UNIQUENESS AND SUPERPOSITION OF THE SPACE-DISTRIBUTION DEPENDENT ZAKAI EQUATIONS*

MEIQI LIU\(^1\) AND HUIJIE QIAO\(^{1,2}\)

\(^1\) Department of Mathematics, Southeast University, Nanjing, Jiangsu 211189, P.R.China
\(^2\) Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract. The work concerns the space-distribution dependent Zakai equations from nonlinear filtering problems of McKean-Vlasov stochastic differential equations with correlated noises. First of all, we establish the space-distribution dependent Kushner-Stratonovich equations and the space-distribution dependent Zakai equations. Then, the pathwise uniqueness of their strong solutions is shown. Finally, we prove a superposition principle between the space-distribution dependent Zakai equations and space-distribution dependent Fokker-Planck equations. As a by-product, we give some conditions under which space-distribution dependent Fokker-Planck equations have weak solutions.

1. Introduction

McKean-Vlasov (distribution-dependent or mean-field) stochastic differential equations (SDEs for short) describe the evolution rules of particle systems perturbed by noises. The difference between McKean-Vlasov SDEs and general SDEs is that the former depend on the positions and probability distributions of these particles. Therefore, McKean-Vlasov SDEs are widely applied in many fields, such as biology, game theory and control theory. Moreover, more and more results about McKean-Vlasov SDEs appear. We mention some results associated with our work. Ding and Qiao \([3, 4]\) investigated the well-posedness and stability of weak solutions for McKean-Vlasov SDEs under non-Lipschitz conditions. Lacker, Shkolnikov and Zhang \([9]\) studied superposition principles for conditional McKean-Vlasov equations. Ren and Wang \([15]\) proved that additive functionals of McKean-Vlasov SDEs have path-independence.

Nonlinear filtering problems are to extract some useful information of unobservable phenomenon from observable ones, and estimate and predict them (c.f. \([1, 7, 8, 11, 12, 13, 14, 16, 17]\)). Thus, nonlinear filtering theory plays an important role in many areas including stochastic control, financial modeling, speech and image processing, and Bayesian networks. Although McKean-Vlasov SDEs have widespread applications, the result about nonlinear filtering problems of McKean-Vlasov SDEs is seldom. Only Sen

\[\text{AMS Subject Classification(2020): } 60G35; 35K55.\]

Keywords: McKean-Vlasov SDEs; the space-distribution dependent Zakai equations; the pathwise uniqueness; space-distribution dependent Fokker-Planck equations; a superposition principle.

*This work was partly supported by NSF of China (No. 11001051, 11371352, 12071071) and China Scholarship Council under Grant No. 201906095034.

Corresponding author: Huijie Qiao, hjqiaoge@seu.edu.cn.

\(1\)
and Caines [18 19] studied nonlinear filtering problems of McKean-Vlasov SDEs with independent noises.

In the paper, we focus on nonlinear filtering problems of McKean-Vlasov SDEs with correlated noises. We explain them in detail. Fix \( T > 0 \). Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})\) be a complete filtered probability space and \( \{W_t, t \geq 0\}, \{V_t, t \geq 0\} \) be \( d \)-dimensional and \( m \)-dimensional standard Brownian motions defined on \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})\), respectively. Moreover, \( W \) and \( V \) are mutually independent. Consider the following McKean-Vlasov signal-observation system (\( \mathbf{I} \)) on \( \mathbb{R}^n \times \mathbb{R}^m \):

\[
\begin{align*}
    dX_t &= b_1(t, X_t, \mathcal{L}_{X_t}^p)dt + \sigma_0(t, X_t, \mathcal{L}_{X_t}^p)dW_t + \sigma_1(t, X_t, \mathcal{L}_{X_t}^p)dV_t, \\
    dY_t &= b_2(t, X_t, \mathcal{L}_{X_t}^p, Y_t)dt + \sigma_2(t, Y_t)dV_t, \quad 0 \leq t \leq T,
\end{align*}
\]

where \( \mathcal{L}_{X_t}^p \) denotes the distribution of \( X_t \) under the probability measure \( \mathbb{P} \), and these coefficients \( b_1 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^n \), \( \sigma_0 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^{n \times d} \), \( \sigma_1 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^{n \times m} \), \( b_2 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m \mapsto \mathbb{R}^m \) and \( \sigma_2 : [0, T] \times \mathbb{R}^m \mapsto \mathbb{R}^{m \times m} \) are Borel measurable. The initial value \( X_0 \) is assumed to be a \( p \)-order (\( p > 2 \)) integrable random variable independent of \( Y_0, W, V \). The system \( \mathbf{I} \) is called a model with a correlated noise. Then we deduce the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation about the system \( \mathbf{I} \). Next, we view the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation as two SDEs and define their strong solutions. Moreover, we prove that strong solutions of the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation both have the pathwise uniqueness. Finally, we define weak solutions of the space-distribution dependent Zakai equation and a space-distribution dependent Fokker-Planck equation, and set up a correspondence between weak solutions of the space-distribution dependent Zakai equation and weak solutions of a space-distribution dependent Fokker-Planck equation.

Moreover, our methods can be applied to study nonlinear filtering problems of McKean-Vlasov SDEs with correlated sensor noises. Concretely speaking, consider the following signal-observation system \( (\mathbf{II}) \) on \( \mathbb{R}^n \times \mathbb{R}^m \):

\[
\begin{align*}
    d\tilde{X}_t &= \tilde{b}_1(t, \tilde{X}_t, \mathcal{L}_{\tilde{X}_t}^p)dt + \tilde{\sigma}_1(t, \tilde{X}_t, \mathcal{L}_{\tilde{X}_t}^p)dV_t, \\
    d\tilde{Y}_t &= \tilde{b}_2(t, \tilde{X}_t, \mathcal{L}_{\tilde{X}_t}^p, \tilde{Y}_t)dt + \tilde{\sigma}_2dW_t + \tilde{\sigma}_3dV_t, \quad 0 \leq t \leq T,
\end{align*}
\]

where the initial value \( \tilde{X}_0 \) is assumed to be a \( p \)-order (\( p > 2 \)) integrable random variable independent of \( \tilde{Y}_0, W, V \). The mappings \( \tilde{b}_1 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^n \), \( \tilde{\sigma}_1 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^{n \times d} \) and \( \tilde{b}_2 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m \mapsto \mathbb{R}^m \) are all Borel measurable. \( \tilde{\sigma}_2, \tilde{\sigma}_3 \) are \( m \times d \) and \( m \times m \) real matrices, respectively. By the way similar to that for the system \( \mathbf{I} \), we can also establish another space-distribution dependent Zakai equation, and study its pathwise uniqueness and superposition principles.

The paper is arranged as follows. In Section \( \mathbf{II} \) we introduce notation and \( L \)-derivative for functions on \( \mathcal{P}_2(\mathbb{R}^n) \) used in the sequel. After this, we introduce nonlinear filtering problems for McKean-Vlasov signal-observation systems with correlated noises, and derive the space-distribution dependent Kushner-Stratonovich equations and the space-distribution dependent Zakai equations. In Section \( \mathbf{III} \) the pathwise uniqueness for strong solutions to the space-distribution dependent Kushner-Stratonovich equations and the
space-distribution dependent Zakai equations is shown. We place a superposition principle for the space-distribution dependent Zakai equation in Section 3. Finally, in Section 6 we summarize our results and apply our methods to the system (2).

The following convention will be used throughout the paper: $C$, with or without indices, will denote different positive constants whose values may change from one place to another.

2. Preliminary

In the section, we introduce notation and $L$-derivative for functions on $\mathcal{P}_2(\mathbb{R}^n)$.

2.1. Notation. In the subsection, we introduce notation used in the sequel.

For convenience, we shall use $| \cdot |$ and $\| \cdot \|$ for norms of vectors and matrices, respectively. Let $A^*$ denote the transpose of the matrix $A$.

Let $\mathcal{B}(\mathbb{R}^n)$ be the Borel $\sigma$-field on $\mathbb{R}^n$. Let $\mathcal{B}_b(\mathbb{R}^n)$ denote the set of all real-valued uniformly bounded $\mathcal{B}(\mathbb{R}^n)$-measurable functions on $\mathbb{R}^n$. $C^2(\mathbb{R}^n)$ stands for the space of continuous functions on $\mathbb{R}^n$ which have continuous partial derivatives of order up to 2, and $C^2_0(\mathbb{R}^n)$ stands for the subspace of $C^2(\mathbb{R}^n)$, consisting of functions whose derivatives up to order 2 are bounded. $C^\infty_c(\mathbb{R}^n)$ is the collection of all functions in $C^2(\mathbb{R}^n)$ with compact supports and $C^\infty_c(\mathbb{R}^n)$ denotes the collection of all real-valued $C^\infty$ functions of compact supports.

Let $\mathcal{M}(\mathbb{R}^n)$ be the set of all bounded Borel measures defined on $\mathcal{B}(\mathbb{R}^n)$ carrying the usual topology of weak convergence. Let $\mathcal{P}(\mathbb{R}^n)$ be the space of all probability measures defined on $\mathcal{B}(\mathbb{R}^n)$ and $\mathcal{P}_2(\mathbb{R}^n)$ be the collection of all the probability measures $\mu$ on $\mathcal{B}(\mathbb{R}^n)$ satisfying

$$\|\mu\|_2^2 := \int_{\mathbb{R}^n} |x|^2 \mu(dx) < \infty.$$ 

We put on $\mathcal{P}_2(\mathbb{R}^n)$ a topology induced by the following 2-Wasserstein metric:

$$\mathbb{W}_2^2(\mu_1, \mu_2) := \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \pi(dx, dy), \quad \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^n),$$

where $\mathcal{C}(\mu_1, \mu_2)$ denotes the set of all the probability measures whose marginal distributions are $\mu_1, \mu_2$, respectively. It is known that $(\mathcal{P}_2(\mathbb{R}^n), \mathbb{W}_2)$ is a Polish space (c.f. [20]).

2.2. $L$-derivative for functions on $\mathcal{P}_2(\mathbb{R}^n)$. In the subsection we recall the definition of $L$-derivative for functions on $\mathcal{P}_2(\mathbb{R}^n)$. The definition was first introduced by Lions (c.f. [2]). Moreover, he used some abstract probability spaces to describe the $L$-derivatives. Here, for the convenience to understand the definition, we apply a straight way to state it (c.f. [15]). Let $I$ be the identity map on $\mathbb{R}^n$. For $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ and $\phi \in L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n)$, $\langle \mu, \phi \rangle := \int_{\mathbb{R}^n} \phi(x) \mu(dx)$. Moreover, by simple calculation, it holds that $\mu \circ (I + \phi)^{-1} \in \mathcal{P}_2(\mathbb{R}^n)$.

**Definition 2.1.** (i) A function $f : \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ is called $L$-differentiable at $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, if the functional

$$L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n) \ni \phi \mapsto f(\mu \circ (I + \phi)^{-1})$$

is differentiable at $\phi$. (ii) If the functional

$$L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \nu; \mathbb{R}^n) \ni \phi \mapsto f(\nu \circ (I + \phi)^{-1})$$

is differentiable at $\phi$, then $\frac{df(\mu \circ (I + \phi)^{-1})}{d\phi}$ exists and is equal to $\frac{df(\nu \circ (I + \phi)^{-1})}{d\phi}$ for all $\phi \in L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n)$. Moreover, the functional $\phi \mapsto f(\mu \circ (I + \phi)^{-1})$ is differentiable at $\phi$ if and only if the functional $\phi \mapsto f(\nu \circ (I + \phi)^{-1})$ is differentiable at $\phi$. (iii) $\frac{df(\mu \circ (I + \phi)^{-1})}{d\phi}$ is differentiable at $\phi$ if and only if $f(\mu \circ (I + \phi)^{-1})$ is differentiable at $\phi$, and

$$\frac{d}{d\phi} \frac{df(\mu \circ (I + \phi)^{-1})}{d\phi} = \frac{df(\nu \circ (I + \phi)^{-1})}{d\phi}$$

for all $\phi \in L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n)$. Finally, we have

$$\frac{d}{d\phi} \frac{df(\mu \circ (I + \phi)^{-1})}{d\phi} = \frac{df(\nu \circ (I + \phi)^{-1})}{d\phi}$$

for all $\phi \in L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n)$.
is Fréchet differentiable at $\phi = 0$; that is, there exists a unique $\gamma \in L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu; \mathbb{R}^n)$ such that

$$
\lim_{\phi \to 0} \frac{f(\mu \circ (I + \phi)^{-1}) - f(\mu) - \langle \mu, \gamma \cdot \phi \rangle}{\sqrt{\langle \mu, |\phi|^2 \rangle}} = 0.
$$

In the case, we denote $\partial_\mu f(\mu) = \gamma$ and call it the L-derivative of $f$ at $\mu$.

(i) A function $f : \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ is called L-differentiable on $\mathcal{P}_2(\mathbb{R}^n)$ if L-derivative $\partial_\mu f(\mu)$ exists for all $\mu \in \mathcal{P}_2(\mathbb{R}^n)$.

(ii) By the same way, $\partial^2_\mu f(\mu)(y, y')$ for $y, y' \in \mathbb{R}^n$ can be defined.

Next, we introduce some related spaces.

**Definition 2.2.** The function $f$ is said to be in $C^2(\mathcal{P}_2(\mathbb{R}^n))$, if $\partial_\mu f$ is continuous, for any $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, $\partial_\mu f(\mu)(\cdot)$ is differentiable, and its derivative $\partial_\mu \partial_\mu f : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \mapsto \mathbb{R}^n \otimes \mathbb{R}^n$ is continuous, and for any $y \in \mathbb{R}^n$, $\partial_\mu f(\cdot)(y)$ is differentiable, and its derivative $\partial^2_\mu f : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^n \otimes \mathbb{R}^n$ is continuous.

**Definition 2.3.** (i) The function $F : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ is said to be in $C^{2,2}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$, if $F(x, \mu)$ is $C^2$ in $x \in \mathbb{R}^n$ and $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ respectively, and its derivatives

$$
\partial_x F(x, \mu), \partial^2_x F(x, \mu), \partial_\mu F(x, \mu)(y), \partial_y \partial_\mu F(x, \mu)(y), \partial^2_\mu F(x, \mu)(y, y')
$$

are jointly continuous in the corresponding variable family $(x, \mu), (x, y, \mu)$ or $(x, y, y')$.

(ii) The function $F : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ is said to be in $C^{2,2}_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$, if $F$ belongs to $C^{2,2}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$ and is uniformly continuous with respect to $(x, \mu)$, and its derivatives and itself are bounded.

(iii) The function $F : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ is said to be in $\mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$, if $F \in C^{2,2}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$ and for any compact set $K \subset \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$,

$$
\sup_{(x, \mu) \in K} \int_{\mathbb{R}^n} \left( \| \partial_y \partial_\mu F(x, \mu)(y) \|^2 + |\partial_\mu F(x, \mu)(y)|^2 \right) \mu(dy) < \infty.
$$

(iv) The function $\Phi : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m \mapsto \mathbb{R}$ is said to be in $C^{2,2}_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m)$, if for $y \in \mathbb{R}^m$, $\Phi(\cdot, \cdot, y) \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$ and for $(x, \mu) \in \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$, $\Phi(x, \mu, \cdot) \in C^2(\mathbb{R}^m)$.

3. Nonlinear filtering problems for McKean-Vlasov signal-observation systems with correlated noises

In this section, we introduce nonlinear filtering problems for McKean-Vlasov signal-observation systems with correlated noises, and derive the space-distribution dependent Kushner-Stratonovich equations and the space-distribution dependent Zakai equations.

3.1. The framework. In the subsection, we introduce McKean-Vlasov signal-observation systems.

Consider the system (M), i.e.

$$
\begin{cases}
   dX_t = b_1(t, X_t, \mathcal{L}^{\mathcal{P}}_{X_t}) dt + \sigma_0(t, X_t, \mathcal{L}^{\mathcal{P}}_{X_t}) dW_t + \sigma_1(t, X_t, \mathcal{L}^{\mathcal{P}}_{X_t}) dV_t, \\
   dY_t = b_2(t, X_t, \mathcal{L}^{\mathcal{P}}_{X_t}, Y_t) dt + \sigma_2(t, Y_t) dV_t, & 0 \leq t \leq T.
\end{cases}
$$

We assume the following:

**H^1_{b_1,\sigma_0,\sigma_1}** For $t \in [0, T]$ and $x_1, x_2 \in \mathbb{R}^n$, $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^n)$,

$$
|b_1(t, x_1, \mu_1) - b_1(t, x_2, \mu_2)| \leq L_1(t) \left( |x_1 - x_2| \kappa_1(|x_1 - x_2|) + \mathbb{W}_2(\mu_1, \mu_2) \right),
$$

where $L_1(t) : \mathbb{R} \mapsto \mathbb{R}$ is a non-decreasing function.
\[ \|\sigma_0(t, x_1, \mu_1) - \sigma_0(t, x_2, \mu_2)\|^2 \leq L_1(t) \left( |x_1 - x_2|^2 \kappa_2(|x_1 - x_2|) + \mathbb{W}_2^2(\mu_1, \mu_2) \right), \]
\[ \|\sigma_1(t, x_1, \mu_1) - \sigma_1(t, x_2, \mu_2)\|^2 \leq L_1(t) \left( |x_1 - x_2|^2 \kappa_3(|x_1 - x_2|) + \mathbb{W}_3^2(\mu_1, \mu_2) \right), \]

where \( L_1(t) > 0 \) is an increasing function and \( \kappa_i \) is a positive continuous function, bounded on \([1, \infty)\) and satisfies
\[
\lim_{x \to 0} \frac{\kappa_i(x)}{\log x} < \infty, \quad i = 1, 2, 3.
\]

\((H_{b_1,\sigma_0,\sigma_1}^2)\) For \( t \in [0, T] \) and \( x \in \mathbb{R}^n, \mu \in \mathcal{P}_2(\mathbb{R}^n), \)
\[ |b_1(t, x, \mu)|^2 + \|\sigma_0(t, x, \mu)\|^2 + \|\sigma_1(t, x, \mu)\|^2 \leq K_1(t)(1 + |x| + \|\mu\|)^2, \]
where \( K_1(t) > 0 \) is an increasing function.

\((H_{\sigma_2}^1)\) For \( t \in [0, T] \) and \( y_1, y_2 \in \mathbb{R}^m, \)
\[ \|\sigma_2(t, y_1) - \sigma_2(t, y_2)\|^2 \leq L_2(t)|y_1 - y_2|^2, \]
where \( L_2(t) > 0 \) is an increasing function.

\((H_{b_2,\sigma_2}^2)\) For \( t \in [0, T], y \in \mathbb{R}^m, \sigma_2(t, y) \) is invertible, and
\[ |b_2(t, x, \mu, y)| \vee \|\sigma_2(t, 0)\| \vee \|\sigma_2(t, y)\| \leq K_2, \] for all \( t \in [0, T], x \in \mathbb{R}^n, \mu \in \mathcal{P}_2(\mathbb{R}^n), y \in \mathbb{R}^m, \)
where \( K_2 > 0 \) is a constant.

Under the assumptions \((H_{b_1,\sigma_0,\sigma_1}^1), (H_{b_1,\sigma_0,\sigma_1}^2), (H_{\sigma_2}^1), (H_{b_2,\sigma_2}^2),\) by Theorem 3.1 in [3], it holds that the system \((1)\) has a pathwise unique strong solution denoted as \((X_t, Y_t)\). Set
\[ h(t, x, \mu, y) := \sigma_2^{-1}(t, y)b_2(t, x, \mu, y), \]
\[ \Gamma_t^{-1} := \exp \left\{ - \int_0^t h(s, X_s, \mathcal{L}_{X_s}^\mathbb{R}, Y_s) dV_s - \frac{1}{2} \int_0^t \left| h(s, X_s, \mathcal{L}_{X_s}^\mathbb{R}, Y_s) \right|^2 ds \right\}. \]

Here and hereafter, we use the convention that repeated indices imply summation. By \((H_{b_2,\sigma_2}^2),\) we know that
\[ \mathbb{E} \left[ \exp \left\{ \int_0^T \left| h(s, X_s, \mathcal{L}_{X_s}^\mathbb{R}, Y_s) \right|^2 ds \right\} \right] < \infty. \]

Thus, the Novikov condition holds, and \( \Gamma_t^{-1} \) is an exponential martingale. Define a measure \( \tilde{\mathbb{P}} \) via
\[ \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \Gamma_t^{-1}, \]
and under the measure \( \tilde{\mathbb{P}}, \)
\[ \tilde{V}_t := V_t + \int_0^t h(s, X_s, \mathcal{L}_{X_s}^\mathbb{R}, Y_s) ds \quad (3) \]
is an \( (\mathcal{F}_t) \)-adapted Brownian motion. Moreover, the \( \sigma \)-algebra \( \mathcal{F}_t^Y \) generated by \( \{Y_s, 0 \leq s \leq t\} \), can be characterized as
\[ \mathcal{F}_t^Y = \mathcal{F}_t^\tilde{V} \vee \mathcal{F}_0^Y, \]
where \( \mathcal{F}_t^\tilde{V} \) denotes the \( \sigma \)-algebra generated by \( \{\tilde{V}_s, 0 \leq s \leq t\} \). We augment \( \mathcal{F}_t^Y \) in a usual sense and still denote the augmentation of \( \mathcal{F}_t^\tilde{V} \) as \( \mathcal{F}_t^\tilde{V}. \)
3.2. The space-distribution dependent Kushner-Stratonovich equation. Set
\[
< \Lambda_t, F > := \mathbb{E}[F(X_t, \mathcal{L}^F_{X_t}) | \mathcal{F}^Y_t], \quad F \in \mathcal{B}_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)),
\]
where \( \mathcal{B}_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \) denotes the set of all bounded measurable functions on \( \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \), and then \( \Lambda_t \) is called the nonlinear filtering of \((X_t, \mathcal{L}^F_{X_t})\) with respect to \( \mathcal{F}^Y_t \). Moreover, the equation \( \Lambda_t \) satisfying is called the space-distribution dependent Kushner-Stratonovich equation. In order to derive the space-distribution dependent Kushner-Stratonovich equation, we need the following result (c.f. Lemma 2.2 in \([11]\)).

**Lemma 3.1.** Under the measure \( \mathbb{P} \), \( V_t := \bar{V}_t - \int_0^t < \Lambda_s, h(s, \cdot, \cdot, Y_s) > ds \) is an \( (\mathcal{F}^Y_t) \)-adapted Brownian motion.

Now, it is the position to establish the space-distribution dependent Kushner-Stratonovich equation.

**Theorem 3.2.** (The space-distribution dependent Kushner-Stratonovich equation)
For \( F \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \), the space-distribution dependent Kushner-Stratonovich equation of the system \( \{X_t\} \) is given by
\[
< \Lambda_t, F > = < \Lambda_0, F > + \int_0^t < \Lambda_s, \mathbb{L}_s F > ds + \int_0^t < \Lambda_s, \partial_x F \sigma^{ij}_0 (s, \cdot, \cdot) > d\bar{V}^j_s
\]
\[
+ \int_0^t < \Lambda_s, F h^j (s, \cdot, \cdot, Y_s) > - < \Lambda_s, F > < \Lambda_s, h^j (s, \cdot, \cdot, Y_s) > d\bar{V}^j_s,
\]
where \( t \in [0, T] \), \( \mathbb{L}_s \) is defined as
\[
(\mathbb{L}_s F)(x, \mu) = \partial_{x_i} F(x, \mu) b^i_1 (s, x, \mu) + \frac{1}{2} \partial_{x_i x_j}^2 F(x, \mu) (\sigma_0 \sigma_0^{*})^{ij} (s, x, \mu)
\]
\[
+ \frac{1}{2} \partial_{x_i x_j} F(x, \mu) (\sigma_1 \sigma_1^{*})^{ij} (s, x, \mu) + \int_{\mathbb{R}^n} (\partial_{\mu} F)_i (x, \mu) (u) b^i_1 (s, u, \mu) \mu(du)
\]
\[
+ \frac{1}{2} \int_{\mathbb{R}^n} \partial_{\mu} (\partial_{\mu} F)_j (x, \mu) (u) (\sigma_0 \sigma_0^{*})^{ij} (s, u, \mu) \mu(du)
\]
\[
+ \frac{1}{2} \int_{\mathbb{R}^n} \partial_{\mu} (\partial_{\mu} F)_j (x, \mu) (u) (\sigma_1 \sigma_1^{*})^{ij} (s, u, \mu) \mu(du).
\]

**Proof.** By the extended Itô’s formula in \([4]\) Proposition 2.9], we know that for \( F \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \)
\[
F(X_t, \mathcal{L}^F_{X_t}) = F(X_0, \mathcal{L}^F_{X_0}) + \int_0^t (\mathbb{L}_s F)(X_s, \mathcal{L}^F_{X_s}) ds
\]
\[
+ \int_0^t \partial_{x_i} F(X_s, \mathcal{L}^F_{X_s}) \sigma_0^{ij} (s, X_s, \mathcal{L}^F_{X_s}) dW^j_s
\]
\[
+ \int_0^t \partial_{x_i} F(X_s, \mathcal{L}^F_{X_s}) \sigma_1^{ik} (s, X_s, \mathcal{L}^F_{X_s}) dV^k_s
\]
\[
=: F(X_0, \mathcal{L}^F_{X_0}) + \int_0^t (\mathbb{L}_s F)(X_s, \mathcal{L}^F_{X_s}) ds + \Pi_t,
\]
where \( \Pi_t \) is a martingale, we define
\[
\mathbb{L}_s := \partial_{x_i} b^i_1 (s, x, \mu) + \frac{1}{2} \partial_{x_i x_j}^2 F(x, \mu) (\sigma_0 \sigma_0^{*})^{ij} (s, x, \mu)
\]
\[
+ \frac{1}{2} \partial_{x_i x_j} F(x, \mu) (\sigma_1 \sigma_1^{*})^{ij} (s, x, \mu) + (\partial_{\mu} F)_i (x, \mu) (u) b^i_1 (s, u, \mu) \mu(du)
\]
\[
+ \partial_{\mu} (\partial_{\mu} F)_j (x, \mu) (u) (\sigma_0 \sigma_0^{*})^{ij} (s, u, \mu) \mu(du)
\]
\[
+ \partial_{\mu} (\partial_{\mu} F)_j (x, \mu) (u) (\sigma_1 \sigma_1^{*})^{ij} (s, u, \mu) \mu(du).
\]
where \((\Pi_t)\) is an \((\mathcal{F}_t)\)-adapted local martingale. Thus, by taking the conditional expectation with respect to \(\mathcal{F}^Y_t\) on two sides of the above equality, it holds that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] + \mathbb{E} \left[ \int_0^t (\mathbb{L}_s F)(X_s, \mathcal{L}^P_{X_0}) ds | \mathcal{F}^Y_t \right] + \mathbb{E}[\Pi_t|\mathcal{F}^Y_t].
\]
We rewrite the above equality to furthermore obtain that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] - \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \int_0^t \mathbb{E} \left[ (\mathbb{L}_s F)(X_s, \mathcal{L}^P_{X_s}) d\mathcal{F}^Y_s \right] ds
\]
and that the right hand side of the above equality is an \((\mathcal{F}^Y_t)\)-adapted local martingale (c.f. [11] Lemma 2.4 and 2.5). Hence, by Corollary III.4.27 in [5] we have that there exists an \(m\)-dimensional \((\mathcal{F}^Y_t)\)-adapted process \((\Phi_t)\) such that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] - \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \int_0^t \mathbb{E} \left[ (\mathbb{L}_s F)(X_s, \mathcal{L}^P_{X_s}) d\mathcal{F}^Y_s \right] ds = \int_0^t \Phi_s d\mathbb{V}_s,
\]
where \((\Pi_t)\) is an \((\mathcal{F}^Y_t)\)-adapted local martingale. Thus, by taking the conditional expectation with respect to \(\mathcal{F}^Y_t\) on two sides of the above equality, it holds that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] + \int_0^t \mathbb{E} \left[ (\mathbb{L}_s F)(X_s, \mathcal{L}^P_{X_s}) d\mathcal{F}^Y_s \right] ds + \mathbb{E}[\Pi_t|\mathcal{F}^Y_t].
\]
Since \(X_0\) is independent of \((\mathcal{F}^Y_t)\), it holds that
\[
\mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})] = \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_0].
\]
From this, it follows that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] = \mathbb{E}[F(X_0, \mathcal{L}^P_{X_0})|\mathcal{F}^Y_t] + \int_0^t \mathbb{E} \left[ (\mathbb{L}_s F)(X_s, \mathcal{L}^P_{X_s}) d\mathcal{F}^Y_s \right] ds + \int_0^t \Phi_s d\mathbb{V}_s. \quad (7)
\]
In the following, we determine the process \((\Phi_t)\). On one side, one can apply the Itô formula to \(\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0}) V^j_t] \) and obtain that
\[
\mathbb{E}[F(X_t, \mathcal{L}^P_{X_0}) | \mathcal{F}^Y_t] V^j_t = \int_0^t \mathbb{E}[F(X_s, \mathcal{L}^P_{X_s}) | \mathcal{F}^Y_s] d\mathbb{V}_s^j + \int_0^t \mathbb{V}_s^j d\mathbb{E}[F(X_s, \mathcal{L}^P_{X_s}) | \mathcal{F}^Y_s] + \int_0^t \Phi_s ds
\]
and that the right hand side of the above equality is an \((\mathcal{F}^Y_t)\)-adapted local martingale.
\[
\int_0^t \mathbb{V}_s^j d\mathbb{E}[F(X_s, \mathcal{L}^P_{X_s}) | \mathcal{F}^Y_s] ds + \int_0^t \Phi_s ds + I^1_t, \quad (8)
\]
where \((I^1_t)\) is an \((\mathcal{F}^Y_t)\)-adapted local martingale.

On the other side, by [3], [6] and the Itô formula for \(F(X_t, \mathcal{L}^P_{X_0}) V^j_t\), we get that for \(j = 1, 2, \cdots, m\),
\[
F(X_t, \mathcal{L}^P_{X_0}) V^j_t = \int_0^t F(X_s, \mathcal{L}^P_{X_s}) d\mathbb{V}_s^j + \int_0^t \mathbb{V}_s^j dF(X_s, \mathcal{L}^P_{X_s}) + \int_0^t \partial_{x_j} F(X_s, \mathcal{L}^P_{X_s}) \sigma_{ij}^2(s, X_s, \mathcal{L}^P_{X_s}) ds.
\]

\[ \begin{align*}
&= \int_0^t F(X_s, \mathcal{L}^p_{X_s}) \phi_j(s, X_s) ds + \int_0^t \hat{V}^j_s (\mathbb{L} F)(X_s, \mathcal{L}^p_{X_s}) ds \\
&\quad + \int_0^t \partial_s F(X_s, \mathcal{L}^p_{X_s}) \sigma_1^j(s, X_s, \mathcal{L}^p_{X_s}) ds + \int_0^t F(X_s, \mathcal{L}^p_{X_s}) dV^j_s + \int_0^t \hat{V}^j_s d\Pi_s.
\end{align*} \]

Note that \( \hat{V}^j_s \) is measurable with respect to \( \mathcal{F}^Y_t \). Thus, by taking the conditional expectation with respect to \( \mathcal{F}^Y_t \) on two sides of the above equality, it holds that

\[ \mathbb{E}[F(X_t, \mathcal{L}^p_{X_t})|\mathcal{F}^Y_t] \hat{V}^j_t = \int_0^t \mathbb{E} \left[ F(X_s, \mathcal{L}^p_{X_s}) \phi_j(s, X_s) | \mathcal{F}^Y_s \right] ds \\
\quad + \int_0^t \hat{V}^j_s \mathbb{E} \left[ (\mathbb{L} F)(X_s, \mathcal{L}^p_{X_s}) | \mathcal{F}^Y_s \right] ds \\
\quad + \int_0^t \mathbb{E} \left[ \partial_s F(X_s, \mathcal{L}^p_{X_s}) \sigma_1^j(s, X_s, \mathcal{L}^p_{X_s}) | \mathcal{F}^Y_s \right] ds + I^j_t, \quad (9)
\]

where \( (I^j_t) \) denotes an \( (\mathcal{F}^Y_t) \)-adapted local martingale.

Since the left side of (8) is the same to that of (9), bounded variation parts of their right sides should be the same. Therefore,

\[
\Phi^j_t = \mathbb{E} \left[ F(X_s, \mathcal{L}^p_{X_s}) \phi_j(s, X_s, \mathcal{L}^p_{X_s}, Y_s)|\mathcal{F}^Y_s \right] - \mathbb{E}[F(X_s, \mathcal{L}^p_{X_s})|\mathcal{F}^Y_s] \mathbb{E}[\phi_j(s, X_s, \mathcal{L}^p_{X_s}, Y_s)|\mathcal{F}^Y_s] \\
+ \mathbb{E} \left[ \partial_s F(X_s, \mathcal{L}^p_{X_s}) \sigma_1^j(s, X_s, \mathcal{L}^p_{X_s}) | \mathcal{F}^Y_s \right], \quad \text{a.s. } \mathbb{P}. \quad (10)
\]

Inserting (10) in (7) and noting \(< \Lambda_t, F > = \mathbb{E}[F(X_t, \mathcal{L}^p_{X_t})|\mathcal{F}^Y_t] \), we get (4). Thus, the proof is complete. \( \square \)

3.3. The space-distribution dependent Zakai equation. Set

\[ < \tilde{\Lambda}_t, F > := \mathbb{E}^\tilde{P}[F(X_t, \mathcal{L}^p_{X_t}) \Gamma_t | \mathcal{F}^Y_t], \quad F \in \mathcal{B}_0(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), \]

where \( \mathbb{E}^\tilde{P} \) denotes the expectation under the probability measure \( \tilde{P} \). Then the space-distribution dependent Zakai equation \( \tilde{\Lambda} \) satisfying is presented as follows.

**Theorem 3.3.** (The space-distribution dependent Zakai equation)

The space-distribution dependent Zakai equation of the system (1) is given by

\[ < \tilde{\Lambda}_t, F > = < \tilde{\Lambda}_0, F > + \int_0^t < \tilde{\Lambda}_s, \mathbb{L} F > ds + \int_0^t < \tilde{\Lambda}_s, F \phi_j(s, \cdot, \cdot, Y_s) > d\hat{V}^j_s \\
+ \int_0^t < \tilde{\Lambda}_s, \partial_s F \sigma_1^j(s, \cdot, \cdot) > d\hat{V}^j_s, F \in \mathcal{L}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), t \in [0, T]. \quad (11) \]

**Proof.** Although the deduction of the space-distribution dependent Zakai equation (11) is the same to that in [11] Theorem 2.8], we give the proof to the readers’ convenience.

By the Kallianpur-Striebel formula, it holds that

\[ < \Lambda_t, F > = \mathbb{E}[F(X_t, \mathcal{L}^p_{X_t})|\mathcal{F}^Y_t] = \frac{\mathbb{E}^\tilde{P}[F(X_t, \mathcal{L}^p_{X_t}) \Gamma_t | \mathcal{F}^Y_t]}{\mathbb{E}^\tilde{P}[\Gamma_t | \mathcal{F}^Y_t]} = < \tilde{\Lambda}_t, F >, \]

and \(< \tilde{\Lambda}_t, F > = < \Lambda_t, F > < \tilde{\Lambda}_t, 1 >. \) Thus, to establish the space-distribution dependent Zakai equation (11), we investigate \(< \tilde{\Lambda}_t, 1 >. \)
First of all, by the Itô formula, it holds that
\[ \Gamma_t = 1 + \int_0^t \Gamma_s h^i(s, X_s, \mathcal{L}_{X_s}^s, Y_s) d\tilde{V}_i^s. \]

Taking the conditional expectation with respect to \( \mathcal{F}_t^Y \) under the probability measure \( \tilde{\mathbb{P}} \), one can have that
\[
\mathbb{E}^{\tilde{\mathbb{P}}}[\Gamma_t | \mathcal{F}_t^Y] = 1 + \int_0^t \mathbb{E}^{\tilde{\mathbb{P}}}[\Gamma_s h^i(s, X_s, \mathcal{L}_{X_s}^s, Y_s) | \mathcal{F}_s^Y] d\tilde{V}_i^s,
\]
namely,
\[ < \tilde{\Lambda}_t, 1 > = 1 + \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, h^i(s, \cdot, \cdot, Y_s) > d\tilde{V}_i^s. \quad (12) \]

Next, combining (4) and (12) and applying the Itô formula to \(< \Lambda_t, F >= < \tilde{\Lambda}_t, 1 >\), we obtain that
\[
< \Lambda_t, F > < \tilde{\Lambda}_t, 1 > = < \Lambda_0, F > < \tilde{\Lambda}_0, 1 > + \int_0^t < \Lambda_s, F > d < \tilde{\Lambda}_s, 1 > + \int_0^t < \tilde{\Lambda}_s, 1 > d < \Lambda_s, F >
+ \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, h^i(s, \cdot, \cdot, Y_s) > \Phi^i_s ds
+ \int_0^t < \tilde{\Lambda}_0, 1 > < \Lambda_s, \Lambda_s F > ds + \int_0^t < \tilde{\Lambda}_s, 1 > \Phi^2_s d\tilde{V}_s^2
+ \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, h^i(s, \cdot, \cdot, Y_s) > \Phi^i_s ds
= < \Lambda_0, F > < \tilde{\Lambda}_0, 1 > + \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, \Lambda_s F > ds
+ \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, \partial_x F \sigma_1^i(s, \cdot, \cdot) > d\tilde{V}_s^j
+ \int_0^t < \tilde{\Lambda}_s, 1 > < \Lambda_s, F h^j(s, \cdot, \cdot, Y_s) > d\tilde{V}_s^j,
\]
where \( \Phi \) is defined in (10). Thus, the above equality together with \(< \tilde{\Lambda}_t, F > = < \Lambda_t, F > < \tilde{\Lambda}_t, 1 >\) yields that
\[
< \tilde{\Lambda}_t, F > = < \Lambda_0, F > + \int_0^t < \tilde{\Lambda}_s, \Lambda_s F > ds + \int_0^t < \tilde{\Lambda}_s, \partial_x F \sigma_1^i(s, \cdot, \cdot) > d\tilde{V}_s^j
+ \int_0^t < \tilde{\Lambda}_s, F h^j(s, \cdot, \cdot, Y_s) > d\tilde{V}_s^j;
\]
which is just the Zakai equation (11). The proof is complete. \( \Box \)

**Remark 3.4.** If \( \sigma_1 = 0 \), Eq. (4), (17) are similar to Eq. (4.36), (4.31) in [18], respectively. If \( b_1, \sigma_0, \sigma_1, b_2 \) are independent of the distribution of \( X_t \), Eq. (4), (17) are the same to Eq. (6), (15) without jumps in [11], respectively. Therefore, our results are more general.
4. The pathwise uniqueness for strong solutions to the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation

In the section we require that \( b_2(t, x, \mu, y), \sigma_2(t, y) \) are independent of \( y \). That is, \( h(t, x, \mu, y) = h(t, x, \mu) \). First of all, we define strong solutions of the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation. After this, a filtered martingale problem is introduced and applied to show the pathwise uniqueness for strong solutions to the space-distribution dependent Kushner-Stratonovich equation and the space-distribution dependent Zakai equation.

**Definition 4.1.** A strong solution for the space-distribution dependent Kushner-Stratonovich equation is a \( (\mathcal{F}_t^Y)_{t \in [0, T]} \)-adapted, continuous and \( \mathcal{P}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \)-valued process \((\Pi_t)_{t \in [0, T]}\) such that \((\Pi_t)_{t \in [0, T]}\) solves the space-distribution dependent Kushner-Stratonovich equation \((11)\), that is,

\[
\begin{align*}
< \Pi_t, F > & = < \Lambda_0, F > + \int_0^t < \Pi_s, \mathbb{L}_s F > ds + \int_0^t < \Pi_s, \partial_x F \sigma_1^{ij}(s, \cdot, \cdot) > d\hat{V}^j_s \\
& + \int_0^t < \Pi_s, F h^j(s, \cdot, \cdot) > - < \Pi_s, F > < \Pi_s, h^j(s, \cdot, \cdot) > d\hat{V}^j_s, \\
F & \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), \quad t \in [0, T],
\end{align*}
\]

where \( \hat{V}_t := \tilde{V}_t - \int_0^t < \Pi_s, h(s, \cdot, \cdot) > ds \).

**Remark 4.2.** By the deduction in Section 3, it is obvious that \((\Lambda_t)_{t \in [0, T]}\) is a strong solution of the space-distribution dependent Kushner-Stratonovich equation \((11)\).

**Definition 4.3.** A strong solution for the space-distribution dependent Zakai equation is a \( (\mathcal{F}_t^Y)_{t \in [0, T]} \)-adapted, continuous and \( \mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \)-valued process \((\Sigma_t)_{t \in [0, T]}\) such that \((\Sigma_t)_{t \in [0, T]}\) solves the space-distribution dependent Zakai equation \((11)\), that is,

\[
\begin{align*}
< \Sigma_t, F > & = < \tilde{\Lambda}_0, F > + \int_0^t < \Sigma_s, \mathbb{L}_s F > ds + \int_0^t < \Sigma_s, F h^j(s, \cdot, \cdot) > d\tilde{V}^j_s \\
& + \int_0^t < \Sigma_s, \partial_x F \sigma_1^{ij}(s, \cdot, \cdot) > d\tilde{V}^j_s, \quad F \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), t \in [0, T]
\end{align*}
\]

**Remark 4.4.** By the deduction in Section 3, it is obvious that \((\tilde{\Lambda}_t)_{t \in [0, T]}\) is a strong solution of the space-distribution dependent Zakai equation \((11)\).

Next, we introduce an operator used in the sequel. For \( \Phi \in C^{2,2,2}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m) \), define

\[
\begin{align*}
(\mathcal{L}_X^{X,Y} \Phi)(x, \mu, y) & := \partial_x \Phi(x, \mu, y) b_1^1(s, x, \mu) + \frac{1}{2} \partial^2_{x,x} \Phi(x, \mu, y) (\sigma_0 \sigma_0^*)^{ij}(s, x, \mu) \\
& + \frac{1}{2} \partial^2_{x,x} \Phi(x, \mu, y) (\sigma_1 \sigma_1^*)^{ij}(s, x, \mu) \\
& + \int_{\mathbb{R}^n} (\partial_\mu \Phi_i)(x, \mu, y)(u)b_1^1(s, u, \mu) \mu(du) \\
& + \frac{1}{2} \int_{\mathbb{R}^n} \partial_u (\partial_\mu \Phi)_j(x, \mu, y)(u)(\sigma_0 \sigma_0^*)^{ij}(s, u, \mu) \mu(du)
\end{align*}
\]
and then we present the concept of filtered martingale problems with respect to $\mathcal{L}^{X,Y}$. 

**Definition 4.5.** A process $(\hat{\Pi}, \hat{U})$ defined on a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t \in [0,T]}, \hat{\mathbb{P}})$, with continuous trajectories and values in $\mathcal{P}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \times \mathbb{R}^m$, is a solution of the filtered martingale problem (FMP for short) $(\mathcal{L}^{X,Y}, X_0, \mathcal{L}_X^p, Y_0)$ if

(i) $\hat{\Pi}$ is $\mathcal{F}_t$-adapted,
(ii) for all $\Phi \in \mathcal{D}(\mathcal{L}^{X,Y})$, \[
<\hat{\Pi}_t, \Phi(\cdot, \cdot, \hat{U}_t) > - \int_0^t <\hat{\Pi}_s, \mathcal{L}^{X,Y}_s \Phi(\cdot, \cdot, \hat{U}_s) > ds
\]
is a $(\hat{\mathbb{P}}, (\hat{\mathcal{F}}_t)_{t \in [0,T]})$-martingale,
(iii) for all $\Phi \in \mathcal{B}_0(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^m)$, $\mathbb{E}^{\hat{\mathbb{P}}}[<\hat{\Pi}_0, \Phi(\cdot, \cdot, \hat{U}_0)>] = \mathbb{E}[\Phi(X_0, \mathcal{L}_X^p, Y_0)]$.

**Remark 4.6.** By the deduction in Section 3, we see that $(\Lambda, Y)$ is a solution of the FMP $(\mathcal{L}^{X,Y}, X_0, \mathcal{L}_X^p, Y_0)$.

**Definition 4.7.** The uniqueness for the FMP $(\mathcal{L}^{X,Y}, X_0, \mathcal{L}_X^p, Y_0)$ means that if $(\hat{\Pi}^1, \hat{U}^1)$, $(\hat{\Pi}^2, \hat{U}^2)$ defined on these probability spaces $(\hat{\Omega}^1, (\hat{\mathcal{F}}^1_t)_{t \in [0,T]}, \hat{\mathbb{P}}^1)$, $(\hat{\Omega}^2, (\hat{\mathcal{F}}^2_t)_{t \in [0,T]}, \hat{\mathbb{P}}^2)$, respectively, are two solutions of the FMP $(\mathcal{L}^{X,Y}, X_0, \mathcal{L}_X^p, Y_0)$, for $t \in [0,T]$ there exists a Borel measurable $\Psi_t : C([0,T], \mathbb{R}^m) \mapsto \mathcal{P}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$ such that
\[
\hat{\Pi}^1 = \Psi_t(U^1), \quad \hat{\mathbb{P}}^1 - a.s., \quad \hat{\Pi}^2 = \Psi_t(U^2), \quad \hat{\mathbb{P}}^2 - a.s.
\]

Here we state and prove the pathwise uniqueness of strong solutions for the space-distribution dependent Kushner-Stratonovich equations by means of filtered martingale problems.

**Theorem 4.8.** Suppose that the uniqueness holds for the FMP $(\mathcal{L}^{X,Y}, X_0, \mathcal{L}_X^p, Y_0)$. If $\{\Pi_t\}_{t \in [0,T]}$ is a strong solution of the space-distribution dependent Kushner-Stratonovich equation [4]. Then $\Pi_t = \Lambda_t, \mathbb{P}$-a.s. for all $t \in [0,T]$.

**Proof.** First of all, for $Y$ in the system $(\Pi)$, i.e.
\[
Y_t = Y_0 + \int_0^t b_2(s, X_s, \mathcal{L}^p_{X_s}) ds + \int_0^t \sigma_2(s) dV_s,
\]
we apply the Itô formula to $G(Y_t)$ for any $G \in C_c^\infty(\mathbb{R}^m)$, and obtain that
\[
G(Y_t) = G(Y_0) + \int_0^t \partial_y G(Y_s) b_2(s, X_s, \mathcal{L}^p_{X_s}) ds + \frac{1}{2} \int_0^t \partial^2_{y,y} G(Y_s) (\sigma_2(x))^2(s) ds + \int_0^t \partial_y G(Y_s) \sigma_2^j(s) dV_s^j.
\]
Besides, note that \( \{\Pi_t\}_{t \in [0,T]} \) satisfies Eq. (13), i.e.

\[
< \Pi_t, F > = < \Lambda_0, F > + \int_0^t < \Pi_s, \mathbb{L}_s F > ds + \int_0^t < \Pi_s, \partial_x \sigma_1^{ij}(s, \cdot, \cdot) > d\hat{V}^j_s
\]

\[
+ \int_0^t \left( < \Pi_s, F h^j(s, \cdot, \cdot) > - < \Pi_s, F > < \Pi_s, h^j(s, \cdot, \cdot) > \right) d\hat{V}^j_s,
\]

\( F \in \mathcal{F}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), \quad t \in [0,T]. \)

Thus, it follows from the Itô formula that

\[
< \Pi_t, F > G(Y_t) = < \Lambda_0, F > G(Y_0) + \int_0^t < \Pi_s, \mathcal{L}^{X,\mathcal{L},Y}_s (F(\cdot, \cdot) G(Y_s)) > ds
\]

\[
+ \int_0^t < \Pi_s, F > \partial_y G(Y_s) \sigma_2^{ij}(s) d\hat{V}^j_s + \int_0^t G(Y_s) < \Pi_s, \partial_x \sigma_1^{ij}(s, \cdot, \cdot) > d\hat{V}^j_s
\]

\[
+ \int_0^t G(Y_s) \left[ < \Pi_s, F h^j(s, \cdot, \cdot) > - < \Pi_s, F > < \Pi_s, h^j(s, \cdot, \cdot) > \right] d\hat{V}^j_s. \tag{15}
\]

Now, we observe that

\[
\hat{V}_t = \hat{V}_t - \int_0^t < \Pi_s, h(s, \cdot, \cdot) > ds
\]

\[
= V_t + \int_0^t h(s, X_s, \mathcal{L}^P_X) ds - \int_0^t < \Pi_s, h(s, \cdot, \cdot) > ds
\]

\[
= \hat{V}_t - \int_0^t ( < \Pi_s, h(s, \cdot, \cdot) > - < \Lambda_s, h(s, \cdot, \cdot) > ) ds.
\]

Set

\[
g(s) := < \Pi_s, h(s, \cdot, \cdot) > - < \Lambda_s, h(s, \cdot, \cdot) >,
\]

\[
\tau_N := T \wedge \inf \left\{ t > 0 : \int_0^t |g(s)|^2 ds > N \right\},
\]

and then \( \tau_N \) is a \( (\mathcal{F}^Y_t)_{t \in [0,T]} \)-stopping time and \( \tau_N \to T \) as \( N \to \infty \) by \( (H^2_{b,\sigma^2}) \). Define the probability measure

\[
\frac{d\mathbb{P}_N}{d\mathbb{P}} = \exp \left\{ \int_0^{\tau_N} g(s) d\hat{V}_s - \frac{1}{2} \int_0^{\tau_N} |g(s)|^2 ds \right\}.
\]

Thus, by Lemma 3.1 and the Girsanov theorem, it holds that \( \hat{V}_t \) is a \( (\mathcal{F}^Y_t)_{t \in [0,T]} \)-Brownian motion under \( \mathbb{P}_N \).

Next, for \( (15) \) we know that under \( \mathbb{P}_N \),

\[
< \Pi_{\tau_N \wedge t}, F > G(Y_{\tau_N \wedge t}) - \int_0^{\tau_N \wedge t} < \Pi_s, \mathcal{L}^{X,\mathcal{L},Y}_s (F(\cdot, \cdot) G(Y_s)) > ds
\]

is a \( (\mathcal{F}^Y_t)_{t \in [0,T]} \)-martingale. Thus, by the appropriate approximation it holds that for \( \Phi \in \mathcal{D}(\mathcal{L}^{X,\mathcal{L},Y}) \),

\[
< \Pi_{\tau_N \wedge t}, \Phi(\cdot, \cdot, Y_{\tau_N \wedge t}) > - \int_0^{\tau_N \wedge t} < \Pi_s, \mathcal{L}^{X,\mathcal{L},Y}_s \Phi(\cdot, \cdot, Y_s) > ds
\]
is a \((\mathbb{P}_N, (\mathcal{F}_t^Y)_{t \in [0,T]})\)-martingale. Therefore, \((\Pi, Y)\) is a solution of the FMP \((\mathcal{L}_t^{X,Y}, X_0, \mathcal{L}_0^P, Y_0)\) on the probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P}_N)\). Besides, by Remark 4.6 we see that \((\Lambda, Y)\) is also a solution of the FMP \((\mathcal{L}_t^{X,Y}, X_0, \mathcal{L}_0^P, Y_0)\) on the probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})\). So, by Corollary 3.4 in [7] and uniqueness for the FMP \((\mathcal{L}_t^{X,Y}, X_0, \mathcal{L}_0^P, Y_0)\), there exists a Borel measurable \(\Psi_t : C([0,T], \mathbb{R}^m) \mapsto \mathcal{P}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))\) such that

\[
\Pi_t 1_{t < \tau_N} = \Psi_t(Y) 1_{t < \tau_N}, \quad \mathbb{P}_N \text{-a.s.,} \quad \Lambda_t = \Psi_t(Y), \quad \mathbb{P} \text{-a.s.,}
\]

and furthermore by the equivalence of \(\mathbb{P}_N\) and \(\mathbb{P}\)

\[
\Pi_t 1_{t < \tau_N} = \Lambda_t 1_{t < \tau_N}, \quad \mathbb{P} \text{-a.s.}
\]

Taking the limits on two sides as \(N \to \infty\), we have

\[
\Pi_t = \Lambda_t, \quad \mathbb{P} \text{-a.s.}
\]

The proof is complete. \(\Box\)

In the following, we prove the pathwise uniqueness of strong solutions for the space-distribution dependent Zakai equations by means of the above theorem.

**Theorem 4.9.** Suppose that the uniqueness holds for the FMP \((\mathcal{L}_t^{X,Y}, X_0, \mathcal{L}_0^P, Y_0)\). Let \(\{\Sigma_t\}_{t \in [0,T]}\) be a strong solution of the space-distribution dependent Zakai equation (11). Then \(\Sigma_t = \Lambda_t\), \(\mathbb{P} - a.s.\) for all \(t \in [0,T]\).

**Proof.** First of all, since \(\{\Sigma_t\}_{t \in [0,T]}\) is a strong solution of the space-distribution dependent Zakai equation (11), by Definition 4.3 it holds that

\[
\begin{aligned}
\langle \Sigma_t, F \rangle &= \langle \tilde{\Lambda}_0, F \rangle + \int_0^t \langle \Sigma_s, L_s F \rangle ds + \int_0^t \langle \Sigma_s, F h^j(s, \cdot, \cdot) \rangle d\tilde{V}_s^j \\
&\quad + \int_0^t \langle \Sigma_s, \partial_x F \sigma_1^{ij}(s, \cdot, \cdot) \rangle d\tilde{V}_s^j, \quad F \in \mathcal{S}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)), t \in [0,T],
\end{aligned}
\]

and

\[
\langle \Sigma_t, 1 \rangle = 1 + \int_0^t \langle \Sigma_s, h^j(s, \cdot, \cdot) \rangle d\tilde{V}_s^j.
\]

Next, for \(\varepsilon > 0\), define the stopping time

\[
\tau_\varepsilon := \inf\{t > 0 : \langle \Sigma_t, 1 \rangle > \varepsilon\} \wedge T,
\]

and then \(\langle \Sigma_{t \wedge \tau_\varepsilon}, 1 \rangle \geq \varepsilon\). And set

\[
\begin{aligned}
\langle \Pi_{t \wedge \tau_\varepsilon}, F \rangle := \langle \Sigma_{t \wedge \tau_\varepsilon}, F \rangle \\
\langle \Sigma_{t \wedge \tau_\varepsilon}, 1 \rangle
\end{aligned}
\]

and then \(\Pi_{t \wedge \tau_\varepsilon}\) is a \((\mathcal{F}_t^Y)\)-adapted \(\mathcal{P}(\mathbb{R}^n)\)-valued process. Applying the Itô formula to \(\langle \Sigma_{t \wedge \tau_\varepsilon}, F \rangle\) and \(\langle \Sigma_{t \wedge \tau_\varepsilon}, 1 \rangle\), we obtain that

\[
\begin{aligned}
\langle \Pi_{t \wedge \tau_\varepsilon}, F \rangle &= \langle \Lambda_0, F \rangle + \int_0^{t \wedge \tau_\varepsilon} \langle \Pi_s, L_s F \rangle ds + \int_0^{t \wedge \tau_\varepsilon} \langle \Pi_s, \partial_x F \sigma_1^{ij}(s, \cdot, \cdot) \rangle d\tilde{V}_s^j \\
&\quad + \int_0^{t \wedge \tau_\varepsilon} \left( \langle \Pi_s, F h^j(s, \cdot, \cdot) \rangle - \langle \Pi_s, F \rangle \langle \Pi_s, h^j(s, \cdot, \cdot) \rangle \right) d\tilde{V}_s^j.
\end{aligned}
\]
From this, it follows that \( \{ \Pi_{t \wedge \tau} \}_{t \in [0, T]} \) is a strong solution of the space-distribution dependent Kushner-Stratonovich equation (11). By Theorem 4.8, we have that
\[
\Pi_t 1_{\{t < \tau \}} = \Lambda_t 1_{\{t < \tau \}}, \quad \mathbb{P} \text{ - a.s..} \tag{16}
\]
In the following, we compare \( \langle \Sigma_{t \wedge \tau}, 1 \rangle \) with \( \langle \Lambda_{t \wedge \tau}, 1 \rangle \). First of all, by (12), it holds that
\[
\langle \Lambda_t, 1 \rangle = \exp \left\{ \int_0^t < \Lambda_s, h^j(s, \cdot, \cdot) > d\bar{V}^j_s - \frac{1}{2} \sum_{j=1}^m \int_0^t < \Lambda_s, h^j(s, \cdot, \cdot) >^2 ds \right\}.
\]
Therefore, \( \langle \Lambda_t, 1 \rangle > 0 \). And applying the Itô formula to \( \langle \Lambda_{t \wedge \tau}, 1 \rangle \), we know that
\[
\frac{\langle \Sigma_{t \wedge \tau}, 1 \rangle}{\langle \Lambda_{t \wedge \tau}, 1 \rangle} = 1 + \int_0^{t \wedge \tau} \left[ \frac{< \Sigma_s, h^j(s, \cdot, \cdot) >}{< \Lambda_s, 1 >} - \frac{< \Sigma_s, 1 >}{< \Lambda_s, 1 >} < \Lambda_s, h^j(s, \cdot, \cdot) > \right] d\bar{V}^j_s.
\]
Note that for \( t < \tau \),
\[
\frac{< \Sigma_s, h^j(s, \cdot, \cdot) >}{< \Lambda_s, 1 >} - \frac{< \Sigma_s, 1 >}{< \Lambda_s, 1 >} < \Lambda_s, h^j(s, \cdot, \cdot) > = \frac{< \Sigma_s, 1 >}{< \Lambda_s, 1 >} < \Pi_s, h^j(s, \cdot, \cdot) > - \frac{< \Sigma_s, 1 >}{< \Lambda_s, 1 >} < \Lambda_s, h^j(s, \cdot, \cdot) > = 0,
\]
where (16) is used in the last equality. So,
\[
\frac{< \Sigma_{t \wedge \tau}, 1 >}{< \Lambda_{t \wedge \tau}, 1 >} = 1, \quad \mathbb{P} \text{ - a.s..} \tag{17}
\]
Based on (16) and (17), it holds that
\[
\Sigma_t 1_{\{t < \tau \}} = < \Sigma_t, 1 > \Pi_t 1_{\{t < \tau \}} = < \Lambda_t, 1 > \Lambda_t 1_{\{t < \tau \}} = < \Lambda_t, 1 > \Lambda_t 1_{\{t < \tau \}}, \quad \mathbb{P} \text{ - a.s.,}
\]
and furthermore by the equivalence between \( \bar{\mathbb{P}} \) and \( \mathbb{P} \),
\[
\Sigma_t 1_{\{t < \tau \}} = < \Lambda_t, 1 > \Lambda_t 1_{\{t < \tau \}}, \quad \bar{\mathbb{P}} \text{ - a.s..}
\]
From this, it follows that \( \tau > \inf \{ t > 0 : < \Lambda_t, 1 > < \varepsilon \} \wedge T \). Note that \( \inf \{ t > 0 : < \Lambda_t, 1 > < \varepsilon \} \wedge T = T \) when \( \varepsilon \) is small enough. So, \( \tau = T \) and \( \Sigma = \Lambda_t, \bar{\mathbb{P}} \text{ - a.s..} \). The proof is complete. \( \square \)

**Remark 4.10.** About the conditions under which the uniqueness holds for the FMP \((\mathcal{L}^X, \mathcal{L}^P, X_0, \mathcal{L}^P_{X_0}, Y_0)\), please refer to Theorem 3.2 in [7].

5. **Superposition between the space-distribution dependent Zakai equation and a space-distribution dependent Fokker-Planck equation**

In the section we also require that \( b_2(t, x, \mu, y), \sigma_2(t, y) \) are independent of \( y \). First of all, we define weak solutions of the space-distribution dependent Zakai equation (11). Then a type of space-distribution dependent Fokker-Planck equations and their weak solutions are introduced. Finally, we prove a superposition principle between weak solutions of the space-distribution dependent Zakai equation (11) and weak solutions of a space-distribution dependent Fokker-Planck equation.
5.1. The space-distribution dependent Zakai equation. In the subsection, we view the space-distribution dependent Zakai equation (11) as a SDE and define its weak solutions and the uniqueness in law.

**Definition 5.1.** \( \{ (\hat{\Omega}, \hat{\mathcal{F}}; \{ \hat{F}_t \}_{t \in [0,T]}, \hat{\mathbb{P}}), (\hat{\Sigma}_t, \hat{V}_t) \} \) is called a weak solution of the space-distribution dependent Zakai equation (11), if the following holds:

(i) \( (\hat{\Omega}, \hat{\mathcal{F}}; \{ \hat{F}_t \}_{t \in [0,T]}, \hat{\mathbb{P}}) \) is a complete filtered probability space;

(ii) \( \hat{\Sigma}_t \) is a \( \mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \)-valued \( (\hat{\mathcal{F}}_t) \)-adapted continuous process and \( \hat{\Sigma}_0 \in \mathcal{P}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \);

(iii) \( \hat{V}_t \) is a \( m \)-dimensional \( (\hat{\mathcal{F}}_t) \)-adapted Brownian motion;

(iv) For any \( t \in [0,T] \),
\[
\hat{\mathbb{P}} \left( \int_0^t \int_{\mathbb{R}^n} \left( \sum_{i=1}^{r_\text{in}} |b_1(1, x, \mu)| + |h(1, x, \mu)| + \| \sigma_1(1, x, \mu) \|^2 \right) \right. \\
+ \left. \sum_{i=1}^{r_\text{in}} |b_1(i, u, \mu)| + \sum_{i=1}^{r_\text{in}} \| \sigma_i(0, u, \mu) \|^2 \mu(du) \right) \, \hat{\Sigma}_t = 1;
\]

(v) \( (\hat{\Sigma}_t, \hat{V}_t) \) satisfies the following equation
\[
< \hat{\Sigma}_t, F > = < \hat{\Sigma}_0, F > + \int_0^t < \hat{\Sigma}_s, (\hat{L}_t F)(\cdot, \cdot) > \, ds + \int_0^t < \hat{\Sigma}_s, \hat{F} h^1_\mathcal{S}(s, \cdot, \cdot) > \, d\hat{V}_s^1 \\
+ \int_0^t < \hat{\Sigma}_s, \partial_x \hat{F} \sigma_1^{ij}(s, \cdot, \cdot) > \, d\hat{V}_s^j, t \in [0,T], F \in \mathcal{F}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)). \tag{18}
\]

**Remark 5.2.** By the deduction in Section 3, it is obvious that \( \{ (\Omega, \mathcal{F}, \{ F_t \}_{t \in [0,T]}, \hat{\mathbb{P}}), (\hat{\Lambda}_t, \hat{V}_t) \} \) is a weak solution of the space-distribution dependent Zakai equation (11).

**Definition 5.3.** The uniqueness in law of weak solutions for the space-distribution dependent Zakai equation (11) means that if there exist two weak solutions \( \{ (\hat{\Omega}_1, \hat{\mathcal{F}}_1; \{ \hat{F}_t^1 \}_{t \in [0,T]}, \hat{\mathbb{P}}^1), (\hat{\Sigma}_1^1, \hat{V}_1^1) \} \) and \( \{ (\hat{\Omega}_2, \hat{\mathcal{F}}_2; \{ \hat{F}_t^2 \}_{t \in [0,T]}, \hat{\mathbb{P}}^2), (\hat{\Sigma}_2^2, \hat{V}_2^2) \} \) with \( \hat{\mathbb{P}}^1 \circ (\hat{\Sigma}_0^1)^{-1} = \hat{\mathbb{P}}^2 \circ (\hat{\Sigma}_0^2)^{-1} \), then \( \hat{\mathbb{P}}^1 \circ (\hat{\Sigma}_1^1)^{-1} = \hat{\mathbb{P}}^2 \circ (\hat{\Sigma}_2^2)^{-1} \).

5.2. The space-distribution dependent Fokker-Planck equation. In the subsection, we introduce the space-distribution dependent Fokker-Planck equation associated with the space-distribution dependent Zakai equation (11) and define its weak solutions.

First of all, set
\[
\mathcal{G} := \left\{ \Sigma \in \mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \mapsto G(\Sigma) = g(< \Sigma, F_1 >, \cdots, < \Sigma, F_k >) : k \in \mathbb{N}, \right. \\
\left. g \in C^2_b(\mathbb{R}^k), F_1, \cdots, F_k \in C^2_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)) \right\}.
\]

Then for any \( G(\Sigma) = g(< \Sigma, F_1 >, \cdots, < \Sigma, F_k >) =: g(< \Sigma, F >) \in \mathcal{G} \), we define the operator \( L_t \) on \( \mathcal{G} \):
\[
L_t G(\Sigma) = \frac{1}{2} \partial_{yu} g(< \Sigma, F >) < \Sigma, F_u h^1(t, \cdot, \cdot) > + \partial_{x_i} F_u \sigma_1^{il}(t, \cdot, \cdot) > \\
\times < \Sigma, F_i h^1(t, \cdot, \cdot) > + \partial_{x_i} F_u \sigma_1^{il}(t, \cdot, \cdot) > \\
+ \partial_{yu} g(< \Sigma, F >) < \Sigma, (L_t F_u)(\cdot, \cdot) >
\]
\[
\frac{1}{2} \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
\times \langle \Sigma, F \rangle + \partial_x F \sigma^*_i(t, \cdot, \cdot) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
+ \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
+ \frac{1}{2} \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
+ \frac{1}{2} \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
+ \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) + \partial_x F \sigma^*_i(t, \cdot, \cdot) \\
+ \partial_y g(\langle \Sigma, F \rangle) + \partial_x F \sigma^*_i(t, \cdot, \cdot) + \partial_x F \sigma^*_i(t, \cdot, \cdot) > 0.
\]

Consider the following Fokker-Planck equation (FPE for short):
\[
\partial_t \Xi = L_t \Xi_t,
\]
where \((\Xi_t)_{t \in [0,T]}\) is a family of probability measures on \(\mathcal{B}(\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)))\). Then, we define weak solutions of Eq. (19).

**Definition 5.4.** A measurable family \((\Xi_t)_{t \in [0,T]}\) of probability measures on \(\mathcal{B}(\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)))\) is called a weak solution of Eq. (19) starting from \(\Xi_0\) at time 0 if
\[
\int_0^T \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} G(\Sigma) \Xi_t(d\Sigma) = \int_0^T \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} L_t G(\Sigma) \Xi_t(d\Sigma) dr.
\]

The uniqueness of the weak solutions to Eq. (19) means that, if for any \(s \in [0, T]\) and any \(\Xi_s \in \mathcal{P}(\mathcal{M}(\mathbb{R}^n))\), \((\Xi_t)_{t \in [s,T]}\) and \((\tilde{\Xi}_t)_{t \in [s,T]}\) are two weak solutions to Eq. (19) starting from \(\Xi_s\) at time \(s\), then \(\Xi_t = \tilde{\Xi}_t\) for any \(t \in [s, T]\).

It is easy to see that under the condition (20), two integrals in the right side of Eq. (21) are well defined.

5.3. **A superposition principle between Eq. (11) and Eq. (19).** In the subsection, we prove the following superposition principle between Eq. (11) and Eq. (19).
Theorem 5.5. (The superposition principle on $\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$)

(i) If $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}_{t \in [0,T]}, \hat{\mathbb{P}}), (\hat{\Sigma}_t, \hat{V}_t)\}$ is a weak solution for Eq. (11), then $(\mathcal{L}^\hat{\mathbb{P}}_{\hat{\Sigma}_t})$ is a weak solution for Eq. (19).

(ii) If weak solutions for Eq. (19) are unique, then weak solutions for Eq. (11) have the uniqueness in law.

Proof. Since $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}_{t \in [0,T]}, \hat{\mathbb{P}}), (\hat{\Sigma}_t, \hat{V}_t)\}$ is a weak solution for Eq. (11), by Definition 5.1 (iv), it holds that

$$
\int_0^T \int_{\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)} \left( |b_1(r, x, \mu)| + |h(r, x, \mu)|^2 + \|\sigma_1(r, x, \mu)\|^2 \right.
+ \|\sigma_0\sigma_0^*(r, x, \mu)|| + \int_{\mathbb{R}^n} |b_1(r, u, \mu)|\mu(du) + \int_{\mathbb{R}^n} \|\sigma_0\sigma_0^*(r, u, \mu)||\mu(du)
+ \int_{\mathbb{R}^n} \|\sigma_1\sigma_1^*(r, u, \mu)||\mu(du)\right) \hat{\Sigma}_t(\text{d}x \text{d}\mu) \text{d}r < \infty, \text{a.s.,}
$$

and

$$
\hat{\mathbb{E}} \int_0^T \int_{\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)} \left( |b_1(r, x, \mu)| + |h(r, x, \mu)|^2 + \|\sigma_1(r, x, \mu)\|^2 \right.
+ \|\sigma_0\sigma_0^*(r, x, \mu)|| + \int_{\mathbb{R}^n} |b_1(r, u, \mu)|\mu(du) + \int_{\mathbb{R}^n} \|\sigma_0\sigma_0^*(r, u, \mu)||\mu(du)
+ \int_{\mathbb{R}^n} \|\sigma_1\sigma_1^*(r, u, \mu)||\mu(du)\right) \hat{\Sigma}_t(\text{d}x \text{d}\mu) \text{d}r
$$

$$=
\int_0^T \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} \int_{\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)} \left( |b_1(r, x, \mu)| + |h(r, x, \mu)|^2 + \|\sigma_1(r, x, \mu)\|^2 \right.
+ \|\sigma_0\sigma_0^*(r, x, \mu)|| + \int_{\mathbb{R}^n} |b_1(r, u, \mu)|\mu(du) + \int_{\mathbb{R}^n} \|\sigma_0\sigma_0^*(r, u, \mu)||\mu(du)
+ \int_{\mathbb{R}^n} \|\sigma_1\sigma_1^*(r, u, \mu)||\mu(du)\right) \Sigma(\text{d}x \text{d}\mu) \mathcal{L}^\hat{\mathbb{P}}_{\hat{\Sigma}_t}(\text{d}\Sigma) \text{d}r < \infty,
$$

where $\hat{\mathbb{E}}$ denotes the expectation under the probability $\hat{\mathbb{P}}$. So, $(\mathcal{L}^\hat{\mathbb{P}}_{\hat{\Sigma}_t})$ satisfies (20).

Next, from Definition 5.1 (v), it follows that for $F \in C^{2,2}_b(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))$,

$$
< \hat{\Sigma}_t, F > = < \hat{\Sigma}_0, F > + \int_0^t < \hat{\Sigma}_s, (\mathbb{L}_a F)(\cdot, \cdot) > \text{d}s + \int_0^t < \hat{\Sigma}_s, F h^j(s, \cdot, \cdot) > \text{d}\hat{V}^j_s
+ \int_0^t < \hat{\Sigma}_s, \partial_x F \sigma^i_1(s, \cdot, \cdot) > \text{d}\hat{V}^j_s, \quad t \in [0, T].
$$

For $G(\Sigma) \in \mathcal{G}$, applying the Itô formula to $G(\hat{\Sigma}_t)$, we know that

$$
G(\hat{\Sigma}_t) = G(\hat{\Sigma}_0) + \int_0^t \partial_{y_0} g(< \hat{\Sigma}_s, F >) < \hat{\Sigma}_s, (\mathbb{L}_a F_u)(\cdot, \cdot) > \text{d}s
+ \int_0^t \partial_{y_u} g(< \hat{\Sigma}_s, F >) < \hat{\Sigma}_s, F_u h^i(s, \cdot, \cdot) + \partial_x F_u \sigma^i_1(s, \cdot, \cdot) > \text{d}\hat{V}^j_s
+ \frac{1}{2} \int_0^t \partial_{y a} y_v g(< \hat{\Sigma}_s, F >) < \hat{\Sigma}_s, F_a h^i(s, \cdot, \cdot) + \partial_x F_a \sigma^i_1(s, \cdot, \cdot) >
$$
\[ \times < \dot{\Sigma}_s, F_u h^I(s, \cdot, \cdot) + \partial_x F_u \sigma_1^u(s, \cdot, \cdot) > ds. \]

By taking the expectation on two sides, it holds that
\[ \dot{\mathbb{E}} G(\dot{\Sigma}_t) = \mathbb{E} G(\hat{\Sigma}_0) + \int_{0}^{t} \mathbb{E} \partial_y u g(< \hat{\Sigma}_s, \mathbf{F} >) < \hat{\Sigma}_s, (\mathbb{L}_s F_u)(\cdot, \cdot) > ds \]
\[ + \frac{1}{2} \int_{0}^{t} \mathbb{E} \partial_y u_\gamma g(< \hat{\Sigma}_s, \mathbf{F} >) < \hat{\Sigma}_s, F_u h^I(s, \cdot, \cdot) + \partial_x F_u \sigma_1^u(s, \cdot, \cdot) > \]
\[ \times < \hat{\Sigma}_s, F_u h^I(s, \cdot, \cdot) + \partial_x F_u \sigma_1^u(s, \cdot, \cdot) > ds, \]
and furthermore
\[ \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} G(\Sigma) \mathcal{L}_{\Sigma_t}^G(d\Sigma) \]
\[ = \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} G(\Sigma) \mathcal{L}_{\Sigma_0}^G(d\Sigma) + \int_{0}^{t} \int_{\mathcal{M}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))} \mathbf{L}_s G(\Sigma) \mathcal{L}_{\Sigma_s}^G(d\Sigma) ds. \]

Thus, \((\mathcal{L}_{\Sigma_t}^G)\) satisfies (21). The proof of (i) is complete.

For (ii), we assume that \(\{\hat{\Sigma}_t, \dot{\Sigma}_t, \{\dot{\Sigma}_t\}_{t \in [0,T]}, \mathbb{P}_1, \mathbb{P}_2\}\) and \(\{\hat{\Sigma}_t, \dot{\Sigma}_t, \{\dot{\Sigma}_t\}_{t \in [0,T]}, \mathbb{P}_1, \mathbb{P}_2\}\) are two weak solutions for Eq. (11) with \(\mathbb{P}_1 \circ (\Sigma_0^1)^{-1} = \mathbb{P}_2 \circ (\Sigma_0^2)^{-1}\) and \(\mathbb{P}_1 \circ (\Sigma_t^1)^{-1} = \mathbb{P}_2 \circ (\Sigma_t^2)^{-1}\). Then it holds that for any \(t \in (0, T]\), \(\mathbb{P}_1 \circ (\Sigma_t^1)^{-1} = \mathbb{P}_2 \circ (\Sigma_t^2)^{-1}\).

Besides, by (i), we know that \(\mathbb{P}_1 \circ (\Sigma_t^1)^{-1}, \mathbb{P}_2 \circ (\Sigma_t^2)^{-1}\) are two weak solutions for Eq. (19). And the uniqueness of weak solutions for Eq. (19) implies that \(\mathbb{P}_1 \circ (\Sigma_t^1)^{-1} = \mathbb{P}_2 \circ (\Sigma_t^2)^{-1}\). This contradicts the above conclusion. Therefore, (ii) holds. The proof is complete. \(\square\)

Remark 5.6. Since \(C_b^{2,2}(\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n))\) is not separable in the usual norms, we can’t obtain the converse conclusion.

Combining Remark 5.2 and Theorem 5.3 (i), we have the following conclusion.

Corollary 5.7. Under the assumptions \(\{\mathbf{H}_{b_1, \sigma_0, \sigma_1}\}, \{\mathbf{H}_{b_2, \sigma_0, \sigma_1}\}, \{\mathbf{H}_{b_2, \sigma_2}\}\), Eq. (19) has a weak solution.

6. CONCLUSION

In the paper, we consider the space-distribution dependent Zakai equations from non-linear filtering problems of McKean-Vlasov SDEs with correlated noises. Firstly, we establish the space-distribution dependent Kushner-Stratonovich equations and the space-distribution dependent Zakai equations. Secondly, the pathwise uniqueness of strong solutions for the space-distribution dependent Kushner-Stratonovich equations and the space-distribution dependent Zakai equations is shown. Finally, we prove a superposition principle between the space-distribution dependent Zakai equations and space-distribution dependent Fokker-Planck equations. As a by-product, we give some conditions under which space-distribution dependent Fokker-Planck equations have weak solutions.

Our methods also can be used to solve nonlinear filtering problems of McKean-Vlasov SDEs with correlated sensor noises. Concretely speaking, consider the slow-fast system in the system (22), i.e.
\[
\begin{align*}
\text{d}X_t &= \tilde{b}_1(t, X_t, \mathcal{L}_{\dot{X}_t}^{\mathbb{P}_1}) dt + \tilde{\sigma}_1(t, X_t, \mathcal{L}_{\dot{X}_t}^{\mathbb{P}_1}) dV_t, \\
\text{d}Y_t &= \tilde{b}_2(t, X_t, \mathcal{L}_{\dot{X}_t}^{\mathbb{P}_1}, Y_t) dt + \tilde{\sigma}_2 dW_t + \tilde{\sigma}_3 dV_t, \quad 0 \leq t \leq T. 
\end{align*}
\]
We assume the following:

(i) \( \hat{b}_1, \hat{\sigma}_1 \) satisfy (H\(_{b_1, \sigma_0, \sigma_1}^1\))\((H_{b_1, \sigma_0, \sigma_1}^2)\) where \( \hat{b}_1, \hat{\sigma}_1 \) replace \( b_1, \sigma_1 \), respectively;
(ii) \( \hat{b}_2(t, x, \mu, y) \) is bounded for all \( t \in [0, T], x \in \mathbb{R}^n, \mu \in \mathcal{P}_2(\mathbb{R}^n), y \in \mathbb{R}^m; \)
(iii) \( \hat{\sigma}_2 \hat{\sigma}_2^* + \hat{\sigma}_3 \hat{\sigma}_3^* = \mathbf{I}_m \), where \( \hat{\sigma}_2^* \) stands for the transpose of the matrix \( \hat{\sigma}_2 \) and \( \mathbf{I}_m \) is the \( m \)-order unit matrix.

Under the above assumptions, it holds that the system (22) has a unique strong solution denoted as \((\check{X}_t, \check{Y}_t)\). Set

\[
U_t := \hat{\sigma}_2 W_t + \hat{\sigma}_3 V_t,
\]

\[
\hat{\Gamma}_t^{-1} := \exp \left\{ - \int_0^t \hat{b}_2^2(s, \check{X}_s, \mathcal{L}_{X_s}^\varnothing, \check{Y}_s) dU_s^j - \frac{1}{2} \int_0^t \left| \hat{b}_2(s, \check{X}_s, \mathcal{L}_{X_s}^\varnothing, \check{Y}_s) \right|^2 dt \right\},
\]

and then \( \hat{\Gamma}_t^{-1} \) is an exponential martingale. Moreover, define the probability measure

\[
\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} := \hat{\Gamma}_T^{-1},
\]

and set

\[
< \tilde{\Lambda}_t, F > := \mathbb{E}^\tilde{\mathbb{P}} [F(\check{X}_t, \mathcal{L}_{X_t}^\varnothing) \hat{\Gamma}_t | \mathcal{F}_t^Y],
\]

where \( \mathbb{E}^\tilde{\mathbb{P}} \) stands for the expectation under the probability measure \( \tilde{\mathbb{P}} \). By deduction the same as to that in Theorem 4.3 we obtain the following space-distribution dependent Zakai equation.

**Corollary 6.1.** (The space-distribution dependent Zakai equation)

The space-distribution dependent Zakai equation of the system (2) is given by

\[
< \tilde{\Lambda}_t, F > = < \tilde{\Lambda}_0, F > + \int_0^t < \tilde{\Lambda}_s, \tilde{\mathcal{L}}_s F > ds + \int_0^t < \tilde{\Lambda}_s, F(\cdot, \cdot) \hat{b}_2^2(s, \cdot, \cdot, \check{Y}_s) > d\tilde{U}_s^j
\]

\[
+ \int_0^t < \tilde{\Lambda}_s, (\partial_x F)(\cdot, \cdot) \hat{\sigma}_1^{jk}(s, \cdot, \cdot) \hat{\sigma}_2^{jk} > d\tilde{U}_s^j, \quad t \in [0, T],
\]

(23)

where

\[
(\tilde{\mathcal{L}}_s F)(x, \mu) := \partial_x F(x, \mu) \hat{b}_1^j(s, x, \mu) + \frac{1}{2} \partial^2_{x_i x_j} F(x, \mu)(\hat{\sigma}_1 \hat{\sigma}_1^*)^{ij}(s, x, \mu)
\]

\[
+ \int_{\mathbb{R}^n} (\partial_x F)_i(x, \mu)(y) \hat{b}_1^j(s, y, \mu) \mu(dy)
\]

\[
+ \frac{1}{2} \int_{\mathbb{R}^n} \partial_y (\partial_x F)_j(x, \mu)(y) (\hat{\sigma}_1 \hat{\sigma}_1^*)^{ij}(s, y, \mu) \mu(dy),
\]

and \( \tilde{U}_t := U_t + \int_0^t \hat{b}_2(s, \check{X}_s, \mathcal{L}_{X_s}^\varnothing, \check{Y}_s) ds \).

Of course, we can study the pathwise uniqueness and superposition principles for Eq. (23) by means the same as to that in Theorem 4.9 and Theorem 5.5.

**Acknowledgements:**

The second named author would like to thank Professor Renming Song for providing her an excellent environment to work in the University of Illinois at Urbana-Champaign.
References

[1] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering*. Springer, New York, 2009.

[2] P. Cardaliaguet: *Notes on mean field games (from P.L. Lion’s lectures at College de France)*. https://www.ceremade.dauphine.fr/cardalia/MFG100629.pdf, 2013.

[3] X. Ding, H. Qiao: Euler-Maruyama approximations for stochastic McKean-Vlasov equations with non-Lipschitz coefficients, *Journal of Theoretical Probability*, 34(2021)1408-1425.

[4] X. Ding, H. Qiao: Stability for stochastic McKean-Vlasov equations with non-Lipschitz coefficients, *SIAM J. Control Optim.*, 59(2021)887-905.

[5] J. Jacod and A.N. Shiryaev: *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 1987.

[6] I. Karatzas and S. E. Shreve: *Brownian motion and stochastic calculus*. 2nd ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 2005.

[7] T. G. Kurtz and D. L. Ocone: Unique characterization of conditional distributions in nonlinear filtering, *Ann. Prob.*, 16(1988)80-107.

[8] T. G. Kurtz and J. Xiong: Particle representations for a class of nonlinear SPDEs, *Stochastic Process. Appl.*, 83(1999)103-126.

[9] D. Lacker, M. Shkolnikov and J. Zhang: Superposition and mimicking theorems for conditional McKean-Vlasov equations, [http://arxiv.org/abs/2004.00099](http://arxiv.org/abs/2004.00099).

[10] K. R. Parthasarathy: *Probability measures on metric spaces*. Academic Press, 1967.

[11] H. Qiao: Nonlinear filtering of stochastic differential equations driven by correlated Lévy noises, *Stochastics*, 93(2021)1156-1185.

[12] H. Qiao: Superposition principles for the Zakai equations and the Fokker-Planck equations on measure spaces, *Bulletin des sciences mathématiques*, 174(2022)103084.

[13] H. Qiao: Uniqueness for measure-valued equations of nonlinear filtering for stochastic dynamical systems with Lévy noises, *Advances in Applied Probability*, 50(2018)396-413.

[14] H. Qiao and J. Duan: Nonlinear filtering of stochastic dynamical systems with Lévy noises, *Advances in Applied Probability*, 47(2015)902-918.

[15] P. Ren and F.-Y. Wang: Space-distribution dependent PDEs for path independent additive functionals of McKean-Vlasov SDEs, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 23(2020)2050018.

[16] L. C. G. Rogers and D. Williams: *Diffusions, Markov processes and Martingales*, vol 2: *Itô calculus*, 2nd Ed. Cambridge University Press, 2000.

[17] B. L. Rozovskii: *Stochastic Evolution Systems, Linear Theory and Applications to Nonlinear Filtering*, Math. Appl. (Sov. Ser.), vol. 35, Kluwer Academic Publishers, 1990.

[18] N. Sen and P. E. Caines: Nonlinear filtering theory for McKean-Vlasov type stochastic differential equations, *SIAM J. Control Optim.*, 54(2016)153-174.

[19] N. Sen and P. E. Caines: Nonlinear filtering for McKean-Vlasov type SDEs with application to mean field games, in *Proceedings of the 21st International symposium on Mathematical Theory of Networks and Systems*, Groningen, The Netherlands, 2014, pp. 1344-1351.

[20] C. Villani: *Topics in optimal transportation*, No. 58. American Mathematical Soc., 2003.