Comparison Theorem for Distribution Dependent Neutral SFDEs*

Xing Huang[a)] **and Chenggui Yuan**[b)]

a)Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

b)Department of Mathematics, Swansea University, Swansea, SA1 8EN, UK.

xinghuang@tju.edu.cn, C.Yuan@swansea.ac.uk.

March 27, 2020

Abstract

In this paper, the existence and uniqueness of strong solutions to distribution dependent neutral SFDEs are proved. We give the conditions such that the order preservation of these equations holds. Moreover, we show these conditions are also necessary when the coefficients are continuous. Under sufficient conditions, the result extends the one in the distribution independent case, and the necessity of these conditions is new even in distribution independent case.

AMS Subject Classification: 65C30, 65L20
Keywords: Comparison Theorem, Distribution dependent neutral SFDEs, Existence and uniqueness, Order preservation, Wasserstein distance.

1 Introduction

It is well-known that the order preservation is always an important topic in every field of mathematics. In the theory of stochastic processes, the order preservation is called "comparison theorem". There are order preservations in the distribution sense and in the pathwise sense, and the pathwise one implies the distribution one. There is a lot of literature to investigate the comparison theorem. For example: Ikeda and Watanabe [8], O'Brien [11], Skorohod [14] and Yamada [18] for one dimensional stochastic differential equations (SDEs) in the pathwise sense, respectively; Chen and Wang [4] for multidimensional diffusion processes in the distribution sense; Gal'cuk and Davis [5], and Mao [10] for one dimensional SDEs driven by semimartingales in the pathwise sense, to name a few, see also [15, 16].

*Supported by NNSFC (11801406).

Moreover, the comparison theorem has been extended to stochastic functional (delay) differential equations (SFDEs), SDEs driven by jump processes and backward SDEs. We refer readers to see [2, 7, 12, 13, 19, 20], and the references therein.

Recently, in their paper [1], Bai and Jiang made a contribution on the comparison theorem for *neutral* SFDEs, where they give sufficient conditions such that the comparison theorem holds for this class of stochastic equations. In present paper, we shall study the comparison theorem for *distribution dependent neutral* SFDEs. Our results cover the ones in [1]. Furthermore, we find the conditions are also necessary.

2 Preliminaries

Throughout the paper, we let $(\mathbb{R}^n, \langle\cdot,\cdot\rangle, |\cdot|)$ be an n-dimensional Euclidean space. Denote $\mathbb{R}^{n\times m}$ by the set of all $n\times m$ matrices endowed with Hilbert-Schmidt norm $\|A\|_{HS} := \sqrt{\text{trace}(A^*A)}$ for every $A\in\mathbb{R}^{n\times m}$, in which A^* denotes the transpose of A. For fixed $r_0>0$, let $\mathscr{C} = C([-r_0,0];\mathbb{R}^n)$ denote the family of all continuous functions $h:[-r_0,0]\to\mathbb{R}^n$, equipped with the uniform norm $\|h\|_\infty := \sup_{-r_0\le\theta\le 0}|h(\theta)|$. Let $\mathscr{P}(\mathscr{C})$ denote all probability measures on $\mathscr{C}$ equipped with the weak topology. For any continuous map $f:[-r_0,\infty)\to\mathbb{R}^n$ and $t\ge 0$, let $f_t\in\mathscr{C}$ be such that $f_t(\theta)=f(\theta+t)$ for $\theta\in[-r_0,0]$. We call $(f_t)_{t\ge 0}$ the segment of $(f(t))_{t\ge -r_0}$. For $p\ge 2$, let $\mathscr{P}_p(\mathscr{C})$ denote all probability measures on $\mathscr{C}$ with finite $p-$moment, i.e. $\mu(\|\cdot\|_\infty^p)=\int_{\mathscr{C}}\|\xi\|_\infty^p\mu(\mathrm{d}\xi)<\infty$. It is well-known that $\mathscr{P}_p(\mathscr{C})$ is a polish space under the L^p-Wasserstein distance

$$\mathbb{W}_p(\mu_1,\mu_2) = \inf_{\pi\in\mathbf{C}(\mu_1,\mu_2)}\left(\int_{\mathscr{C}\times\mathscr{C}}\|\xi-\eta\|_\infty^p\pi(\mathrm{d}\xi,\mathrm{d}\eta)\right)^{1/p},$$

where $\mathbf{C}(\mu_1,\mu_2)$ denotes the class of couplings of μ_1 and μ_2. Let $(\Omega,\mathscr{F},\{\mathscr{F}_t\}_{t\ge 0},\mathbb{P})$ be a complete filtered probability space, and $\{W(t)\}_{t\ge 0}$ be an m-dimensional standard Brownian motion defined on this probability space. For any real numbers a,b, we denote $a\vee b=\max\{a,b\}, a\wedge b=\min\{a,b\}, a^+=a\vee 0$ and $a^-=-(a\wedge 0)$, where $a^+(a^-)$ is called the positive (negative) part of a. For a random variable Y on some probability space $(E,\mathscr{E},\mathbb{Q})$, we denote $\mathscr{L}_Y|\mathbb{Q}$ the distribution of Y under $\mathbb{Q}$. In this paper, we consider the following distribution dependent neutral stochastic functional differential equations (NSFDEs) on $\mathbb{R}^n$:

$$\mathrm{d}\big[X(t)-D(X_t)\big] = b(t,X_t,\mathscr{L}_{X_t})\mathrm{d}t + \sigma(t,X_t,\mathscr{L}_{X_t})\mathrm{d}W(t) \tag{2.1}$$

and

$$\mathrm{d}\big[\bar{X}(t)-D(\bar{X}_t)\big] = \bar{b}(t,\bar{X}_t,\mathscr{L}_{\bar{X}_t})\mathrm{d}t + \bar{\sigma}(t,\bar{X}_t,\mathscr{L}_{\bar{X}_t})\mathrm{d}W(t), \tag{2.2}$$

where $D:\mathscr{C}\to\mathbb{R}^n$, which is called neutral term, $b,\bar{b}:[0,\infty)\times\mathscr{C}\times\mathscr{P}(\mathscr{C})\to\mathbb{R}^n$, $\sigma,\bar{\sigma}:[0,\infty)\times\mathscr{C}\times\mathscr{P}(\mathscr{C})\to\mathbb{R}^{n\times m}$ are measurable, and $\mathscr{L}_{X_t}$ denotes the distribution of X_t.

Definition 2.1. (1) For any $s\ge 0$, a continuous adapted process $(X_{s,t})_{t\ge s}$ on $\mathscr{C}$ is called a (strong) solution of (2.1) from time s, if

$$\mathbb{E}|D(X_{s,t})|^2 + \mathbb{E}\|X_{s,t}\|_\infty^2 + \int_s^t \mathbb{E}\big\{|b(r,X_{s,r},\mathscr{L}_{X_{s,r}})| + \|\sigma(r,X_{s,r},\mathscr{L}_{X_{s,r}})\|^2\big\}\mathrm{d}r<\infty,\quad t\ge s,$$

and $(X_{s,}(t) := X_{s,t}(0))_{t\geq s}$ satisfies $\mathbb{P}$-a.s.

$$X_{s,}(t) - D(X_{s,t}) = X_{s,}(s) - D(X_{s,s}) + \int_s^t b(r, X_{s,r}, \mathscr{L}_{X_{s,r}})\mathrm{d}r \\ + \int_s^t \sigma(r, X_{s,r}, \mathscr{L}_{X_{s,r}})\mathrm{d}W(r), \quad t \geq s.$$

We say that (2.1) has (strong) existence and uniqueness, if for any $s \geq 0$ and $\mathscr{F}_s$-measurable random variable $X_{s,s}$ with $\mathbb{E}\|X_{s,s}\|_\infty^2 < \infty$, the equation from time s has a unique solution $(X_{s,t})_{t\geq s}$. When $s = 0$ we simply denote $X_{0,} = X$; i.e. $X_{0,}(t) = X(t), X_{0,t} = X_t, t \geq 0$.

(2) A couple $(\tilde{X}_{s,t}, \tilde{W}(t))_{t\geq s}$ is called a weak solution to (2.1) from time s, if $\tilde{W}(t)$ is an m-dimensional standard Brownian motion on a complete filtered probability space $(\tilde{\Omega}, \{\tilde{\mathscr{F}}_t\}_{t\geq s}, \tilde{\mathbb{P}})$, and $\tilde{X}_{s,t}$ solves

$$\mathrm{d}(\tilde{X}_{s,}(t) - D(\tilde{X}_{s,t})) = b(t, \tilde{X}_{s,t}, \mathscr{L}_{\tilde{X}_{s,t}}|_{\tilde{\mathbb{P}}})\mathrm{d}t + \sigma(t, \tilde{X}_{s,t}, \mathscr{L}_{\tilde{X}_{s,t}}|_{\tilde{\mathbb{P}}})\mathrm{d}\tilde{W}(t), \quad t \geq s.$$

(3) (2.1) is said to satisfy weak uniqueness, if for any $s \geq 0$, the distribution of a weak solution $(X_{s,t})_{t\geq s}$ to (2.1) from $s \geq 0$ is uniquely determined by $\mathscr{L}_{X_{s,s}}$.

For future use, we need the following assumptions.

(A1) $D(0) = 0$ and $D(\xi) \leq D(\eta)$ for $\xi \leq \eta$.

(A2) There exists a constant $L > 0$ such that

$$|b(t, \xi, \mu) - b(t, \eta, \nu)|^2 + |\bar{b}(t, \xi, \mu) - \bar{b}(t, \eta, \nu)|^2 \\ \leq L(\|\xi - \eta\|_\infty^2 + \mathbb{W}_2(\mu, \nu)^2), \quad t \geq 0, \xi, \eta \in \mathscr{C}; \mu, \nu \in \mathscr{P}_2(\mathscr{C}).$$

(A3) Let L be as in **(A2)** and assume

$$\|\sigma(t, \xi, \mu) - \sigma(t, \eta, \nu)\|_{HS}^2 + \|\bar{\sigma}(t, \xi, \mu) - \bar{\sigma}(t, \eta, \nu)\|_{HS}^2 \\ \leq L(\|\xi - \eta\|_\infty^2 + \mathbb{W}_2(\mu, \nu)^2), \quad t \geq 0, \xi, \eta \in \mathscr{C}; \mu, \nu \in \mathscr{P}_2(\mathscr{C}).$$

(A4) There exists an increasing function $\beta(t) \geq 0$ such that

$$|b(t, 0, \delta_0)|^2 + |\bar{b}(t, 0, \delta_0)|^2 + |\sigma(t, 0, \delta_0)|^2 + |\bar{\sigma}(t, 0, \delta_0)|^2 \leq \beta(t), t \geq 0,$$

where δ_0 is the Dirac measure at point $0 \in \mathscr{C}$.

(A5) There exists a constant $\kappa \in (0, 1)$ such that

$$|D(\xi) - D(\eta)| \leq \kappa \max_{1\leq i\leq n} \|\xi^i - \eta^i\|_\infty.$$

3 Existence and Uniqueness

In this section, we investigate the existence and uniqueness of the solution to (2.1). To this end, we use conditions which are weaker than the assumptions above.

(A5') There exists a constant $\kappa \in (0,1)$ such that

$$|D(\xi) - D(\eta)| \le \kappa \|\xi - \eta\|_\infty.$$

Theorem 3.1. *Assume* **(A2)**-**(A4)** *and* **(A5')***, then the equation* (2.1) *has a unique strong solution. Moreover, the weak uniqueness holds.*

We will prove this result by using the argument of [6] and [17], and we only need to consider the first equation in (2.1). For fixed $s \ge 0$ and $\mathscr{F}_s$-measurable $\mathscr{C}$-valued random variable $X_{s,s}$ with $\mathbb{E}\|X_{s,s}\|_\infty^2 < \infty$, we construct the first equation in (2.1) by iteration w.r.t. the measure component as follows. Firstly, let

$$X_{s,t}^{(0)} = X_{s,s}, \quad \mu_{s,t}^{(0)} = \mathscr{L}_{X_{s,t}^{(0)}}, \quad t \ge s.$$

For any $n \ge 1$, let $(X_{s,t}^{(n)})_{t \ge s}$ solve the classical neutral SFDE

$$\mathrm{d}(X_{s,}^{(n)}(t) - D(X_{s,t}^{(n)})) = b(t, X_{s,t}^{(n)}, \mu_{s,t}^{(n-1)})\mathrm{d}t + \sigma(t, X_{s,t}^{(n)}, \mu_{s,t}^{(n-1)})\,\mathrm{d}W(t), \quad t \ge s, \tag{3.1}$$

with $X_{s,s}^{(n)} = X_{s,s}$, where $\mu_{s,t}^{(n-1)} := \mathscr{L}_{X_{s,t}^{(n-1)}}$ and $X_{s,t}^{(n)}(\theta) := X_{s,}^{(n)}(t+\theta)$ for $\theta \in [-r_0, 0]$.

Lemma 3.2. *Assume* **(A2)**-**(A4)** *and* **(A5')**. *Then, for every $n \ge 1$, the neutral SFDE* (3.1) *has a unique strong solution $X_{s,t}^{(n)}$ with*

$$\mathbb{E} \sup_{t \in [s-r_0, T]} |X_{s,}^{(n)}(t)|^2 < \infty, \quad T > s, n \ge 1. \tag{3.2}$$

Moreover, for any $T > 0$, there exists $t_0 > 0$ such that for all $s \in [0,T]$ and $X_{s,s} \in L^2(\Omega \to \mathscr{C}; \mathscr{F}_s)$,

$$\mathbb{E} \sup_{t \in [s, s+t_0]} |X_{s,}^{(n+1)}(t) - X_{s,}^{(n)}(t)|^2 \le 4\mathrm{e}^{-n} \mathbb{E} \sup_{t \in [s, s+t_0]} |X_{s,}^{(1)}(t)|^2, \quad n \ge 1. \tag{3.3}$$

Proof. We shall use the approximation method in [17, Lemma 2.1] and [6, Lemma 3.2], however, we need to overcome the difficulties caused by the neutral term. Without loss of generality, we may assume that $s = 0$ and simply denote $X_{0,}(t) = X(t), X_{0,t} = X_t, t \ge 0$.

(1) We first prove that the SDE (3.1) has a unique strong solution and (3.2) holds.

For $n = 1$, let

$$\check{b}(t,\xi) = b(t,\xi,\mu_t^{(0)}), \quad \check{\sigma}(t,\xi) = \sigma(t,\xi,\mu_t^{(0)}), \quad t \ge 0, \xi \in \mathscr{C}.$$

Then (3.1) reduces to

$$\mathrm{d}(X^{(1)}(t) - D(X_t^{(1)})) = \check{b}(t, X_t^{(1)})\mathrm{d}t + \check{\sigma}(t, X_t^{(1)})\mathrm{d}W(t), \quad X_0^{(1)} = X_0, t \ge 0. \tag{3.4}$$

By **(A2)**-**(A4)** and **(A5')**, the coefficients $\check{b}$ and $\check{\sigma}$ satisfy the standard monotonicity condition which implies strong existence, uniqueness and non-explosion for neutral SFDE (3.4), see e.g. [1, Theorem 2.1]. By **(A2)**-**(A4)** and **(A5')**, there exists an increasing function $H:\mathbb{R}_+\to\mathbb{R}_+$ such that

$$\begin{aligned}
&|b(t,\xi,\mu_t^{(0)})|^2+\|\sigma(t,\xi,\mu_t^{(0)})\|_{HS}^2\\
&\le 2|b(t,\xi,\mu_t^{(0)})-b(t,0,\mu_t^{(0)})|^2+2|b(t,0,\mu_t^{(0)})|^2\\
&+2\|\sigma(t,\xi,\mu_t^{(0)})-\sigma(t,0,\mu_t^{(0)})\|_{HS}^2+2\|\sigma(t,0,\mu_t^{(0)})\|_{HS}^2\\
&\le H(t)\big\{1+\|\xi\|_\infty^2+\mu_t^{(0)}(\|\cdot\|_\infty^2)\big\},\quad t\ge 0,\xi\in\mathscr{C},
\end{aligned}$$

here $\mu_t^{(0)}(\|\cdot\|_\infty^2)=\int_{\mathscr{C}}\|\xi\|_\infty^2\mu_t^{(0)}(\mathrm{d}\xi)$. For any $N\in[1,\infty)$ and $\tau_N:=\inf\{t\ge 0:|X^{(1)}(t)|\ge N\}$, we arrive at

$$\begin{aligned}
&|X^{(1)}(t\wedge\tau_N)-D(X^{(1)}_{t\wedge\tau_N})|^2\\
&\le 3|X^{(1)}(0)-D(X^{(1)}_0)|^2+3\left|\int_0^{t\wedge\tau_N}\sigma(s,X_s^{(1)},\mu_s^{(0)})\mathrm{d}W(s)\right|^2\\
&\quad+3\left|\int_0^{t\wedge\tau_N}b(s,X_s^{(1)},\mu_s^{(0)})\mathrm{d}s\right|^2.
\end{aligned}$$

Applying inequality $(x+y)^2\le\frac{x^2}{p}+\frac{y^2}{1-p}$ for $p\in(0,1)$ and $x,y\ge 0$, we have

$$\begin{aligned}
&|X^{(1)}(t\wedge\tau_N)|^2\le\frac{|X^{(1)}(t\wedge\tau_N)-D(X^{(1)}_{t\wedge\tau_N})+D(0)|^2}{1-\kappa}+\frac{|D(X^{(1)}_{t\wedge\tau_N})-D(0)|^2}{\kappa}\\
&\le\kappa\|X^{(1)}_{t\wedge\tau_N}\|_\infty^2+c|D(0)|^2+c\|X_0^{(1)}\|_\infty^2+c\left|\int_0^{t\wedge\tau_N}\sigma(s,X_s^{(1)},\mu_s^{(0)})\mathrm{d}W(s)\right|^2\\
&\quad+c\left|\int_0^{t\wedge\tau_N}b(s,X_s^{(1)},\mu_s^{(0)})\mathrm{d}s\right|^2,\quad t\le\tau_N
\end{aligned}$$

for some constant $c>0$. Noting $\kappa\in(0,1)$, combining this with **(A4)** and applying the BDG inequality we have

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[-r_0,t\wedge\tau_N]}|X^{(1)}(s)|^2\le c\mathbb{E}\|X_0^{(1)}\|_\infty^2+c|D(0)|^2\\
&\qquad+H(t)\mathbb{E}\int_0^{t\wedge\tau_N}\big(1+\|X_s^{(1)}\|_\infty^2+\mu_s^{(0)}(\|\cdot\|_\infty^2)\big)\mathrm{d}s,\quad t\ge 0.
\end{aligned}$$

This implies

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[-r_0,t\wedge\tau_N]}|X^{(1)}(s)|^2\le c\mathbb{E}\|X_0^{(1)}\|_\infty^2+c|D(0)|^2\\
&\qquad+H(t)\int_0^t\Big\{1+\mathbb{E}\sup_{r\in[-r_0,s\wedge\tau_N]}|X^{(1)}(r)|^2+\mu_s^{(0)}(\|\cdot\|_\infty^2)\Big\}\mathrm{d}s,\quad t\ge 0.
\end{aligned}$$

By first applying Gronwall's Lemma then letting $N \to \infty$, we arrive at

$$\mathbb{E} \sup_{s\in[-r_0,t]} |X^{(1)}(s)|^2 < \infty, \quad t \geq 0.$$

Therefore, (3.2) holds for $n = 1$.

Now, assuming that the assertion holds for $n = k$ for some $k \geq 1$, one can show it for $n = k+1$. Since the proof is similar to the argument above with $(X_\cdot^{(k+1)}, \mu_\cdot^{(k)}, X_\cdot^{(k)})$ replacing $(X_\cdot^{(1)}, \mu_\cdot^{(0)}, X_\cdot^{(0)})$, we omit it here.

(2) To prove (3.3), let

$$\begin{aligned}
&\xi^{(n)}(t) = X^{(n+1)}(t) - X^{(n)}(t),\\
&\Lambda_t^{(n)} = \sigma(t, X_t^{(n+1)}, \mu_t^{(n)}) - \sigma(t, X_t^{(n)}, \mu_t^{(n-1)}),\\
&B_t^{(n)} = b(t, X_t^{(n+1)}, \mu_t^{(n)}) - b(t, X_t^{(n)}, \mu_t^{(n-1)}).
\end{aligned}$$

By **(A2)** and Itô's formula, there exists an increasing function $K_1 : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\begin{aligned}
|\xi^{(n)}(t) - (D(X_t^{(n+1)}) - D(X_t^{(n)}))|^2 \leq 2 \left| \int_0^t \Lambda_s^{(n)} \mathrm{d}W(s) \right|^2 \\
+ K_1(t) \int_0^t \left\{ \|\xi_s^{(n)}\|_\infty^2 + \mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \right\} \mathrm{d}s.
\end{aligned}$$

Again using inequality $(x+y)^2 \leq \frac{x^2}{\kappa} + \frac{y^2}{1-\kappa}$, we have

$$\begin{aligned}
|\xi^{(n)}(t)|^2 &\leq \kappa \|\xi_t^{(n)}\|_\infty^2 + \frac{2}{1-\kappa} \left| \int_0^t \Lambda_s^{(n)} \mathrm{d}W(s) \right|^2 \\
&\quad + \frac{K_1(t)}{1-\kappa} \int_0^t \left\{ \|\xi_s^{(n)}\|_\infty^2 + \mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \right\} \mathrm{d}s.
\end{aligned}$$

By the BDG inequality and noting $\kappa \in (0,1)$, we obtain

$$\begin{aligned}
\mathbb{E} \sup_{s\in[0,t]} |\xi^{(n)}(s)|^2 &\leq K_2(t) \int_0^t \left\{ \mathbb{E}\|\xi_s^{(n)}\|_\infty^2 + \mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \right\} \mathrm{d}s \\
&\leq K_2(t) \int_0^t \left\{ \mathbb{E} \sup_{r\in[0,s]} |\xi^{(n)}(r)|^2 + \mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \right\} \mathrm{d}s, \quad t \geq 0
\end{aligned}$$

for some increasing function $K_2 : \mathbb{R}_+ \to \mathbb{R}_+$.

By Gronwall's Lemma and $\mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \leq \mathbb{E}\|\xi_s^{(n-1)}\|_\infty^2$, we obtain

$$\begin{aligned}
\mathbb{E} \sup_{s\in[0,t]} |\xi^{(n)}(s)|^2 &\leq tK_2(t)\mathrm{e}^{tK_2(t)} \sup_{s\in[0,t]} \mathbb{W}_2(\mu_s^{(n)}, \mu_s^{(n-1)})^2 \\
&\leq tK_2(t)\mathrm{e}^{tK_2(t)} \mathbb{E} \sup_{s\in[0,t]} |\xi^{(n-1)}(s)|^2, \quad t \geq 0.
\end{aligned}$$

Taking $t_0 > 0$ such that $t_0 K_2(T) \mathrm{e}^{t_0 K_2(T)} \le \mathrm{e}^{-1}$, we arrive at

$$\mathbb{E} \sup_{s\in[0,t_0]} |\xi^{(n)}(s)|^2 \le \mathrm{e}^{-1} \mathbb{E} \sup_{s\in[0,t_0]} |\xi^{(n-1)}(s)|^2, \quad n \ge 1.$$

In view of

$$\mathbb{E} \sup_{s\in[0,t_0]} |\xi^{(0)}(s)|^2 \le 2\mathbb{E}\Big\{ |X(0)|^2 + \sup_{s\in[0,t_0]} |X^{(1)}(s)|^2 \Big\} \le 4\mathbb{E} \sup_{s\in[0,t_0]} |X^{(1)}(s)|^2,$$

we obtain (3.3). □

Proof of Theorem 3.1. (**Existence**) For simplicity, we only consider $s = 0$ and denote $X_{0,} = X$; i.e. $X_{0,}(t) = X(t), X_{0,t} = X_t, t \ge 0$.

Let $(X_t)_{t\in[0,t_0]}$ be the unique limit of $(X_t^{(n)})_{t\in[0,t_0]}$ in Lemma 3.2, then $(X_t)_{t\in[0,t_0]}$ is an adapted continuous process and satisfies

$$\lim_{n\to\infty} \sup_{t\in[0,t_0]} \mathbb{W}_2(\mu_t^{(n)}, \mu_t)^2 \le \lim_{n\to\infty} \mathbb{E} \sup_{t\in[0,t_0]} |X^{(n)}(t) - X(t)|^2 = 0, \tag{3.5}$$

where μ_t is the distribution of X_t. Rewriting (3.1), we have

$$X^{(n)}(t) - D(X_t^{(n)}) = X(0) - D(X_0) + \int_0^t b(s, X_s^{(n)}, \mu_s^{(n-1)}) \mathrm{d}s + \int_0^t \sigma(s, X_s^{(n)}, \mu_s^{(n-1)}) \mathrm{d}W(s).$$

Then (3.5), **(A2)**, **(A3)**, **(A5')** and the dominated convergence theorem imply $\mathbb{P}$-a.s.

$$X(t) - D(X_t) = X(0) - D(X_0) + \int_0^t b(s, X_s, \mu_s) \mathrm{d}s + \int_0^t \sigma(s, X_s, \mu_s) \mathrm{d}W(s), \quad t \in [0, t_0].$$

Therefore, $(X_t)_{t\in[0,t_0]}$ solves (2.1) up to time t_0. Moreover, $\mathbb{E}\sup_{s\in[0,t_0]} |X(s)|^2 < \infty$ follows by (3.5). The same holds for $(X_{s,t})_{t\in[s,(s+t_0)\wedge T]}$ and $s \in [0, T]$. So, by solving the equation piecewise in time, and using the arbitrariness of $T > 0$, we conclude that (2.1) has a strong solution $(X_t)_{t\ge 0}$ with

$$\mathbb{E} \sup_{s\in[0,t]} |X(s)|^2 < \infty, \quad t \ge 0. \tag{3.6}$$

Uniqueness Let X and Y be two solutions to (2.1), i.e.

$$\mathrm{d}\big[X(t) - D(X_t)\big] = b(t, X_t, \mathscr{L}_{X_t}) \mathrm{d}t + \sigma(t, X_t, \mathscr{L}_{X_t}) \mathrm{d}W(t),$$

and

$$\mathrm{d}\big[Y(t) - D(Y_t)\big] = b(t, Y_t, \mathscr{L}_{Y_t}) \mathrm{d}t + \sigma(t, Y_t, \mathscr{L}_{Y_t}) \mathrm{d}W(t).$$

By **(A2)**, we have

$$\begin{aligned} |X(t) - Y(t) - (D(X_t) - D(Y_t))|^2 \le & 2 \left| \int_0^t \{\sigma(s, X_s, \mathscr{L}_{X_s}) - \sigma(s, Y_s, \mathscr{L}_{Y_s})\} \mathrm{d}W(s) \right|^2 \\ & + \beta_1(t) \int_0^t \big\{ \|X_s - Y_s\|_\infty^2 + \mathbb{W}_2(\mathscr{L}_{X_s}, \mathscr{L}_{Y_s})^2 \big\} \mathrm{d}s \end{aligned}$$

for an increasing function $\beta_1 : [0, \infty) \to [0, \infty)$. Applying inequality $(x+y)^2 \le \frac{x^2}{\kappa} + \frac{y^2}{1-\kappa}$, we get

$$
\begin{aligned}
|X(t) - Y(t)|^2 \le & \kappa \|X_t - Y_t\|_\infty^2 + \frac{2}{1-\kappa} \left| \int_0^t \{\sigma(s, X_s, \mathscr{L}_{X_s}) - \sigma(s, Y_s, \mathscr{L}_{Y_s})\} \mathrm{d}W(s) \right|^2 \\
& + \frac{\beta_1(t)}{1-\kappa} \int_0^t \left\{ \|X_s - Y_s\|_\infty^2 + \mathbb{W}_2(\mathscr{L}_{X_s}, \mathscr{L}_{Y_s})^2 \right\} \mathrm{d}s.
\end{aligned}
$$

Noting that $\mathbb{W}_2(\mathscr{L}_{X_t}, \mathscr{L}_{Y_t})^2 \le \mathbb{E}\|X_t - Y_t\|_\infty^2$, **(A3)** and the BDG inequality imply that $\gamma_t := \sup_{s \in [-r_0, t]} |X(s) - Y(s)|^2$ satisfies

$$
\mathbb{E}\gamma_t \le \beta_2(t) \int_0^t \mathbb{E}\gamma_s \mathrm{d}s, \quad t \ge 0
$$

for an increasing function $\beta_2 : [0, \infty) \to [0, \infty)$. So, applying Gronwall's inequality implies

$$
\mathbb{E}\gamma_t = 0, \quad t \ge 0.
$$

(**Weak uniqueness**) Since the proof is similar to that of [17, Theorem 2.1], we omit it here. □

4 Comparison Theorem

In order to obtain the comparison theorem for distribution dependent NSFDEs, we introduce the partial order on $\mathscr{C}$.

- For $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{R}^n$, we call $x \le y$ if and only if $x_i \le y_i, i = 1, \dots, n$; $x < y$ if and only if $x \le y$ and $x \ne y$; $x \ll y$ if and only if $x_i < y_i, i = 1, \dots, n$.
- For $\xi = (\xi_1, \cdots, \xi_n), \eta = (\eta_1, \cdots, \eta_n) \in \mathscr{C}$, we call $\xi \le \eta$ if and only if $\xi(\theta) \le \eta(\theta), \theta \in [-r_0, 0]$; $\xi < \eta$ if and only if $\xi \le \eta$ and $\xi \ne \eta$; $\xi \ll \eta$ if and only if $\xi(\theta) < \eta(\theta), \theta \in [-r_0, 0]$; for any $\xi, \eta \in \mathscr{C}$, $\xi \wedge \eta$ is defined by $(\xi \wedge \eta)_i = \xi_i \wedge \eta_i, i = 1, \dots, n$.
- We also define the following partial order associated with the neutral term $D(\cdot)$, that is: $\xi \le_D \eta$ if and only if $\xi \le \eta$ and $\xi(0) - D(\xi) \le \eta(0) - D(\eta)$; $\xi <_D \eta$ if and only if $\xi \le_D \eta$ and $\xi \ne \eta$.
- A function h on $\mathscr{C}$ is called increasing if $h(\xi) \le h(\eta)$ for $\xi \le \eta$. Letting $\mu_1, \mu_2 \in \mathscr{P}(\mathscr{C})$, we call $\mu_1 \le \mu_2$ if and only if $\mu_1(h) \le \mu_2(h)$ holds for all increasing functions $h \in C_b(\mathscr{C})$, where the latter denotes all bounded continuous functions on $\mathscr{C}$.

Denote by $(X(s, \xi; t), \bar{X}(s, \bar{\xi}; t))_{t \ge s}$ the solutions to (2.1)-(2.2) with $(X_s(s, \xi), \bar{X}_s(s, \bar{\xi})) = (\xi, \bar{\xi})$. Let $(X_t(s, \xi), \bar{X}_t(s, \bar{\xi}))_{t \ge s}$ be the segment process.

Definition 4.1. The distribution dependent NSFDE (2.1)-(2.2) is called D-order-preserving, if for any $s \geq 0$ and $\xi, \bar{\xi} \in L^2(\Omega \to \mathscr{C}, \mathscr{F}_s, \mathbb{P})$ with $\mathbb{P}(\xi \leq_D \bar{\xi}) = 1$, one has

$$\mathbb{P}(X_t(s,\xi) \leq_D \bar{X}_t(s,\bar{\xi}), \quad t \geq s) = 1.$$

Definition 4.2. A function $f : \mathscr{C} \to \mathbb{R}^1$ is called D-increasing, if for any $\xi \leq_D \eta$, it holds $f(\xi) \leq f(\eta)$. If two probability measures μ, ν on $\mathscr{C}$ satisfy $\mu(f) \leq \nu(f)$ for any D-increasing function $f \in C_b(\mathscr{C})$, we denote $\mu \leq_D \nu$.

Remark 4.1. *In fact, if $\mu \leq_D \nu$, by [9, Theorem 1 (i) $\Rightarrow$ (ii)], there exists $\pi \in \mathbf{C}(\mu,\nu)$ with*

$$\pi(\{(\xi_1,\xi_2), \xi_1 \leq_D \xi_2\}) = 1.$$

4.1 Sufficient Conditions for the Comparison Theorem

In this subsection, we provide sufficient conditions such that the comparison theorem holds. Comparing with the result in [1], we shall cope with the difficulties caused by the distribution dependence.

Theorem 4.2. *Let* **(A1)**-**(A5)** *hold and $b, \bar{b}$ and $\sigma, \bar{\sigma}$ are continuous on $[0,\infty) \times \mathscr{C} \times \mathscr{P}_2(\mathscr{C})$. Assume $\xi \leq_D \bar{\xi}$ and the following conditions hold.*

(i) *The drift terms $b = (b_1, \ldots, b_n)$ and $\bar{b} = (\bar{b}_1, \ldots, \bar{b}_n)$ satisfy $b_i(t,\eta,\mu) \leq \bar{b}_i(t,\bar{\eta},\bar{\mu})$ for any $1 \leq i \leq n$ provided $\mu, \bar{\mu} \in \mathscr{P}_2(\mathscr{C})$ with $\mu \leq_D \bar{\mu}$, $\eta, \bar{\eta} \in \mathscr{C}$ with $\eta \leq_D \bar{\eta}$ and $\eta^i(0) - D^i(\eta) = \bar{\eta}^i(0) - D^i(\bar{\eta})$.*

(ii) *The diffusion terms $\sigma = (\sigma_{ij})$ and $\bar{\sigma} = (\bar{\sigma}_{ij})$ satisfy $\sigma = \bar{\sigma}$. Moreover, $\sigma_{ij}(t,\eta,\mu)$ only depends on t and $\eta^i(0) - D^i(\eta)$.*

Then it holds $\mathbb{P}(X_t(s,\xi) \leq_D \bar{X}_t(s,\bar{\xi}), t \geq s) = 1$.

In the following, for simplicity, let $s = 0$, $X(t) = X(s,\xi;t), \bar{X}(t) = \bar{X}(s,\bar{\xi};t), X_D(t) = X(t) - D(X_t), \bar{X}_D(t) = \bar{X}(t) - D(\bar{X}_t)$.

Define the following stopping times:

$$\rho_i = \inf\{t > 0 : X^i(t) > \bar{X}^i(t)\}, \quad i = 1,2,\ldots,n,$$
$$\Upsilon_i = \inf\{t > 0 : X^i_D(t) > \bar{X}^i_D(t)\}, \quad i = 1,2,\ldots,n.$$

Let $\rho = \min\{\rho_1,\ldots,\rho_n\}$ and $\Upsilon = \min\{\Upsilon_1,\ldots,\Upsilon_n\}$. We firstly give a modified proof of [1, Proposition 3.1] which extends the result there to the case that D is nonlinear.

Proposition 4.3. *Assume* **(A1)** *and* **(A5)** *hold, then we have*

$$\Upsilon \leq \rho \text{ on } \Omega. \tag{4.1}$$

Proof. Set

$$\rho_i^l := \inf\{t > 0 : X^i(t) > \bar{X}^i(t) + \frac{1}{l}\}, \quad i = 1, 2, \ldots, n; l \geq 1, \tag{4.2}$$

and

$$\rho^l := \min\{\rho_1^l, \ldots, \rho_n^l\}.$$

Then it is easy to see for $i = 1, \cdots, n$ and $0 \leq t \leq \rho_i^l$,

$$X^i(\rho_i^l) = \bar{X}^i(\rho_i^l) + \frac{1}{l}, \quad X^i(t) \leq \bar{X}^i(t) + \frac{1}{l}, \tag{4.3}$$

and $\rho = \inf\{\rho^l : l \geq 1\}$. Moreover, by the definition of ρ_i and Υ_i, one has

$$X^i(\rho_i) = \bar{X}^i(\rho_i), \quad X^i(t) \leq \bar{X}^i(t), \quad i = 1, \cdots, n; 0 \leq t \leq \rho_i, \tag{4.4}$$

and

$$X_D^i(\Upsilon_i) = \bar{X}_D^i(\Upsilon_i), \quad X_D^i(t) \leq \bar{X}_D^i(t), \quad i = 1, \cdots, n; 0 \leq t \leq \Upsilon_i. \tag{4.5}$$

We only need to prove $\Upsilon \leq \rho^l$ for any $l \geq 1$ and $\omega \in \Omega$. To this end, we assume that there exists a $l \geq 1$ and $\omega_0 \in \Omega$ such that $\rho^l(\omega_0) < \Upsilon(\omega_0)$. Then there exists a $1 \leq n_0 = n_0(\omega_0) \leq n$ such that $\rho_{n_0}^l(\omega_0) = \rho^l(\omega_0)$. Then by (4.5), we have $[X_D^{n_0}(\rho_{n_0}^l)](\omega_0) \leq [\bar{X}_D^{n_0}(\rho_{n_0}^l)](\omega_0)$. This together with (4.3) implies $\frac{1}{l} + D^{n_0}(\bar{X}_{\rho_{n_0}^l}(\omega_0)) - D^{n_0}(X_{\rho_{n_0}^l}(\omega_0)) \leq 0$. This combined with (4.3) and the monotonicity of D yields

$$\frac{1}{l} + D^{n_0}(\bar{X}_{\rho_{n_0}^l}(\omega_0)) - D^{n_0}(\bar{X}_{\rho_{n_0}^l}(\omega_0) + \frac{1}{l}) \leq 0.$$

By **(A5)**, we obtain $\frac{1}{l} - \frac{\kappa}{l} \leq 0$. In view of $\kappa \in (0, 1)$, this is a contradiction. Thus, we finish the proof. □

Remark 4.4. *Under* **(A1)** *and* **(A5)**, *the following two conditions* **(C1)** *and* **(C2)** *are equivalent.*

(C1) **(A3)** *together with Theorem 4.2 (ii).*

(C2) $\sigma = \bar{\sigma}$ *and there exists a constant* $\tilde{L} > 0$ *such that for any* $i = 1, \ldots, n$,

$$\begin{aligned} &\sum_{j=1}^{m} \left(|\sigma_{ij}(t, \xi, \mu) - \sigma_{ij}(t, \eta, \nu)|^2 + |\bar{\sigma}_{ij}(t, \xi, \mu) - \bar{\sigma}_{ij}(t, \eta, \nu)|^2 \right) \\ &\leq \tilde{L} |\xi^i(0) - D^i(\xi) - \eta^i(0) + D^i(\eta)|^2, \quad t \geq 0, \xi, \eta \in \mathscr{C}; \mu, \nu \in \mathscr{P}_2(\mathscr{C}). \end{aligned} \tag{4.6}$$

It should be point out that condition **(C2)** *in the distribution independent case has been used in [1, Theorem 3.1].*

In fact, it is clear that (4.6) *implies* **(A3)** *by using* **(A5)**. *By taking* $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$, (4.6) *and* $\sigma = \bar{\sigma}$ *lead to Theorem 4.2 (ii). So* **(C2)** *implies* **(C1)**. *It suffices* *to show that* **(C1)** *implies* **(C2)**. *To this end, let* $\mathbf{e} \in \mathscr{C}$ *be defined by* $\mathbf{e}^k(s) = 1, s \in [-r_0, 0], 1 \le k \le n$. *Let* i, $\xi, \eta \in \mathscr{C}$ *and* $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$ *be fixed. Consider the function* $g_i(r) = r - D^i(r\mathbf{e}), r \in \mathbb{R}$. **(A1)** *and* **(A5)** *imply*

$$g_i(0) = 0; \quad g_i(r) \ge r - \kappa r, \quad r > 0; \quad g_i(r) \le r - \kappa r, \quad r < 0.$$

This together with the continuity of D *leads to that there exist* $r_i, \bar{r}_i \in \mathbb{R}$ *such that*

$$g_i(r_i) = r_i - D^i(r_i\mathbf{e}) = \xi^i(0) - D^i(\xi), \quad g_i(\bar{r}_i) = \bar{r}_i - D^i(\bar{r}_i\mathbf{e}) = \eta^i(0) - D^i(\eta).$$

Thus, it follows from **(A5)** *that*

$$|\xi^i(0) - D^i(\xi) - \eta^i(0) + D^i(\eta)| \ge |r_i - \bar{r}_i| - |D^i(r_i\mathbf{e}) - D^i(\bar{r}_i\mathbf{e})| \ge |r_i - \bar{r}_i| - \kappa|r_i - \bar{r}_i|.$$

Noting that $\{r_i\mathbf{e}\}^i(0) = r_i$ *and* $\{\bar{r}_i\mathbf{e}\}^i(0) = \bar{r}_i$, *by Theorem 4.2 (ii), we have* $\sigma_{ij}(t, \xi, \mu) = \sigma_{ij}(t, r_i\mathbf{e}, \mu)$ *and* $\sigma_{ij}(t, \eta, \nu) = \sigma_{ij}(t, \bar{r}_i\mathbf{e}, \mu)$. *This combined with* $\sigma = \bar{\sigma}$ *by Theorem 4.2 (ii) and* **(A3)** *yields*

$$\begin{aligned}
&\sum_{j=1}^m \left(|\sigma_{ij}(t, \xi, \mu) - \sigma_{ij}(t, \eta, \nu)|^2 + |\bar{\sigma}_{ij}(t, \xi, \mu) - \bar{\sigma}_{ij}(t, \eta, \nu)|^2\right) \\
&= \sum_{j=1}^m \left(|\sigma_{ij}(t, r_i\mathbf{e}, \mu) - \sigma_{ij}(t, \bar{r}_i\mathbf{e}, \mu)|^2 + |\bar{\sigma}_{ij}(t, r_i\mathbf{e}, \mu) - \bar{\sigma}_{ij}(t, \bar{r}_i\mathbf{e}, \mu)|^2\right) \\
&\le L|r_i - \bar{r}_i|^2 \le L(1-\kappa)^{-2}|\xi^i(0) - D^i(\xi) - \eta^i(0) + D^i(\eta)|^2.
\end{aligned}$$

As a result, (4.6) *holds for* $\tilde{L} = L(1-\kappa)^{-2}$.

Remark 4.5. *With Proposition 4.3 and Remark 4.4 in hand, repeating the proof of [1, Theorem 3.1], we obtain the following result: If* $b, \bar{b}$ *and* $\sigma, \bar{\sigma}$ *do not depend on the distribution, Theorem 4.2 holds if the condition* $b_i(t, \eta) \le \bar{b}_i(t, \bar{\eta})$ *in (i) is replaced by* $b_i(t, \eta) < \bar{b}_i(t, \bar{\eta})$.

Now we intend to prove the distribution dependent case.

Proof of Theorem 4.2. We first prove the result in Theorem 4.2 holds when the condition $b_i(t, \eta, \mu) \le \bar{b}_i(t, \bar{\eta}, \bar{\mu})$ in (i) is replaced by $b_i(t, \eta, \mu) < \bar{b}_i(t, \bar{\eta}, \bar{\mu})$. For any $n \ge 0$, let $(X_{s,t}^{(n)})_{t \ge s}$ solve (3.1) with $X_{s,s}^{(n)} = \xi$ and $X_{s,t}^{(0)} = \xi, t \ge s$. Similarly, let

$$\bar{X}_{s,t}^{(0)} = \bar{\xi}, \quad t \ge s,$$

and $(\bar{X}_{s,t}^{(n)})_{t \ge s}$ solve (3.1) with $\bar{b}$ and $\bar{\sigma}$ in place of b and σ and $\bar{X}_{s,s}^{(n)} = \bar{\xi}$. Denote $\bar{\mu}_{s,t}^{(n-1)} := \mathscr{L}_{\bar{X}_{s,t}^{(n-1)}}$. We should remark that $\{\bar{X}_{s,t}^{(0)}\}_{t \ge s}$ and $\{X_{s,t}^{(0)}\}_{t \ge s}$ are continuous $\mathscr{C}$-valued processes. Without loss of generality, we assume $s = 0$ and omit the subscript s. Put

$$b^n(t, \eta) = b(t, \eta, \mu_t^{(n-1)}), \quad \sigma^n(t, \eta) = \sigma(t, \eta, \mu_t^{(n-1)}),$$

and

$$\bar{b}^n(t, \eta) = \bar{b}(t, \eta, \bar{\mu}_t^{(n-1)}), \quad \bar{\sigma}^n(t, \eta) = \bar{\sigma}(t, \eta, \bar{\mu}_t^{(n-1)}).$$

For $n = 1$, thanks to $\mu_t^{(0)} \le_D \bar{\mu}_t^{(0)}$, by (i) and (ii) in Theorem 4.2, we have

(1) $b_i^1(t,\eta) < \bar{b}_i^1(t,\bar\eta)$ for any $1 \le i \le n$ provided $\eta, \bar\eta \in \mathscr{C}$ with $\eta \le_D \bar\eta$ and $\eta^i(0) - D^i(\eta) = \bar\eta^i(0) - D^i(\bar\eta)$.

(2) $\sigma^1 = \bar\sigma^1$ and $\sigma_{ij}^1(t,\eta)$ only depends on t and $\eta^i(0) - D^i(\eta)$.

Then by Remark 4.5, it holds $\mathbb{P}$-a.s.

$$X_t^{(1)} \le_D \bar{X}_t^{(1)}, \quad t \ge 0.$$

Next, assume $\mathbb{P}$-a.s.

$$X_t^{(n-1)} \le_D \bar{X}_t^{(n-1)}, \quad t \ge 0.$$

This together with [9, Theorem 1 (iv) $\Rightarrow$ (i)] yields $\mathscr{L}_{X_t^{(n-1)}} \le_D \mathscr{L}_{\bar{X}_t^{(n-1)}}$. Repeating the proof for b^n, σ^n, $\bar{b}^n$, $\bar\sigma^n$, $X^{(n-1)}$ in place of b^1, σ^1, $\bar{b}^1$, $\bar\sigma^1$, $X^{(0)}$, we can prove $\mathbb{P}$-a.s.

$$X_t^{(n)} \le_D \bar{X}_t^{(n)}, \quad t \ge 0.$$

By (3.5), we conclude $\mathbb{P}$-a.s.

$$X_t \le_D \bar{X}_t, \quad t \ge 0.$$

Then the required assertion follows.

In general, if the Assumption (i) in Theorem 4.2 holds, then let $(\bar{b}_\varepsilon, \bar{X}^\varepsilon)$ be as in Lemma 4.6 below. By the above conclusion, we have $\mathbb{P}$-a.s.

$$X_t \le_D \bar{X}_t^\varepsilon, \quad t \ge 0.$$

Letting ε goes to 0, it follows from Lemma 4.6 below and the continuity of D that $\mathbb{P}$-a.s.

$$X_t \le_D \bar{X}_t, \quad t \ge 0.$$

This completes the proof. $\square$

Lemma 4.6. *Let $\bar{b}_\varepsilon = \bar{b} + \vec\varepsilon$, here $\vec\varepsilon = (\varepsilon, \varepsilon, \cdots, \varepsilon) \in \mathbb{R}^n$ and $\varepsilon > 0$. Let $\bar{X}^\varepsilon(t)$ solve* (2.2) *with $\bar{X}_0^\varepsilon = \bar{X}_0$ and $\bar{b}_\varepsilon$ in place of $\bar{b}$. If the conditions in Theorem 4.2 hold, then for any $T > 0$, it holds*

$$\lim_{\varepsilon \to 0^+} \mathbb{E} \sup_{t \in [0,T]} |\bar{X}^\varepsilon(t) - \bar{X}(t)| = 0.$$

The proof is standard, and we omit it here.

4.2 Necessary Conditions for the Comparison Theorem

In this subsection, we show the conditions in Theorem 4.2 are also necessary. To this end, we firstly introduce a lemma.

Lemma 4.7. *Assume* **(A1)** *and* **(A5)**. *Then the following assertions hold.*

(1) For any $1 \le i \le n$, $\xi, \eta \in \mathscr{C}$ with $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$, there exists $\zeta \in \mathscr{C}$ such that $\zeta \le \xi \wedge \eta$ and $\zeta^i(0) - D^i(\zeta) = \xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$.

(2) For $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$, there exists $\tilde{\mu} \in \mathscr{P}_2(\mathscr{C})$ such that $\tilde{\mu} \leq_D \mu$ and $\tilde{\mu} \leq_D \nu$.

Proof. (1) Fix $1 \leq i \leq n$, $\xi, \eta \in \mathscr{C}$ with $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$. Without loss of generality, assume $\xi^i(0) \leq \eta^i(0)$. If $D^i(\xi \wedge \eta) = D^i(\xi)$, let $\zeta = \xi \wedge \eta$. Otherwise, (**A1**) implies $D^i(\xi \wedge \eta) < D^i(\xi)$. Recall that $\mathbf{e} \in \mathscr{C}$ is defined by $\mathbf{e}^k(s) = 1, s \in [-r_0, 0], 1 \leq k \leq n$. Define

$$h^i(r) = r + D^i(\xi \wedge \eta - r\mathbf{e}) - D^i(\xi), \quad r \in \mathbb{R}. \tag{4.7}$$

We have $h^i(0) < 0$ and **(A5)** implies

$$h^i(r) \geq r - \kappa \max_{1 \leq k \leq n} \|(\xi \wedge \eta)^k - r\mathbf{e}^k - \xi^k\|_\infty \geq r(1-\kappa) - \kappa \max_{1 \leq k \leq n} \|\xi^k - (\xi \wedge \eta)^k\|_\infty, \quad r > 0.$$

This yields $h^i(r) > 0$ for large enough $r > 0$ due to $\kappa \in (0, 1)$. The continuity of D implies that there exists a constant $v > 0$ such that

$$h^i(v) = v + D^i(\xi \wedge \eta - v\mathbf{e}) - D^i(\xi) = 0. \tag{4.8}$$

Let $\zeta = \xi \wedge \eta - v\mathbf{e}$, then it is clear that $\zeta \leq \xi \wedge \eta$. Moreover, it follows from (4.8) that

$$\zeta^i(0) - D^i(\zeta) = (\xi \wedge \eta)^i(0) - v - D^i(\xi \wedge \eta - v\mathbf{e}) = \xi^i(0) - D^i(\xi).$$

(2) Fix $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$. Let two $\mathscr{C}$-valued random variables (Γ_1, Γ_2) on $(\mathscr{C}^2, \mathscr{B}(\mathscr{C}^2), \mu \times \nu)$ be defined as $\Gamma_k(\xi_1, \xi_2) = \xi_k, k = 1, 2$. Then it is clear that $\mathscr{L}_{\Gamma_1}|\mu \times \nu = \mu$ and $\mathscr{L}_{\Gamma_2}|\mu \times \nu = \nu$. Based on this, we can construct a $\mathscr{C}$-valued random variable $\tilde{\Gamma}$ on $(\mathscr{C}^2, \mathscr{B}(\mathscr{C}^2), \mu \times \nu)$ such that $\tilde{\Gamma} \leq \Gamma_k$ and $\tilde{\Gamma}(0) - D(\tilde{\Gamma}) \leq \Gamma_k(0) - D(\Gamma_k)$, $k = 1, 2$. Let $\tilde{\mu} = \mathscr{L}_{\tilde{\Gamma}}|\mu \times \nu$. Then we have $\tilde{\mu} \leq_D \mu$ and $\tilde{\mu} \leq_D \nu$.

In fact, for any $(\xi_1, \xi_2) \in \mathscr{C}^2$, let

$$\alpha = -\left| \min_{s \in [-r_0, 0]} (\xi_1 \wedge \xi_2)(s) \right|,$$

and

$$\alpha_i = \frac{-|(\xi^i(0) - D^i(\xi)) \wedge (\eta^i(0) - D^i(\eta))|}{1 - \kappa}, \quad 1 \leq i \leq n.$$

Take $\tilde{\alpha} = \alpha \wedge \min_i \alpha_i$ and $\tilde{\Gamma}(\xi_1, \xi_2) = \tilde{\alpha}\mathbf{e}, (\xi_1, \xi_2) \in \mathscr{C}^2$. Then we have $\tilde{\alpha}\mathbf{e} \leq \alpha\mathbf{e} \leq \xi_1 \wedge \xi_2$. Moreover, for any $1 \leq i \leq n$, **(A1)** and **(A5)** yield

$$\tilde{\Gamma}^i(0) - D^i(\tilde{\Gamma}) = \tilde{\alpha} - D^i(\tilde{\alpha}\mathbf{e}) \leq \tilde{\alpha} - \kappa\tilde{\alpha} \leq \alpha_i - \kappa\alpha_i \leq (\xi^i(0) - D^i(\xi)) \wedge (\eta^i(0) - D^i(\eta)).$$

Thus, the proof is finished. $\square$

Theorem 4.8. *Assume* **(A1)**-**(A5)** *and that* (2.1)-(2.2) *is D-order-preserving for any complete filtered probability space $(\Omega, \{\mathscr{F}_t\}_{t \geq 0}, \mathbb{P})$ and m-dimensional Brownian motion $W(t)$ thereon. Then for any $1 \leq i \leq n$, $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$ with $\mu \leq_D \nu$, and $\xi, \eta \in \mathscr{C}$ with $\xi \leq_D \eta$ and $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$, the following assertions hold:*

(i') $b_i(t,\xi,\mu) \le \bar{b}_i(t,\eta,\nu)$ *if* b_i *and* $\bar{b}_i$ *are continuous at points* (t,ξ,μ) *and* (t,η,ν) *respectively.*

(ii') *For any* $1 \le j \le m$, $\sigma_{ij}(t,\xi,\mu) = \bar{\sigma}_{ij}(t,\eta,\nu)$ *if* σ_{ij} *and* $\bar{\sigma}_{ij}$ *are continuous at points* (t,ξ,μ) *and* (t,η,ν) *respectively.*

Consequently, when $b, \bar{b}$ *and* $\sigma, \bar{\sigma}$ *are continuous on* $[0,\infty) \times \mathscr{C} \times \mathscr{P}_2(\mathscr{C})$, *conditions* (i) *and* (ii) *hold.*

We first observe that when $b, \bar{b}$ are continuous on $[0,\infty) \times \mathscr{C} \times \mathscr{P}_2(\mathscr{C})$, (i') implies (i). Next, we prove when $\sigma, \bar{\sigma}$ are continuous on $[0,\infty) \times \mathscr{C} \times \mathscr{P}_2(\mathscr{C})$, (ii') implies (ii).

Firstly, taking $\xi = \eta$ and $\mu = \nu$, by the continuity of σ and $\bar{\sigma}$, (ii') implies $\sigma = \bar{\sigma}$.

Let ζ and $\tilde{\mu}$ be associated to ξ, η and μ, ν as in Lemma 4.7. Applying (ii') twice we obtain

$$\sigma_{ij}(t,\xi,\mu) = \sigma_{ij}(t,\zeta,\tilde{\mu}) = \sigma_{ij}(t,\eta,\nu).$$

This together with $\sigma = \bar{\sigma}$ implies (ii).

Now, let $t_0 \ge 0$, $1 \le i \le n$, $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$ with $\mu \le_D \nu$, and $\xi, \eta \in \mathscr{C}$ with $\xi \le \eta$ and $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$. To prove (i') and (ii') for $t = t_0$, we construct a family of complete filtered probability spaces $(\Omega, \{\mathscr{F}_t^\varepsilon\}_{t\ge 0}, \mathbb{P}^\varepsilon)_{\varepsilon\in[0,1)}$, m-dimensional Brownian motion $W(t)$, and initial random variables $X_{t_0} \le \bar{X}_{t_0}$ as follows.

Firstly, since $\mu \le_D \nu$, by Remark 4.1, we may take $\pi_0 \in \mathbf{C}(\mu,\nu)$ such that

$$\pi_0(\{(\xi_1,\xi_2) \in \mathscr{C}^2 : \xi_1 \le_D \xi_2\}) = 1. \tag{4.9}$$

For any $\varepsilon \in [0,1)$, let

$$\pi_\varepsilon = (1-\varepsilon)\pi_0 + \varepsilon \delta_{(\xi,\eta)}, \tag{4.10}$$

where $\delta_{(\xi,\eta)}$ is the Dirac measure at point (ξ,η). Let $\mathbb{P}_0$ be the standard Wiener measure on $\Omega_0 := C([0,\infty) \to \mathbb{R}^m)$, and let $\mathscr{F}_{0,t}$ be the completion of $\sigma(\omega_0 \mapsto \omega_0(s) : s \le t)$ with respect to the Wiener measure. Then the coordinate process $\{W_0(t)\}(\omega_0) := \omega_0(t)$, $\omega_0 \in \Omega_0, t \ge 0$ is an m-dimensional Brownian motion on the filtered probability space $(\Omega_0, \{\mathscr{F}_{0,t}\}_{t\ge0}, \mathbb{P}_0)$.

Next, for any $\varepsilon \in [0,1)$, let $\Omega = \Omega_0 \times \mathscr{C}^2$, $\mathbb{P}^\varepsilon = \mathbb{P}_0 \times \pi_\varepsilon$ and $\mathscr{F}_t^\varepsilon$ be the completion of $\mathscr{F}_{0,t} \times \mathscr{B}(\mathscr{C}^2)$ under the probability measure $\mathbb{P}^\varepsilon$. Then the process

$$\{W(t)\}(\omega) := \{W_0(t)\}(\omega_0) = \omega_0(t), \quad t \ge 0, \omega = (\omega_0; \xi_1, \xi_2) \in \Omega = \Omega_0 \times \mathscr{C}^2$$

is the m-dimensional Brownian motion on the filtered probability space $(\Omega, \{\mathscr{F}_t^\varepsilon\}_{t\ge0}, \mathbb{P}^\varepsilon)$.

Finally, let

$$X_{t_0}(\omega) := \xi_1, \quad \bar{X}_{t_0}(\omega) := \xi_2, \quad \omega = (\omega_0; \xi_1, \xi_2) \in \Omega = \Omega_0 \times \mathscr{C}^2.$$

They are $\mathscr{F}_{t_0}^\varepsilon$-measurable random variables with

$$\mathscr{L}_{X_{t_0}}|_{\mathbb{P}^\varepsilon} = \mu_\varepsilon := \pi_\varepsilon(\cdot \times \mathscr{C}), \quad \mathscr{L}_{\bar{X}_{t_0}}|_{\mathbb{P}^\varepsilon} = \nu_\varepsilon := \pi_\varepsilon(\mathscr{C} \times \cdot). \tag{4.11}$$

By $\xi \le \eta$ with $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$ and (4.9), (4.10), we have

$$\mathbb{P}^\varepsilon(X_{t_0} \le_D \bar{X}_{t_0}) = \pi_\varepsilon\big(\{(\xi_1, \xi_2) \in \mathscr{C}^2 : \ \xi_1 \le_D \xi_2\}\big) = 1, \quad \varepsilon \in [0, 1). \tag{4.12}$$

So, letting $(X_t, \bar{X}_t)_{t \ge t_0}$ be the segment process of the solution to (2.1) and (2.2) with initial value $(X_{t_0}, \bar{X}_{t_0})$, the D-order preservation implies

$$\mathbb{P}^\varepsilon\big(X_t \le_D \bar{X}_t, \ t \ge t_0\big) = 1, \quad \varepsilon \in [0, 1). \tag{4.13}$$

Let $\mathbb{E}^\varepsilon$ be the expectation for $\mathbb{P}^\varepsilon$. With the above preparations, we are able to prove (i') and (ii') as follows.

Proof of (i'). Let $b_i, \bar{b}_i$ be continuous at points (t_0, ξ, μ) and (t_0, η, ν) respectively. We intend to prove $b_i(t_0, \xi, \mu) \le \bar{b}_i(t_0, \eta, \nu)$. Otherwise, there exists a constant $c_0 > 0$ such that

$$b_i(t_0, \xi, \mu) \ge c_0 + \bar{b}_i(t_0, \eta, \nu). \tag{4.14}$$

Let $\mu_\varepsilon, \nu_\varepsilon$ be as in (4.11). It is clear that as $\varepsilon \to 0$, $\mu_\varepsilon \to \mu, \ \nu_\varepsilon \to \nu$ weakly. Moreover, for large enough n, we have

$$\begin{aligned}
&\sup_{\varepsilon \in [0,1)} \int_{\mathscr{C}} \|z\|_\infty^2 1_{\|z\|_\infty^2 \ge n} \mu_\varepsilon(\mathrm{d}z) + \sup_{\varepsilon \in [0,1)} \int_{\mathscr{C}} \|z\|_\infty^2 1_{\|z\|_\infty^2 \ge n} \nu_\varepsilon(\mathrm{d}z) \\
&= \int_{\mathscr{C}} \|z\|_\infty^2 1_{\|z\|_\infty^2 \ge n} \mu(\mathrm{d}z) + \int_{\mathscr{C}} \|z\|_\infty^2 1_{\|z\|_\infty^2 \ge n} \nu_(\mathrm{d}z).
\end{aligned}$$

Noting $\mu, \nu \in \mathscr{P}_2(\mathscr{C})$, the dominated convergence theorem implies that $\{\mu_\varepsilon, \nu_\varepsilon\}_{\varepsilon \in [0,1)}$ are L^2-uniformly integrable in $\mathscr{P}_2(\mathscr{C})$. Consequently, it follows from [3, Theorem 5.5] that

$$\lim_{\varepsilon \downarrow 0} \{\mathbb{W}_2(\mu_\varepsilon, \mu) + \mathbb{W}_2(\nu_\varepsilon, \nu)\} = 0.$$

Combining this with (4.14) and the continuity of b and $\bar{b}$, there exists $\varepsilon \in (0, 1)$ such that

$$b_i(t_0, \xi, \mu_\varepsilon) \ge \frac{1}{2} c_0 + \bar{b}_i(t_0, \eta, \nu_\varepsilon) > \bar{b}_i(t_0, \eta, \nu_\varepsilon). \tag{4.15}$$

Now, consider the event

$$A := \{X_{t_0} = \xi, \bar{X}_{t_0} = \eta\} \in \mathscr{F}_{t_0}^\varepsilon. \tag{4.16}$$

Then we get

$$\mathbb{P}^\varepsilon(A) \ge \varepsilon \delta_{(\xi, \eta)}(\{(\xi, \eta)\}) = \varepsilon > 0. \tag{4.17}$$

By (2.1), (2.2) and (4.13), for any $s \ge 0$, it holds $\mathbb{P}^\varepsilon$-a.s.

$$\begin{aligned}
0 &\ge X^i(t_0 + s) - D^i(X_{t_0+s}) - (\bar{X}^i(t_0 + s) - D^i(\bar{X}_{t_0+s})) \\
&= X^i(t_0) - D^i(X_{t_0}) - (\bar{X}^i(t_0) - D^i(\bar{X}_{t_0})) \\
&\quad + \int_{t_0}^{t_0+s} b_i(r, X_r, \mathscr{L}_{X_r}) \, \mathrm{d}r - \int_{t_0}^{t_0+s} \bar{b}_i(r, \bar{X}_r, \mathscr{L}_{\bar{X}_r}) \, \mathrm{d}r \\
&\quad + \sum_{j=1}^m \int_{t_0}^{t_0+s} \sigma_{ij}(r, X_r, \mathscr{L}_{X_r}) \, \mathrm{d}W_j(r) - \int_{t_0}^{t_0+s} \bar{\sigma}_{ij}(r, \bar{X}_r, \mathscr{L}_{\bar{X}_r}) \, \mathrm{d}W_j(r).
\end{aligned} \tag{4.18}$$

By **(A2)** and the non-explosion of the solution to (2.1) and (2.2), taking conditional expectation in (4.18) with respect to $\mathscr{F}^{\varepsilon}_{t_0}$, we obtain $\mathbb{P}^{\varepsilon}$-a.s.

$$\mathbb{E}^{\varepsilon}\bigg(\int_{t_0}^{t_0+s} b_i(r,X_r,\mathscr{L}_{X_r})\,\mathrm{d}r\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg)\le \mathbb{E}^{\varepsilon}\bigg(\int_{t_0}^{t_0+s} \bar{b}_i(r,\bar{X}_r,\mathscr{L}_{\bar{X}_r})\,\mathrm{d}r\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg),\quad s>0.$$

This implies

$$\mathbb{E}^{\varepsilon}\bigg(\frac{1}{s}\int_{t_0}^{t_0+s} b_i(r,X_r,\mathscr{L}_{X_r})\,\mathrm{d}r\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg)\le \mathbb{E}^{\varepsilon}\bigg(\frac{1}{s}\int_{t_0}^{t_0+s} \bar{b}_i(r,\bar{X}_r,\mathscr{L}_{\bar{X}_r})\,\mathrm{d}r\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg),\quad s>0.$$

Combining this with the fact that b_i and $\bar{b}_i$ are continuous at points (t_0,ξ,μ) and (t_0,η,ν) respectively, and using **(A2)**, the non-explosion and continuity of the solution to (2.1) and (2.2), taking $s\downarrow 0$ we derive $\mathbb{P}^{\varepsilon}$-a.s.

$$\mathbb{E}^{\varepsilon}\big(b_i(t_0,X_{t_0},\mathscr{L}_{X_{t_0}})\big|\mathscr{F}^{\varepsilon}_{t_0}\big)\le \mathbb{E}^{\varepsilon}\big(\bar{b}_i(t_0,\bar{X}_{t_0},\mathscr{L}_{\bar{X}_{t_0}})\big|\mathscr{F}^{\varepsilon}_{t_0}\big).$$

This together with (4.16) and (4.11) leads to $\mathbb{P}^{\varepsilon}$-a.s.

$$b_i(t_0,\xi,\mu_\varepsilon)1_A\le \bar{b}_i(t_0,\eta,\nu_\varepsilon)1_A,$$

which is impossible according to (4.15) and (4.17). Therefore, $b_i(t_0,\xi,\mu)\le\bar{b}_i(t_0,\eta,\nu)$ has to be true. $\square$

Proof of (*ii′*). Let σ_{ij} and $\bar{\sigma}_{ij}$ be continuous at points (t_0,ξ,μ) and (t_0,η,ν) respectively. If $\sigma_{ij}(t_0,\xi,\mu)\ne\bar{\sigma}_{ij}(t_0,\eta,\nu)$, by **(A3)**, there exist constants $c_1>0$ and $\varepsilon\in(0,1)$ such that

$$|\sigma_{ij}(t_0,\xi,\mu_\varepsilon)-\bar{\sigma}_{ij}(t_0,\eta,\nu_\varepsilon)|^2\ge 2c_1>0. \tag{4.19}$$

For any $n,l\ge 1$, let

$$\begin{aligned}
&\tau=\inf\big\{t\ge t_0: |\sigma_{ij}(t,X_t,\mathscr{L}_{X_t}|_{\mathbb{P}^{\varepsilon}})-\bar{\sigma}_{ij}(t,\bar{X}_t,\mathscr{L}_{\bar{X}_t}|_{\mathbb{P}^{\varepsilon}})|^2\le c_1\big\},\\
&\tau_n=\inf\big\{t\ge t_0: X^i(t)-D^i(X_t)-(\bar{X}^i(t)-D^i(\bar{X}_t))\le -\frac{1}{n}\big\},\\
&\tau_{n,l}=\tau\wedge\tau_n\wedge\inf\big\{t\ge t_0: |b_i(t,X_t,\mathscr{L}_{X_t}|_{\mathbb{P}^{\varepsilon}})-\bar{b}_i(t,\bar{X}_t,\mathscr{L}_{\bar{X}_t}|_{\mathbb{P}^{\varepsilon}})|\ge l\big\}.
\end{aligned}$$

Let $g_n(s)=\mathrm{e}^{ns}-1$. Then $g_n\in C_b^2((-\infty,0])$. By the D-order-preservation we have $X^i_t\le_D \bar{X}^i_t,\ t\ge t_0$. So, letting $Z^i(s)=(X^i-\bar{X}^i)(s)-D^i(X_s)+D^i(\bar{X}_s)$ and applying Itô's formula, we obtain $\mathbb{P}^{\varepsilon}$-a.s.

$$\begin{aligned}
0&\ge \mathbb{E}^{\varepsilon}\big(g_n(Z^i(t\wedge\tau_{n,l}))|\mathscr{F}^{\varepsilon}_{t_0}\big)-g_n(Z^i(t_0))\\
&=\mathbb{E}^{\varepsilon}\bigg(\sum_{j=1}^m\int_{t_0}^{t\wedge\tau_{n,l}} g_n'(Z^i(s))\big(\sigma_{ij}(s,X_s,\mathscr{L}_{X_s}|_{\mathbb{P}^{\varepsilon}})-\bar{\sigma}_{ij}(s,\bar{X}_s,\mathscr{L}_{\bar{X}_s}|_{\mathbb{P}^{\varepsilon}})\big)\mathrm{d}W_j(s)\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg)\\
&+\mathbb{E}^{\varepsilon}\bigg(\int_{t_0}^{t\wedge\tau_{n,l}}\bigg\{g_n'(Z^i(s))\big(b_i(s,X_s,\mathscr{L}_{X_s}|_{\mathbb{P}^{\varepsilon}})-\bar{b}_i(s,\bar{X}_s,\mathscr{L}_{\bar{X}_s}|_{\mathbb{P}^{\varepsilon}})\big)\\
&+\frac{g_n''(Z^i(s))}{2}\sum_{j=1}^m\big|\sigma_{ij}(s,X_s,\mathscr{L}_{X_s}|_{\mathbb{P}^{\varepsilon}})-\bar{\sigma}_{ij}(s,\bar{X}_s,\mathscr{L}_{\bar{X}_s}|_{\mathbb{P}^{\varepsilon}})\big|^2\bigg\}\mathrm{d}s\bigg|\mathscr{F}^{\varepsilon}_{t_0}\bigg)\\
&\ge g_n(Z^i(t_0))+\Big(\frac{n^2c_1}{2\mathrm{e}}-nl\Big)\mathbb{E}^{\varepsilon}(t\wedge\tau_{n,l}-t_0|\mathscr{F}^{\varepsilon}_{t_0}),\quad n,l\ge 1.
\end{aligned}$$

By (4.16) and $\xi^i(0) - D^i(\xi) = \eta^i(0) - D^i(\eta)$, this implies

$$1_A\Big(\frac{n^2c_1}{2\mathrm{e}} - nl\Big)\mathbb{E}^\varepsilon(t\wedge\tau_{n,l} - t_0|\mathscr{F}^\varepsilon_{t_0}) \le -1_A g_n(Z^i(t_0)) = -1_A g_n(0) = 0$$

for all $n, l \ge 1$ and $t > t_0$.

Taking $l \ge 2|b_i(t_0,\xi,\mu_\varepsilon) - \bar{b}_i(t_0,\eta,\nu_\varepsilon)|$ and $n > \frac{2el}{c_1}$, we obtain

$$1_A\mathbb{E}^\varepsilon(t\wedge\tau_{n,l} - t_0|\mathscr{F}^\varepsilon_{t_0}) = 0, \quad t > t_0. \tag{4.20}$$

But by **(A3)**, (4.19) and the continuity of the solution, on the set A we have

$$\tau_{n,l} > t_0.$$

So, (4.20) implies $\mathbb{P}^\varepsilon(A) = 0$, which contradicts (4.17). Hence, it holds

$$\sigma_{ij}(t_0,\xi,\mu) = \bar{\sigma}_{ij}(t_0,\eta,\nu).$$

□

Acknowledgement. The authors would like to thank the associated editor and referees for their helpful comments and suggestions.

References

[1] X. Bai, J. Jiang, *Comparison theorems for neutral stochastic functional differential equations,* J. Differential Equ. 260(2016), 7250–7277.

[2] J. Bao, C. Yuan, *Comparison theorem for stochastic differential delay equations with jumps,* Acta Appl. Math. 116(2011), 119–132.

[3] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I,* Springer, 2018.

[4] M.-F. Chen, F.-Y. Wang, *On order-preservation and positive correlations for multidimensional diffusion processes,* Prob. Theory. Relat. Fields 95(1993), 421–428.

[5] L. Gal'cuk, M. Davis, *A note on a comparison theorem for equations with different diffusions,* Stochastics 6(1982), 147–149.

[6] X. Huang, M. Röckner, F.-Y. Wang, *Nonlinear Fokker–Planck equations for probability measures on path space and path-distribution dependent SDEs,* Discrete Contin. Dyn. Syst. 39(2019), 3017–3035.

[7] X. Huang, F.-Y. Wang, *Order-preservation for multidimensional stochastic functional differential equations with jumps,* J. Evol. Equat. 14(2014), 445–460.

[8] N. Ikeda, S. Watanabe, *A comparison theorem for solutions of stochastic differential equations and its applications,* Osaka J. Math. 14(1977), 619–633.

[9] T. Kamae, U.Krengel, G. L. O'Brien, *Stochastic inequalities on partially ordered spaces,* Ann. Probab. 5(1977), 899–912.

[10] X. Mao, *A note on comparison theorems for stochastic differential equations with respect to semimartingales,* Stochastics 37(1991), 49–59.

[11] G. L. O'Brien, *A new comparison theorem for solution of stochastic differential equations,* Stochastics 3(1980), 245–249.

[12] S. Peng, Z. Yang, *Anticipated backward stochastic differential equations,* Ann. Probab. 37(2009), 877–902.

[13] S. Peng, X. Zhu, *Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations,* Stoch. Proc. Appl. 116(2006), 370–380.

[14] A.V. Skorohod, *Studies in the theory of random process,* Addison-Wesley, 1965.

[15] J.-M. Wang, *Stochastic comparison for Lévy-type processes,* J. Theor. Probab. 26(2013), 997–1019.

[16] F.-Y. Wang, *The stochastic order and critical phenomena for superprocesses,* Infin. Dimens. Anal. Quantum Probab. Relat. Top. 9(2006), 107–128.

[17] F.-Y. Wang, *Distribution-dependent SDEs for Landau type equations,* Stoch. Proc. Appl. 128(2018), 595-621.

[18] T. Yamada, *On comparison theorem for solutions of stochastic differential equations and its applications,* J. Math. Kyoto Univ. 13(1973), 495–512.

[19] Z. Yang, X. Mao, C. Yuan, *Comparison theorem of one-dimensional stochastic hybrid systems,* Systems Control Lett. 57(2008), 56–63.

[20] X. Zhu, *On the comparison theorem for multi-dimensional stochastic differential equations with jumps (in Chinese),* Sci. Sin. Math. 42(2012), 303–311.