equation for the probability measure of the player’s state. We construct a convex and compact subset of $C^0([0,T], P_1(\mathbb{R}^n))$ and use the Schauder fixed point theorem to establish the existence of a solution to (4)-(6). The proof relies on some results of quasilinear equations for parabolic type [20].

Next we discuss issue (ii). Lions [6] shows that for a solution $(u, m)$ of the MFE (4), the feedback strategy $\bar{\alpha}(t,x) := -D_x u(t,x)$ is optimal for the optimal stochastic control problem

$$\min_{\alpha} J(\alpha) := \mathbb{E} \left[ \int_0^T \frac{1}{2} |\alpha_s|^2 + F(X_s, m_s) \right] dt + G(X_T, m_T);$$

$$X_t^u = X_0 + \int_0^t \alpha_s ds + \sqrt{2} dW_s, \quad t \in [0,T].$$

Pham and Wei [24, 27, 28] study the dynamic programming for optimal control of stochastic McKean-Vlasov dynamics and prove a verification theorem in the McKean-Vlasov framework. In Section 4 of our paper, we generalize Lions’ result [6 Lemma 3.7, pp.15] and show that for a solution $(u, m)$ of the MFE (4)-(6), the feedback strategy $\bar{\alpha}(t,x) := \phi(t,x, Du(t,x))$ is optimal for the control problem (11)-(12) with $m_t$ being the distribution of the state corresponding to the optimal strategy $\bar{\alpha}(t,x)$.

Then we discuss issue (iii). Buckdahn et al. [5] consider a process $X^{t,x,\xi}$ as the solution of a mean-field SDE with initial data $(t,x)$ whose coefficients depend on the law of $X^{t,\xi}$. They characterize the function $V(t,x,P_\xi) = \mathbb{E}[\Phi(X_T^{t,x,P_\xi}, P_{X_T^{t,\xi}})]$ under appropriate regularity conditions as the unique classical solution of a PDE for measure-dependent fields by studying the first- and second-order derivatives of the solution of the mean-field SDE with respect to the initial probability law. Chassagneux et al. [13] show that the master equations admit a classical solution on sufficiently small time intervals by studying the differentiability with respect to the initial condition of the flow generated by a forward-backward stochastic system of McKean-Vlasov type.
Summarizing our contributions in this paper, we
– show the existence of solutions of MFEs (4)-(6) in the class $H^{1+\frac{1}{2}+\frac{1}{4}}([0,T] \times \mathbb{R}^n)$;
– give the optimal strategy of the control problem (1) and (2);
– give an example to illustrate the role of our main result.

The paper is organized as follows. In Section 2, we discuss mean field SDEs and associated Fokker-Plank equations. In Section 3, we give the existence result of the solution to mean field equations (4)-(6) in Theorem 3.1. In Section 4, we use the solution $(u, m)$ to equations (4)-(6) to construct an optimal strategy of mean field stochastic control problem (1) and (2). Finally, in Section 5 we give an example to illustrate the role of our Theorem 3.1.

1.1 Notations

Let $(\Omega_0, \mathcal{F}, \{\mathcal{F}_t, 0 \leq t \leq T\}, \mathbb{P})$ 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 $\mathcal{S}^2_{\mathcal{F}}(0, T)$ 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
\[ \mathbb{E}\left[ \sup_{0 \leq t \leq T} |\beta_t|^2 dt \right] < +\infty. \]

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
\[ \int_{\mathbb{R}^n} |x| dm(x) < \infty. \]

The Kantorovitch-Rubinstein distance is defined on $\mathcal{P}_1(\mathbb{R})$ by
\[ d_1(m_1, m_2) = \inf_{\gamma \in \Pi(m_1, m_2)} \int_{\mathbb{R}^2n} |x - y| d\gamma(x, y), \quad m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^n), \]
where $\Pi(m_1, m_2)$ denotes the collection of all probability measures on $\mathbb{R}^{2n}$ with marginals $m_1$ and $m_2$. The space $(\mathcal{P}_1(\mathbb{R}), d_1)$ is a complete separable metric space.

$\Omega$ is a domain in $\mathbb{R}^n$, i.e. an arbitrary open connected subset of $\mathbb{R}^n$. In this paper, $\Omega$ is considered to be a bounded domain. $S$ is the boundary of $\Omega$.

$Q_T$ is the cylinder $(0, T) \times \Omega$. $S_T$ is the lateral surface of $Q_T$, or more precisely the set of points $(t, x) \in [0, T] \times S$.

$H_+^l(Q_T)$ is the Banach space of all continuous functions $u(\cdot, \cdot)$ in $Q_T$, having all the continuous derivatives $D^r_t D^s_x$ with $2r + s < l$, and having a finite norm
\[ |u|^{(l)}_{Q_T} = \langle u \rangle^{(l)}_{Q_T} + \sum_{j=0}^{[l]} \langle u \rangle^{(j)}_{Q_T}, \]
In this section, we discuss the mean field SDE and associated Fokker-Planck equation. Let

\[
\langle u \rangle_{Q_T}^{(j)} = \sum_{2r+s=j} |D^r_tD^su|_{Q_T}^{(0)}, \quad \langle u \rangle_{Q_T}^{(0)} = \sup_{Q_T} |u|;
\]

where

\[
\langle u \rangle_{Q_T}^{(j)} = \langle u \rangle_{t,Q_T}^{(j)} + \langle u \rangle_{x,Q_T}^{(j)},
\]

\[
\langle u \rangle_{t,Q_T}^{(j)} = \sum_{0 \leq t-2r-s \leq 2} \langle D^r_tD^su \rangle_{t,Q_T}^{(j-2r-s)}, \quad \langle u \rangle_{x,Q_T}^{(j)} = \sum_{2r+s=\|f\|} \langle D^r_tD^su \rangle_{x,Q_T}^{(j-|f|)};
\]

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

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

\(C^{1,2}(\bar{Q}_T)\) (or \(C^{1,2}(Q_T)\)) is the set of all continuous functions in \(\bar{Q}_T\) (in \(Q_T\)) having continuous derivatives \(u_x, u_{xx}, u_t\) in \(\bar{Q}_T\) (in \(Q_T\)).

\(O^{1,2}(\bar{Q}_T)\) (or \(O^{1,2}(Q_T)\)) is the set of all differentiable functions \(i\) in \(\bar{Q}_T\) (in \(Q_T\)) whose spatial derivative \(u_x\) is differentiable in \(x\) at each point of \(\bar{Q}_T\) (of \(Q_T\)) and whose derivatives \(u_x, u_{xx}\) and \(u_t\) are bounded in \(\bar{Q}_T\) (in \(Q_T\)).

\section{Mean field SDE and Fokker-Planck equation}

In this section, we discuss the mean field SDE

\[
\begin{cases}
  dX_t = b(t, X_t, \mathcal{L}(X_t))dt + \sigma(t, X_t, \mathcal{L}(X_t))dW_t, \quad t \in (0, T]; \\
  X_0 = \xi_0 \sim m_0,
\end{cases}
\]

and associated Fokker-Planck equation

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

where

\[
b : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}^n, \quad b(t, x, m) = (b_1(t, x, m), \ldots, b_n(t, x, m))^T;
\]

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

Here are our main assumptions to ensure the existence of the solution of equation (8). For notational convenience, we use the same constant \(L\) for all the conditions below.

\textbf{(A1)} There exist \(0 < \gamma_1 \leq \gamma_2 < +\infty\), such that

\[
\gamma_1 \xi^2 \leq \sum_{i,j=1}^n a_{ij}(t, x, m)\xi_i \xi_j \leq \gamma_2 \xi^2, \quad \forall \xi \in \mathbb{R}^n, \quad (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n).
\]
The functions $b$ and $\sigma$ are uniformly bounded by $L$ over $[0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)$. The functions $b_i(t, \cdot, m) : \mathbb{R}^n \to \mathbb{R}$ are differentiable for all $(t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)$ and $1 \leq i \leq n$, with the derivatives $\partial b_i \partial x_i$ being bounded by $L$. The functions $a_{ij}(t, \cdot, m) : \mathbb{R}^n \to \mathbb{R}$ are twice differentiable for all $(t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)$ and $1 \leq i, j \leq n$, with the derivatives $\partial a_{ij} \partial x_i$ and $\partial^2 a_{ij} \partial x_i \partial x_j$ being bounded by $L$. That is,

$$|b| + |\sigma| + \left| \frac{\partial b_i}{\partial x_i} \right| + \left| \frac{\partial a_{ij}}{\partial x_i} \right| + \left| \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \right| \leq L, \quad (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \quad 1 \leq i, j \leq n.$$

The functions $(b, \frac{\partial b}{\partial x}, a_{ij}, \frac{\partial a_{ij}}{\partial x_i}, \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}) : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)$ are $L$-Lipschitz continuous in $(x, m)$ and $\frac{1}{2}$-Hölder continuous in $t$. That is,

$$|\langle b, \frac{\partial b}{\partial x_i}, a_{ij}, \frac{\partial a_{ij}}{\partial x_i}, \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \rangle(t', x', m') - \langle b, \frac{\partial b}{\partial x_i}, a_{ij}, \frac{\partial a_{ij}}{\partial x_i}, \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \rangle(t, x, m)|$$

$$\leq L(|t' - t|^{\frac{1}{2}} + |x' - x| + d_1(m', m)), \quad (t, x, m), (t', x', m') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n), \quad 1 \leq i, j \leq n.$$
3 Analysis of second order MFEs

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

\[
\begin{aligned}
\frac{\partial u}{\partial t}(t, x) + \langle Du(t, x), b(t, x, m(t, \cdot), \phi(t, x, Du(t, x))) \rangle + \sum_{i,j=1}^{n} a_{ij}(t, x, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) \\
+ f(t, x, m(t, \cdot), \phi(t, x, Du(t, x))) = 0, \quad (t, x) \in (0, T) \times \mathbb{R}^n; \\
\frac{\partial m}{\partial t}(t, x) - \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x, m(t, \cdot))m(t, x)] \\
+ \sum_{i=1}^{n} \frac{\partial}{\partial x_i} [b_i(t, x, m(t, \cdot), \phi(t, x, Du(t, x)))m(t, x)] = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n; \\
m(0, x) = m_0(x), \quad u(T, x) = g(x, m(T, \cdot)), \quad x \in \mathbb{R}^n.
\end{aligned}
\]  

(9)

Here,

\[
b : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}^n, \quad b(t, x, m, \alpha) = (b_1(t, x, m, \alpha), \ldots, b_n(t, x, m, \alpha))^T; \\
\sigma : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}^n; \quad a_{ij}(t, x, m) = \frac{1}{2} [\sigma \sigma^T(t, x, m)]_{ij}, \quad 1 \leq i, j \leq n; \\
f : [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R} \to \mathbb{R}; \quad g : \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}; \quad m_0 : \mathbb{R}^n \to \mathbb{R}.
\]

Our aim in this section is to prove the existence of classical solutions of equation (9). First we state our main assumptions in this section. For notational convenience, we use the same constant \( L \) for all the conditions below.

(B1) The functions \( b \) and \( f \) are of the forms

\[
b(t, x, m, \alpha) = b^0(t, x, m) + b^1(t, x, \alpha), \quad f(t, x, m, \alpha) = f^0(t, x, m) + f^1(t, x, \alpha),
\]

\((t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n.
\]

(B2) There exists \( 0 < \gamma_1 \leq \gamma_2 < +\infty \), such that

\[
\gamma_1 \xi^2 \leq \sum_{i,j=1}^{n} a_{ij}(t, x, m)\xi_i \xi_j \leq \gamma_2 \xi^2, \quad \forall \xi \in \mathbb{R}^n, \quad (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n).
\]

(B3) The functions \( a_{ij} \) are uniformly bounded by \( L \) for \( 1 \leq i, j \leq n \). The functions \( b_i \) satisfy a linear growth condition in \( \alpha \) for \( 1 \leq i \leq n \). The function \( f \) satisfies a quadratic growth condition in \( \alpha \). Moreover, for \( 1 \leq i, j \leq n \), the functions \( a_{ij}(t, \cdot, \cdot, \cdot) : \mathbb{R}^n \to \mathbb{R} \) are twice differentiable for all \((t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)\), with the derivatives \( \frac{\partial a_{ij}}{\partial x_i} \) and \( \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \) being bounded by \( L \). For \( 1 \leq i \leq n \), the functions \( b_i(t, \cdot, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) are differentiable for all \((t, m) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^n)\), with the derivatives \( \frac{\partial b_i}{\partial x_i} \) satisfying a linear growth condition in \( \alpha \) and \( \frac{\partial b_i}{\partial \alpha} \) being bounded by \( L \). The function \( f(t, x, m, \cdot) : \mathbb{R}^n \to \mathbb{R} \) is differentiable for all \((t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)\), with the derivatives \( \frac{\partial f}{\partial \alpha} \) satisfying a linear growth condition in \( \alpha \). That is,

\[
|a_{ij}| + \left| \frac{\partial a_{ij}}{\partial x_i} \right| + \left| \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} \right| \leq L, \quad (t, x, m) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \quad 1 \leq i, j \leq n,
\]

\[
\frac{|b_i(t, x, m, \alpha)|}{1 + |\alpha|} + \left| \frac{\partial b_i}{\partial x_i}(t, x, m, \alpha) \right| + \left| \frac{\partial b_i}{\partial \alpha} \right| \leq L, \quad (t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n, \quad 1 \leq i \leq n,
\]

\[
\frac{|f(t, x, m, \alpha)|}{1 + |\alpha|^2} + \left| \frac{\partial f}{\partial \alpha}(t, x, m, \alpha) \right| \leq L, \quad (t, x, m, \alpha) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n.
\]
Theorem 3.1. Suppose that Assumptions (B1)-(B7) hold. Then, equation (12) has at least one classical solution $(u, m) \in H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T)$.

Remark 3.2. Theorem 3.1 includes Lions’ existence result of equation (12) [Theorem 3.1, pp. 10], with the optimal feedback control $\phi(t, x, p) = -p$. 

(B4) The functions $F_{ij} := (a_{ij}, \frac{\partial a_{ij}}{\partial x_i}, \frac{\partial a_{ij}}{\partial x_j}, b_i, \frac{\partial b_i}{\partial x_i}, \frac{\partial a_i}{\partial a_j}, g)$ for $1 \leq i, j \leq n$ and $f$ satisfy

$$|F_{ij}(t', x', m', \alpha') - F_{ij}(t, x, m, \alpha)| \leq L(|t' - t|^{\frac{1}{2}} + |x' - x| + d_1(m', m) + |\alpha' - \alpha|),$$

$$(t, x, m, \alpha), (t', x', m', \alpha') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n, \quad 1 \leq i, j \leq n;$$

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

$$(t, x, m, \alpha), (t', x', m', \alpha') \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n.$$

(B5) The function $g(\cdot, m) \in H^{2+\frac{1}{2}}(\mathbb{R}^n)$ for all $m \in \mathcal{P}_1(\mathbb{R}^n)$. There exists $0 < \beta < 1$ such that

$$\sup_{m \in \mathcal{P}_1(\mathbb{R}^n)} |g(x, m)|^{(1+\beta)} \leq L.$$

(B6) The distribution $m_0$ is absolutely continuous with respect to the Lebesgue measure, has a Hölder continuous density (still denoted by $m_0$) such that $m_0 \in H^{2+\frac{1}{2}}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} |x|^2 m_0(dx) < +\infty$.

We define the Hamiltonian:

$$H(t, x, m, \alpha, p) := \langle p, b(t, x, m, \alpha) \rangle + f(t, x, m, \alpha)$$

$$= \langle [p, b^0(t, x, m)] + f^0(t, x, m) \rangle + \langle [p, b^1(t, x, \alpha)] + f^1(t, x, \alpha) \rangle,$$

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

and the minimizing control function

$$\phi(t, x, p) := \arg\min_{\alpha \in A} \langle p, b^1(t, x, \alpha) \rangle + f^1(t, x, \alpha), \quad (t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n,$$

where $A \subset \mathbb{R}^n$ is the control space. We make the following assumption on $\phi$.

(B7) For each $(t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$, there exists a unique vector $\phi(t, x, p) \in \mathbb{R}^n$ satisfying (13) and growing linearly in $p$. Moreover, the function $\phi(t, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is differentiable for all $t \in [0, T]$ with the derivatives $\frac{\partial \phi}{\partial x_i}$ growing linearly in $p$ and $\frac{\partial \phi}{\partial p}$ being bounded by $L$. That is,

$$\frac{|\phi(t, x, p)|}{1 + |p|} + \frac{|\frac{\partial \phi}{\partial x_i}(t, x, p)|}{1 + |p|} + \frac{|\frac{\partial \phi}{\partial p}(t, x, p)|}{1 + |p|} \leq L, \quad (t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n, \quad 1 \leq i \leq n.$$

The functions $(\phi, \frac{\partial \phi}{\partial x_i}, \frac{\partial \phi}{\partial p})(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ are $L$-Lipschitz continuous in the last two arguments and $\frac{1}{2}$-Hölder continuous in the first argument. That is,

$$|\phi(t, x, p), \phi(t', x', p') - \phi(t, x, p) \langle \phi(t, x, p), \phi(t', x', p') \rangle \leq L(|t' - t|^\frac{1}{2} + |x' - x| + |p' - p|),$$

$$(t, x, p), (t', x', p') \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n, \quad 1 \leq i \leq n.$$

We have the following main result in this section:

Theorem 3.1. Suppose that Assumptions (B1)-(B7) hold. Then, equation (12) has at least one classical solution $(u, m) \in H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T) \times H^{1+\frac{1}{2}}(0, T)$.
The proof of Theorem 3.1 is divided into several parts and relies on the use of Schauder fixed point theorem. We first define $D_{C_1} \subseteq C^0([0, T], \mathcal{P}_1(\mathbb{R}^n))$ as follows:

$$D_{C_1} := \{\mu \in C^0([0, T], \mathcal{P}_1(\mathbb{R}^n)) : \sup_{s \neq t} \frac{d_1(\mu(s), \mu(t))}{|t - s|^2} \leq C_1, \sup_{t \in [0, T]} \int_{\mathbb{R}^n} |x|^2 \mu(t)(dx) \leq C_1\}, \quad (12)$$

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, see [6, Lemma 5.7]. Now we define a map $\Phi : D_{C_1} \rightarrow D_{C_1}$. For any $\mu \in D_{C_1}$, let $u$ be the unique solution to the following PDE:

$$\begin{align*}
\frac{\partial u}{\partial t}(t, x) + \langle Du(t, x), b(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) \rangle + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) \\
+ f(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n; \\
u(T, x) = g(x, \mu(T, \cdot)), \quad x \in \mathbb{R}^n.
\end{align*} \quad (13)$$

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

$$\begin{align*}
\frac{\partial m}{\partial t}(t, x) - \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(t, x, \mu(t, \cdot)) m(t, x)) \\
+ \sum_{i=1}^n \frac{\partial}{\partial x_i}(b_i(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) m(t, x)) = 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
m(0, x) = m_0(x), \quad x \in \mathbb{R}^n.
\end{align*} \quad (14)$$

We first show that $\Phi$ is well-defined, i.e. equations (13) and (14) have a unique solution, and further that the solution of (14) belongs to $D_{C_1}$. We then show the continuing of $\Phi$, and Theorem 3.1 as a consequence of Schauder fixed point theorem.

### 3.1 Existence and uniqueness of the solution of equation (13)

To show the existence and uniqueness of the solution to equation (13), we need the following theorem, which is available in [20, Theorem 8.1, pp.495; Theorem 2.9, pp.23].

**Theorem 3.3.** Consider the quasi-linear equation:

$$\begin{align*}
u_t(t, x) - \sum_{i=1}^n \frac{\partial}{\partial x_i} a_i(t, x, u, u_x)(t, x) + a(t, x, u, u_x)(t, x) &= 0, \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
u(0, x) &= \phi_0(x), \quad x \in \mathbb{R}^n.
\end{align*} \quad (15)$$

Suppose that the following conditions hold

(a) $\phi_0(\cdot) \in H^{2+l}(\mathbb{R}^n)$ for some $l \in (0, 1)$.

(b) For $(t, x, u, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, the functions $a_i(t, x, u, p)$ and $a(t, x, u, p)$ are continuous, $a_i(t, x, u, p)$ are differentiable with respect to $(x, u, p)$. Define

$$a_{ij}(t, x, u, u_x) := \frac{\partial a_i(t, x, u, p)}{\partial p_j},$$

$$A(t, x, u, p) := a(t, x, u, p) - \sum_{i=1}^n \frac{\partial a_i(t, x, u, p)}{\partial u} p_i - \sum_{i=1}^n \frac{\partial a_i(t, x, u, p)}{\partial x_i}. \quad (16)$$

9
Inequalities

\[ \sum_{i,j=1}^{n} a_{ij}(t, x, u, p)\xi_{i}\xi_{j} \geq 0, \quad \forall \xi \in \mathbb{R}^{n}, \]  
\[ A(t, x, u, 0)u \geq -b_{1}u^{2} - b_{2} \]  

are fulfilled for some constants \( b_{1}, b_{2} \in [0, +\infty) \).

(c) For \((t, x, p) \in [0, T] \times \mathbb{R}^{n} \times \mathbb{R}^{n} \) and \(|u| \leq M \) where

\[ M := \inf_{\lambda > b_{1}} e^{\lambda T} \left[ \max_{x \in \mathbb{R}^{n}} |\phi_{0}(x)| + \sqrt{\frac{b_{2}}{\lambda - b_{1}}} \right], \]  

the inequalities

\[ \gamma \xi^{2} \leq \sum_{i,j=1}^{n} a_{ij}(t, x, u, p)\xi_{i}\xi_{j} \leq \mu \xi^{2}, \quad \forall \xi \in \mathbb{R}^{n}, \]  
\[ \sum_{i=1}^{n}(|a_{i}| + |\frac{\partial a_{i}}{\partial u}|)(1 + |p|) + \sum_{i,j=1}^{n} |\frac{\partial a_{i}}{\partial x_{j}}| + |a| \leq \mu(1 + |p|)^{2} \]  

are fulfilled for some constants \( \gamma, \mu \in (0, +\infty) \).

(d) For \( 1 \leq i, j \leq n, (t, x) \in [0, T] \times \mathbb{R}^{n}, |u| \leq M \) and \(|p| \leq M_{1} \), where \( M_{1} \in (0, +\infty) \) is a constant, the functions \( (a, a_{i}, a_{ij}, \frac{\partial a_{i}}{\partial u}, \frac{\partial a_{i}}{\partial x_{j}}) \) are continuous functions satisfying a Hölder condition in \((t, x, u, p)\) with exponents \( (\frac{1}{2}, 1, 1, 1) \) respectively.

Then, quasi-linear PDE (15) has a solution \( u \in H^{1+\frac{1}{2}, 2+l}([0, T] \times \mathbb{R}^{n}) \) such that \( u(t, x) \leq M \) for all \((t, x) \in [0, T] \times \mathbb{R}^{n} \).

If in addition the functions \( a_{ij}(t, x, u, p) \) and \( A(t, x, u, p) \) defined in (16) are differentiable with respect to \((u, p)\) such that

\[ \left| \frac{\partial a_{ij}(t, x, u, p)}{\partial u} \right| + \left| \frac{\partial a_{ij}(t, x, u, p)}{\partial p} \right| + \left| \frac{\partial A(t, x, u, p)}{\partial p} \right| \leq \mu_{1}(N), \]
\[ (t, x) \in [0, T] \times \mathbb{R}^{n}, \quad |(u, p)| \leq N, \quad \forall N \in \mathbb{N}, \]
\[ \frac{\partial A(t, x, u, p)}{\partial u} \geq -\mu_{2}(N), \quad (t, x) \in [0, T] \times \mathbb{R}^{n}, \quad |(u, p)| \leq N, \quad \forall N \in \mathbb{N}, \]

with some constants \( \mu_{1}(N) \) and \( \mu_{2}(N) \) for an arbitrary \( N \). Then (15) has no more than one classical solution \( u(\cdot, \cdot) \) that is bounded in \([0, T] \times \mathbb{R}^{n} \) and has bounded derivatives of first and second orders.

In view of above theorem, to show the existence of solution of (13), we only need to check whether the conditions in Theorem 3.3 are fulfilled in our model. We set

\[ v(t, x) = u(T - t, x), \quad (t, x) \in [0, T] \times \mathbb{R}^{n}, \]

then equation (13) reads

\begin{equation}
\begin{cases}
v_{t}(t, x) - \sum_{i=1}^{n} \frac{\partial}{\partial x_{i}}a_{t}(t, x, v_{x})(t, x) + a(t, x, v_{x})(t, x) = 0, & (t, x) \in (0, T] \times \mathbb{R}^{n}; \\
v(0, x) = g(x, \mu(T, \cdot)), & x \in \mathbb{R}^{n},
\end{cases}
\end{equation}
where

\[ a_i(t, x, p) = \sum_{j=1}^{n} a_{ij}(t, x, \mu(t, \cdot))p_j, \]

\[ a(t, x, p) = \sum_{i,j=1}^{n} \frac{\partial a_{ij}}{\partial x_i}(t, x, \mu(t, \cdot))p_j - \langle p, b(t, x, \mu(t, \cdot), \phi(t, x, p)) \rangle - f(t, x, \mu(t, \cdot), \phi(t, x, p)). \]

From Assumption (B5), we know that \( g(\cdot, \mu(T)) \in H^{2+\frac{1}{2}}(\mathbb{R}^n) \). So Condition (a) in Theorem 3.3 is fulfilled with \( l = \frac{1}{2} \).

To check (b), we note that in our model,

\[ a_{ij}(t, x) = a_{ij}(t, x, \mu(t, \cdot)), \]

\[ A(t, x, p) = -\langle p, b(t, x, \mu(t, \cdot), \phi(t, x, p)) \rangle - f(t, x, \mu(t, \cdot), \phi(t, x, p)). \]

From Assumption (B2), we have (17). From Assumptions (B3) and (B7), we have

\[ A(t, x, p) = -\langle p, b(t, x, \mu(t, \cdot), \phi(t, x, p)) \rangle - f(t, x, \mu(t, \cdot), \phi(t, x, p)). \]

which yields (18).

Now we check (c). From Assumption (B2), we have (20). From Assumptions (B3) and (B7), we know that for \((t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n\),

\[ \sum_{i=1}^{n} (|a_i| + \left| \frac{\partial a_i}{\partial v} \right|)(1 + |p|) + \sum_{i,j=1}^{n} \left| \frac{\partial a_i}{\partial x_j} \right| + |a| \]

\[ \leq \sum_{i,j=1}^{n} |a_{ij}| |p_j|(1 + |p|) + \sum_{i,j=1}^{n} (\left| \frac{\partial a_{ij}}{\partial x_i} \right| + \left| \frac{\partial a_{ij}}{\partial x_j} \right|)|p| \]

\[ + |p| |b(t, x, \mu(t, \cdot), \phi(t, x, p))| + |f(t, x, \mu(t, \cdot), \phi(t, x, p))| \]

\[ \leq C(L)(1 + |p|^2) + L|p|(1 + |\phi(t, x, p)|) + L(1 + |\phi(t, x, p)|)^2 \]

\[ \leq C(L)(1 + |p|^2) + L|p|(1 + L(1 + |p|)) + L(1 + L^2(1 + |p|^2)) \]

which implies (21). Here, the notation \( C(L) \) stands for a constant depending only on \( L \). Moreover, in view of the definition of \( M \) in (19) and Assumption (B5), we have \( M < +\infty \) depending only on \( L \).

Now we check (d). For any arbitrary \((t', x'), (t, x) \in [0, T] \times \mathbb{R}^n\) and \(|p'|, |p| \leq M_1\), where \( M_1 \geq 1 \) is a constant, from Assumption (B4), we know that

\[ |a_i(t', x', p) - a_i(t, x, p)| = \left| \sum_{j=1}^{n} a_{ij}(t', x', \mu(t', \cdot))p_j - \sum_{j=1}^{n} a_{ij}(t, x, \mu(t, \cdot))p_j \right| \]

\[ \leq M_1 \sum_{j=1}^{n} |a_{ij}(t', x', \mu(t', \cdot)) - a_{ij}(t, x, \mu(t))| \]

\[ \leq M_1 L n(|t' - t| + |x' - x| + d_1(\mu(t'), \mu(t))). \]
Since \( \mu \in D_{C_1} \), we know that \( d_1(\mu(t'), \mu(t)) \leq C_1 |t' - t|^{\frac{3}{4}} \) and

\[
|a_i(t', x', p') - a_i(t, x, p)| \leq M_1 Ln(C_1 + 1)(|t' - t|^{\frac{3}{4}} + |x' - x|). \quad (23)
\]

From Assumption (B3), we know that

\[
|a_i(t', x', p') - a_i(t', x', p)| = \sum_{j=1}^{n} a_{ij}(t', x', \mu(t'))(p_j' - p_j) \leq nL|p'| - p|.
\]

We deduce from (23), (24) and Assumption (B3) that

\[
\frac{|a_i(t', x', p') - a_i(t, x, p)|}{|t' - t|^{\frac{3}{4}} + |x' - x|^{\frac{3}{4}} + |p' - p|^{\frac{3}{4}}} \leq (2nL1)^{\frac{1}{2}} \left( \frac{|a_i(t', x', p') - a_i(t, x, p)|}{|t' - t|^{\frac{3}{4}} + |x' - x| + |p' - p|} \right)^{\frac{1}{2}}
\]

\[
\leq (2n^2L^2M^1(1 + C_1))^\frac{1}{2}.
\]

So we know that \( a_i(t, x, p) \) satisfies a Hölder condition in \((t, x, p)\) with exponents \((\frac{1}{4}, \frac{1}{2}, \frac{1}{2})\) respectively. Similarly, we can show that \((a_{ij}, \frac{\partial a_{ij}}{\partial x_i}(t, x, p))\) satisfy a Hölder condition in \((t, x, p)\) with exponents \((\frac{1}{4}, \frac{1}{2}, \frac{1}{2})\) respectively. Now we check \( a(t, x, p) \). Recall that

\[
a(t, x, p) = \sum_{i,j=1}^{n} \frac{\partial a_{ij}}{\partial x_i}(t, x, \mu(t, \cdot))(p_j - p(b(t, x, \mu(t, \cdot), \phi(t, x, p)))) - f(t, x, \mu(t, \cdot), \phi(t, x, p)).
\]

For the function

\[
k(t, x, p) := (p, b(t, x, \mu(t, \cdot), \phi(t, x, p))),
\]

from Assumptions (B3), (B4), (B7) and the fact that \( \mu \in D_{C_1} \), we have

\[
|k(t', x', p') - k(t, x, p)| \leq |b(t, x, \mu(t', \cdot), \phi(t, x, p))| |p' - p| + M_1 |b(t, x', \mu(t', \cdot), \phi(t', x', p')) - b(t, x, \mu(t, \cdot), \phi(t, x, p))| \leq L(1 + |\phi(t, x, p)|) |p' - p| + M_1 L(|t' - t|^{\frac{3}{4}} + |x' - x| + d_1(\mu(t'), \mu(t)) + |\phi(t', x', p') - \phi(t, x, p)|) \leq L(1 + L(1 + M_1)) |p' - p| + M_1 L((1 + C_1 + L)|t' - t|^{\frac{3}{4}} + (1 + L)|x' - x| + L |p' - p|) \leq C(L, M_1, C_1)(|t' - t|^{\frac{3}{4}} + |x' - x| + |p' - p|).
\]

Here, the notation \( C(L, M_1, C_1) \) stands for a constant depending only on \((L, M_1, C_1)\). So, we have the following estimate

\[
\frac{|k(t', x', p') - k(t, x, p)|}{|t' - t|^{\frac{3}{4}} + |x' - x|^{\frac{3}{4}} + |p' - p|^{\frac{3}{4}}} \leq (\frac{|k(t', x', p')| + |k(t, x, p)|}{|t' - t|^{\frac{3}{4}} + |x' - x|^{\frac{3}{4}} + |p' - p|^{\frac{3}{4}}})^{\frac{1}{2}} \leq (2M_1 L(1 + L(1 + M_1)))^{\frac{1}{2}} \leq C(L, M_1, C_1).
\]

12
So, $k(t,x,p)$ is Hölder continuous in $(t,x,p)$ with exponents $(\frac{1}{4}, \frac{1}{2}, \frac{1}{2})$, respectively. Again from Assumptions (B3), (B4), (B7), we have

$$
|f(t',x',\mu(t',\cdot),\phi(t',x',p')) - f(t,x,\mu(t,\cdot),\phi(t,x,p))| \\
\leq L(|t' - t|^\frac{1}{2} + |x' - x| + d_1(\mu(t'),\mu(t)) + (1 + |\phi(t,x,p)| + |\phi(t',x',p')|)|\phi(t',x',p') - \phi(t,x,p))| \\
\leq L(1 + C_1 + (1 + 2L(1 + M_1))L(|t' - t|^\frac{1}{2} + |x' - x| + |p' - p|).
$$

So, we have

$$
|f(t',x',\mu(t',\cdot),\phi(t',x',p')) - f(t,x,\mu(t,\cdot),\phi(t,x,p))| \\
\leq \left( |f(t,x,\mu,\phi)(t',x',p')| + |f(t,x,\mu,\phi)(t,x,p)| \right)^{\frac{1}{2}} \frac{|f(t,x,\mu,\phi)(t',x',p') - f(t,x,\mu,\phi)(t,x,p)|^{\frac{1}{2}}}{|t' - t|^\frac{1}{2} + |x' - x| + |p' - p|^{\frac{1}{2}}} \\
\leq (2L(1 + (L(1 + M_1))^2)^{\frac{1}{2}}(f(t,x,\mu,\phi)(t',x',p') - f(t,x,\mu,\phi)(t,x,p)|^{\frac{1}{2}}) \\
\leq C(L, M_1, C_1).
$$

Thus $f(t,x,\mu(t,\cdot),\phi(t,x,p))$ is Hölder continuous in $(t,x,p)$ with exponents $(\frac{1}{4}, \frac{1}{2}, \frac{1}{2})$, respectively. Up to now, we know that $a(t,x,p)$ satisfies a Hölder condition in $(t,x,p)$ with exponents $(\frac{1}{4}, \frac{1}{2}, \frac{1}{2})$, respectively.

Up to now, we have already checked conditions (a)-(d) in Theorem 3.3. So, PDE (22) has a solution $v \in H^{1+\frac{1}{2},2+\frac{1}{2}}([0,T] \times \mathbb{R}^n)$ such that $v(t,x) \leq M$ for all $(t,x) \in [0,T] \times \mathbb{R}^n$, where $M$ depends only on $L$.

Now we check the uniqueness. First note that

$$
\frac{\partial a_{ij}(t,x)}{\partial v} = \frac{\partial a_{ij}(t,x)}{\partial p} = \frac{\partial A(t,x,p)}{\partial v} = 0.
$$

And for $(t,x) \in [0,T] \times \mathbb{R}^n$ and $|p| \leq N$, from Assumptions (B3) and (B7), we have

$$
|\frac{\partial A(t,x,p)}{\partial p}| \\
\leq |b(t,x,\mu(t,\cdot),\phi(t,x,p))| + N\left|\frac{\partial b(t,x,\mu(t,\cdot),\phi(t,x,p))}{\partial \alpha}\right| \left|\frac{\partial \phi}{\partial p}(t,x,p)\right| \\
+ \left|\frac{\partial f}{\partial \alpha}(t,x,\mu(t,\cdot),\phi(t,x,p))\right| \left|\frac{\partial \phi}{\partial p}(t,x,p)\right| \\
\leq C(n,L)(1 + N).
$$

Here, the notation $C(n,L)$ stands for a constant depending only on $n$ and $L$. Then from Theorem 3.3 we know that equation (22) has no more than one classical solution $v(\cdot,\cdot)$ that is bounded in $[0,T] \times \mathbb{R}^n$ and has bounded derivatives of first and second orders.

In the next subsection, we will prove the existence of the solution of equation (14), which is based on the estimate of the gradient $Du$ in equation (13). Now we give some estimates of the solution $u$ of (13). The following theorem is borrowed from [20] Theorem 3.1, pp.437].
**Theorem 3.4.** Suppose that \( u \in O^{1,2}(Q_T) \) is a solution of equation

\[
u_t(t, x) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} a_i(t, x, u, u_x)(t, x) + a(t, x, u, u_x)(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n,\]

with \( \max_{Q_T} |u| \leq M. \) Suppose that for \( 1 \leq i \leq n \) and \( (t, x, u, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \), the functions \( a_i(t, x, u) \) and \( a(t, x, u, p) \) satisfy conditions

\[
\begin{align*}
\gamma |\xi|^2 & \leq \sum_{i,j=1}^{n} \frac{\partial a_i(t, x, u, p)}{\partial p_j} \xi_i \xi_j \leq \mu |\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \\
|a_i(t, x, u, p)| + \left| \frac{\partial a_i(t, x, u, p)}{\partial u} \right| & \leq \mu |p| + \varphi_1(t, x), \\
\left| \frac{\partial a_i(t, x, u, p)}{\partial x_j} \right| & \leq \mu |p|^2 + \varphi_2(t, x), \\
|a(t, x, u, p)| & \leq \mu |p|^2 + \varphi_3(t, x),
\end{align*}
\]

in which

\[
\| (\varphi_1, \varphi_2, \varphi_3) \|_{2r, 2q, Q_T} \leq \mu_1,
\]

for some constants \( \gamma, \mu, \mu_1 \in (0, +\infty) \), and with

\[
\begin{cases}
\frac{1}{r} + \frac{n}{2q} = 1 - h, \\
(q, r, h) \in \left[ \frac{n}{2(1-h)}, \infty \right] \times \left[ \frac{1}{1-h}, \infty \right] \times (0, 1), & \text{for } n \geq 2, \\
(q, r, h) \in \left[ 1, \infty \right] \times \left[ \frac{1}{1-h}, \frac{2}{1-2h} \right] \times \left( 0, \frac{1}{2} \right), & \text{for } n = 1.
\end{cases}
\]

Furthermore, suppose that \( |u_x(0, \cdot)|_\Omega^{(\beta)} < \infty \) for some \( \beta \in (0, 1) \). Then, for any domain \( Q' \subset Q_T \) separated from the boundary \( S_T \) by a positive distance \( d \), the quantity \( |u_x(\cdot)|_{Q'}^{(\alpha)} \) with some \( \alpha > 0 \) and the norms \( \| u_{xx}, u_t \|_{2, Q'} \) are bounded by a constant depending only on \( (n, M, \gamma, \mu, \mu_1, q, r, d, \beta, |u_x(0, \cdot)|_\Omega^{(\beta)}) \).

The number \( \alpha > 0 \) depends only on \( (n, \gamma, \mu, q, r, \beta) \).

Now we estimate the gradient \( Dv(t, x) \). We have already known that \( v \in H^{1+\frac{1}{q}, 2+\frac{1}{q}}([0, T] \times \mathbb{R}^n) \subset O^{1,2}([0, T] \times \mathbb{R}^n) \) that does not exceed \( M \), where \( M \) is a constant depending only on \( L \). We choose

\[
\Omega_1 \subset \subset \Omega_2 \subset \subset \ldots \subset \subset \Omega_k \subset \subset \ldots, \quad \bigcup_{k=1}^{\infty} \Omega_k = \mathbb{R}^n,
\]

and set

\[
Q_T^k := [0, T] \times \Omega_k, \quad S_T^k := [0, T] \times \partial \Omega_k.
\]

We choose \( d = 1 \) which is the same for all \( k \geq 1 \), and choose \( q = +\infty \) and \( r = \frac{4}{3} \). Now we apply Theorem 3.4 for \( Q_T^k \). Note that

\[
v|_{Q_T^k} \in H^{1+\frac{1}{q}, 2+\frac{1}{q}}(Q_T^k) \subset O^{1,2}(Q_T^k), \quad \max_{Q_T^k} |v| \leq \max_{[0, T] \times \mathbb{R}^n} |v| \leq M.
\]
Further, (25) is true in view of Assumption (B2), and (30) is true for our choice \((q, r, h) = (+\infty, 4, 3, 2)\) for \(n \geq 1\). From assumptions (B3) and (B7), we have
\[
\begin{align*}
|a_i(t, x, v, p)| &+ |\frac{\partial a_i(t, x, v, p)}{\partial v}| \leq C(n, L)|p|, \\
|\frac{\partial a_i(t, x, v, p)}{\partial x_j}| &\leq C(n, L)|p| \leq C(n, L)p^2 + C(n, L), \\
|a(t, x, v, p)| &\leq C(n, L)|p| + |p|L(1 + |\phi(t, x, p)|) + L(1 + |\phi(t, x, p)|^2) \\
&\leq C(n, L)|p| + |p|L(1 + L(1 + |p|)) + L(1 + L^2(1 + |p|)^2) \\
&\leq C(n, L)p^2 + C(n, L).
\end{align*}
\]

Here, the notation \(C(n, L)\) stands for a constant depending only on \(n\) and \(L\). To prove the four inequalities (26)-(29), we can choose \(\varphi_1(t, x) = 0\) and constants \(\varphi_2(t, x)\) and \(\varphi_3(t, x)\) depending only on \((n, L)\), which satisfies
\[
\|\varphi_1, \varphi_2, \varphi_3\|_{8, +\infty, Q^k_T} \leq C(n, L)T^3 = C(n, L, T).
\]

Moreover, since \(v_x(0, x) = g_x(x, \mu(T, \cdot))\) for \(x \in \mathbb{R}^n\), we know from Assumption (B5) that \(|v_x(0, \cdot)|_{\mathbb{R}^n}^{(\beta)} \leq L\). So from Theorem 3.3 we know that, for each \(k \geq 1\) and \(\tilde{Q}^k \subset Q^k_T\) separated from the boundary \(S^k_T\) by 1, the quantity \(|v_x|^{(\alpha)}_{\tilde{Q}^k}\) with some \(\alpha > 0\) is bounded by a constant depending only on \((n, \gamma_1, \gamma_2, L, T, \beta)\). The number \(\alpha > 0\) depends only on \((n, \gamma_1, \gamma_2, L, T, \beta)\). Since \(\bigcup_{k=1}^{\infty} \tilde{Q}^k = [0, T] \times \mathbb{R}^n\), and the constants above are independent of \(k\), so we have that the quantity \(|u_x|^{(\alpha)}_{[0, T] \times \mathbb{R}^n}\) with some \(\alpha > 0\) is bounded by a constant depending only on \((n, \gamma_1, \gamma_2, L, T, \beta)\). The number \(\alpha > 0\) depends only on \((n, \gamma_1, \gamma_2, L, \beta)\).

As a summary of the subsection, we have the following:

**Theorem 3.5.** Let Assumptions (B1)-(B7) be satisfied. For \(\mu \in \mathcal{D}_{C_1}\), equation (13) has a unique solution \(u \in H^{1+\frac{1}{4}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n)\) that does not exceed \(M\), where \(M \in (0, +\infty)\) is a constant depending only on \(L\). Moreover, the quantity \(|u_x|^{(\alpha)}_{[0, T] \times \mathbb{R}^n}\) with some \(\alpha > 0\) is bounded by a constant depending only on \((n, \gamma_1, \gamma_2, L, T, \beta)\). The number \(\alpha > 0\) depends only on \((n, \gamma_1, \gamma_2, L, \beta)\).
3.2 Existence and uniqueness of the solution of equation \(14\)

In this subsection, we show the existence and uniqueness of the solution of equation \(14\), which also reads

\[
\begin{align*}
\frac{\partial m}{\partial t}(t, x) & - \sum_{i,j=1}^{n} a_{ij}(t, x, \mu(t, \cdot)) \frac{\partial^2 m}{\partial x_i \partial x_j}(t, x) \\
+ \sum_{i=1}^{n} \left[ b_i(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) - \sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j}(t, x, \mu(t, \cdot)) \right] \frac{\partial m}{\partial x_i}(t, x) \\
+ \sum_{i=1}^{n} \left[ - \sum_{j=1}^{n} \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}(t, x, \mu(t, \cdot)) + \frac{\partial b_i}{\partial x_i}(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) \\
+ \frac{\partial b_i}{\partial \alpha}(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) (\frac{\partial \phi}{\partial x_i}(t, x, Du(t, x))) \\
+ \frac{\partial \phi}{\partial p}(t, x, Du(t, x)) \frac{\partial Du}{\partial x_i}(t, x) \right] m(t, x) = 0,
\end{align*}
\]

\(m(0, x) = m_0(x), \quad x \in \mathbb{R}^n.\)

The following theorem is available in [20, Theorem 5.1, pp.320].

**Theorem 3.6.** Consider the equation

\[
\begin{align*}
\frac{\partial u(t, x)}{\partial t} - \mathcal{L}(t, x)u(t, x) &= f(t, x), \quad (t, x) \in (0, T] \times \mathbb{R}^n; \\
u(0, x) &= \varphi(x), \quad x \in \mathbb{R}^n,
\end{align*}
\]

where the differential operator \(\mathcal{L}(t, x)\) is defined as follows:

\[
\mathcal{L}(t, x)u = \sum_{i,j=1}^{n} a_{ij}(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^{n} a_i(t, x) \frac{\partial u}{\partial x_i} - a(t, x)u, \quad (t, x) \in [0, T] \times \mathbb{R}^n.
\]

Suppose that \(l > 0\) is a non-integral number and the coefficients of the operator \(\mathcal{L}\) belongs to the class \(H^{\frac{l}{2},l}([0, T] \times \mathbb{R}^n)\). Then, for any \(f \in H^{\frac{l}{2},l}([0, T] \times \mathbb{R}^n)\) and \(\varphi \in H^{2+l}(\mathbb{R}^n)\), equation \(33\) has a unique solution in the class \(H^{1+\frac{l}{2},2+l}([0, T] \times \mathbb{R}^n)\). It satisfies the inequality

\[
|u|_{[0,T] \times \mathbb{R}^n}^{(2+l)} \leq C(|f|_{[0,T] \times \mathbb{R}^n}^{(l)} + |\varphi|_{\mathbb{R}^n}^{(2+l)}),
\]

for a constant \(C\) not depending on \(f\) and \(\varphi\).

Now we turn to \(32\). In view of the preceding theorem, we need to check the Hölder condition of the coefficients of \(32\). First we note from Assumption (B6) that \(m_0 \in H^{2+\frac{l}{2}}(\mathbb{R}^n)\). From Assumptions (B3), (B4) and the fact that \(\mu \in \mathcal{D}_{C_1}\), we know that for \(1 \leq i, j \leq n\),

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

From Theorem \(3.5\) we know that

\[
Du \in H^{\frac{l}{2},\frac{l}{2}}([0, T] \times \mathbb{R}^n), \quad D^2u \in H^{\frac{l}{2},\frac{l}{2}}([0, T] \times \mathbb{R}^n).
\]

\(16\)
So from (34), Assumptions (B3), (B4), (B7) and the fact that $\mu \in \mathcal{D}_{C_1}$, we know that for $1 \leq i \leq n$,

\[
(b_i, \frac{\partial b_i}{\partial x_i}, \frac{\partial b_i}{\partial \alpha})(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) \in H^{\frac{1}{4}, \frac{1}{2}}([0, T] \times \mathbb{R}^n);
\]

\[
\left(\frac{\partial \phi}{\partial x_i}, \frac{\partial \phi}{\partial p}, \frac{\partial \phi}{\partial t}, \phi(t, x, Du(t, x))\right) \in H^{\frac{1}{4}, \frac{1}{2}}([0, T] \times \mathbb{R}^n).
\]

It remains to show that

\[
h_i(t, x) := \frac{\partial b_i}{\partial \alpha}(t, x, \mu(t, \cdot), \phi(t, x, Du(t, x))) \left(\frac{\partial \phi}{\partial x_i}(t, x, Du(t, x)) + \frac{\partial \phi}{\partial p}(t, x, Du(t, x)) \frac{\partial Du}{\partial x_i}(t, x)\right)
\]

\[
\in H^{\frac{1}{4}, \frac{1}{2}}([0, T] \times \mathbb{R}^n), \quad 1 \leq i \leq n.
\]

Actually,

\[
|h_i(t', x') - h_i(t, x)|
\]

\[
\leq \left| \frac{\partial b_i}{\partial \alpha}(t, x, \mu, \phi)(t', x')) \right| \left| \frac{\partial \phi}{\partial x_i}(t', x', Du(t', x')) - \frac{\partial \phi}{\partial x_i}(t, x, Du(t, x)) \right| + |Du(t, x)| \left| \frac{\partial \phi}{\partial p}(t, x, Du(t, x)) \right| \left| \frac{\partial Du}{\partial x_i}(t, x) \right|
\]

\[
+ \left| \frac{\partial b_i}{\partial \alpha}(t, x, \mu, \phi)(t', x') \right| \left| \frac{\partial \phi}{\partial x_i}(t', x', Du(t', x')) - \frac{\partial \phi}{\partial x_i}(t, x, Du(t, x)) \right| \left| \frac{\partial Du}{\partial x_i}(t, x) \right|
\]

\[
+ \left| \frac{\partial b_i}{\partial \alpha}(t, x, \mu, \phi)(t', x') - \frac{\partial b_i}{\partial \alpha}(t, x, \mu, \phi)(t, x) \right| \left| \frac{\partial \phi}{\partial x_i}(t, x, Du(t, x)) \right| \left| \frac{\partial Du}{\partial x_i}(t, x) \right|
\]

\[
+ |Du(t, x)| \left| \frac{\partial \phi}{\partial p}(t, x, Du(t, x)) \right| \left| \frac{\partial Du}{\partial x_i}(t, x) \right|.
\]

From Assumptions (B3), (B4), (B7) and the fact that $\mu \in \mathcal{D}_{C_1}$, we have

\[
|h_i(t', x') - h_i(t, x)| \leq \mathcal{C}(n, L, C_1)(1 + |Du(t, x)| + \left| \frac{\partial Du}{\partial x_i}(t, x) \right|)
\]

\[
\cdot \left( |t' - t| + |x' - x| + |Du(t', x') - Du(t, x)| + \left| \frac{\partial Du}{\partial x_i}(t', x') - \frac{\partial Du}{\partial x_i}(t, x) \right| \right).
\]

Here, the notation $\mathcal{C}(n, L, C_1)$ stands for a constant depending only on $(n, L, C_1)$. Then from (34), we know that $h_i \in H^{\frac{1}{4}, \frac{1}{2}}([0, T] \times \mathbb{R}^n)$ for $1 \leq i \leq n.$

Up to now, we have shown that all the coefficients in (32) belongs to the class $H^{\frac{1}{4}, \frac{1}{2}}([0, T] \times \mathbb{R}^n)$.

Then from Theorem 3.6, there is a unique solution in the class $H^{1+\frac{1}{4}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n)$ of equation (32), equivalently, of equation (13).

**Theorem 3.7.** Let Assumptions (B1)-(B7) be satisfied. For $\mu \in \mathcal{D}_{C_1}$ and $u \in H^{1+\frac{1}{4}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n)$ as the unique solution of equation (13), equation (14) has a unique solution $m \in H^{1+\frac{1}{4}, 2+\frac{1}{2}}([0, T] \times \mathbb{R}^n)$.

### 3.3 The solution of (14) belongs to $\mathcal{D}_{C_1}$

In this subsection, we show that the solution $m$ to equation (14) belongs to the set $\mathcal{D}_{C_1}$. Consider the following SDE:

\[
\begin{cases}
    dX_t = b(t, X_t, \mu(t, \cdot), \phi(t, X_t, Du(t, X_t)))dt + \sigma(t, X_t, \mu(t, \cdot))dW_t, \quad t \in (0, T]; \\
    X_0 = \xi_0 \sim m_0.
\end{cases}
\]
From Assumptions (B4), (B6), (B7), the fact that \( Du \in H^{1+\frac{1}{2},1+\frac{1}{2}}([0,T] \times \mathbb{R}^n) \) and standard arguments of SDE, we know that \([55]\) has a unique solution \( X = \{X_t, 0 \leq t \leq T\} \in S^2_{\mathbb{F}}(0,T) \). We set \( \tilde{m}(t) = \mathcal{L}(X_t) \). We give the definition of a weak solution of \([14]\).

**Definition 3.8.** \([6\] Definition 3.2, pp.11\]. We say that \( m \) is a weak solution of \([14]\), if \( m \in L^1([0,T], \mathcal{P}_1(\mathbb{R}^n)) \) such that for any test function \( \varphi \in C^\infty_c([0,T] \times \mathbb{R}^n) \), we have

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

The following proposition shows that, under our assumptions, there is a unique weak solution of \([14]\), and it belongs to \( H^{1+\frac{1}{2},2+\frac{1}{2}}([0,T] \times \mathbb{R}^n) \). The proposition is available in \([6\] pp.12\].

**Proposition 3.9.** If the coefficients of equation \([14]\) belongs to the class \( H^{l,\frac{1}{2}}([0,T] \times \mathbb{R}^n) \) for some \( l \in (0,1) \), and \( m_0 \in H^l(\mathbb{R}^n) \), then, equation \([14]\) has a unique weak solution.

In view of Proposition 3.9, we know that \( m \) is the unique weak solution of equation \([14]\). The next lemma shows that \( \tilde{m} \) is also a weak solution of equation \([14]\), which shows that \( m(t) = \tilde{m}(t) = \mathcal{L}(X_t), \quad t \in [0,T] \).

**Lemma 3.10.** Let Assumptions (B1)-(B7) be satisfied. Then, \( \tilde{m} \) is a weak solution of equation \([14]\).

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

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

\[
+ \int_0^t \langle D\varphi(s,X_s), \sigma(s,X_s,\mu(s,\cdot))dW_s \rangle, \quad t \in [0,T].
\]

Taking the expectation on both sides and noting that \( \varphi \in C^\infty_c([0,T] \times \mathbb{R}^n) \), we have

\[
\mathbb{E}[\varphi(0,\xi_0)] + \mathbb{E}\left[ \int_0^T (\partial_t \varphi(s,X_s) + \langle D\varphi(s,X_s), b(s,X_s,\mu(s,\cdot),\phi(s,X_s,Du(s,X_s))) \rangle + \frac{1}{2} \text{Tr} [\sigma \sigma^T(s,X_s,\mu(s,\cdot))D^2 \varphi(s,X_s)] )ds \right] = 0.
\]

By definition of \( \tilde{m} \), we know that \( \tilde{m} \) is a weak solution of equation \([14]\). \qed

Now we can prove that the solution of \([14]\) belongs to \( \mathcal{D}_{C_1} \).

**Theorem 3.11.** Let Assumptions (B1)-(B7) be satisfied. There is \( C_1 \in (0,+\infty) \) depending only on \( (n,M,\gamma_1,\gamma_2,L,\beta,T,m_0) \), such that the solution \( m \) of \([14]\) belongs to \( \mathcal{D}_{C_1} \).
Proof. From the definition of $d_1$, we have for $0 \leq s < t \leq T$,

$$d_1(m(s), m(t)) \leq \mathbb{E}|X_s - X_t|$$

$$= \mathbb{E}\left| \int_s^t b(\tau, X_\tau, \mu(\tau, \cdot), \phi(\tau, X_\tau, Du(\tau, X_\tau)))d\tau + \int_s^t \sigma(\tau, X_\tau, \mu(\tau, \cdot))dW_\tau \right|$$

$$\leq |t - s|L(1 + L(1 + \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} |Du(t, x)|)) + L|t - s|^{\frac{1}{2}}.$$ 

From Theorem 3.5 we have

$$\sup_{(t, x) \in [0, T] \times \mathbb{R}^n} |Du(t, x)| \leq C(n, \gamma_1, \gamma_2, L, T, \beta),$$

and therefore,

$$\sup_{s \neq t} \frac{d_1(m(s), m(t))}{|t - s|^{\frac{1}{2}}} \leq C(n, \gamma_1, \gamma_2, L, T, \beta).$$

Here, the notation $C(n, \gamma_1, \gamma_2, L, T, \beta)$ stands for a constant depending only on $(n, \gamma_1, \gamma_2, L, T, \beta)$. We also have

$$\int_{\mathbb{R}^n} |x|^2dm(t)(x) = \mathbb{E}|X_t|^2 \leq 2\mathbb{E}|\xi_0|^2 + T^2(1 + L(1 + \sup_{(t, x) \in [0, T] \times \mathbb{R}^n} |Du(t, x)|)^2 + L^2T^2)$$

$$\leq C(n, \gamma_1, \gamma_2, L, T, \beta, m_0).$$

Therefore,

$$\sup_{t \in [0, T]} \int_{\mathbb{R}^n} |x|^2dm(t)(x) \leq C(n, \gamma_1, \gamma_2, L, T, \beta, m_0).$$

In summary, we have $m \in D_{C_1}$ with $C_1 = C(n, \gamma_1, \gamma_2, L, T, \beta, m_0)$. The proof is complete. \qed

### 3.4 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. In this subsection, 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 $\{\mu_n, n \geq 1\} \subset D_{C_1}$ converge to some $\mu \in D_{C_1}$, with respect to the norm $\rho$. Let $\{(u_n, m_n), n \geq 1\}$ and $(u, m)$ be the corresponding solutions. From Assumptions (B3), (B4) and (B7), we know that the coefficients in $[13]$ corresponding to $\mu_n$ uniformly converge to the coefficients in $[13]$ corresponding to $\mu$. Then we get the local uniform convergence of $\{u_n, n \geq 1\}$ to $u$ by standard arguments. From Theorem 3.5 we know that the gradients $\{Du_n, n \geq 1\}$ are uniformly bounded and uniformly Hölder continuous and therefore locally uniformly converges to $Du$.

For any converging subsequence $\{m_{n_k}, k \geq 1\}$ of the relatively compact sequence $\{m_n, n \geq 1\}$ (since $D_{C_1}$ is compact), we assume that $\{m_{n_k}, k \geq 1\}$ converge to some $\tilde{m} \in D_{C_1}$. For any $\varphi \in C^\infty_c([0, T] \times \mathbb{R}^n)$, since $m_{n_k}$ is a weak solution of $[12]$ corresponding to $(\mu_{n_k}, u_{n_k})$, we have

$$\int_{x \in \mathbb{R}^n} \phi(0, x)m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu_{n_k}(t, \cdot))\partial x_i \partial x_j \varphi(t, x) \right. \left. + \sum_{i=1}^n b_i(t, x, \mu_{n_k}(t, \cdot), \phi(t, x, Du_{n_k}(t, x)))\partial x_i \varphi(t, x) \right] m_{n_k}(t)(dx)dt = 0. \tag{36}$$
Therefore, we have

\[
\left| \left( \sum_{i,j=1}^{n} a_{ij}(t,x,\mu_{nk}(t,\cdot))\partial_{x_i,x_j} \varphi(t,x) + \sum_{i=1}^{n} b_{i}(t,x,\mu_{nk}(t,\cdot),\phi(t,x,Du_{nk}(t,x)))\partial_{x_i} \varphi(t,x) \right) - \left( \sum_{i,j=1}^{n} a_{ij}(t,x,\mu(t,\cdot))\partial_{x_i,x_j} \varphi(t,x) + \sum_{i=1}^{n} b_{i}(t,x,\mu(t,\cdot),\phi(t,x,Du(t,x)))\partial_{x_i} \varphi(t,x) \right) \right|
\]

\[
\leq nL \left( n|\partial_{x_i,x_j} \varphi(t,x)| + |\partial_{x_i} \varphi(t,x)| \right) \rho(\mu_{nk},\mu) + nL^2 |\partial_{x_i} \varphi(t,x)| \sup_{(t,x) \in \text{supp}(\varphi)} |Du_{nk}(t,x) - Du(t,x)|.
\]

From Kantorovich-Rubininstein Theorem [6, Theorem 5.5, pp. 36] and the fact that \( \{m_{nk}, k \geq 1 \} \) converge to \( \hat{m} \) in \( C^0([0,T], \mathcal{P}_1(\mathbb{R}^n)) \), we have

\[
\lim_{k \to +\infty} \int_{0}^{T} \int_{\mathbb{R}^n} |\partial_{x_i,x_j} \varphi(t,x)|m_{nk}(t)(dx)dt = \int_{0}^{T} \int_{\mathbb{R}^n} |\partial_{x_i,x_j} \varphi(t,x)|\hat{m}(t)(dx)dt < +\infty;
\]

\[
\lim_{k \to +\infty} \int_{0}^{T} \int_{\mathbb{R}^n} |\partial_{x_i} \varphi(t,x)|m_{nk}(t)(dx)dt = \int_{0}^{T} \int_{\mathbb{R}^n} |\partial_{x_i} \varphi(t,x)|\hat{m}(dx)dt < \infty.
\]

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

\[
\lim_{k \to +\infty} \rho(\mu_{nk},\mu) = 0;
\]

\[
\lim_{k \to +\infty} \sup_{(t,x) \in \text{supp}(\varphi)} |Du_{nk}(t,x) - Du(t,x)| = 0.
\]
Plugging (38) and (39) into (37), we have
\[
\lim_{k \to +\infty} \left| \int_0^T \int_{\mathbb{R}^n} \left( \sum_{i,j=1}^n a_{ij}(t, x, \mu_{n_k}(t, \cdot)) \partial_{x_i x_j} \varphi(t, x) \\
+ \sum_{i=1}^n b_i(t, x, \mu_{n_k}(t, \cdot), \phi(t, x, D\mu_{n_k}(t, x))) \partial_{x_i} \varphi(t, x) \right) \\
- \left[ \sum_{i,j=1}^n a_{ij}(t, x, \mu(\cdot)) \partial_{x_i x_j} \varphi(t, x) \\
+ \sum_{i=1}^n b_i(t, x, \mu(\cdot), \phi(t, x, D\mu(t, x))) \partial_{x_i} \varphi(t, x) \right] m_{n_k}(t)(dx)dt \right| = 0.
\] (40)

Plugging (40) into (36), we have
\[
\lim_{k \to +\infty} \int_{x \in \mathbb{R}^n} \phi(0, x)m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(\cdot)) \partial_{x_i x_j} \varphi(t, x) \\
+ \sum_{i=1}^n b_i(t, x, \mu(\cdot), \phi(t, x, D\mu(t, x))) \partial_{x_i} \varphi(t, x) \right] m_{n_k}(t)(dx)dt = 0.
\] (41)

Again from Kantorovich-Rubinstein Theorem, we have
\[
\lim_{k \to +\infty} \int_0^T \int_{\mathbb{R}^n} \left[ \sum_{i,j=1}^n a_{ij}(t, x, \mu(\cdot)) \partial_{x_i x_j} \varphi(t, x) \\
+ \sum_{i=1}^n b_i(t, x, \mu(\cdot), \phi(t, x, D\mu(t, x))) \partial_{x_i} \varphi(t, x) \right] m_{n_k}(t)(dx)dt = 0.
\] (42)

Plugging (42) into (41), we have eventually
\[
\int_{x \in \mathbb{R}^n} \phi(0, x)m_0(dx) + \int_0^T \int_{\mathbb{R}^n} \left[ \partial_t \varphi(t, x) + \sum_{i,j=1}^n a_{ij}(t, x, \mu(t)) \partial_{x_i x_j} \varphi(t, x) \\
+ \sum_{i=1}^n b_i(t, x, \mu(t), \phi(t, x, D\mu(t, x))) \partial_{x_i} \varphi(t, x) \right] \hat{m}(t)(dx)dt = 0,
\]
which means that \(\hat{m}\) is a weak solution to (14) corresponding to \(\mu\). Then, from Proposition 3.9 we know that \(m = \hat{m}\). Up to now we can see that, any converging subsequence \(\{m_{n_k}, k \geq 1\}\) of the relatively compact sequence \(\{m_n, n \geq 1\}\) converge to \(m\). So we know that \(\{m_n, n \geq 1\}\) converge to \(m\). Thus, \(\Phi\) is continuous.

We conclude by Schauder fixed point theorem that the continuous map \(\Phi\) has a fixed point in \(D_{C^1}\). This fixed point is a classical solution of MFEs (9). The proof of Theorem 3.1 is complete.
4 Application to mean field games

In this section, we investigate the optimal strategy of mean field controlled stochastic system:

\[
\begin{aligned}
dX_t &= b(t, X_t, \mathcal{L}(X_t^\alpha), \alpha_t)dt + \sigma(t, X_t, \mathcal{L}(X_t^\alpha))dW_t, \quad t \in (0, T]; \\
X_0 &= \xi_0 \sim m_0, \\
\end{aligned}
\]  

(43)

with the cost functional

\[
J(\alpha)(\mathcal{L}(X_t^\alpha))_{0 \leq t \leq T} = \mathbb{E}\left[\int_0^T f(t, X_t, \mathcal{L}(X_t^\alpha), \alpha_t)dt + g(X_T, \mathcal{L}(X_T^\alpha))\right],
\]  

(44)

where \( \hat{\alpha} \) is the optimal control.

We fix a solution \((u, m)\) of the MFEs (11). The following Theorem shows that the feedback strategy \( \tilde{\alpha}(t, x) := \phi(t, x, Du(t, x)) \) is optimal for the above mean field stochastic control problem.

**Theorem 4.1.** Let Assumptions (B1)-(B7) be satisfied and let \((u, m) \in H^{1+\frac{3}{2}+\frac{1}{2}}(0, T] \times \mathbb{R}^n) \times H^{1+\frac{1}{2}+\frac{1}{2}}(0, T] \times \mathbb{R}^n)\) be a classical solution of MFEs (11). Let \( \hat{X} = \{\hat{X}_t, 0 \leq t \leq T\} \) be the unique solution of the SDE

\[
\begin{aligned}
d\hat{X}_t &= b(t, \hat{X}_t, m(t, \cdot), \tilde{\alpha}(t, \hat{X}_t))dt + \sigma(t, \hat{X}_t, m(t, \cdot))dW_t, \quad t \in (0, T]; \\
\hat{X}_0 &= \xi_0 \sim m_0. \\
\end{aligned}
\]  

(45)

Then, \( m(t) = \mathcal{L}(\hat{X}_t) \) and

\[
\inf_{\alpha} J(\alpha|m) = J(\tilde{\alpha}|m) = \mathbb{E}[u(0, \xi_0)] = \int_{\mathbb{R}^n} u(0, x)m_0(dx),
\]

where \( \tilde{\alpha}(t) = \tilde{\alpha}(t, \hat{X}_t) \) for \( t \in [0, T] \).

**Proof.** From Lemma 3.10, we know that \( \mathcal{L}(\hat{X}_t) \) is a weak solution of the equation

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

(46)

From Proposition 3.9, we know that \( m \) is the unique solution to equation (10). Therefore, \( m(t) = \mathcal{L}(\hat{X}_t) \) for \( t \in [0, T] \).

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

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

(47)

We have from Itô’s formula that

\[
\mathbb{E}[u(T, X_T)] = \mathbb{E}\left[u(0, \xi_0) + \int_0^T \left(\frac{\partial u}{\partial t}(t, X_t) + \langle Du(t, X_t), b(t, X_t, m(t, \cdot), \alpha_t)\rangle \\
+ \sum_{i,j=1}^n a_{ij}(t, X_t, m(t, \cdot)) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X_t)\right)dt\right].
\]
From (49) and the definition of the Hamiltonian in (10), we have

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

Therefore,

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

From Assumptions (B1) and (B7), we know that

\[ H(t, X_t, m(t, \cdot), \alpha_t, Du(t, X_t)) \geq H(t, X_t, m(t, \cdot), \phi(t, X_t, Du(t, X_t)), Du(t, X_t)) \]
\[ = \langle Du(t, X_t), b(t, X_t, m(t, \cdot), \phi(t, X_t, Du(t, X_t))), Du(t, X_t) \rangle + f(t, X_t, m(t, \cdot), \phi(t, X_t, Du(t, X_t))). \]

Plugging (49) into (48), from (9), we have

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

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

**Remark 4.2.** Carmona and Delarue [8, Theorem 4.44, pp.263] consider the MFG problem 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. Carmona and Delarue [8, Theorem 4.29, pp.246] prove existence of a solution to a McKean-Vlasov FBSDE under nondegenerate assumptions [8, Assumption, pp.245]. They make less restrictive assumptions than ours. In our work, we make more differentiability assumptions on \( x \). That is because we use the analytical method to consider the classical solutions to the forward-backward PDE, rather than a probabilistic representation, so the differentiability of coefficients is needed. These differentiability assumptions are expected to be relaxed to more general cases by using the appropriate approximation technique. But this will involve detailed analysis and beyond the scope of this paper, so will not be discussed here.

Following the stochastic maximum principle for optimality, Carmona and Delarue [8] transform the MFG problem into solvability of a distribution dependent FBSDE. They prove the existence of the FBSDE within a linear-convex framework. Moreover, a weak mean-reverting condition is needed. In our work, we do not require any convexity assumption which is usually imposed in Pontryagin principle. We do not need the linear condition, and allow the volatility to depend upon the distribution of the state.
5 Example

Our model in Sections 3 and 4 with Assumptions (B1)-(B7) concerns the following mean field controlled stochastic system:

\[
\begin{align*}
  dX_t &= (B(X_t, L(X_t^\hat{\alpha})) + \alpha_t)dt + \sigma(X_t, L(X_t^\hat{\alpha}))dW_t, \quad t \in (0, T]; \\
  X_0 &= \xi_0 \sim m_0,
\end{align*}
\]

with the cost function

\[
J(\alpha)(L(X_t^\hat{\alpha}))_{0 \leq t \leq T} = \mathbb{E} \left[ \int_0^T \left( \frac{1}{2} |\alpha_t|^2 + F(X_t, L(X_t^\hat{\alpha})) \right) dt + G(X_T, L(X_T^\hat{\alpha})) \right],
\]

where \( \hat{\alpha} \) is the optimal control. Here,

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

Here are assumptions on the functions \((B, \sigma, F, G)\) and the distribution \(m_0\). For notational convenience, we use the same constant \(L\) for all the conditions below.

(C1) There exists \(0 < \gamma < +\infty\), such that

\[
\sigma \sigma^T(x, m) \geq \gamma I_n, \quad (x, m) \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n),
\]

where \(I_n\) is an identity matrix of order \(n\).

(C2) The functions \((B, \sigma, F)\) are uniformly bounded by \(L\). The function \(B(\cdot, m)\) is differentiable for all \(m \in \mathcal{P}_1(\mathbb{R}^n)\), with the derivative being bounded by \(L\). The function \(\sigma(\cdot, m)\) is twice differentiable for all \(m \in \mathcal{P}_1(\mathbb{R}^n)\), with the derivatives being bounded by \(L\).

(C3) The functions \(K(\cdot, \cdot) := (B, B_x, \sigma, \sigma_x, \sigma_{xx}, F, G)(\cdot, \cdot)\) are \(L\)-Lipschitz continuous. That is,

\[
|K(x', m') - K(x, m)| \leq L(|x' - x| + d_i(m', m)), \quad (x, m), (x', m') \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n).
\]

(C4) The function \(G(\cdot, m) \in H^{2+\frac{1}{2}}(\mathbb{R}^n)\) for all \(m \in \mathcal{P}_1(\mathbb{R}^n)\). There exists \(0 < \beta < 1\) such that

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

(C5) The distribution \(m_0\) is absolutely continuous with respect to the Lebesgue measure, has a Hölder continuous density (still denoted by \(m_0\)) such that \(m_0 \in H^{2+\frac{1}{2}}(\mathbb{R}^n)\) and \(\int_{\mathbb{R}^n} |x|^2 m_0(dx) < +\infty\). The preceding five assumptions (C1)-(C5) implies Assumptions (B1)-(B6). Moreover,

\[
H(t, x, m, \alpha, p) = \langle p, B(x, m) + \alpha \rangle + \frac{1}{2} |\alpha|^2 + F(x, m), \\
(t, x, m, \alpha, p) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n.
\]

Here we choose the control space \(A = \mathbb{R}^n\), then,

\[
\phi(t, x, p) = -p, \quad (t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\]

Assumption (B7) is satisfied. In view of Sections 3 and 4 the MFEs associated with this system has a classical solution \((u, m)\) and the optimal control is given by the feedback strategy \(\hat{\alpha}(t, x) = -Du(t, x)\).
Remark 5.1. The preceding example includes as a special case Lions’ results of mean field games [6, Theorem 3.1, pp.10; Lemma 3.7, pp.15] with $B = 0$ and $\sigma = \sqrt{2}$.

Remark 5.2. Bensoussan et al. [3] solve the forward-backward partial PDEs for linear quadratic mean field games under the condition that the volatility $\sigma$ is a constant. Our work cannot include Bensoussan’s results because of our bounded assumption. However, our work also go beyond their framework. We do not need the linear quadratic condition and allow the volatility $\sigma$ to depend on the state and the distribution of the state. Our bounded assumption is expected to be relaxed by using the appropriate approximation technique.

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 functionals and mean field games with common noise, Stoch. Proc. Appl., 129(10), 2019, 3859-3892.
[3] Bensoussan, A., Frehse, J. and Yam, P., The Master equation in mean field theory, Journal de Mathematiques Pures et Appliquees, 103(6), 2015, 1441-1474.
[4] 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.
[5] Buckdahn, R., Li, J., Peng, S. and Rainer, C., Mean-field stochastic differential equations and associated pdes, Ann. Probab., 45(2), 2017, 824-878.
[6] Cardaliaguet, P., Notes on mean field games, From P.-L. Lions’ lectures at College de France, 2010.
[7] Carmona, R. and Delarue, F., Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics, Ann. Probab., 43(5), 2015, 2647-2700.
[8] Carmona, R. and Delarue, F., Probabilistic analysis of mean-field games, SIAM J. Control Optim., 51(4), 2013, 2705-2734.
[9] Carmona, R. and Delarue, F., Probabilistic Theory of Mean Field Games: Vol. I, Mean Field FBSDEs, Control, and Games, Stochastic Analysis and Applications, Springer Verlag, 2017.
[10] Carmona, R., Delarue, F. and Lacker, D., Mean field games with common noise, Ann. Probab., 44(6), 2016, 3740-3803.
[11] Carmona, R. and Lacker, D., A probabilistic weak formulation of mean field games and applications, Ann. Probab., 25(3), 2015, 1189–1231.
[12] 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.
[13] Chassagneux, J. F., Crisan, D. and Delarue, F., A probabilistic approach to classical solutions of the master equation for large population equilibria, Mathematics, 2014.
[14] Du, K. and Wu, Z., Linear quadratic mean-field-game of backward stochastic differential systems, Math. Control Relat. F., 8(3&4), 2018, 653.
[15] Du, K. and Wu, Z., Linear-quadratic stackelberg game for mean-field backward stochastic differential system and application, Math. Probl. Eng., 2019.
[16] Hu, Y. and Peng, S., Maximum principle for semilinear stochastic evolution control systems, Stochastics, 33(3), 1990, 159–180.
[17] Huang, M., Caines, P. E. and Malhamé, R. P., Large-population cost-coupled lqg problems with nonuniform agents: Individual-mass behavior and decentralized $\epsilon$-nash equilibria, IEEE T. Automat. Contr., 52(9), 2017, 1560–1571.
[18] 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.
[19] Kolokoltsov, V. N., Li, J. and Yang, W., Mean field games and nonlinear Markov processes, *Mathematics*, 2011.

[20] 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.

[21] Lacker, D., A general characterization of the mean field limit for stochastic differential games, *Probab. Theory Rel.*, 165(3), 2016, 581-648.

[22] 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.

[23] 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.

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

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

[26] 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.

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

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

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

[30] 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.