J Dyn Diff Equat (2017) 29:1–24
DOI 10.1007/s10884-014-9406-x

Comparison Theorem for Stochastic Functional Differential Equations and Applications

Xiaoming Bai · Jifa Jiang

Received: 16 September 2014 / Accepted: 23 September 2014 / Published online: 15 October 2014

Abstract The comparison theorem is proved for stochastic functional differential equations whose drift term satisfies the quasimonotone condition and diffusion term is independent of delay. Application is given to stochastic neutral networks with delays.

Keywords Comparison theorem · Stochastic functional differential equations · Quasimonotone condition · Stochastic neutral network

1 Introduction

The study of comparison results for stochastic differential equations started with Skorohod [20]. After that, comparison theorems for solutions of two one-dimensional It$\hat{o}$'s stochastic ordinary differential equations with the same diffusion coefficients were intensively studied, and many applications, including stochastic optimal control and test for explosions were presented, see [9–11,19,22,23] and references therein. It is worth to mention that Peng and Zhu [18] have given a necessary and sufficient condition of the comparison theorem in this case. Yan [24] gave some conclusion about equations driven by general continuous local martingale, continuous increasing process and general increasing process but still based on the same diffusion coefficients. The first comparison theorem for two multi-dimensional It$\hat{o}$'s stochastic ordinary differential equations with the same diffusion coefficients was proved by Geiβ and Manthey [8]. An additional condition called quasimonotone must be imposed on. It is mainly based on this comparison theorem that Cheushov [5] established stochastic

X. Bai
Mathematics and Science College, Hefei University of Technology, Hefei 230009, Anhui, People's Republic of China
e-mail: bxm_96@sina.com

J. Jiang (✉)
Mathematics and Science College, Shanghai Normal University, Shanghai 200234, People's Republic of China
e-mail: jiangjf@shnu.edu.cn

monotone dynamical systems and investigated the structure of random attractor and long-term behavior for some special quasimonotone stochastic ordinary differential equations.

In contrast to the theory of stochastic ordinary differential equations, the theory of stochastic partial differential equations lacks an important tool, namely, we do not have a well-applicable It$\hat{o}$'s formula. One technique to handle the difficulties arising from the missing It$\hat{o}$'s formula is comparison technique. Pardoux and his collaborators [3,7] began with the study of comparison theorem for parabolic stochastic partial differential equations. Kotelenez [14] was the first one to consider comparison theorem for a wide class of parabolic stochastic partial differential equations with Lipschitz drift and diffusion coefficient. Manthey and Zausinger [15] extended Kotelenez's result to the case that drift coefficient allows to be polynomial growth using a different method. Applying the method in [15], Assing [2] generalized Manthey and Zausinger's result to systems of parabolic stochastic partial differential equations by a method of approximation.

As far as we know, there is very few comparison theorem for stochastic functional differential equations. The only one we have found was presented in [25] for scalar stochastic functional differential equation. They also claimed that "so far, there is no result for comparison theorem on stochastic differential delay equation" in the introduction of their work. If one tries to investigate monotone random dynamical systems generated by stochastic functional differential equations, he has to prove a comparison theorem as in the paper [8]. Besides, such a comparison theorem has its own independent interest. Motivated by these, we will prove a comparison theorem for stochastic functional differential equations under quasimonotone condition and other regular conditions. To achieve this, we need to prove a global existence and uniqueness theorem under weaker conditions. This comparison theorem lays the foundation to investigate deep long-term dynamical behavior of quasimonotone stochastic functional differential equations, which has been thoroughly exploded in deterministic functional differential equations in [21].

2 Preliminaries

Let $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space satisfying the usual conditions. Fix an arbitrary $\tau > 0$ and two positive integers d and r.

Consider the following two systems of stochastic functional differential equations(SFDEs) in the sense of It$\hat{o}$:

$$\begin{cases} dx_i(t) = f_i(t, x(t), x_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, x(t))dW_t^j, \quad i = 1, 2, \ldots, d, \\ x(\theta) = \phi(\theta),\ \theta \in J \triangleq [-\tau, 0],\ \phi \in C(J, \mathbb{R}^d), \end{cases} \tag{2.1}$$

and

$$\begin{cases} d\widehat{x}_i(t)7 = \widehat{f}_i(t, \widehat{x}(t), \widehat{x}_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, \widehat{x}(t))dW_t^j, \quad i = 1, 2, \ldots, d, \\ \widehat{x}(\theta) = \psi(\theta),\ \theta \in J,\ \psi \in C(J, \mathbb{R}^d) \end{cases} \tag{2.2}$$

where $x(t) = (x_1(t), \ldots, x_i(t), \ldots, x_d(t)), \widehat{x}(t) = (\widehat{x}_1(t), \ldots, \widehat{x}_i(t), \ldots, \widehat{x}_d(t)), x_t(\theta) = x(t+\theta), \widehat{x}_t(\theta) = \widehat{x}(t+\theta), \theta \in J = [-\tau, 0], W_t(\omega) = (W_t^1(\omega), \ldots, W_t^r(\omega))$ is an $r-$dimensional $\{\mathfrak{F}_t\}_{t\geq 0}$-adapted Wiener process with values in $\mathbb{R}^r, \mathcal{C} \triangleq C(J, \mathbb{R}^d)$ is the Banach space of all continuous functions $\phi : J \to \mathbb{R}^d$ with the sup-norm $\|\phi\| = \sup\{|\phi(s)|, s \in J\}$ and $|\cdot|$ denotes the Euclidean norm.

We make the following hypotheses:

(H1) The *drift terms* $f = (f_1, \ldots, f_d), \widehat{f} = (\widehat{f_1}, \ldots, \widehat{f_d}) : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}^d$ are continuous and the inequality $f_i(t, \phi(0), \phi) < \widehat{f_i}(t, \psi(0), \psi)$ is fulfilled whenever $\phi \leq_{\mathcal{C}} \psi$ and $\phi_i(0) = \psi_i(0)$ holds for some i and $t \geq 0, \phi, \psi \in \mathcal{C}$, the notation $\leq_{\mathcal{C}}$ is given in the next section.

(H2) The drift term satisfies *global Lipschitz condition*, that is, there exists a constant $L > 0$ such that for each $i = 1, 2, \ldots, d$,

$$|f_i(t, x, \phi) - f_i(t, x', \psi)|^2 \leq L\big(|x - x'|^2 + \|\phi - \psi\|^2\big)$$

for all $t \geq 0, x, x' \in \mathbb{R}^d$ and $\phi, \psi \in \mathcal{C}$.

(H3) The *diffusion term* $\sigma(t, x) = (\sigma_{ij}(t, x)) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}, i = 1, 2, \ldots, d, j = 1, 2, \ldots, r$ is continuous and there exists a nondecreasing continuous concave function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ with $\rho(0) = 0, \rho(x) > 0$ for $x > 0$, and $\int_{0+} \frac{dx}{\rho(x)} = \infty$ such that for each $i = 1, 2, \ldots, d$,

$$\sum_{j=1}^{r} |\sigma_{ij}(t, x) - \sigma_{ij}(t, x')|^2 \leq \rho\big(|x_i - x'_i|^2\big),$$

for all $t \geq 0, x, x' \in \mathbb{R}^d$.

If $\rho(|x_i - x'_i|^2)$ is replaced by $\rho(|x - x'|^2)$, then we use (H3*) to denote the corresponding hypothesis.

(H4) The drift and diffusion terms have *linear growth*, that is, there is a constant $\gamma > 0$ such that $f = (f_1, \ldots, f_d) : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}^d$, and $\sigma(t, x) = (\sigma_{ij}(t, x)) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}, i = 1, 2, \ldots, d, j = 1, 2, \ldots, r$ satisfy

$$|f(t, x, \phi)|^2 \leq \gamma\big(1 + |x|^2 + \|\phi\|^2\big) \text{ and}$$
$$|\sigma(t, x)|^2 \leq \gamma\big(1 + |x|^2\big)$$

for all $t \geq 0, x \in \mathbb{R}^d, \phi \in \mathcal{C}$.

Theorem 1 *Assume that* (H2), (H3*) *and* (H4) *hold for system* (2.1). *Then the system* (2.1) *has a strong solution $x(t, \phi)$ for all $t > 0$, and the strong uniqueness holds. Furthermore $x_t(\phi)$ is a $\mathcal{C}$-valued process adapted to $\{\mathfrak{F}_t\}_{t\geq 0}$ with continuous sample paths.*

This proof is presented in appendix and extends Mao's technique ([16]) for backward stochastic differential equations to stochastic functional differential equations.

Now we review the general Gronwall inequality in [1], which is useful in the subsequent sections. Consider the following inequality

$$u(t) \leq a(t) + \int_0^t \lambda(t, s)\eta(u(s))ds, \quad 0 \leq t \leq t_1, \tag{2.3}$$

which satisfies the following properties:

(S1) η is continuous and nondecreasing function on $[0, \infty)$ and is positive on $(0, \infty)$,

(S2) $a(t)$ is continuously differentiable in t and nonnegative on $[0, t_1]$, where $t_1 > 0$ is a constant,

(S3) $\lambda(t, s)$ is continuous and nonnegative function on $[0, t_1] \times [0, t_1]$.

Theorem 2 *Suppose that* (S1)–(S3) *hold and* $u(t)$ *is a continuous and nonnegative function on* $[0, t_1]$ *satisfying* (2.3). *Then*

$$u(t) \leq W^{-1}\Big[W(r(t)) + \int_0^t \max_{0\leq\chi\leq t} \lambda(\chi, s)ds\Big], \quad 0 \leq t \leq t_c, \tag{2.4}$$

where $W(u) \triangleq \int_{\widehat{u}}^u \frac{dz}{\eta(z)}$, $\widehat{u} > 0$ *is constant and* $r(t)$ *is determined by*

$$r(t) \triangleq a(0) + \int_0^t |a'(s)|ds,$$

$t_c \leq t_1$ *is the largest number such that*

$$W(r(t_c)) + \int_0^{t_c} \max_{0\leq\chi\leq t_c} \lambda(\chi, s)ds \leq \int_{\widehat{u}}^{\infty} \frac{dz}{\eta(z)}.$$

The proof can be found in [1].

We apply Theorem 2 to study such an inequality :

$$u(t) \leq \int_0^t \lambda_1\rho(u(s))ds + \int_0^t \lambda_2 u(s)ds, \ t \geq 0. \tag{2.5}$$

Corollary 3 *If* ρ *satisfies the hypothesis* (H3), λ_1 *and* λ_2 *are two positive constants and* $u(t)$ *is a continuous and nonnegative function on* $[0, \infty)$ *satisfying* (2.5), *then* $u(t) = 0$, $t \geq 0$.

Proof Let $t_1 > 0$ be arbitrary and consider the following inequality:

$$u(t) \leq \int_0^t \lambda_1\rho(u(s))ds + \int_0^t \lambda_2 u(s)ds, \quad 0 \leq t \leq t_1.$$

Let $\lambda = \max\{\lambda_1, \lambda_2\}$ and $\varrho(u) = u + \rho(u)$. Then we have

$$u(t) \leq \int_0^t \lambda\varrho(u(s))ds, \quad 0 \leq t \leq t_1.$$

By Theorem 2, it is easy to see that $W(u) = \int_{\widehat{u}}^u \frac{dz}{\varrho(z)}$, $r(t) \equiv 0$,

$$u(t) \leq W^{-1}\Big[W(r(t)) + \int_0^t \lambda ds\Big], \quad 0 \leq t \leq t_c, \tag{2.6}$$

where t_c is the largest number such that

$$W(r(t_c)) + \int_0^{t_c} \lambda ds \leq \int_{\widehat{u}}^{\infty} \frac{dz}{\varrho(z)}. \tag{2.7}$$

Since ρ is a concave function with $\rho(0) = 0$, we have

$$\rho(z) \geq \rho(1)z, \quad \text{for} \quad 0 \leq z \leq 1.$$

So

$$\int_{0+} \frac{dz}{\varrho(z)} = \int_{0+} \frac{dz}{z + \rho(z)} \geq \frac{\rho(1)}{\rho(1) + 1} \int_{0+} \frac{dz}{\rho(z)} = \infty, \tag{2.8}$$

which implies that (2.7) is true for $t_c = t_1$. Therefore, $u(t) = 0$ for $0 \leq t \leq t_1$ by (2.6). Since t_1 is arbitrarily chosen, we have $u(t) = 0, \ t \geq 0$. The proof is complete. □

Before finishing this section, we provide a criterion for stopping times.

Proposition 4 *Suppose that ς is an $\mathfrak{F}_t-$stopping time and Q is progressively measurable. Then*

$$D_Q(\omega) \triangleq \inf\{s \mid s > \varsigma(\omega), (s,\omega) \in Q\}$$

is an $\mathfrak{F}_t-$stopping time.

Proof Let

$$\widetilde{Q} \triangleq ((\varsigma, \infty)) \triangleq \{(s,\omega) \in \mathbb{R}_+ \times \Omega \mid \varsigma(\omega) < s\}.$$

Then we will prove that $\widetilde{Q}$ is a predictable set. It suffices to show that $I_{\widetilde{Q}}$ is an $\mathfrak{F}_t-$adapted left-continuous process. It is easy to check that the sample path $I_{\widetilde{Q}}(t,\omega)$ is left-continuous for a fixed ω. Since $I_{\widetilde{Q}}(t,\omega) = I_{\{\varsigma<t\}}(\omega)$ and the filtration $\{\mathfrak{F}_t\}_{t\geq 0}$ is right-continuous, $I_{\widetilde{Q}}$ is adapted to the filtration $\{\mathfrak{F}_t\}_{t\geq 0}$.

By the definition of $\widetilde{Q}$, it is easy to see that

$$D_Q(\omega) = \inf\{s \mid (s,\omega) \in Q \cap \widetilde{Q}\}.$$

The direct proof shows that for a fixed t,

$$\{\omega \in \Omega \mid D_Q(\omega) < t\} = \pi(Q \cap \widetilde{Q} \cap ((0,t) \times \Omega)) \tag{2.9}$$

where

$$\pi : \mathbb{R}_+ \times \Omega \to \Omega$$

is the projection mapping.

Since Q is progressively measurable and $\widetilde{Q}$ is predictable, $Q \cap \widetilde{Q}$ is progressively measurable, which implies that $Q \cap \widetilde{Q} \cap ([0,t] \times \Omega) \in \mathcal{B}([0,t]) \times \mathfrak{F}_t$. Thus,

$$Q \cap \widetilde{Q} \cap ((0,t) \times \Omega) = (Q \cap \widetilde{Q} \cap ([0,t] \times \Omega)) \bigcap ((0,t) \times \Omega) \in \mathcal{B}([0,t]) \times \mathfrak{F}_t.$$

It follows from (2.9) and projection theorem that $\{\omega \in \Omega \mid D_Q(\omega) < t\} \in \mathfrak{F}_t$. This completes the proof. □

Corollary 5 *Suppose that $F(t,\omega) : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ is a progressively measurable process and ς is an $\mathfrak{F}_t-$stopping time. If $E \in \mathcal{B}(\mathbb{R}^d)$, then*

$$D_E(\omega) \triangleq \inf\{s \mid s > \varsigma(\omega), F(s,\omega) \in E\}$$

is an $\mathfrak{F}_t-$stopping time.

Proof Let $Q = F^{-1}(E)$. Then Q is progressively measurable. The conclusion follows immediately from Proposition 4. □

3 Comparison Theorems for SFDEs

To obtain the comparison results of SFDEs, we need the partial orders in $\mathbb{R}^d$ and $\mathcal{C}$. The positive cone in $\mathbb{R}^d$, denoted by $\mathbb{R}^d_+$, is the set of all d tuples with nonnegative coordinates. It gives rise to a partial order on $\mathbb{R}^d$ in the following way:

$$\begin{aligned} x \leq y &\Longleftrightarrow x_i \leq y_i, \quad \text{for} \quad i = 1, \ldots, d, \\ x < y &\Longleftrightarrow x \leq y \quad \text{and } x_i < y_i, \quad \text{for some } i \in \{1, \ldots, d\}, \\ x \ll y &\Longleftrightarrow x_i < y_i, \quad \text{for } i = 1, \ldots, d. \end{aligned}$$

Let

$$\mathcal{C}_+ = \{\phi \in \mathcal{C} : \phi(s) \geq 0, \quad s \in J\}.$$

Then $\mathcal{C}_+$ is a positive cone of the Banach space $\mathcal{C}$. Hence the partial order on $\mathcal{C}$ is given as follows

$$\begin{aligned}
&\phi \leq_{\mathcal{C}} \psi \Longleftrightarrow \phi(s) \leq \psi(s), \quad s \in J,\\
&\phi <_{\mathcal{C}} \psi \Longleftrightarrow \phi \leq_{\mathcal{C}} \psi, \text{ and } \quad \phi \neq \psi,\\
&\phi \ll_{\mathcal{C}} \psi \Longleftrightarrow \phi(s) \ll \psi(s), \quad s \in J.
\end{aligned}$$

Now we present our first comparison result.

Theorem 6 *Suppose that the drift terms $f, \widehat{f}$ and the diffusion term $\sigma(t, x)$ satisfy the hypotheses* (H1)–(H4). *If $\phi, \psi \in \mathcal{C}$ satisfying $\phi \leq_{\mathcal{C}} \psi$, then $\mathbb{P}(\{x(t,\phi) \leq \widehat{x}(t,\psi), t \geq 0\}) = 1$ and hence $\mathbb{P}(\{x_t(\phi) \leq_{\mathcal{C}} \widehat{x}_t(\psi), t \geq 0\}) = 1$.*

Proof Let $\widehat{X}(t) = \widehat{x}(t,\psi)$ and $X(t) = x(t,\phi)$. Then we have

$$\begin{aligned}
d\widehat{X}_i(t) &= \widehat{f}_i\left(t, \widehat{X}(t), \widehat{X}_t\right)dt + \sum_{j=1}^{r} \sigma_{ij}\left(t, \widehat{X}(t)\right) dW_t^j, \quad i = 1, 2, \ldots, d,\\
\widehat{X}(t) &= \psi(t), \quad t \in [-\tau, 0]
\end{aligned}$$

and

$$\begin{aligned}
dX_i(t) &= f_i(t, X(t), X_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, X(t))dW_t^j, \quad i = 1, 2, \ldots, d,\\
X(t) &= \phi(t), \quad t \in [-\tau, 0].
\end{aligned}$$

By Theorem 1, $X(t)$ and $\widehat{X}(t)$ are $\mathfrak{F}_t-$adapted continuous processes.

Set $Y_i(t) = \widehat{X}_i(t) - X_i(t),\ i = 1, 2, \ldots, d$. Then we have

$$\begin{aligned}
dY_i(t) = &\Big[\widehat{f}_i(t, \widehat{X}(t), \widehat{X}_t) - f_i(t, X(t), X_t)\Big]dt\\
&+ \sum_{j=1}^{r}(\sigma_{ij}(t, \widehat{X}(t)) - \sigma_{ij}(t, X(t)))dW_t^j, \quad i = 1, 2, \ldots, d.
\end{aligned}$$

Now we introduce the following function which was first presented in [4]:

$$\varphi_\epsilon(y) = \begin{cases} y^2, & y \leq 0,\\ y^2 - \frac{y^3}{6\epsilon}, & 0 < y \leq 2\epsilon,\\ 2\epsilon y - \frac{4}{3}\epsilon^2, & y > 2\epsilon. \end{cases}$$

It is easy to see that $\varphi_\epsilon(y) \in C^2(\mathbb{R})$, $\varphi_\epsilon'(y) \to 2y^-$ uniformly with respect to y, $\varphi_\epsilon''(y) \to 2I_{(y\leq 0)}$ and $\varphi_\epsilon(y) \to |y^-|^2$ provided that $\epsilon \to 0$, where $y^- = y \wedge 0$.

Define the stopping times

$$\begin{aligned}
\varGamma_N &\triangleq \inf \left\{t > 0 : |X(t)| > N, |\widehat{X}(t)| > N\right\} \wedge N, \quad \text{forevery} \quad N > 0,\\
\varLambda_i &\triangleq \inf\left\{t > 0 : X_i(t) > \widehat{X}_i(t)\right\}, \quad i = 1, 2, \ldots, d, \qquad (3.1)
\end{aligned}$$

and

$$\Lambda = \Lambda_1 \wedge \cdots \wedge \Lambda_d.$$

Obviously $\Gamma_N \uparrow \infty$ as $N \uparrow \infty$, and either $\Lambda_i = +\infty$ or

$$X_i(\Lambda_i) = \widehat{X}_i(\Lambda_i) \text{ and } X_i(t) \le \widehat{X}_i(t), \quad \text{for} \quad 0 \le t \le \Lambda_i, \quad i = 1, \dots, d. \tag{3.2}$$

In order to verify the conclusion, we have to prove that

$$\mathbb{P}(\{\Lambda < \infty\}) = 0,$$

it suffices to show that $\mathbb{P}(\{\Lambda < \Gamma_N\}) = 0$ for every $N > 0$.

For $i = 1, 2, \dots, d$, let

$$\vartheta_i \triangleq \inf\Big\{t > \Lambda_i : Z_i(t,\omega) \triangleq f_i(t, X(t), X_t) - \widehat{f}_i(t, \widehat{X}_1(t), \dots, \widehat{X}_{i-1}(t), \widehat{X}_i(t) - Y_i^-(t), \widehat{X}_{i+1}(t), \dots, \widehat{X}_d(t), \widehat{X}_{1,t}, \dots, \widehat{X}_{i-1,t}, \widehat{X}_{i,t} - \widetilde{Y}_{i,t}^-, \widehat{X}_{i+1,t} \dots, \widehat{X}_{d,t}) > 0\Big\}, \tag{3.3}$$

where $Y_i^-(t) = Y_i(t) \wedge 0$ and $\widetilde{Y}_{i,t}^-(\theta) = \widetilde{Y}_i^-(t+\theta) = Y_i^-(t), -\tau \le \theta \le 0$. It is easy to see that the sample path for $Z_i(t,\omega)$ is continuous for a fixed ω. By monotone class theorem, one can verify that $Z_i(t,\omega)$ is $\mathfrak{F}_t$−adapted. Then applying Corollary 5, we get that ϑ_i is an $\mathfrak{F}_t$−stopping time.

We claim that

$$\Lambda_i < \vartheta_i \text{ on } \{\Lambda_i = \Lambda < \infty\}, \quad i = 1, 2, \dots, d.$$

By the definition of ϑ_i, $\Lambda_i \le \vartheta_i$ holds for every $i = 1, 2, \dots, d$. Again by the definition of ϑ_i, together with the continuity of f_i and $\widehat{f}_i$, $i = 1, 2, \dots, d$ and the pathwise continuity of $X$and $\widehat{X}$ we have

$$\begin{aligned} &f_i(\vartheta_i, X(\vartheta_i), X_{\vartheta_i}) - \widehat{f}_i(\vartheta_i, \widehat{X}_1(\vartheta_i), \dots, \widehat{X}_{i-1}(\vartheta_i), \widehat{X}_i(\vartheta_i) - Y_i^-(\vartheta_i), \widehat{X}_{i+1}(\vartheta_i), \\ &\dots, \widehat{X}_d(\vartheta_i), \widehat{X}_{1,\vartheta_i}, \dots, \widehat{X}_{i-1,\vartheta_i}, \widehat{X}_{i,\vartheta_i} - \widetilde{Y}_{i,\vartheta_i}^-, \widehat{X}_{i+1,\vartheta_i}, \dots, \widehat{X}_{d,\vartheta_i}) \ge 0. \end{aligned} \tag{3.4}$$

In order to prove this claim, let us assume the contrary. Then $\vartheta_i = \Lambda_i$ on $\{\Lambda_i = \Lambda < \infty\}$. Since $X_i(\Lambda_i) = \widehat{X}_i(\Lambda_i)$, we have $Y_i^-(\vartheta_i) = 0$, $\widetilde{Y}_{i,\vartheta_i}^-(\theta) = 0$, $-\tau \le \theta \le 0$. In view of (3.2) and the hypothesis (H1), it is easy to see that

$$\begin{aligned} &f_i(\vartheta_i, X(\vartheta_i), X_{\vartheta_i}) - \widehat{f}_i(\vartheta_i, \widehat{X}_1(\vartheta_i), \dots, \widehat{X}_{i-1}(\vartheta_i), \widehat{X}_i(\vartheta_i) - Y_i^-(\vartheta_i), \\ &\widehat{X}_{i+1}(\vartheta_i), \dots, \widehat{X}_d(\vartheta_i), \widehat{X}_{1,\vartheta_i}, \dots, \widehat{X}_{i-1,\vartheta_i}, \widehat{X}_{i,\vartheta_i} - \widetilde{Y}_{i,\vartheta_i}^-, \widehat{X}_{i+1,\vartheta_i}, \dots, \widehat{X}_{d,\vartheta_i}) < 0 \end{aligned}$$

on $\{\Lambda_i = \Lambda < \infty\}$, which contradicts (3.4). Hence this claim holds.

By (3.3) and the above claim, it can be seen that for all $s \in [\Lambda_i, \vartheta_i]$

$$\begin{aligned} &f_i(s, X(s), X_s) - \widehat{f}_i(s, \widehat{X}_1(s), \dots, \widehat{X}_{i-1}(s), \widehat{X}_i(s) - Y_i^-(s), \widehat{X}_{i+1}(s), \\ &\dots, \widehat{X}_d(s), \widehat{X}_{1,s}, \dots, \widehat{X}_{i-1,s}, \widehat{X}_{i,s} - \widetilde{Y}_{i,s}^-, \widehat{X}_{i+1,s} \dots, \widehat{X}_{d,s}) \le 0 \end{aligned} \tag{3.5}$$

on $\{\Lambda_i = \Lambda < \infty\}$.

Now our purpose is to prove that for every $N > 0$, $\mathbb{P}(\{\Lambda < \Gamma_N\}) = 0$. To this end, we assume that

$$\mathbb{P}(\{\Lambda < \Gamma_N\}) > 0$$

for some N. It follows that there exists an $i \in \{1, 2, \dots, d\}$ such that

$$\mathbb{P}(A) > 0$$

where $A = \{\Lambda_i = \Lambda < \Gamma_N\}$.

Since $Y_i(t),\ i = 1, 2, \ldots, d$ is a continuous semimartingale ([17]), applying the Itô formula, we have

$$\begin{aligned}
&\varphi_\epsilon(Y_i((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N))\\
&\quad=\varphi_\epsilon(Y_i(\Lambda\wedge\vartheta_i\wedge\Gamma_N))\\
&\qquad+\int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}\varphi'_\epsilon(Y_i(s))[\widehat{f}_i(s,\widehat{X}(s),\widehat{X}_s)-f_i(s,X(s),X_s)]ds\\
&\qquad+\int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}\varphi'_\epsilon(Y_i(s))\sum_{j=1}^{r}(\sigma_{ij}(s,\widehat{X}(s))-\sigma_{ij}(s,X(s)))dW_s^j\\
&\qquad+\int_{\Lambda\wedge\wedge\vartheta_i\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}\frac{1}{2}\varphi''_\epsilon(Y_i(s))\sum_{j=1}^{r}(\sigma_{ij}(s,\widehat{X}(s))-\sigma_{ij}(s,X(s)))^2ds\\
&\quad\triangleq\Delta_1+\Delta_2+\Delta_3+\Delta_4.
\end{aligned}\tag{3.6}$$

Note that I_A is $\mathcal{F}_\Lambda-$ measurable (see Lemma 1.2.16 in [13]). Hence

$$E[\Delta_3 I_A]=E[E[\Delta_3 I_A|\mathcal{F}_\Lambda]]=E[I_A E[\Delta_3|\mathcal{F}_\Lambda]]=0.\tag{3.7}$$

In fact, let

$$M_\chi(\omega)\triangleq\int_0^\chi\varphi'_\epsilon(Y_i(s))\sum_{j=1}^{r}(\sigma_{ij}(s,\widehat{X}(s))-\sigma_{ij}(s,X(s)))dW_s^j.$$

Then $M_\chi(\omega)$ is a continuous martingale. Thus

$$\begin{aligned}
E[\Delta_3|\mathcal{F}_\Lambda]&=E[M_{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}-M_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}|\mathcal{F}_\Lambda]\\
&=E[M_{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}|\mathcal{F}_\Lambda]-M_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}\\
&=E[M_{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}|\mathcal{F}_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}]-M_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}=0.
\end{aligned}$$

We have used the optional sampling theorem (see [13]) in the last equality and (ii) of Problem 1.2.17 in [13] in the second equality. This proves (3.7).

Multiplied by the indicator function I_A to the two sides of (3.6), and then taking expectation and setting $\epsilon\to 0$, we obtain that

$$\begin{aligned}
&E\Big(|Y_i^-((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N)|^2 I_A\Big)\\
&=EI_A\int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}2\Big(Y_i^-(s)\Big)\Big[\widehat{f}_i(s,\widehat{X}(s),\widehat{X}_s)-f_i(s,X(s),X_s)\Big]ds\\
&\quad+EI_A\int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N}I_{\{Y_i(s)\le 0\}}\sum_{j=1}^{r}\Big(\sigma_{ij}(s,\widehat{X}(s))-\sigma_{ij}(s,X(s))\Big)^2ds\\
&\triangleq EI_A\Sigma_1+EI_A\Sigma_2.
\end{aligned}\tag{3.8}$$

Relation (3.5) and $y^- \le 0$ imply that

$$
\begin{aligned}
I_A \Sigma_1 = I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} & 2(Y_i^-(s)) \big[\widehat{f}_i(s, \widehat{X}(s), \widehat{X}_s) - \widehat{f}_i(s, \widehat{X}_1(s), \ldots, \widehat{X}_{i-1}(s), \\
& \widehat{X}_i(s) - Y_i^-(s), \widehat{X}_{i+1}(s), \ldots, \widehat{X}_d(s), \widehat{X}_{1,s}, \ldots, \widehat{X}_{i-1,s}, \widehat{X}_{i,s} - \widetilde{Y}_{i,s}^-, \widehat{X}_{i+1,s}, \ldots, \widehat{X}_{d,s}) \big] ds \\
& + I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} 2(Y_i^-(s)) \big[\widehat{f}_i(s, \widehat{X}_1(s), \ldots, \widehat{X}_{i-1}(s), \widehat{X}_i(s) - Y_i^-(s), \widehat{X}_{i+1}(s), \\
& \ldots, \widehat{X}_d(s), \widehat{X}_{1,s}, \ldots, \widehat{X}_{i-1,s}, \widehat{X}_{i,s} - \widetilde{Y}_{i,s}^-, \widehat{X}_{i+1,s}, \ldots, \widehat{X}_{d,s}) - f_i(s, X(s), X_s) \big] ds \\
& \le I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} 2(Y_i^-(s)) \big[\widehat{f}_i(s, \widehat{X}(s), \widehat{X}_s) - \widehat{f}_i(s, \widehat{X}_1(s), \ldots, \widehat{X}_{i-1}(s), \widehat{X}_i(s) \\
& - Y_i^-(s), \widehat{X}_{i+1}(s), \ldots, \widehat{X}_d(s), \widehat{X}_{1,s}, \ldots, \widehat{X}_{i-1,s}, \widehat{X}_{i,s} - \widetilde{Y}_{i,s}^-, \widehat{X}_{i+1,s}, \ldots, \widehat{X}_{d,s}) \big] ds.
\end{aligned}
$$

By the global Lipschitz condition for the drift $\widehat{f}$, there exists a constant $L^* > 0$ such that

$$
I_A \Sigma_1 \le I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} 2|Y_i^-(s)| \times L^* \Big(|Y_i^-(s)| + \|\widetilde{Y}_{i,s}^-\| \Big) ds. \tag{3.9}
$$

Since $\widetilde{Y}_{i,s}^-(\theta) = Y_i^-(s)$, $-\tau \le \theta \le 0$, by (3.9) we have

$$
I_A \Sigma_1 \le I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} 4L^* |Y_i^-(s)|^2 ds. \tag{3.10}
$$

From (H3) it follows that

$$
\begin{aligned}
I_A \Sigma_2 &= I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} I_{\{Y_i(s)\le 0\}} \sum_{j=1}^{r} \Big(\sigma_{ij}(s, \widehat{X}(s)) - \sigma_{ij}(s, X(s)) \Big)^2 ds \\
&\le I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} I_{\{Y_i(s)\le 0\}} \times \rho\Big(|Y_i(s)|^2 \Big) ds \\
&\le I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} \rho\Big(|Y_i^-(s)|^2 \Big) ds.
\end{aligned}
\tag{3.11}
$$

By (3.8), (3.10) and (3.11), we have

$$
\begin{aligned}
E\Big(|Y_i^-((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N)|^2 I_A \Big) &\le E I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} 4L^* |Y_i^-(s)|^2 ds \\
&\quad + E I_A \int_{\Lambda\wedge\vartheta_i\wedge\Gamma_N}^{(\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N} \rho\Big(|Y_i^-(s)|^2 \Big) ds \\
&\le E \int_0^t 4L^* |Y_i^-\Big((s+\Lambda)\wedge\vartheta_i\wedge\Gamma_N \Big)|^2 I_A ds \\
&\quad + E \int_0^t \rho\Big(|Y_i^-((s+\Lambda)\wedge\vartheta_i\wedge\Gamma_N)|^2 \Big) I_A ds \\
&\le \int_0^t 4L^* E\Big(|Y_i^-((s+\Lambda)\wedge\vartheta_i\wedge\Gamma_N)|^2 I_A \Big) ds \\
&\quad + \int_0^t \rho\Big(E(|Y_i^-((s+\Lambda)\wedge\vartheta_i\wedge\Gamma_N)|^2 I_A) \Big) ds
\end{aligned}
$$

where the last inequality has applied the Jensen inequality because of the concavity of ρ. Note that $t \to E|Y_i^-((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N)|^2$ is continuous (see Remark 14) and using Corollary 3, we have

$$E\Big(|Y_i^-((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N\Big)|^2 I_A) = 0,$$

which implies that

$$X_i\Big((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N) \le \widehat{X}_i((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N\Big) \text{ a.s. } \mathbb{P}$$

for every $t \ge 0$ on $\{\Lambda_i = \Lambda < \Gamma_N\}$. It follows from the continuity of $X_i(t)$, $\widehat{X}_i(t)$ that

$$X_i\Big((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N) \le \widehat{X}_i((\Lambda+t)\wedge\vartheta_i\wedge\Gamma_N\Big), \quad 0 \le t < \infty \text{ a.s. } \mathbb{P}$$

on $\{\Lambda_i = \Lambda < \Gamma_N\}$. This contradicts (3.1), which shows that $\mathbb{P}(\{\Lambda < \Gamma_N\}) = 0$ for every $N > 0$. Hence we have $\mathbb{P}(\{\Lambda = \infty\}) = 1$. Therefore

$$\mathbb{P}\Big(\{X(t) \le \widehat{X}(t), t \ge 0\}\Big) = 1,$$

i.e.,

$$\mathbb{P}\Big(\{x(t,\phi) \le \widehat{x}(t,\psi), t \ge 0\}\Big) = 1,$$

which shows that

$$\mathbb{P}\Big(\{x_t(\phi) \le_{\mathcal{C}} \widehat{x}_t(\psi), t \ge 0\}\Big) = 1.$$

The proof is complete. □

Now we give a condition which makes it possible to replace $<$ in (H1) by $\le$. For this purpose we review the following definition given in [21].

Definition 7 A mapping $g : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}^d$ is called quasi-monotonously increasing, if for every $i = 1, 2, \ldots, d$

$$g_i\Big(t, \phi(0), \phi\Big) \le g_i\Big(t, \psi(0), \psi\Big),$$

wherever $\phi \le_{\mathcal{C}} \psi$ with $\phi_i(0) = \psi_i(0)$ and $t > 0$.

Theorem 8 *If either drift term f to system* (2.1) *or $\widehat{f}$ to system* (2.2) *is quasi-monotonously increasing, then the condition* (H1) *of Theorem 6 can be replaced by*

(H1*) *The function $f_i : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}$ and $\widehat{f}_i : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}$, $i = 1, 2, \ldots, d$ are continuous and satisfy the following condition: $f_i(t, \phi(0), \phi) \le \widehat{f}_i(t, \phi(0), \phi)$ for every $\phi \in \mathcal{C}, t \ge 0$.*

Proof Assume that the drift coefficient f to system (2.1) is quasi-monotonously increasing, a similar argument holds if the drift coefficient $\widehat{f}$ to system (2.2) is quasi-monotonously increasing. Let $\zeta > 0$ be arbitrarily chosen and define

$$f_i^\zeta \triangleq f_i - \zeta, \quad i = 1, 2, \ldots, d.$$

Consider the following auxiliary system

$$\begin{cases} dx_i(t) = f_i^\zeta(t, x(t), x_t)dt + \sum_{j=1}^r \sigma_{ij}(t, x(t))dW_t^j, \quad i = 1, 2, \ldots, d, \\ x(\theta) = \phi(\theta), \ \theta \in J, \ \phi \in C(J, \mathbb{R}^d). \end{cases} \tag{3.12}$$

By the hypotheses (H2)–(H4) and Theorem 1, system (3.12) has a unique strong solution $X^{\zeta}(t), t \geq 0$. From (H1*) and the quasi-monotonicity of f it follows that the pair $(f_i^{\zeta}, \widehat{f}_i), i = 1, 2, \ldots, d$, satisfies (H1). Hence, applying Theorem 6, we get that

$$X_i^{\zeta}(t) \leq \widehat{X}_i(t) \quad \text{for all } t \geq 0, \quad \text{a.s. } \mathbb{P}$$

for $i = 1, 2, \ldots, d$. Choose a strictly decreasing sequence $\zeta_n, n \geq 1$ with $\lim_{n\to\infty} \zeta_n = 0$. By the same arguments as above we have

$$X_i^{\zeta_1}(t) \leq X_i^{\zeta_2}(t) \leq \cdots \leq \widehat{X}_i(t) \quad \text{for all } t \geq 0, \ \text{a.s. } \mathbb{P}$$

as well as

$$X_i^{\zeta_1}(t) \leq X_i^{\zeta_2}(t) \leq \cdots \leq X_i(t) \quad \text{for all } t \geq 0, \ \text{a.s. } \mathbb{P}$$

for $i = 1, 2, \ldots, d$. Define

$$\mathbb{X}_i(t) \triangleq \lim_{n\to\infty} X_i^{\zeta_n}(t), \tag{3.13}$$

for $t \geq 0, i = 1, 2, \ldots, d$. Then

$$\mathbb{X}_i(t) \leq \widehat{X}_i(t) \quad \text{for all } t \geq 0, \ \text{a.s. } \mathbb{P}$$

for $i = 1, 2, \ldots, d$. To complete the proof, we shall show that $\mathbb{X}(t)$ is a modification of the solution $X(t)$. By the strong solution uniqueness it suffices to check that $\mathbb{X}(t)$ satisfies (2.1) a.s. $\mathbb{P}$ for every $t \geq 0$.

First we prove that $X^{\zeta_n}(t)$ converges to $\mathbb{X}(t)$ uniformly on $t \in [0, T]$ almost surely as $n \to \infty$. In terms of the Hölder inequality and (3.12) we have

$$\begin{aligned}
\sup_{0\leq\chi\leq t} |X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)|^2 &= \sup_{0\leq\chi\leq t} \Big|\chi(\zeta_n - \zeta_{n+1})\overrightarrow{e} \\
&\quad + \int_0^{\chi} \big(f(s, X^{\zeta_{n+1}}(s), X_s^{\zeta_{n+1}}) - f(s, X^{\zeta_n}(s), X_s^{\zeta_n})\big)ds \\
&\quad + \int_0^{\chi} \big(\sigma(s, X^{\zeta_{n+1}}(s)) - \sigma(s, X^{\zeta_n}(s))\big)dW_s\Big|^2 \\
&\leq 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +3\Big(\int_0^t |f(s, X^{\zeta_{n+1}}(s), X_s^{\zeta_{n+1}}) - f(s, X^{\zeta_n}(s), X_s^{\zeta_n})|ds\Big)^2 \\
&\quad +3 \sup_{0\leq\chi\leq t} \Big|\int_0^{\chi} (\sigma(s, X^{\zeta_{n+1}}(s)) - \sigma(s, X^{\zeta_n}(s)))dW_s\Big|^2 \\
&\leq 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +3T \int_0^t \Big|f(s, X^{\zeta_{n+1}}(s), X_s^{\zeta_{n+1}}) - f(s, X^{\zeta_n}(s), X_s^{\zeta_n})\Big|^2 ds \\
&\quad +3 \sup_{0\leq\chi\leq t} \Big|\int_0^{\chi} (\sigma(s, X^{\zeta_{n+1}}(s)) - \sigma(s, X^{\zeta_n}(s)))dW_s\Big|^2,
\end{aligned}$$

where $\overrightarrow{e} = (1, 1, \ldots, 1)^\tau$ is a d dimensional vector. According to (H2) and (H3) and the Burkholder–Davis–Gundy inequality, for $t \in [0, T]$, we have

$$\begin{aligned}
E \sup_{0\le\chi\le t} |X^{\zeta_{n+1}}(t) - X^{\zeta_n}(\chi)|^2 &\le 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +3TE\Big(\int_0^t |f(s, X^{\zeta_{n+1}}(s), X_s^{\zeta_{n+1}}) - f(s, X^{\zeta_n}(s), X_s^{\zeta_n})|^2 ds\Big) \\
&\quad +3E \sup_{0\le\chi\le t}\Big|\int_0^\chi (\sigma(s, X^{\zeta_{n+1}}(s)) - \sigma(s, X^{\zeta_n}(s)))dW_s\Big|^2 \\
&\le 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +3TdLE\Big(\int_0^t (|X^{\zeta_{n+1}}(s) - X^{\zeta_n}(s)|^2 + \|X_s^{\zeta_{n+1}} - X_s^{\zeta_n}\|^2)ds\Big) \\
&+ 12E\int_0^t \Big|\sigma(s, X^{\zeta_{n+1}}(s)) - \sigma(s, X^{\zeta_n}(s))\Big|^2 ds \\
&\le 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +6TdL\int_0^t \Big(E\sup_{0\le\chi\le s} |X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)|^2\Big)ds \\
&\quad +12d\int_0^t \rho\Big(E\sup_{0\le\chi\le s} |X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)|^2\Big)ds \\
&\le 3d(\zeta_n - \zeta_{n+1})^2 T^2 \\
&\quad +(6TdL + 12d)\int_0^t \varrho\Big(E\sup_{0\le\chi\le s} |X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)|^2\Big)ds.
\end{aligned}$$

Note that $t \to E\sup_{0\le\chi\le t} |X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)|^2$ is continuous (see Remark 14) and hence by Theorem 2, we have

$$E\sup_{0\le\chi\le t}\Big|X^{\zeta_{n+1}}(\chi) - X^{\zeta_n}(\chi)\Big|^2 \le W^{-1}\Big[W(r(t)) + (6TdL + 12d)t\Big],\ 0 \le t \le t_c,$$

where $W(u) = \int_{\widehat{u}}^u \frac{dz}{\varrho(z)}$, $\widehat{u}$ is any given positive constant, $r(t) = 3d(\zeta_n - \zeta_{n+1})^2 T^2$, and $t_c \le T$ is the largest number such that

$$W(r(t_c)) + (6TdL + 12d)t_c \le \int_{\widehat{u}}^\infty \frac{dz}{\varrho(z)}, \tag{3.14}$$

By (2.8) it is easy to see that (3.14) holds if n is sufficiently large, i.e., there exists $N_1 > 0$ such that if $n \ge N_1$, then (3.14) holds. Hence, as $n \ge N_1$, $t_c = T$. Then

$$E\sup_{0\le t\le T}\Big|X^{\zeta_{n+1}}(t) - X^{\zeta_n}(t)\Big|^2 \le W^{-1}\Big[W(r(T)) + 6dLT^2 + 12dT\Big]. \tag{3.15}$$

Furthermore, since $\int_{0+} \frac{1}{\varrho(z)}dz = \infty$, there exist $\mu_n, n \ge 1$ such that for every n, $\mu_n < \frac{1}{8^n}$ and $\int_{\mu_n}^{\frac{1}{8^n}} \frac{1}{\varrho(z)}dz = 6dLT^2 + 12dT$. Let $r(T) = \mu_n$, then $\zeta_n - \zeta_{n+1} = \sqrt{\frac{\mu_n}{3dT^2}}$, and hence $\zeta_n = \sum_{k=n}^\infty \sqrt{\frac{\mu_k}{3dT^2}}$, $n \ge 1$. Moreover by (3.15) we have that for $n \ge N_1$,

$$E\sup_{0\le t\le T}\Big|X^{\zeta_{n+1}}(t) - X^{\zeta_n}(t)\Big|^2 \le \frac{1}{8^n}.$$

By Chebyshev's inequality we have

$$\sum_{n=N_1}^{\infty} \mathbb{P}\Big(\sup_{0\le t\le T} |X^{\zeta_{n+1}}(t) - X^{\zeta_n}(t)| > 2^{-(n+1)}\Big) \le \sum_{n=N_1}^{\infty} 4\frac{1}{2^n} < \infty.$$

Then Borel–Cantelli Lemma implies that there exists an event $\Omega^* \in \mathfrak{F}$ with $\mathbb{P}(\Omega^*) = 1$ such that for every $\omega \in \Omega^*$ there is an integer $N_2(\omega) \ge N_1$ satisfying that

$$\sup_{0\le t\le T} \Big|X^{\zeta_{n+1}}(t) - X^{\zeta_n}(t)\Big| \le 2^{-(n+1)},\ n \ge N_2(\omega).$$

Consequently

$$\sup_{0\le t\le T} \big|X^{\zeta_{n+p}}(t) - X^{\zeta_n}(t)\big| \le 2^{-n}, \quad p \ge 1,\ n \ge N_2(\omega),$$

that is, $X^{\zeta_n}(s)$ converges to $\mathbb{X}(s)$ uniformly on $[0, T]$ almost surely. The continuity of $\mathbb{X}(s)$ follows.

Now define

$$T_N \triangleq \inf\big\{t > 0 : \| X_t \| > N \text{ or } \|X_t^{\zeta_1}\| > N\big\} \wedge N, \quad \text{forevery} \quad N > 0.$$

In terms of (3.13) and Lebesgue's Theorem on dominated convergence, we have

$$\int_0^{t\wedge T_N} f_i^{\zeta_n}\big(s, X^{\zeta_n}(s), X_s^{\zeta_n}\big)ds \to \int_0^{t\wedge T_N} f_i\big(s, \mathbb{X}(s), \mathbb{X}_s\big)ds,\ \text{a.s. } \mathbb{P},$$

and

$$E\int_0^{t\wedge T_N} \big|\sigma_{ij}(s, X^{\zeta_n}(s)) - \sigma_{ij}(s, \mathbb{X}(s))\big|^2 ds \to 0,$$

as $n \to \infty$.

Thus, by Proposition 3.2.10 in [13], we have

$$\int_0^{t\wedge T_N} \sigma_{ij}(s, X^{\zeta_n}(s))dW_s^j \xrightarrow{L^2} \int_0^{t\wedge T_N} \sigma_{ij}(s, \mathbb{X}(s))dW_s^j$$

as $n \to \infty$. Then taking a subsequence if necessary, we obtain that

$$\lim_{n\to\infty} X_i^{\zeta_n}(t \wedge T_N) = \phi_i(0) + \int_0^{t\wedge T_N} f_i(s, \mathbb{X}(s), \mathbb{X}_s)ds + \sum_{j=1}^{r} \int_0^{t\wedge T_N} \sigma_{ij}(s, \mathbb{X}(s))dW_s^j,$$

that is,

$$\mathbb{X}_i(t \wedge T_N) = \phi_i(0) + \int_0^{t\wedge T_N} f_i(s, \mathbb{X}(s), \mathbb{X}_s)ds + \sum_{j=1}^{r} \int_0^{t\wedge T_N} \sigma_{ij}(s, \mathbb{X}(s))dW_s^j.$$

Note that as $N \uparrow \infty$, $T_N \uparrow \infty$. Setting $N \uparrow \infty$, from the path continuity of $\mathbb{X}(t)$ we conclude that

$$\mathbb{X}_i(t) = \phi_i(0) + \int_0^t f_i(s, \mathbb{X}(s), \mathbb{X}_s)ds + \sum_{j=1}^{r} \int_0^t \sigma_{ij}(s, \mathbb{X}(s))dW_s^j.$$

The proof is complete. □

4 Application

As an application of our comparison theorem, we will show that the solutions of a class of stochastic neutral networks with delays possess monotonicity and sublinearity. Consider the following stochastic neutral networks described by Stratonovich type of stochastic differential equations with delay:

$$\begin{cases} dx_i(t) = \big[-a_i x_i(t) + \sum\limits_{j=1}^{d} b_{ij} h_j(x_j(t-\tau_{ij}))\big]dt + \sum\limits_{k=1}^{r} \sigma_{ik} x_i(t) \circ dW_t^k, \\ x_i(\theta) = \phi_i(\theta), \theta \in J \triangleq [-\tau, 0], \tau = \max\limits_{1\le i,j\le d} \tau_{ij}, \phi \in \mathcal{C} \triangleq C(J, \mathbb{R}_+^d), \quad i = 1, 2, \ldots, d, \end{cases} \tag{4.1}$$

where the activation functions satisfy the following:

(A1) $h_i \in C^1(\mathbb{R}), \quad h_i(0) = 0, i = 1, 2, \ldots, d,$
(A2) $\lim\limits_{s\to\infty} \frac{h_i(s)}{s} = 0, \quad i = 1, 2, \ldots, d,$
(A3) $0 < h_i'(s) \le 1, \quad i = 1, 2, \ldots, d,$
(A4) $b_{ij} > 0$ for all $i, j = 1, 2, \ldots, d.$

Theorem 9 *Besides* (A1)–(A4)*, we further assume that the system* (4.1) *satisfies*
(A5) *each $h_i(s)$ is a sublinear function from $\mathbb{R}_+$ into $\mathbb{R}$ in the sense that*

$$\lambda h_i(s) \le h_i(\lambda s) \quad \text{for all } 0 < \lambda < 1 \text{ and s} > 0.$$

Then the system (4.1) *has a unique strong solution $\Phi(t, \phi)$ for $t \ge 0$, where $\phi = (\phi_1, \ldots, \phi_d)$. Moreover, it generates a monotone random dynamical system, which means that*

$$\phi \le_{\mathcal{C}} \psi \Longrightarrow \Phi(t, \phi) \le_{\mathcal{C}} \Phi(t, \psi), \ t \ge 0$$

and $\Phi(t, \phi)$ is sublinear in the sense that for every $\phi \ge 0$,

$$\lambda\Phi(t, \phi) \le_{\mathcal{C}} \Phi(t, \lambda\phi), \quad \text{for all } t \ge 0 \text{ and } 0 \le \lambda \le 1.$$

Proof First, we shall illustrate the strong solutions for (4.1) generate a random dynamical system by the conjugacy technique (see Imkeller and Schmalfuss [12]).

Denote by $z(\omega)$ the random variable in $\mathbb{R}^r$ such that

$$z(t, \omega) := z(\theta_t\omega) = (z_1(t, \omega), z_2(t, \omega), ..., z_r(t, \omega))$$

is *Stationary Ornstein-Uhlenbeck Process* in $\mathbb{R}^r$ which solves the equations

$$dz_k = -\mu z_k dt + dW_t^k, \quad \mu > 0, k = 1, \ldots, r.$$

Let us first introduce a conjugate transformation:

$$T(\omega, y) = \big(y_1.e_1^{\sigma}(\omega), \ldots, y_d.e_d^{\sigma}(\omega)\big)$$

where

$$e_i^\sigma(\omega) = \exp\{z_i^\sigma(\omega)\}, \quad z_i^\sigma(\omega) = \sum_{j=1}^{r} \sigma_{ij} z_j(\omega).$$

Applying Itô formula to the function $y_i(t,\omega) = x_i(t,\omega)\exp\{-z_i^\sigma(\theta_t\omega)\}$, we transform the system (4.1) into

$$\begin{cases} \frac{dy_i(t)}{dt} = -a_i y_i(t) + \mu y_i(t) z_i^\sigma(\theta_t\omega) + \exp\{-z_i^\sigma(\theta_t\omega)\}\times \\ \sum_{j=1}^{d} b_{ij} h_j(y_j(t-\tau_{ij})\exp\{-z_j^\sigma(\theta_{t-\tau_{ij}}\omega)\}) \\ y(\theta) = \phi(\theta), \theta \in J \triangleq [-\tau, 0], \quad i = 1, \ldots, d. \end{cases} \tag{4.2}$$

For each i, define

$$\begin{cases} F_i(\omega,\phi) := -a_i\phi_i(0) + \mu z_i^\sigma(\omega)\phi_i(0) + \exp\{-z_i^\sigma(\omega)\}\times \\ \sum_{j=1}^{d} b_{ij} h_j(\phi_j(-\tau_{ij})\exp\{-z_j^\sigma(\theta_{-\tau_{ij}}\omega)\}). \end{cases} \tag{4.3}$$

Then it follows from (A1) and (A3) that

(C1) $F_i(\omega, 0) \equiv 0$;
(C2) F_i satisfies the global Lipschitz condition in the sense that for any $\phi, \psi \in \mathcal{C}$,

$$|F_i(\omega,\phi) - F_i(\omega,\psi)| \le L_i(\omega) \parallel \phi - \psi \parallel$$

where $L_i(\omega) = a_i + \mu|z_i^\sigma(\omega)| + \exp\{-z_i^\sigma(\omega)\} \times \sum_{j=1}^{d} b_{ij}\exp\{-z_j^\sigma(\theta_{-\tau_{ij}}\omega)\}$. The system (4.2) can be written as a type of *Random Functional Differential Equations:*

$$\begin{cases} \frac{dy(t)}{dt} &= F(\theta_t\omega, y_t) \\ y_0 &= \phi \in \mathcal{C}. \end{cases} \tag{4.4}$$

One can shows that the solutions for (4.4) satisfying (C1) and (C2) generate a random dynamical system by the fundamental theory of deterministic functional differential equations, we omit the detail.

Stratonovich stochastic differential equation (4.1) can be rewritten in the Itô form

$$\begin{cases} dx_i(t) = \overline{f}_i(x_i(t), x_1(t-\tau_{i1}), \ldots, x_d(t-\tau_{id}))dt + \sum_{k=1}^{r} \sigma_{ik} x_i(t) dW_t^k, \\ x_i(\theta) \;\; = \phi_i(\theta_i), \quad i = 1, 2, \ldots, d, \end{cases} \tag{4.5}$$

where $\overline{f}_i = f_i(x_i(t), x_1(t-\tau_{i1}), \ldots, x_d(t-\tau_{id})) + \frac{1}{2}\sum_{k=1}^{r} \sigma_{ik}^2 x_i(t)$ with

$$f_i(x_i(t), x_1(t-\tau_{i1}), \ldots, x_d(t-\tau_{id})) = -a_i x_i(t) + \sum_{j=1}^{d} b_{ij} h_j(x_j(t-\tau_{ij})).$$

Then we only need to prove that the results in Theorem 9 hold for system (4.5). It is obvious that under (A1)–(A4), the global existence and uniqueness strong solution $\Phi(t,\phi)$ of the system (4.5) can be obtained by Theorem 1. In addition, observe that the drift terms of equation (4.5) is quasi-monotonously increasing. Then monotonicity of the solutions of the system (4.5) is obtained by applying Theorem 8.

The idea to prove sublinearity of the solution of the equations (4.1) follows from [5]. Let $x_i^\lambda(t) = \lambda \Phi_i(t, \phi)$. Then

$$\begin{cases} dx_i^\lambda(t) = f_i^\lambda\big(x_i^\lambda(t), x_1^\lambda(t-\tau_{i1}), \ldots, x_d^\lambda(t-\tau_{id})\big)dt + \sum_{k=1}^r \sigma_{ik} x_i^\lambda(t) \circ dW_t^k, \\ x_i^\lambda(\theta) \ \ = \lambda \phi_i(\theta_i), \quad i = 1, 2, \ldots, d, \end{cases}$$

where

$$f_i^\lambda\big(x_i^\lambda(t), x_1^\lambda(t-\tau_{i1}), \ldots, x_d^\lambda(t-\tau_{id})\big) = -a_i x_i^\lambda(t) + \sum_{j=1}^d \lambda b_{ij} h_j\big(\lambda^{-1} x_j^\lambda(t-\tau_{ij})\big).$$

Using (A5), it is easy to see that for every $\phi = (\phi_1, \ldots, \phi_d) \geq 0,\ i = 1, \ldots, d$,

$$f_i^\lambda\big(\phi_i(0), \phi_1(-\tau_{i1}), \ldots, \phi_d(-\tau_{id})\big) \leq f_i\big(\phi_i(0), \phi_1(-\tau_{i1}), \ldots, \phi_d(-\tau_{id})\big),$$

which implies that (H1*) in Theorem 8 holds. The last paragraph shows that the comparison theorem is still valid for Stratonovich type of stochastic functional differential equation. Again by the quasi-monotonicity of the drift terms of the system (4.1), we can apply Theorem 8 to deduce the sublinearity for $\Phi(t, \phi)$. The proof is complete. □

Remark 10 Chueshov and Scheutzow [6] obtained comparison principle under the assumptions that the drift terms are decomposed into a non-delay part and a delay part and the non-delay drift terms are at least continuously differentiable and diffusion terms are at least twice continuously differentiable. When the drift terms are replaced by this non-delay drift part, the SDEs generate a stochastic flow of diffeomorphisms in $\mathbb{R}^d$, which helps them to represent the SFDEs as a random FDEs. Our comparison result can apply to the case that the diffusion terms are even not non-Lipschitz (see Example 11). Thus the corresponding SDEs only generate a stochastic flow of hemeomorphisms, rather than diffeomorphisms in $\mathbb{R}^d$. However, the comparison principle in Chueshov and Scheutzow [6] is valid for more general noise, exactly, they considered SFDEs driven by a Kunita-type martingale field. Their idea is to represent the SFDEs as a random FDE, while our technique is to combine the method of [4] and stopping times with the nonlinear Gronwall inequality.

Example 11 Let $d = r = 1$ and consider the following two stochastic delay differential equations.

$$\begin{cases} dx(t) = \frac{1}{2}x(t-\tau)dt + \sigma(x(t))dW_t, \\ x(\theta) \ \ = \phi(\theta),\ \theta \in J \triangleq [-\tau, 0],\ \phi \in C(J, \mathbb{R}), \end{cases} \tag{4.6}$$

and

$$\begin{cases} d\widehat{x}(t) = [\frac{1}{2}\widehat{x}(t-\tau) + 1]dt + \sigma(\widehat{x}(t))dW_t, \\ \widehat{x}(\theta) \ \ = \psi(\theta),\ \theta \in J,\ \psi \in C(J, \mathbb{R}), \end{cases} \tag{4.7}$$

where $\sigma(x) = \begin{cases} \frac{1}{2}|x|\sqrt[4]{\ln\frac{1}{|x|}}, & |x| \leq \frac{1}{e}, \\ \frac{1}{2e}, & |x| > \frac{1}{e}. \end{cases}$

First it is obvious that the assumptions (H1), (H2) and (H4) hold for systems (4.6) and (4.7), and then it is easy to see that for every $x, x' \in \mathbb{R}$,

$$|\sigma(x) - \sigma(x')|^2 \leq \rho\big(|x - x'|^2\big),$$

where $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ and $\rho(x) = \begin{cases} \frac{1}{2}x\sqrt{\ln \frac{1}{x}}, & x \le \frac{1}{e}, \\ \frac{1}{2e}, & x > \frac{1}{e}. \end{cases}$ Note that $\rho(0) = 0$, $\int_{0^+} \frac{dx}{\rho(x)} = \infty$, and $\rho(x)$ is a nondecreasing continuous concave function, hence (H3) holds. Then Theorem 6 can apply to the equations (4.6) and (4.7). However, since σ is not differentiable at the origin and comparison principle in Chueshov and Scheutzow [6] needs the diffusion term σ to be at least twice continuously differentiable, their comparison principle cannot applied to the systems (4.6) and (4.7).

Acknowledgements The first author is partly supported by the Natural Science Research Project of High Education of Anhui province under Grant No. 2012AJZR0323. The second author was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 11371252, Research and Innovation Project of Shanghai Education Committee under Grant No. 14zz120, and the Program of Shanghai Normal University (DZL121).

Appendix: The Proof of Theorem 1

The idea for the proof is partially borrowed from [16].

Since ρ is concave with $\rho(0) = 0$, there are two positive constants a_1, a_2 such that

$$\rho(x) \le a_1 x + a_2 \text{ for } x \ge 0. \tag{5.1}$$

As before, $\varrho(x) \triangleq x + \rho(x)$ for $x \ge 0$.

Lemma 12 *For any $T > 1$ and $\xi \in L^2(\Omega, C(I, \mathbb{R}^d))$, where I is an interval, let*

$$\begin{aligned} \tilde{C}_1(\xi) &\triangleq 3\big[(1 + \gamma T^2)\mathrm{E} \parallel \xi \parallel^2 + \gamma T^2 + 4\gamma T\big] e^{6\gamma T(T+2)}, \\ C_1 &\triangleq 3\gamma T(T+4)e^{6\gamma T(T+2)}, \\ C_2 &\triangleq 4d(TL+2), \\ C_3 &\triangleq C_2 \varrho(4\tilde{C}_1(\xi)). \end{aligned}$$

Then $T_1 \triangleq \frac{4C_1}{C_2(4C_1 + 4C_1 a_1 + a_2)}$, depending only on T and ρ rather than on ξ, satisfies that

$$C_2 \varrho(C_3 T_1) \le C_3. \tag{5.2}$$

Proof $\tilde{T}_1 \triangleq \frac{4\tilde{C}_1(\xi)}{C_2(4\tilde{C}_1(\xi) + 4\tilde{C}_1(\xi)a_1 + a_2)}$. Then $T_1 < \tilde{T}_1 < 1$. Since ϱ is increasing on $[0, \infty)$, it suffices to show that

$$C_2 \varrho\big(C_3 \tilde{T}_1\big) \le C_3 = C_2 \varrho\big(4\tilde{C}_1(\xi)\big).$$

Equivalently, $(C_3 \tilde{T}_1) \le 4\tilde{C}_1(\xi)$, i.e.,

$$C_2 \varrho\big(4\tilde{C}_1(\xi)\big)\tilde{T}_1 \le 4\tilde{C}_1(\xi).$$

By (5.1),

$$C_2 \varrho\big(4\tilde{C}_1(\xi)\big)\tilde{T}_1 \le C_2 \tilde{T}_1\big(4\tilde{C}_1(\xi) + 4\tilde{C}_1(\xi)a_1 + a_2\big) = 4\tilde{C}_1(\xi).$$

This proves (5.2). □

From now on, fix $T > 1$ and let k be an integer with $\frac{T}{k} < T_1$. Denote by

$$J_i \triangleq \left[\frac{iT}{k} - \tau, \frac{iT}{k}\right] \text{ and } I_i \triangleq \left[\frac{iT}{k}, \frac{(i+1)T}{k}\right] \quad \text{for } i = 0, 1, \ldots, k-1.$$

Consider the stochastic functional differential equations

$$\begin{cases} x(t) = \xi\left(\frac{iT}{k}\right) + \int\limits_{\frac{iT}{k}}^{t} f\left(s, x(s), x_s\right)ds + \int_{\frac{iT}{k}}^{t} \sigma\left(s, x(s)\right)dW_s, \quad t \in I_i \\ x(\theta) = \xi(\theta), \quad \theta \in J_i. \end{cases} \tag{5.3}$$

Proposition 13 *Let $\xi \in L^2(\Omega, C(J_i, \mathbb{R}^d))$ be $\{\mathfrak{F}_t\}_{t\in J_i}$ adapted. Then (5.3) has a solution $x \in L^2(\Omega, C([\frac{iT}{k} - \tau, \frac{(i+1)T}{k}], \mathbb{R}^d))$ adapted to $\{\mathfrak{F}_t\}_{t\in J_i \cup I_i}$ and with initial process ξ.*

Proof To prove the existence of solution to (5.3), let us construct the following sequence of successive approximations by setting

$$\begin{cases} x^n(t) = \xi\left(\frac{iT}{k}\right) + \int\limits_{\frac{iT}{k}}^{t} f\left(s, x^{n-1}(s), x_s^{n-1}\right)ds + \int_{\frac{iT}{k}}^{t} \sigma\left(s, x^{n-1}(s)\right)dW_s, \quad t \in I_i \\ x^n(\theta) = \xi(\theta), \quad \theta \in J_i, \end{cases} \tag{5.4}$$

for $n \geq 1$, and

$$\begin{cases} x^0(t) = \xi\left(\frac{iT}{k}\right), \quad t \in I_i \\ x^0(\theta) = \xi(\theta), \quad \theta \in J_i. \end{cases}$$

For every $t \in I_i$, by (5.4) and the Hölder inequality we have

$$\begin{aligned} \sup_{\frac{iT}{k} \leq \chi \leq t} |x^n(\chi)|^2 \leq& \sup_{\frac{iT}{k} \leq \chi \leq t} \Bigg(3|\xi\left(\frac{iT}{k}\right)|^2 + 3|\int\limits_{\frac{iT}{k}}^{\chi} f(s, x^{n-1}(s), x_s^{n-1})ds|^2 \\ &+ 3|\int_{\frac{iT}{k}}^{\chi} \sigma(s, x^{n-1}(s))dW_s|^2\Bigg) \\ \leq& 3\left|\xi\left(\frac{iT}{k}\right)\right|^2 + 3\left(t - \frac{iT}{k}\right)\int\limits_{\frac{iT}{k}}^{t} \left|f(s, x^{n-1}(s), x_s^{n-1})\right|^2 ds \\ &+ 3\sup_{\frac{iT}{k} \leq \chi \leq t}\left|\int_{\frac{iT}{k}}^{\chi} \sigma(s, x^{n-1}(s))dW_s\right|^2. \end{aligned}$$

By the Burkholder–Davis–Gundy inequality we have

$$\begin{aligned} E\sup_{\frac{iT}{k} \leq \chi \leq t} |x^n(\chi)|^2 \leq& 3E\left|\xi\left(\frac{iT}{k}\right)\right|^2 + 3\left(t - \frac{iT}{k}\right)E\Bigg(\int\limits_{\frac{iT}{k}}^{t} |f(s, x^{n-1}(s), x_s^{n-1})|^2 ds\Bigg) \\ &+ 12E\int_{\frac{iT}{k}}^{t} \left|\sigma(s, x^{n-1}(s))\right|^2 ds. \end{aligned}$$

It follows from (H4) and $t - \frac{iT}{k} \leq T$ for $t \in I_i$ that

$$E \sup_{\frac{iT}{k} \leq \chi \leq t} |x^n(\chi)|^2 \leq 3E \left|\xi(\frac{iT}{k})\right|^2 + 3\gamma \left(t - \frac{iT}{k}\right) \left(\int_{\frac{iT}{k}}^{t} (1 + E|x^{n-1}(s)|^2 + E\|x_s^{n-1}\|^2) ds \right)$$
$$+ 12\gamma \int_{\frac{iT}{k}}^{t} \left(1 + E|x^{n-1}(s)|^2\right) ds$$
$$\leq 3\left[(1 + \gamma T^2) E\|\xi\|^2 + \gamma T^2 + 4\gamma T)\right]$$
$$+ 6\gamma(T+2) \int_{\frac{iT}{k}}^{t} \left(E \sup_{\frac{iT}{k} \leq \chi \leq s} |x^{n-1}(\chi)|^2 \right) ds,$$

which shows

$$\sup_{0 \leq j \leq n} E \sup_{\frac{iT}{k} \leq \chi \leq t} |x^j(\chi)|^2 \leq 3\left[(1 + \gamma T^2) E\|\xi\|^2 + \gamma T^2 + 4\gamma T)\right]$$
$$+ 6\gamma(T+2) \int_{\frac{iT}{k}}^{t} \left(\sup_{0 \leq j \leq n} E \sup_{\frac{iT}{k} \leq \chi \leq s} |x^j(\chi)|^2 \right) ds.$$

Using the classical Gronwall inequality we get that for $t \in I_i$

$$\sup_{0 \leq j \leq n} E \sup_{\frac{iT}{k} \leq \chi \leq t} |x^j(\chi)|^2 \leq 3\left[(1 + \gamma T^2) E\|\xi\|^2 + \gamma T^2 + 4\gamma T)\right] e^{6\gamma(T+2)(t - \frac{iT}{k})}$$
$$\leq 3\left[(1 + \gamma T^2) E\|\xi\|^2 + \gamma T^2 + 4\gamma T)\right] e^{6\gamma T(T+2)}.$$

In particular,

$$E \sup_{t \in I_i} |x^n(t)|^2 \leq 3\left[(1 + \gamma T^2) E\|\xi\|^2 + \gamma T^2 + 4\gamma T)\right] e^{6\gamma T(T+2)} = \tilde{C}_1(\xi), \ n \geq 1. \tag{5.5}$$

From (5.4) it follows that $\{x^n(t), n \geq 1, t \in I_i\}$ are adapted to $(\mathfrak{F}_t)_{t \in I_i}$ with continuous sample paths. Again by (5.4) and the Hölder inequality we get that for any $n \geq 1$, $m \geq 1$ and $t \in I_i$

$$\sup_{\frac{iT}{k} \leq \chi \leq t} |x^{n+m}(\chi) - x^n(\chi)|^2 \leq \sup_{\frac{iT}{k} \leq \chi \leq t} \left(2\left| \int_{\frac{iT}{k}}^{\chi} (f(s, x^{n+m-1}(s), x_s^{n+m-1}) - f(s, x^{n-1}(s), \right.\right.$$
$$\times x_s^{n-1}))ds|^2$$
$$\left. + 2\left| \int_{\frac{iT}{k}}^{\chi} (\sigma(s, x^{n+m-1}(s)) - \sigma(s, x^{n-1}(s)))dW_s \right|^2 \right)$$
$$\leq 2T \int_{\frac{iT}{k}}^{t} \left| f(s, x^{n+m-1}(s), x_s^{n+m-1}) - f(s, x^{n-1}(s), x_s^{n-1}) \right|^2 ds$$
$$+ 2 \sup_{\frac{iT}{k} \leq \chi \leq t} \left| \int_{\frac{iT}{k}}^{\chi} (\sigma(s, x^{n+m-1}(s)) - \sigma(s, x^{n-1}(s)))dW_s \right|^2.$$

By (H2), (H3*), the concavity of ρ and the Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned}E\sup_{\frac{iT}{k}\le\chi\le t}|x^{n+m}(\chi)-x^{n}(\chi)|^2 &\le 2TdLE\int_{\frac{iT}{k}}^{t}\left(|x^{n+m-1}(s)-x^{n-1}(s)|^2+\|x_s^{n+m-1}-x_s^{n-1}\|^2\right)ds\\ &\quad+8E\int_{\frac{iT}{k}}^{t}\left|\sigma(s,x^{n+m-1}(s))-\sigma(s,x^{n-1}(s))\right|^2ds\\ &\le 4TdL\int_{\frac{iT}{k}}^{t}\left(E\sup_{\frac{iT}{k}\le\chi\le s}|x^{n+m-1}(\chi)-x^{n-1}(\chi)|^2\right)ds\\ &\quad+8d\int_{\frac{iT}{k}}^{t}\rho\left(E\sup_{\frac{iT}{k}\le\chi\le s}|x^{n+m-1}(\chi)-x^{n-1}(\chi)|^2\right)ds\\ &\le C_2\int_{\frac{iT}{k}}^{t}\varrho\left(E\sup_{\frac{iT}{k}\le\chi\le s}|x^{n+m-1}(\chi)-x^{n-1}(\chi)|^2\right)ds, \end{aligned}\tag{5.6}$$

where $C_2=4TdL+8d$. The above inequality and (5.5) show that

$$E\sup_{\frac{iT}{k}\le\chi\le t}\left|x^{n+m}(\chi)-x^{n}(\chi)\right|^2\le C_3\Big(t-\frac{iT}{k}\Big),\tag{5.7}$$

where $C_3\triangleq C_2\varrho(4\tilde{C}_1(\xi))$.

Define two sequences of functions $\{\varphi_n(t)\}$ and $\{\widetilde{\varphi}_{n,m}(t)\}$ on I_i as follows:

$$\begin{aligned}\varphi_1(t)&=C_3\left(t-\frac{iT}{k}\right),\\ \varphi_{n+1}(t)&=C_2\int_{\frac{iT}{k}}^{t}\varrho(\varphi_n(s))ds,\quad n\ge1,\\ \widetilde{\varphi}_{n,m}(t)&=E\sup_{\frac{iT}{k}\le\chi\le t}|x^{n+m}(\chi)-x^{n}(\chi)|^2,\quad n\ge1,m\ge1.\end{aligned}$$

We claim that for every $n\ge1$ and $m\ge1$,

$$0\le\widetilde{\varphi}_{n,m}(t)\le\varphi_n(t)\le\varphi_{n-1}(t)\le\cdots\le\varphi_1(t),\quad t\in I_i.\tag{5.8}$$

First of all, By (5.2), the assumption $\frac{T}{k}<T_1$ and the monotonicity for ϱ, we have for any $t\in I_i$,

$$C_2\varrho\left(C_3(t-\frac{iT}{k})\right)\le C_3.\tag{5.9}$$

In view of (5.7) and (5.6), we have

$$\begin{aligned}\widetilde{\varphi}_{1,m}(t)&=E\sup_{\frac{iT}{k}\le\chi\le t}\left|x^{1+m}(\chi)-x^{1}(\chi)\right|^2\le C_3\left(t-\frac{iT}{k}\right)=\varphi_1(t).\\ \widetilde{\varphi}_{2,m}(t)&=E\sup_{\frac{iT}{k}\le\chi\le t}\left|x^{2+m}(\chi)-x^{2}(\chi)\right|^2\end{aligned}$$

$$\leq C_2 \int_{\frac{iT}{k}}^{t} \varrho \left(E \sup_{\frac{iT}{k} \leq \chi \leq s} |x^{1+m}(\chi) - x^1(\chi)|^2 \right) ds$$

$$\leq C_2 \int_{\frac{iT}{k}}^{t} \varrho \left(\varphi_1(s)\right) ds = \varphi_2(t).$$

But by (5.9) we also have

$$\varphi_2(t) = C_2 \int_{\frac{iT}{k}}^{t} \varrho(\varphi_1(s))ds \leq \varphi_1(t).$$

Now we have already shown that for $t \in I_i$,

$$\widetilde{\varphi}_{2,m}(t) \leq \varphi_2(t) \leq \varphi_1(t).$$

Next we assume that (5.8) holds for some $n \geq 2$, then by (5.6)

$$\begin{aligned}\widetilde{\varphi}_{n+1,m}(t) &\leq C_2 \int_{\frac{iT}{k}}^{t} \varrho \left(E \sup_{\frac{iT}{k} \leq \chi \leq s} |x^{n+m}(\chi) - x^n(\chi)|^2 \right) ds \\ &\leq C_2 \int_{\frac{iT}{k}}^{t} \varrho(\varphi_n(s))ds = \varphi_{n+1}(t) \\ &\leq C_2 \int_{\frac{iT}{k}}^{t} \varrho(\varphi_{n-1}(s))ds = \varphi_n(t),\end{aligned}$$

that is, (5.8) holds for $n+1$ as well. Consequently by induction (5.8) must hold for $n \geq 1$.

Now our purpose is to prove

$$E \sup_{t \in I_i} \left| x^l(t) - x^n(t) \right|^2 \to 0 \tag{5.10}$$

as $l, n \to \infty$. Note that for every $n \geq 1$, $\varphi_n(t)$ is increasing on I_i and for each t, $\varphi_n(t)$ is monotonically nonincreasing as $n \to \infty$. Hence we can define the function $\varphi(t)$ by $\varphi_n(t) \downarrow \varphi(t)$. It is easy to see that $\varphi(t)$ is continuous and increasing on I_i. By the definition of $\varphi_n(t)$ and $\varphi(t)$ we have

$$\varphi(t) = C_2 \int_{\frac{iT}{k}}^{t} \varrho(\varphi(s))ds, \quad t \in I_i.$$

The proof of Corollary 2.3 shows that $\varphi(t) = 0,\ t \in I_i$. Clearly $\varphi_n(\frac{(i+1)T}{k}) \downarrow 0$ as $n \to \infty$. Hence for any $\epsilon > 0$, there exists an integer $N \geq 1$ such that $\varphi_n(\frac{(i+1)T}{k}) < \epsilon$ whenever $n > N$. For any $m \geq 1$ and $n > N$, the above claim deduces that

$$E \sup_{t \in I_i} \left| x^{n+m}(t) - x^n(t) \right|^2 = \widetilde{\varphi}_{n,m}\left(\frac{(i+1)T}{k}\right) \leq \varphi_n\left(\frac{(i+1)T}{k}\right) < \epsilon.$$

So (5.10) holds. It follows that

$$\lim_{n,l\to\infty} E \sup_{t\in J_i\bigcup I_i} \left|x^n(t)-x^l(t)\right|^2=0. \tag{5.11}$$

For each $x(t)\in\mathbb{L}^2(\Omega, C(J_i\bigcup I_i,\mathbb{R}^d))$, define $\|x\|_\star \triangleq (E\sup_{t\in J_i\bigcup I_i}|x(t)|^2)^{\frac{1}{2}}$. Then the space $\mathbb{L}^2(\Omega, C(J_i\bigcup I_i,\mathbb{R}^d))$ is a Banach space. Consequently by (5.11) there exists $x(t)\in\mathbb{L}^2(\Omega, C(J_i\bigcup I_i,\mathbb{R}^d))$ such that

$$\lim_{n\to\infty} E \sup_{t\in J_i\bigcup I_i} |x^n(t)-x(t)|^2=0.$$

For any $\delta>0$, by Chebyshev's inequality we have

$$\lim_{n\to\infty}\mathbb{P}\left(\sup_{t\in J_i\bigcup I_i}|x^n(t)-x(t)|\geq\delta\right)=0.$$

By the definition of limit, there is a subsequence $\{n_k\}_{k=1}^{\infty}$ satisfying that

$$\mathbb{P}\left(\sup_{t\in J_i\bigcup I_i}|x^{n_k}(t)-x(t)|\geq\frac{1}{k}\right)\leq\frac{1}{2^k},\, k\geq 1.$$

The Borel–Cantelli Lemma shows that $x^{n_k}(t)$ converges to $x(t)$ uniformly on $J_i\bigcup I_i$ almost surely. It follows that $x(t),\ t\in I_i$ has continuous sample paths and is adapted to $\{\mathfrak{F}_t\}_{t\in I_i}$. Moreover, simply computation shows that

$$E\sup_{t\in I_i}\left|\int_{\frac{iT}{k}}^t \left(f(s,x^{n-1}(s),x_s^{n-1})-f(s,x(s),x_s)\right)ds+\int_{\frac{iT}{k}}^t\left(\sigma(s,x^{n-1}(s))\right.\right.$$
$$\left.\left.-\sigma(s,x(s))\right)dW(s)\right|^2\to 0$$

as $n\to\infty$. Letting $n\to\infty$, we conclude that $x(t),\ t\in I_i$ satisfies system (5.3). This completes the proof. □

Proof of Theorem 2.1 For any $\phi\in\mathcal{C}$ and $T>1$, applying Proposition 13 with $i=0,\ \xi=\phi$, we get the solution $x=x(t)$ for (2.1) on I_0, which is adapted to $\{\mathfrak{F}_t\}_{t\in I_0}$. Then again using Proposition 13 with $i=1,\ \xi=x_{\frac{T}{k}}$, we have a solution for (5.3) on I_1 adapted to $\{\mathfrak{F}_t\}_{t\in I_1}$. In this way, we have extended the solution for (2.1) to the interval $I_0\bigcup I_1$. Repeatedly applying Proposition 13 with $\xi=x_{\frac{iT}{k}}$ for $i=2,\ldots,k-1$ in order, we obtain that the existence for the strong solution for (2.1) on the interval $[0,T]$, which is adapted to $\{\mathfrak{F}_t\}_{t\in[0,T]}$. Since T is arbitrary, the global strong solution exists. Lemma II.2.1 in [17] guarantees that $x_t(\phi)$ is a $\mathcal{C}$-valued process adapted to $\{\mathfrak{F}_t\}_{t\geq 0}$ with continuous sample paths.

Finally we finish the proof of the uniqueness of solution to system (2.1). To this end, we assume that $\{x(t),t\geq 0\}$ and $\{x^*(t),t\geq 0\}$ are solutions to system (2.1). Using the above similar arguments, we have

$$E\sup_{0\leq t\leq T}|x(t)-x^*(t)|^2=0.$$

This shows that $\{x(t),t\geq 0\}$ and $\{x^*(t),t\geq 0\}$ are modifications of one another, and thus are indistinguishable. This completes the proof. □

Remark 14 Using Fatou Lemma and (5.5) one has

$$E \sup_{0 \le t \le T} |x(t)|^2 \le \widetilde{C}_1(\phi).$$

Hence by the pathwise continuity of $x(t)$ and Lebesgue's Theorem on dominated convergence one concludes that $t \to E \sup_{0 \le \chi \le t} |x(\chi)|^2$ is continuous.

Note We independently obtain the comparison theorem for SFDEs. This result was first presented in the Second International Conference on Recent Advances in Random Dynamical Systems, which held in Nanjing Normal University on June 20–23, 2011, and then in several international conferences. We submitted it to Stochastic Analysis and Applications on October 24, 2011 and withdrew the submission on August 18, 2014.

References

1. Agarwal, R.P., Deng, S., Zhang, W.: Generalization of a retarded Gronwall-like inequality and its applications. Appl. Math. Comput. **165**(3), 599–612 (2005)
2. Assing, S.: Comparison of systems of stochastic partial differential equations. Stoch. Proc. Appl. **82**(2), 259–282 (1999)
3. Buckdahn, R., Pardoux, E.: Monotonicity methods for white noise driven SPDE's. In: Pinsky, M. (ed.) Diffusion Processes and Related Problems in Analysis, pp. 219–233. Birkh$\breve{a}$user, Basel (1990)
4. Buckdahn, R., Peng, S.: Ergodic backward stochastic differential equations and associated partial differential equations. Prog. Probab. **45**, 73–85 (1999)
5. Chueshov, I.: Monotone Random Systems: Theory and Applications. Lecture Notes in Mathematics. Springer, Berlin (2002)
6. Chueshov, I., Scheutzow, M.: Invariance and monotonicity for stocastic delay differential equations. Discrete Contin. Dyn. Syst. Ser. B **18**(6), 1533–1554 (2013)
7. Donati-Martin, C., Pardoux, E.: White noise driven SPDEs with reflection. Probab. Theory Relat. Fields **95**(1), 1–24 (1993)
8. Geiß, C., Manthey, R.: Comparison theorems for stochastic differential equations in finite and infinite dimensions. Stoch. Proc. Appl. **53**(1), 23–35 (1994)
9. Hajek, B.: Mean stochastic comparison of diffusions. Z. Wahrscheinlichkeitstheorie verw. Gebiete **68**, 315–329 (1985)
10. Ikeda, N., Watanabe, S.: A comparison theorem for solutions of stochastic differential equations and its applications. Osaka J. Math. **14**(3), 619–633 (1977)
11. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. Noth-Holland, Amsterdam (1981)
12. Imkeller, P., Schmalfuss, B.: The conjugacy of stochastic and random differential equations and the existence of global attractors. J. Dyn. Differ. Equ. **13**(2), 215–249 (2001)
13. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, Berlin (1991)
14. Kotelenez, P.: Comparison methods for a class of function valued stochastic partial differential equations. Probab. Theory Relat. Fields **93**(1), 1–19 (1992)
15. Manthey, R., Zausinger, T.: Stochastic evolution equations in $L_p^{2\nu}$. Stoch. Stoch. Rep. **66**(1), 37–85 (1999)
16. Mao, X.: Adapted solutions of backward stochastic differential equations with non-Lipschitz cofficients. Stoch. Proc. Appl. **58**(2), 281–292 (1995)
17. Mohammed, S.E.A.: Stochastic Functional Differential Equations. Researsch Notes in Mathematics. Pitman, Boston (1984)
18. Peng, S., Zhu, X.: Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations. Stoch. Proc. Appl. **116**(3), 370–380 (2006)
19. Shiga, T.: Diffusion processes in population genetics. J. Math. Kyoto Univ. **21**(1), 133–151 (1981)
20. Skorohod, A.V.: Studies in the Theory of Random Processes. Addison-Wesley, Reading, MA (1965)
21. Smith, H.L.: Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems. Mathematical surveys and monographs, vol 41. American Mathematical Society, Providence, RI (1995)
22. Yamada, T.: On a comparison theorems for solutions of stochastic differential equations and its applications. J. Math. Kyoto Univ. **13**(3), 497–512 (1973)

23. Yamada, T., Ogura, Y.: On the strong comparison theorems for solutions of stochastic differential equations. Z. Wahrscheinlichkeitstheorie verw. Gebiete **56**(1), 3–19 (1981)
24. Yan, J.: A comparison theorem for semimartingale and its applications. Séminaire de Probabilités, XX, Lecture Notes in Mathematics, vol. 1204. Springer, Berlin (1986)
25. Yang, Z., Mao, X., Yuan, C.: Comparison theorem for one-dimensional stochastic hybrid delay systems. Syst. Control Lett. **57**(1), 56–63 (2008)