Mao and Mao *Advances in Difference Equations* (2016) 2016:77
DOI 10.1186/s13662-016-0802-x

Advances in Difference Equations
a SpringerOpen Journal

RESEARCH

Open Access

An averaging principle for neutral stochastic functional differential equations driven by Poisson random measure

Wei Mao[1*] and Xuerong Mao[2]

*Correspondence:
jsjysxx365@126.com
[1]School of Mathematics and Information Technology, Jiangsu Second Normal University, Nanjing, 210013, China
Full list of author information is available at the end of the article

Abstract

In this paper, we study the averaging principle for neutral stochastic functional differential equations (SFDEs) with Poisson random measure. By stochastic inequality, Burkholder-Davis-Gundy's inequality and Kunita's inequality, we prove that the solution of the averaged neutral SFDEs with Poisson random measure converges to that of the standard one in L^p sense and also in probability. Some illustrative examples are presented to demonstrate this theory.

Keywords: averaging principle; neutral SFDEs; L^p convergence; convergence in probability; Poisson random measure

1 Introduction

Since Krylov and Bogolyubov [1] put forward the averaging principles for dynamical systems in 1937, the averaging principles have received great attention and many people have devoted their efforts to the study of averaging principles of nonlinear dynamical systems. For example, the averaging principles for nonlinear ordinary differential equations (ODEs) can be found in [2, 3]. For the averaging principles of nonlinear partial differential equations (PDEs), we refer to [4, 5].

With the developing of stochastic analysis theory, many authors began to study the averaging principle for stochastic differential equations (SDEs). Khasminskii [6] first extended the averaging theory for ODEs to the case of stochastic differential equations (SDEs) and studied the averaging principle of SDEs driven by Brownian motion. After that, there grew an extensive literature on averaging principles for SDEs. Freidlin and Wentzell [7] provided a mathematically rigorous overview of fundamental stochastic averaging method. Golec and Ladde [8], Veretennikov [9], Khasminskii and Yin [10], Givon *et al.* [11] studied the averaging principle to stochastic differential systems in the sense of mean square and probability. On the other hand, Stoyanov and Bainov [12], Kolomiets and Melnikov [13], Givon [14], Xu *et al.* [15] established the averaging principle for stochastic differential equations with Lévy jumps. They proved that the solutions of averaged systems converge to the solutions of original systems in mean square under the Lipschitz conditions.

On the other hand, Yin and Ramachandran [16] studied the asymptotic properties of stochastic differential delay equation (SDDEs) with wideband noise perturbations. By adopting the martingale averaging techniques and the method of weak convergence, they showed that the underlying process $y^\varepsilon(t)$ converges weakly to a random process $x^\varepsilon(t)$ of SDDEs as $\varepsilon \to 0$. Tan and Lei [17] investigated the averaging method for a class of SDDEs with constant delay. Under non-Lipschitz conditions, they showed the convergence between the standard form and the averaged form of SDDEs. Furthermore, Xu *et al.* [18] and Mao *et al.* [19] also extended the convergence results [9, 12, 13, 15] to the case of stochastic functional differential equations (SFDEs) and SDDEs with variable delays, respectively.

SDDEs and SFDEs are well known to model problems from many areas of science and engineering, the future state of which is determined by the present and past states. In fact, many stochastic systems not only depend on the present and past states but also involve the derivatives with delays as well as the function itself. In this case, neutral SDDEs (SFDEs) has been used to described such systems. In the past few years, the theory of neutral SDDEs (SFDEs) has attracted more and more attention (see *e.g.* [20–25]). However, to the best of our knowledge, there is no research about using the averaging methods to obtain the approximate solutions to neutral SDDEs (SFDEs). In order to fill the gap, we will study the averaging principle of neutral SFDEs with Poisson random measure. By using the averaging method, we give the averaged form of neutral SFDEs (1) and show that the pth moment of solution to equation (7) is bounded. Then, applying the stochastic inequality, Burkholder-Davis-Gundy's inequality and Kunita's inequality, we prove that the solution of the averaged neutral SFDEs with Poisson random measure (7) converges to that of the standard one (6) in L^p sense and also in probability under the Lipschitz conditions. Meantime, we relax the Lipschitz condition and obtain the averaging principle for neutral SFDEs with Poisson random measure (1) under non-Lipschitz conditions. It should be pointed out that the previous works [6, 9, 12–15, 17, 18] on averaging principle mainly discussed L^2 strong convergence for stochastic differential equations and they do not imply L^p ($p > 2$) strong convergence. Moreover, since the neutral term is involved, the proof of the main results are much more technical. The results obtained of this paper are a generalization and improvement of some results in [6, 9, 12, 13, 15, 17, 18].

The rest of this paper is organized as follows. In Section 2, we introduce some preliminaries and establish our main results. In Section 3, some lemmas will be given which will be crucial in the proof of the main results, Theorems 2.2 and 2.4. Section 4 is devoted to the proof of the main results. Finally, two illustrative examples will be given in Section 5.

2 Averaging principle and main results

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with some filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. Here $w(t)$ is an m-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Let $\tau > 0$, and $D([-\tau, 0]; R^n)$ denote the family of all right-continuous functions with left-hand limits φ from $[-\tau, 0] \to R^n$. The space $D([-\tau, 0]; R^n)$ is assumed to be equipped with the norm $\|\varphi\| = \sup_{-\tau\leq t\leq 0} |\varphi(t)|$. $D^b_{\mathcal{F}_0}([-\tau, 0]; R^n)$ denotes the family of all almost surely bounded, $\mathcal{F}_0$-measurable, $D([-\tau, 0]; R^n)$ valued random variable $\xi = \{\xi(\theta) : -\tau \leq \theta \leq 0\}$. For any $p \geq 2$, let $\mathcal{L}^p_{\mathcal{F}_0}([-\tau, 0]; R^n)$ denote the family of all $\mathcal{F}_0$ measurable, $D([-\tau, 0]; R^n)$-valued random variables $\varphi = \{\varphi(\theta) : -\tau \leq \theta \leq 0\}$ such that $E\sup_{-\tau\leq\theta\leq 0} |\varphi(\theta)|^p < \infty$.

Let $\{\bar{p} = \bar{p}(t), t \geq 0\}$ be a stationary $\mathcal{F}_t$-adapted and R^n-valued Poisson point process. Then, for $A \in \mathcal{B}(R^n - \{0\})$, $0 \notin$ the closure of A, we define the Poisson counting measure

N associated with $\bar{p}$ by

$$N\big((0,t]\times A\big):=\#\big\{0<s\leq t,\bar{p}(s)\in A\big\}=\sum_{t_0<s\leq t}I_A\big(\bar{p}(s)\big),$$

where # denotes the cardinality of the set $\{\cdot\}$. For simplicity, we denote $N(t,A):=N((0,t]\times A)$. It is well known that there exists a σ-finite measure π such that

$$E\big[N(t,A)\big]=\pi(A)t,\qquad P\big(N(t,A)=n\big)=\frac{\exp(-t\pi(A))(\pi(A)t)^n}{n!}.$$

This measure π is called the Lévy measure. Moreover, by Doob-Meyer's decomposition theorem, there exists a unique $\{\mathcal{F}_t\}$-adapted martingale $\tilde{N}(t,A)$ and a unique $\{\mathcal{F}_t\}$-adapted natural increasing process $\hat{N}(t,A)$ such that

$$N(t,A)=\tilde{N}(t,A)+\hat{N}(t,A),\quad t>0.$$

Here $\tilde{N}(t,A)$ is called the compensated Poisson random measure and $\hat{N}(t,A)=\pi(A)t$ is called the compensator. For more details on Poisson point process and Lévy jumps, see [26–28].

Consider the following neutral SFDEs with Poisson random measure

$$d\big[x(t)-D(x_t)\big]=f(t,x_t)\,dt+g(t,x_t)\,dw(t)+\int_Z h(t,x_t,v)N(dt,dv),\tag{1}$$

where $x_t=\{x(t+\theta):-\tau\leq\theta\leq 0\}$ is regarded as a $D([-\tau,0];R^n)$-valued stochastic process. $f:[0,T]\times D([-\tau,0];R^n)\to R^n$, $g:[0,T]\times D([-\tau,0];R^n)\to R^{n\times m}$ and $h:[0,T]\times D([-\tau,0];R^n)\times Z\to R^n$ are both Borel-measurable functions. The initial condition x_0 is defined by

$$x_0=\xi=\big\{\xi(t):-\tau\leq t\leq 0\big\}\in\mathcal{L}^2_{\mathcal{F}_0}\big([-\tau,0];R^n\big),$$

that is, ξ is an $\mathcal{F}_0$-measurable $D([-\tau,0];R^n)$-valued random variable and $E\|\xi\|^2<\infty$.

To study the averaging method of equation (1), we need the following assumptions.

Assumption 2.1 Let $D(0)=0$ and for all $\varphi,\psi\in D([-\tau,0];R^n)$, there exists a constant $k_0\in(0,1)$ such that

$$\big|D(\varphi)-D(\psi)\big|\leq k_0\|\varphi-\psi\|.\tag{2}$$

Assumption 2.2 For all $\varphi,\psi\in D([-\tau,0];R^n)$ and $t\in[0,T]$, there exist two positive constants k_1, k_2 such that

$$\big|f(t,\varphi)-f(t,\psi)\big|^2\vee\big|g(t,\varphi)-g(t,\psi)\big|^2\leq k_1\|\varphi-\psi\|^2$$

and

$$\int_Z\big|h(t,\varphi,v)-h(t,\varphi,v)\big|^p\pi(dv)\leq k_2\|\varphi-\psi\|^p,\quad p\geq 2.\tag{3}$$

Assumption 2.3 For all $\varphi \in D([-\tau,0];R^n)$ and $t \in [0,T]$, there exist two positive constants k_3, k_4 such that

$$\left|f(t,\varphi)\right|^2 \vee \left|g(t,\varphi)\right|^2 \leq k_3\left(1+\|\varphi\|^2\right)$$

and

$$\int_Z \left|h(t,\varphi,v)\right|^p \pi(dv) \leq k_4\left(1+\|\varphi\|^p\right), \quad p \geq 2. \tag{4}$$

Let $C^{2,1}([-\tau,T]\times R^n;R_+)$ denote the family of all nonnegative functions $V(t,x)$ defined on $[-\tau,T]\times R^n$ which are continuously twice differentiable in x and once differentiable in t. For each $V \in C^{2,1}([-\tau,T]\times R^n;R_+)$, define an operator LV by

$$\begin{aligned} LV(t,x,y) &= V_t\big(t,x-D(y)\big) + V_x\big(t,x-D(y)\big)f(t,y) \\ &\quad + \frac{1}{2}\operatorname{trace}\big[g^\top(t,y)V_{xx}\big(t,x-D(y)\big)g(t,y)\big] \\ &\quad + \int_Z \big[V\big(t,x-D(y)+h(t,y,v)\big) - V\big(t,x-D(y)\big)\big]\pi(dv), \end{aligned} \tag{5}$$

where

$$\begin{aligned} &V_t(t,x) = \frac{\partial V(t,x)}{\partial t}, \qquad V_x(t,x) = \left(\frac{\partial V(t,x)}{\partial x_1},\ldots,\frac{\partial V(t,x)}{\partial x_n}\right), \\ &V_{xx}(t,x) = \left(\frac{\partial^2 V(t,x)}{\partial x_i\,\partial x_j}\right)_{n\times n}. \end{aligned}$$

Similar to the proof of [29], we have the following existence result.

Theorem 2.1 *If Assumptions* 2.1-2.3 *hold*, *equation* (1) *has a unique solution in the sense of* L^p.

Now, we study the averaging principle for neutral SFDEs with Poisson random measure. Let us consider the standard form of equation (1)

$$\begin{aligned} x_\varepsilon(t) &= x(0) + D(x_{\varepsilon,t}) - D(x_0) + \varepsilon\int_0^t f(s,x_{\varepsilon,s})\,ds \\ &\quad + \sqrt{\varepsilon}\int_0^t g(s,x_{\varepsilon,s})\,dw(s) + \sqrt{\varepsilon}\int_0^t\int_Z h(s,x_{\varepsilon,s},v)N(ds,dv), \end{aligned} \tag{6}$$

where the coefficients f, g, and h have the same assumptions as in (3), (4), and $\varepsilon \in [0,\varepsilon_0]$ is a positive small parameter with ε_0 is a fixed number.

Let $\bar{f}(x): D([-\tau,0];R^n) \to R^n$, $\bar{g}(x): D([-\tau,0];R^n) \to R^{n\times m}$ and $\bar{h}(x,v): D([-\tau,0];R^n) \times Z \to R^n$ be measurable functions, satisfying Assumptions 2.2 and 2.3. We also assume that the following condition is satisfied.

Assumption 2.4 For any $\varphi \in D([-\tau,0];R^n)$ and $p \geq 2$, there exist three positive bounded functions $\psi_i(T_1)$, $i = 1,2,3$, such that

$$\frac{1}{T_1}\int_0^{T_1} \left|f(t,\varphi) - \bar{f}(\varphi)\right|^p dt \leq \psi_1(T_1)\left(1+\|\varphi\|^p\right),$$

$$\frac{1}{T_1}\int_0^{T_1}\bigl|g(t,\varphi)-\bar{g}(\varphi)\bigr|^p\,dt \le \psi_2(T_1)\bigl(1+\|\varphi\|^p\bigr),$$

and

$$\frac{1}{T_1}\int_0^{T_1}\int_Z\bigl|h(t,\varphi,v)-\bar{h}(\varphi,v)\bigr|^p\pi(dv)\,dt \le \psi_3(T_1)\bigl(1+\|\varphi\|^p\bigr),$$

where $\lim_{T_1\to\infty}\psi_i(T_1)=0$.

Then we have the averaging form of the standard neutral SFDEs with Poisson random measure

$$\begin{aligned} y_\varepsilon(t) &= y(0)+D(y_{\varepsilon,t})-D(y_0)+\varepsilon\int_0^t \bar{f}(y_{\varepsilon,s})\,ds+\sqrt{\varepsilon}\int_0^t \bar{g}(y_{\varepsilon,s})\,dw(s) \\ &\quad +\sqrt{\varepsilon}\int_0^t\int_Z \bar{h}(y_{\varepsilon,s},v)N(ds,dv), \end{aligned} \tag{7}$$

where $y(0)=x(0)$, $y_0=x_0$.

Obviously, under Assumptions 2.1-2.3, the standard neutral SFDEs with Poisson random measure (6) and the averaged one (7) have a unique solutions in L^p, respectively.

Now, we present our main results which are used for revealing the relationship between the processes $x_\varepsilon(t)$ and $y_\varepsilon(t)$.

Theorem 2.2 *Let Assumptions* 2.1-2.4 *hold. For a given arbitrary small number* $\delta_1>0$ *and* $p\ge 2$, *there exist* $L>0$, $\varepsilon_1\in(0,\varepsilon_0]$, *and* $\beta\in(0,1)$ *such that*

$$E\bigl|x_\varepsilon(t)-y_\varepsilon(t)\bigr|^p \le \delta_1, \quad \forall t\in\bigl[0,L\varepsilon^{-\beta}\bigr], \tag{8}$$

for all $\varepsilon\in(0,\varepsilon_1]$.

The proof of this theorem will be shown in Section 4.

Remark 2.1 In particular, when $p=2$, we see that the solution of the averaged neutral SFDEs with Poisson random measure converges to that of the standard one in second moment.

With Theorem 2.2, it is easy to show the convergence in probability between the processes $x_\varepsilon(t)$ and $y_\varepsilon(t)$.

Corollary 2.1 *Let Assumptions* 2.1-2.4 *hold. For a given arbitrary small number* $\delta_2>0$, *there exists* $\varepsilon_2\in[0,\varepsilon_0]$ *such that for all* $\varepsilon\in(0,\varepsilon_2]$, *we have*

$$\lim_{\varepsilon\to 0}P\Bigl(\sup_{0<t\le L\varepsilon^{-\beta}}\bigl|x_\varepsilon(t)-y_\varepsilon(t)\bigr|>\delta_2\Bigr)=0,$$

where L *and* β *are defined by Theorem* 2.2.

Proof By Theorem 2.2 and the Chebyshev inequality, for any given number $\delta_2 > 0$, we can obtain

$$P\Big(\sup_{0<t\leq L\varepsilon^{-\beta}}\big|x_\varepsilon(t)-y_\varepsilon(t)\big| > \delta_2\Big) \leq \frac{1}{\delta_2^p}E\Big(\sup_{0<t\leq L\varepsilon^{-\beta}}\big|x_\varepsilon(t)-y_\varepsilon(t)\big|^p\Big) \leq \frac{cL\varepsilon^{1-\beta}}{\delta_2^p}.$$

Let $\varepsilon \to 0$, and the required result follows. □

Next, we extend the averaging principle for neutral SFDEs with Poisson random measure to the case of non-Lipschitz condition.

Assumption 2.5 Let $k(\cdot)$, $\rho(\cdot)$ be two concave nondecreasing functions from R_+ to R_+ such that $k(0) = \rho(0) = 0$ and $\int_{0^+} \frac{u^{p-1}}{k^p(u)+\rho^p(u)}\,du = \infty$. For all $\varphi, \psi \in D([-\tau,0];R^n)$, $t \in [0,T]$, and $p \geq 2$, then

$$\begin{aligned} &\big|f(t,\varphi)-f(t,\psi)\big| \vee \big|g(t,\varphi)-g(t,\psi)\big| \leq k\big(\|\varphi-\psi\|\big), \\ &\Big[\int_Z\big|h(t,\varphi,v)-h(t,\psi,v)\big|^p\pi(dv)\Big]^{\frac{1}{p}} \leq \rho\big(\|\varphi-\psi\|\big). \end{aligned} \tag{9}$$

Remark 2.2 As we know, the existence and uniqueness of solution for NSFDEs under the above assumptions were proved by Bao and Hou [30], Ren and Xia [31] and Wei and Cai [32]. If $k(u) = \rho(u) = Lu$, then Assumption 2.5 reduces to the Lipschitz conditions (3). In other words, Assumption 2.5 is much weaker than Assumption 2.2.

Theorem 2.3 *If Assumptions* 2.1 *and* 2.5 *hold, then there exists a unique solution to equation* (1) *in the sense of* L^p.

Proof The proof is similar to Ren and Xia [31] and Wei and Cai [32], and we thus omit here. □

Theorem 2.4 *Let Assumptions* 2.1, 2.4, *and* 2.5 *hold. For a given arbitrary small number* $\delta_3 > 0$, *there exist* $L > 0$, $\varepsilon_3 \in (0,\varepsilon_0]$, *and* $\beta \in (0,1)$ *such that*

$$E\big|x_\varepsilon(t)-y_\varepsilon(t)\big|^2 \leq \delta_3, \quad \forall t \in \big[0, L\varepsilon^{-\beta}\big], \tag{10}$$

for all $\varepsilon \in (0,\varepsilon_3]$.

Proof The proof of this theorem will be shown in Section 4. □

Similarly, with Theorem 2.4, we can show that the convergence in probability of the standard solution of equation (6) and averaged solution of equation (7).

Corollary 2.2 *Let Assumptions* 2.1, 2.4, *and* 2.5 *hold. For a given arbitrary small number* $\delta_4 > 0$, *there exists* $\varepsilon_4 \in [0,\varepsilon_0]$ *such that for all* $\varepsilon \in (0,\varepsilon_4]$, *we have*

$$\lim_{\varepsilon\to 0} P\Big(\sup_{0<t\leq L\varepsilon^{-\beta}}\big|x_\varepsilon(t)-y_\varepsilon(t)\big| > \delta_4\Big) = 0,$$

where L *and* β *are defined by Theorem* 2.4.

Remark 2.3 If jump term $h = \tilde{h} = 0$, then equation (1) and (36) will become neutral SFDEs (SDDEs) which have been investigated by [21–25]. Under our assumptions, we can show that the solution of the averaged neutral SFDEs (SDDEs) converges to that of the standard one in pth moment and in probability.

Remark 2.4 If the neutral term $D(\cdot) = 0$ and $\tilde{D}(\cdot) = 0$, then equation (1) and (36) will reduce to SFDEs (SDDEs) with jumps which have been studied by [18, 19]. Hence, the corresponding results in [18, 19] are generalized and improved.

3 Some useful lemmas

In order to prove our main results, we need to introduce the following lemmas.

Lemma 3.1 *Let $p > 1$ and $a, b \in R^n$. Then, for $\epsilon > 0$,*

$$|a + b|^p \le \bigl[1 + \epsilon^{\frac{1}{p-1}}\bigr]^{p-1}\biggl(|a|^p + \frac{|b|^p}{\epsilon}\biggr).$$

Lemma 3.2 *Let $p > 2$ and $a, b > 0$. Then, for $\epsilon > 0$,*

$$a^{p-1}b \le \frac{\epsilon(p-1)}{p}a^p + \frac{1}{p\epsilon^{p-1}}b^p, \qquad a^{p-2}b^2 \le \frac{\epsilon(p-2)}{p}a^p + \frac{1}{p\epsilon^{\frac{p-2}{2}}}b^p.$$

Lemma 3.3 *Let $p > 2$ and $a, b \in R^n$. Then, for any $\delta \in (0,1)$,*

$$|a + b|^p \le \frac{|a|^p}{(1-\delta)^{p-1}} + \frac{|b|^p}{\delta^{p-1}}.$$

Lemma 3.4 *Let $\phi : R_+ \times Z \to R^n$ and assume that*

$$\int_0^t \int_Z \bigl|\phi(s,v)\bigr|^p \pi(dv)\,ds < \infty, \quad p \ge 2.$$

Then there exists $D_p > 0$ such that

$$E\biggl(\sup_{0\le t\le u}\biggl|\int_0^t \int_Z \phi(s,v)\tilde{N}(ds,dv)\biggr|^p\biggr) \le D_p\biggl\{E\biggl(\int_0^u \int_Z \bigl|\phi(s,v)\bigr|^2 \pi(dv)\,ds\biggr)^{\frac{p}{2}} + E\int_0^u \int_Z \bigl|\phi(s,v)\bigr|^p \pi(dv)\,ds\biggr\}.$$

The proof of Lemma 3.1 and Lemma 3.2 can be found in [33], the proof of Lemma 3.4 can be found in [26, 28] and the proof of Lemma 3.3 can be obtained from Lemma 3.1 by putting $\epsilon = \frac{\delta}{1-\delta}$. The following lemma shows that if the initial data are in L^p ($p \ge 2$) then the solution of averaged neutral SFDEs with Poisson random measure will be in L^p.

Lemma 3.5 *Let Assumptions* 2.1 *and* 2.3 *hold. If the initial data $\xi \in \mathcal{L}^p_{\mathcal{F}_0}([-\tau,0];R^n)$ for some $p \ge 2$, then for any $t \ge 0$, the unique solution $y_\varepsilon(t)$ of equation* (7) *has the property that*

$$E\sup_{-\tau\le s\le t}\bigl|y_\varepsilon(s)\bigr|^p \le C, \tag{11}$$

where $C = [(1+\tilde{C})E\|\xi\|^p + \frac{2\bar{C}}{(1-k_0)^p}T]e^{\frac{2\bar{C}}{(1-k_0)^p}T}$.

Proof By the Itô formula to $V(t, y_\varepsilon(t) - D(y_{\varepsilon,t})) = |y_\varepsilon(t) - D(y_{\varepsilon,t})|^p$, we obtain

$$\begin{aligned}
&\bigl|y_\varepsilon(t) - D(y_{\varepsilon,t})\bigr|^p \\
&\quad = \bigl|y_\varepsilon(0) - D(y_0)\bigr|^p + \int_0^t LV\bigl(y_\varepsilon(s), y_{\varepsilon,s}, s\bigr)\,ds \\
&\qquad + p\sqrt{\varepsilon}\int_0^t \bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^{p-2}\bigl[y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr]^\top \bar{g}(y_{\varepsilon,s})\,dw(s) \\
&\qquad + \int_0^t \int_Z \bigl\{\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s}) + \sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s}, v)\bigr|^p \\
&\qquad - \bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^p\bigr\}\tilde{N}(ds, du),
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
LV(x, y, t) = {}& p\varepsilon\bigl|x - D(y)\bigr|^{p-2}\bigl[x - D(y)\bigr]^\top \bar{f}(y) \\
&+ \frac{p(p-1)}{2}\varepsilon\bigl|x - D(y)\bigr|^{p-2}\bigl|\bar{g}(y)\bigr|^2 \\
&+ \int_Z \bigl[\bigl|x - D(y) + \sqrt{\varepsilon}\bar{h}(y, v)\bigr|^p - \bigl|x - D(y)\bigr|^p\bigr]\pi(dv).
\end{aligned}$$

Taking the expectation on both sides of (12), one gets

$$\begin{aligned}
&E\sup_{0\le s\le t}\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^p \\
&\quad \le E\sup_{0\le s\le t}\bigl|y_\varepsilon(0) - D(y_0)\bigr|^p + E\sup_{0\le s\le t}\int_0^s p\varepsilon\bigl|y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr|^{p-2} \\
&\qquad \times \bigl[y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr]^\top \bar{f}(y_{\varepsilon,\sigma})\,d\sigma \\
&\qquad + E\sup_{0\le s\le t}\int_0^s \frac{p(p-1)}{2}\varepsilon\bigl|y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr|^{p-2}\bigl|\bar{g}(y_{\varepsilon,\sigma})\bigr|^2\,d\sigma \\
&\qquad + E\sup_{0\le s\le t}\int_0^s p\sqrt{\varepsilon}\bigl|y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr|^{p-2}\bigl[y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr]^\top \bar{g}(y_{\varepsilon,\sigma})\,dw(\sigma) \\
&\qquad + E\sup_{0\le s\le t}\int_0^s \int_Z \bigl\{\bigl|y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma}) + \sqrt{\varepsilon}\bar{h}(y_{\varepsilon,\sigma}, v)\bigr|^p \\
&\qquad - \bigl|y_\varepsilon(\sigma) - D(y_{\varepsilon,\sigma})\bigr|^p\bigr\}N(d\sigma, du) \\
&\quad = E\sup_{0\le s\le t}\bigl|y_\varepsilon(0) - D(y_0)\bigr|^p + \sum_{i=1}^4 I_i.
\end{aligned} \tag{13}$$

By Lemma 3.1 and Assumption 2.1, we get

$$\begin{aligned}
E\sup_{0\le s\le t}\bigl|y_\varepsilon(0) - D(y_0)\bigr|^p &\le \bigl[1 + \epsilon_1^{\frac{1}{p-1}}\bigr]^{p-1}\biggl(\bigl|y_\varepsilon(0)\bigr|^p + \frac{|D(y_0)|^p}{\epsilon_1}\biggr) \\
&\le \bigl[1 + \epsilon_1^{\frac{1}{p-1}}\bigr]^{p-1}\biggl(\bigl|y_\varepsilon(0)\bigr|^p + \frac{k_0^p\|y_0\|^p}{\epsilon_1}\biggr).
\end{aligned}$$

Letting $\epsilon_1 = k_0^{p-1}$,

$$E\sup_{0\leq s\leq t}\left|y_\varepsilon(0) - D(y_0)\right|^p \leq (1+k_0)^p E\|\xi\|^p. \tag{14}$$

Recalling Lemma 3.2, there exists a $\epsilon_2 > 0$ such that

$$\begin{aligned} I_1 &\leq p\varepsilon E\int_0^t \left[\frac{\epsilon_2(p-1)}{p}\left|y_\varepsilon(s) - D(y_{\varepsilon,s})\right|^p + \frac{1}{p\epsilon_2^{p-1}}\left|\bar{f}(y_{\varepsilon,s})\right|^p\right] ds \\ &\leq p\varepsilon E\int_0^t \left[\frac{\epsilon_2(p-1)}{p}(1+k_0)^p\|y_{\varepsilon,s}\|^p + \frac{1}{p\epsilon_2^{p-1}}\left|\bar{f}(y_{\varepsilon,s})\right|^p\right] ds. \end{aligned} \tag{15}$$

By Assumption 2.3 and the basic inequality, we get

$$\left|\bar{f}(y_{\varepsilon,s})\right|^p \leq (2k_3)^{\frac{p}{2}}\left(1 + \|y_{\varepsilon,s}\|^p\right). \tag{16}$$

Letting $\epsilon_2 = \frac{\sqrt{2k_3}}{1+k_0}$, it follows from (15) and (16) that

$$I_1 \leq p\varepsilon\sqrt{2k_3}(1+k_0)^{p-1}E\int_0^t \left(1 + \|y_{\varepsilon,s}\|^p\right) ds. \tag{17}$$

By Lemma 3.2 and Assumption 2.3, we obtain

$$\begin{aligned} I_2 &\leq \frac{p(p-1)}{2}\varepsilon E\int_0^t \left[\frac{\epsilon_3(p-2)}{p}|y_\varepsilon(s) - D(y_{\varepsilon,s})|^p + \frac{2}{p\epsilon_3^{\frac{p-2}{2}}}\left|\bar{g}(y_{\varepsilon,s})\right|^p\right] ds \\ &\leq \frac{p(p-1)}{2}\varepsilon E\int_0^t \left[\frac{\epsilon_3(p-2)}{p}(1+k_0)^p\|y_{\varepsilon,s}\|^p + \frac{(2k_3)^{\frac{p}{2}}}{p\epsilon_3^{\frac{p-2}{2}}}\left(1 + \|y_{\varepsilon,s}\|^p\right)\right] ds. \end{aligned}$$

Letting $\epsilon_3 = \frac{2k_3}{(1+k_0)^2}$,

$$I_2 \leq (p-1)^2\varepsilon k_3(1+k_0)^{p-2}E\int_0^t \left(1 + \|y_{\varepsilon,s}\|^p\right) ds. \tag{18}$$

For the estimation of I_3: by the Burkholder-Davis-Gundy's inequality, there exists a positive constant C_p such that

$$I_3 \leq \sqrt{\varepsilon}C_p E\left[\int_0^t \left|y_\varepsilon(s) - D(y_{\varepsilon,s})\right|^{2p-2}\left|\bar{g}(y_{\varepsilon,s})\right|^2 ds\right]^{\frac{1}{2}}.$$

Further, by the Young inequality and Assumption 2.3, we deduce that

$$\begin{aligned} I_3 &\leq \frac{1}{2}E\sup_{0\leq s\leq t}\left|y_\varepsilon(s) - D(y_{\varepsilon,s})\right|^p \\ &\quad + \varepsilon C_p^2 k_3(1+k_0)^{p-2}E\int_0^t \left(1 + \|y_{\varepsilon,s}\|^p\right) ds. \end{aligned} \tag{19}$$

Finally, we will estimate I_4. Note $N(dt,dv) = \tilde{N}(dt,dv) + \pi(dv)\,dt$ and $\tilde{N}(dt,dv)$ is a martingale, one has

$$I_4 \le E\int_0^t \int_Z \bigl[\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s}) + \sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s},v)\bigr|^p - \bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^p\bigr]\pi(dv)\,ds.$$

By the mean value theorem, we obtain

$$I_4 \le pE\int_0^t \int_Z \bigl[\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s}) + \theta\sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s},v)\bigr|^{p-1}\bigl|\sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s},v)\bigr|\bigr]\pi(dv)\,ds,$$

where $|\theta| \le 1$. This, together with the basic inequality $|a+b|^{p-1} \le 2^{p-2}(|a|^{p-1} + |b|^{p-1})$, implies that

$$I_4 \le pCE\int_0^t \int_Z \bigl[\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^{p-1}\bigl|\sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s},v)\bigr| + \bigl|\sqrt{\varepsilon}\bar{h}(y_{\varepsilon,s},v)\bigr|^p\bigr]\pi(dv)\,ds.$$

By Lemma 3.2, Assumptions 2.1 and 2.3, it follows that

$$I_4 \le pC\bigl[\bigl(k_4 + \pi(Z)\bigr)(1+k_0)^{p-1}\sqrt{\varepsilon} + k_4\sqrt{\varepsilon}^p\bigr]E\int_0^t \bigl(1 + \|y_{\varepsilon,s}\|^p\bigr)\,ds. \tag{20}$$

Combining (14), (17)-(20), we obtain

$$E\sup_{0\le s\le t}\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^p \le 2(1+k_0)^p E\|\xi\|^p + 2\bar{C}\int_{t_0}^t E\bigl(1 + \|y_{\varepsilon,s}\|^p\bigr)\,ds,$$

where

$$\begin{aligned}\bar{C} &= \bigl[(p-1)^2 + C_p^2\bigr]\varepsilon k_3(1+k_0)^{p-2} + p\varepsilon\sqrt{2k_3}(1+k_0)^{p-1} \\ &\quad + pC\bigl[\bigl(k_4 + \pi(Z)\bigr)(1+k_0)^{p-1}\sqrt{\varepsilon} + k_4\sqrt{\varepsilon}^p\bigr].\end{aligned}$$

On the other hand, by Lemma 3.1, we have

$$\begin{aligned}E\sup_{0\le s\le t}\bigl|y_\varepsilon(s)\bigr|^p &\le \frac{k_0}{1-k_0}E\|\xi\|^p + \frac{1}{(1-k_0)^p}E\sup_{t_0\le s\le t}\bigl(\bigl|y_\varepsilon(s) - D(y_{\varepsilon,s})\bigr|^p\bigr) \\ &\le \tilde{C}E\|\xi\|^p + \frac{2\bar{C}}{(1-k_0)^p}T + \frac{2\bar{C}}{(1-k_0)^p}\int_{t_0}^t E\|y_{\varepsilon,s}\|^p\,ds,\end{aligned}$$

where $\tilde{C} = \frac{k_0}{1-k_0} + \frac{2(1+k_0)^p}{(1-k_0)^p}$. Consequently,

$$E\sup_{-\tau\le s\le t}\bigl|y_\varepsilon(s)\bigr|^p \le (1+\tilde{C})E\|\xi\|^p + \frac{2\bar{C}}{(1-k_0)^p}T + \frac{2\bar{C}}{(1-k_0)^p}\int_{t_0}^t E\Bigl(\sup_{-\tau\le\sigma\le s}\bigl|y_\varepsilon(\sigma)\bigr|^p\Bigr)\,ds.$$

Therefore, we apply the Gronwall inequality to get

$$E\sup_{-\tau\le s\le t}\bigl|y_\varepsilon(s)\bigr|^p \le \biggl[(1+\tilde{C})E\|\xi\|^p + \frac{2\bar{C}}{(1-k_0)^p}T\biggr]e^{\frac{2\bar{C}}{(1-k_0)^p}T}.$$

The proof is complete. □

4 Proof of main results

Proof of Theorem 2.2 By Lemma 3.3, it follows that

$$\begin{aligned}
\left|x_\varepsilon(t)-y_\varepsilon(t)\right|^p &= \left|x_\varepsilon(t)-y_\varepsilon(t)-\left[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\right]+\left[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\right]\right|^p \\
&\leq \frac{|x_\varepsilon(t)-y_\varepsilon(t)-[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})]|^p}{(1-\delta)^{p-1}} \\
&\quad+\frac{|D(x_{\varepsilon,t})-D(y_{\varepsilon,t})|^p}{\delta^{p-1}}.
\end{aligned}\tag{21}$$

Letting $\delta = k_0$ and taking the expectation on both sides of (21), we have

$$\begin{aligned}
E\sup_{0\leq t\leq u}|x_\varepsilon(t)-y_\varepsilon(t)|^p &\leq \frac{E\sup_{0\leq t\leq u}|x_\varepsilon(t)-y_\varepsilon(t)-[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})]|^p}{(1-k_0)^{p-1}} \\
&\quad+k_0E\sup_{0\leq t\leq u}\left|x_\varepsilon(t)-y_\varepsilon(t)\right|^p.
\end{aligned}$$

Consequently,

$$E\sup_{0\leq t\leq u}\left|x_\varepsilon(t)-y_\varepsilon(t)\right|^p \leq \frac{E\sup_{0\leq t\leq u}|x_\varepsilon(t)-y_\varepsilon(t)-[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})]|^p}{(1-k_0)^p},\tag{22}$$

where $k_0\in(0,1)$. Next, we will estimate $E\sup_{0\leq t\leq u}|x_\varepsilon(t)-y_\varepsilon(t)-[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})]|^p$. From (6) and (7), we have

$$\begin{aligned}
&x_\varepsilon(t)-y_\varepsilon(t)-\left[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\right] \\
&\quad=\varepsilon\int_0^t\left[f(s,x_{\varepsilon,s})-\bar{f}(y_{\varepsilon,s})\right]ds \\
&\qquad+\sqrt{\varepsilon}\int_0^t\left[g(s,x_{\varepsilon,s})-\bar{g}(y_{\varepsilon,s})\right]dw(s) \\
&\qquad+\sqrt{\varepsilon}\int_0^t\int_Z\left[h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right]N(ds,dv).
\end{aligned}$$

Using the elementary inequality $|a+b+c|^p\leq 3^{p-1}(|a|^p+|b|^p+|c|^p)$, it follows that for any $u\in[0,T]$,

$$\begin{aligned}
&E\sup_{0\leq t\leq u}\left|x_\varepsilon(t)-y_\varepsilon(t)-\left[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\right]\right|^p \\
&\quad\leq 3^{p-1}\varepsilon^pE\sup_{0\leq t\leq u}\left|\int_0^t\left[f(s,x_{\varepsilon,s})-\bar{f}(y_{\varepsilon,s})\right]ds\right|^p \\
&\qquad+3^{p-1}\varepsilon^{\frac{p}{2}}E\sup_{0\leq t\leq u}\left|\int_0^t\left[g(s,x_{\varepsilon,s})-\bar{g}(y_{\varepsilon,s})\right]dw(s)\right|^p \\
&\qquad+3^{p-1}\varepsilon^{\frac{p}{2}}E\sup_{0\leq t\leq u}\left|\int_0^t\int_Z\left[h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right]N(ds,dv)\right|^p \\
&\quad=J_1+J_2+J_3.
\end{aligned}\tag{23}$$

By the Hölder inequality, we get

$$J_1 \le 3^{p-1}\varepsilon^p u^{p-1} E\int_0^u \left|f(s,x_{\varepsilon,s}) - \bar{f}(y_{\varepsilon,s})\right|^p ds.$$

By Lemma 3.1 and Assumption 2.2, it follows that for any $\epsilon_4 > 0$

$$\begin{aligned} J_1 &\le 3^{p-1}\varepsilon^p u^{p-1} E\int_0^u \left|f(s,x_{\varepsilon,s}) - f(s,y_{\varepsilon,s}) + f(s,y_{\varepsilon,s}) - \bar{f}(y_{\varepsilon,s})\right|^p ds \\ &\le 3^{p-1}\varepsilon^p u^{p-1}\left[1 + \epsilon_4^{\frac{1}{p-1}}\right]^{p-1} E\int_0^u \left(\frac{\sqrt{k_1}^p \|x_{\varepsilon,s} - y_{\varepsilon,s}\|^p}{\epsilon_4} + \left|f(s,y_{\varepsilon,s}) - \bar{f}(y_{\varepsilon,s})\right|^p\right) ds. \end{aligned}$$

Letting $\epsilon_4 = \sqrt{k_1}^{p-1}$,

$$\begin{aligned} J_1 &\le 3^{p-1}\varepsilon^p u^{p-1}(1+\sqrt{k_1})^p \int_0^u E\sup_{0\le\sigma\le s}\left|x_\varepsilon(\sigma) - y_\varepsilon(\sigma)\right|^p ds \\ &\quad + 3^{p-1}\varepsilon^p u^p (1+\sqrt{k_1})^p E\frac{1}{u}\int_0^u \left|f(s,y_{\varepsilon,s}) - \bar{f}(y_{\varepsilon,s})\right|^p ds. \end{aligned}$$

Then Assumption 2.4 implies that

$$\begin{aligned} J_1 &\le 3^{p-1}\varepsilon^p u^{p-1}(1+\sqrt{k_1})^p \int_0^u E\sup_{0\le\sigma\le s}\left|x_\varepsilon(\sigma) - y_\varepsilon(\sigma)\right|^p ds \\ &\quad + 3^{p-1}\varepsilon^p u^p (1+\sqrt{k_1})^p \psi_1(u)\Bigl(1 + E\sup_{0\le t\le u}\|y_{\varepsilon,t}\|^p\Bigr). \end{aligned} \tag{24}$$

For the second term J_2 of (23): by the Burkholder-Davis-Gundy inequality, there exists a $C_p > 0$ such that

$$\begin{aligned} J_2 &\le 3^{p-1}\varepsilon^{\frac{p}{2}} C_p E\left(\int_0^u \left|g(s,x_{\varepsilon,s}) - \bar{g}(y_{\varepsilon,s})\right|^2 ds\right)^{\frac{p}{2}} \\ &\le 3^{p-1}\varepsilon^{\frac{p}{2}} C_p u^{\frac{p}{2}-1} E\int_0^u \left|g(s,x_{\varepsilon,s}) - \bar{g}(y_{\varepsilon,s})\right|^p ds. \end{aligned}$$

Similar to J_1, we get

$$\begin{aligned} J_2 &\le 3^{p-1}\varepsilon^{\frac{p}{2}} C_p u^{\frac{p}{2}-1}(1+\sqrt{k_1})^p \int_0^u E\sup_{0\le\sigma\le s}\left|x_\varepsilon(\sigma) - y_\varepsilon(\sigma)\right|^p ds \\ &\quad + 3^{p-1}\varepsilon^{\frac{p}{2}} C_p u^{\frac{p}{2}}(1+\sqrt{k_1})^p \psi_2(u)\Bigl(1 + E\sup_{0\le t\le u}\|y_{\varepsilon,t}\|^p\Bigr). \end{aligned} \tag{25}$$

Since $N(dt,dv) = \tilde{N}(dt,dv) + \pi(dv)\,dt$ and using the basic inequality $|a+b|^p \le 2^{p-1}(|a|^p + |b|^p)$, we have

$$\begin{aligned} J_3 &\le 6^{p-1}\varepsilon^{\frac{p}{2}} E\sup_{0\le t\le u}\left|\int_0^t\int_Z \left[h(s,x_{\varepsilon,s},v) - \bar{h}(y_{\varepsilon,s},v)\right]\tilde{N}(ds,dv)\right|^p \\ &\quad + 6^{p-1}\varepsilon^{\frac{p}{2}} E\sup_{0\le t\le u}\left|\int_0^t\int_Z \left[h(s,x_{\varepsilon,s},v) - \bar{h}(y_{\varepsilon,s},v)\right]\pi(dv)\,ds\right|^p \\ &= 6^{p-1}\varepsilon^{\frac{p}{2}}(L_1 + L_2). \end{aligned} \tag{26}$$

By Lemma 3.4, there exists a D_p such that

$$L_1 \le D_p\left\{E\left(\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^2\pi(dv)\,ds\right)^{\frac{p}{2}} + E\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^p\pi(dv)\,ds\right\}. \tag{27}$$

By Assumptions 2.2 and 2.4, we have

$$\begin{aligned}
&E\left(\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^2\pi(dv)\,ds\right)^{\frac{p}{2}}\\
&\quad\le E\left(2k_2\int_0^u\|x_{\varepsilon,s}-y_{\varepsilon,s}\|^2\,ds\right.\\
&\qquad\left.+2u\frac{1}{u}\int_0^u\int_Z\left|h(s,y_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^2\pi(dv)\,ds\right)^{\frac{p}{2}}\\
&\quad\le E\left[2k_2\int_0^u\|x_{\varepsilon,s}-y_{\varepsilon,s}\|^2\,ds+2u\psi_3(u)\bigl(1+\|y_{\varepsilon,s}\|^2\bigr)\right]^{\frac{p}{2}}.
\end{aligned}\tag{28}$$

Using the basic inequality and the Hölder inequality, we obtain

$$\begin{aligned}
&E\left(\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^2\pi(dv)\,ds\right)^{\frac{p}{2}}\\
&\quad\le 3^{\frac{p}{2}-1}\left\{E\left[2k_2\int_0^u\|x_{\varepsilon,s}-y_{\varepsilon,s}|^2\,ds\right]^{\frac{p}{2}}+\bigl[2u\psi_3(u)\bigr]^{\frac{p}{2}}+\bigl[2u\psi_3(u)\bigr]^{\frac{p}{2}}E\|y_{\varepsilon,s}\|^p\right\}\\
&\quad\le 3^{\frac{p}{2}-1}\left\{(2k_2)^{\frac{p}{2}}u^{\frac{p}{2}-1}\int_0^u E\sup_{0\le\sigma\le s}\left|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\right|^p\,ds\right.\\
&\qquad\left.+\bigl[2u\psi_3(u)\bigr]^{\frac{p}{2}}+\bigl[2u\psi_3(u)\bigr]^{\frac{p}{2}}E\|y_{\varepsilon,s}\|^p\right\}.
\end{aligned}\tag{29}$$

Similar to the estimation of J_1, we derive that

$$\begin{aligned}
&E\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^p\pi(dv)\,ds\\
&\quad\le(1+k_2)E\int_0^u\|x_{\varepsilon,s}-y_{\varepsilon,s}\|^p\,ds+(1+k_2)E\int_0^u\int_Z\left|h(s,y_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^p\pi(dv)\,ds\\
&\quad\le(1+k_2)\int_0^u E\sup_{0\le\sigma\le s}\left|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\right|^p\,ds\\
&\qquad+(1+k_2)u\psi_3(u)\Bigl(1+E\sup_{0\le t\le u}\|y_{\varepsilon,t}\|^p\Bigr).
\end{aligned}\tag{30}$$

On the other hand, using the Hölder inequality, it follows that

$$\begin{aligned}
L_2&\le E\sup_{0\le t\le u}\left\{\left(\int_0^t ds\right)^{p-1}\left(\int_0^t\left|\int_Z\bigl[h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\bigr]\pi(dv)\right|^p ds\right)\right\}\\
&\le\bigl[u\pi(Z)\bigr]^{p-1}E\int_0^u\int_Z\left|h(s,x_{\varepsilon,s},v)-\bar{h}(y_{\varepsilon,s},v)\right|^p\pi(dv)\,ds
\end{aligned}$$

$$\leq (1+k_2)\big[u\pi(Z)\big]^{p-1}\bigg[\int_0^u E\sup_{0\leq\sigma\leq s}\big|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\big|^p\,ds$$
$$+\,u\psi_3(u)\Big(1+E\sup_{0\leq t\leq u}\|y_{\varepsilon,t}\|^p\Big)\bigg]. \tag{31}$$

Hence, substituting (29)-(31) into (26), we get

$$J_3 \leq 6^{p-1}\varepsilon^{\frac{p}{2}}\big[D_p 3^{\frac{p}{2}-1}(2k_2)^{\frac{p}{2}}u^{\frac{p}{2}-1}+(1+k_2)\big(D_p+\big(u\pi(Z)\big)^{p-1}\big)\big]$$
$$\times\int_0^u E\sup_{0\leq\sigma\leq s}\big|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\big|^p\,ds+6^{p-1}\varepsilon^{\frac{p}{2}}\big[D_p 3^{\frac{p}{2}-1}\big(2u\psi_3(u)\big)^{\frac{p}{2}}u^{\frac{p}{2}-1}$$
$$+(1+k_2)\big(D_p+\big(u\pi(Z)\big)^{p-1}\big)u\psi_3(u)\big]\Big(1+E\sup_{0\leq t\leq u}\|y_{\varepsilon,t}\|^p\Big). \tag{32}$$

Combining with (24), (25), and (32),

$$E\sup_{0\leq t\leq u}\big|x_\varepsilon(t)-y_\varepsilon(t)-\big[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\big]\big|^p$$
$$\leq C_1\varepsilon E\int_0^u\sup_{0\leq\sigma\leq t}\big|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\big|^p\,dt+C_2\varepsilon u\Big(1+E\sup_{-\tau\leq t\leq u}\big|y_\varepsilon(t)\big|^p\Big), \tag{33}$$

where $C_1=3^{p-1}(1+\sqrt{k_1})^p(\varepsilon^{p-1}u^{p-1}+\varepsilon^{\frac{p}{2}-1}C_pu^{\frac{p}{2}-1})+6^{p-1}\varepsilon^{\frac{p}{2}-1}[D_p3^{\frac{p}{2}-1}(2k_2)^{\frac{p}{2}}u^{\frac{p}{2}-1}+(1+k_2)(D_p+(u\pi(Z))^{p-1})]$, $C_2=3^{p-1}(1+\sqrt{k_1})^p(\varepsilon^{p-1}u^{p-1}\psi_1(u)+\varepsilon^{\frac{p}{2}-1}C_pu^{\frac{p}{2}-1}\psi_2(u))+6^{p-1}\varepsilon^{\frac{p}{2}-1}\times[D_p3^{\frac{p}{2}-1}(2u\psi_3(u))^{\frac{p}{2}}u^{\frac{p}{2}-2}+(1+k_2)(D_p+(u\pi(Z))^{p-1})\psi_3(u)]$. Hence Assumption 2.4 and Lemma 3.5 imply that

$$E\sup_{0\leq t\leq u}\big|x_\varepsilon(t)-y_\varepsilon(t)-\big[D(x_{\varepsilon,t})-D(y_{\varepsilon,t})\big]\big|^p$$
$$\leq C_1\varepsilon\int_0^u E\sup_{0\leq\sigma\leq t}\big|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\big|^p\,dt$$
$$+\,C_2\varepsilon u(1+C). \tag{34}$$

Inserting (34) into (22),

$$E\sup_{0\leq t\leq u}\big|x_\varepsilon(t)-y_\varepsilon(t)\big|^p\leq\frac{C_1\varepsilon}{(1-k_0)^p}\int_0^u E\sup_{0\leq\sigma\leq t}\big|x_\varepsilon(\sigma)-y_\varepsilon(\sigma)\big|^p\,dt$$
$$+\frac{C_2\varepsilon u(1+C)}{(1-k_0)^p}.$$

Finally, by the Gronwall inequality, we have

$$E\sup_{0\leq t\leq u}\big|x_\varepsilon(t)-y_\varepsilon(t)\big|^p\leq\frac{C_2\varepsilon u(1+C)}{(1-k_0)^p}e^{\frac{C_1\varepsilon u}{(1-k_0)^p}}.$$

Choose $\beta\in(0,1)$ and $L>0$ such that for every $t\in[0,L\varepsilon^{-\beta}]\subseteq[0,T]$,

$$E\sup_{t\in[0,L\varepsilon^{-\beta}]}\big|x_\varepsilon(t)-y_\varepsilon(t)\big|^p\leq cL\varepsilon^{1-\beta},$$

where $c = \frac{C_2(1+C)}{(1-k_0)^p} e^{\frac{C_1}{(1-k_0)^p} L\varepsilon^{1-\beta}}$. Consequently, given any number δ_1, we can choose $\varepsilon_1 \in [0, \varepsilon_0]$ such that for each $\varepsilon \in [0, \varepsilon_1]$ and for $t \in [0, L\varepsilon^{-\beta}]$,

$$E \sup_{t\in[0,L\varepsilon^{-\beta}]} \left|x_\varepsilon(t) - y_\varepsilon(t)\right|^p \leq \delta_1.$$

The proof is complete. □

Proof of Theorem 2.4 The key technique to prove this theorem is already presented in the proof of Theorem 2.2, so we here only highlight some parts which need to be modified. By Assumption 2.5, J_1 of (23) should become

$$J_1 \leq 3^{p-1}\varepsilon^p u^{p-1} 2^{p-1} E \int_0^u \left(k^p\left(\|x_{\varepsilon,s} - y_{\varepsilon,s}\|\right) + \left|f(s, y_{\varepsilon,s}) - \bar{f}(y_{\varepsilon,s})\right|^p\right) ds.$$

In fact, since the function $k(\cdot)$ is concave and increasing, there must exist a positive number c_p such that

$$k^p\left(|x|\right) \leq c_p\left(1 + |x|^p\right), \quad \text{for all } p \geq 2.$$

Hence,

$$\begin{aligned} J_1 &\leq c_p C_3 \int_0^u \left(1 + E \sup_{0\leq t\leq u} \left|x_\varepsilon(s) - y_\varepsilon(s)\right|^p\right) ds + c_p C_3 u \\ &\quad + u C_3 \psi_1(u)\left(1 + E \sup_{0\leq t\leq u} \|y_{\varepsilon,t}\|^p\right), \end{aligned} \tag{35}$$

where $C_3 = 3^{p-1}\varepsilon^p u^{p-1} 2^p$. Similarly, J_2 and J_3 can be estimated as J_1. Finally, all of required assertions can be obtained in the same way as the proof of Theorem 2.2. The proof is therefore completed. □

5 Examples

Example 5.1 Consider the following neutral stochastic differential delay equations:

$$\begin{aligned} d\left[x(t) - \tilde{D}\left(x(t-\tau)\right)\right] &= \tilde{f}\left(t, x(t), x(t-\tau)\right) dt + \tilde{g}\left(t, x(t), x(t-\tau)\right) dw(t) \\ &\quad + \int_Z \tilde{h}\left(t, x(t), x(t-\tau), v\right) N(dt, dv), \end{aligned} \tag{36}$$

where $\tau > 0$ is a constant delay and the coefficients of equation (36) satisfy Assumptions 2.1-2.3. Obviously, if we define

$$D(\varphi) = \tilde{D}\left(\varphi(-\tau)\right), \qquad f(t, \varphi) = \tilde{f}\left(t, \varphi(0), \varphi(-\tau)\right),$$

and

$$g(t, \varphi) = \tilde{g}\left(t, \varphi(0), \varphi(-\tau)\right), \qquad h(t, \varphi, v) = \tilde{h}\left(t, \varphi(0), \varphi(-\tau), v\right),$$

then equation (36) will become equation (1). It is naturally seen that equation (36) has a unique solution in the sense of L^p. Meanwhile, similar to (6) and (7), we can get the standard form of equation (36)

$$\begin{aligned} x_\varepsilon(t) &= x(0) + D\big(x_\varepsilon(t-\tau)\big) - D\big(x(-\tau)\big) + \int_0^t f\big(s, x_\varepsilon(s), x_\varepsilon(s-\tau)\big)\,ds \\ &\quad + \sqrt{\varepsilon}\int_0^t g\big(s, x_\varepsilon(s), x_\varepsilon(s-\tau)\big)\,dw(s) \\ &\quad + \sqrt{\varepsilon}\int_0^t \int_Z h\big(s, x_\varepsilon(s), x_\varepsilon(s-\tau), v\big)N(ds, dv), \end{aligned} \tag{37}$$

and the averaging form of equation (36)

$$\begin{aligned} y_\varepsilon(t) &= x(0) + D\big(y_\varepsilon(t-\tau)\big) - D\big(y(-\tau)\big) + \int_0^t \bar{f}\big(y_\varepsilon(s), y_\varepsilon(s-\tau)\big)\,ds \\ &\quad + \sqrt{\varepsilon}\int_0^t \bar{f}\big(y_\varepsilon(s), y_\varepsilon(s-\tau)\big)\,dw(s) \\ &\quad + \sqrt{\varepsilon}\int_0^t \int_Z \bar{h}\big(y_\varepsilon(s), y_\varepsilon(s-\tau), v\big)N(ds, dv). \end{aligned} \tag{38}$$

Similar to the proof of Theorem 2.2 and Corollary 2.1, we can show the convergence of the standard solution of equation (37) and the averaged one of equation (38) in pth moment and in probability.

Example 5.2 Let $N(t)$ be a scalar Poisson processes. Consider neutral SFDEs with Poisson processes of the form

$$d\big[x_\varepsilon(t) - D(x_{\varepsilon,t})\big] = \varepsilon f(t, x_{\varepsilon,t})\,dt + \sqrt{\varepsilon}h(t, x_{\varepsilon,t})\,dN(t), \tag{39}$$

with initial data $x_{\varepsilon,0} = x_0 = \xi(t)$, when $-\tau \le t \le 0$. Here

$$D(x) = 0.1x, \qquad f(t,x) = \cos^2 tx,$$

and

$$h(t,x) = \rho(x) = \begin{cases} 0, & \text{if } x = 0, \\ cx(\log x^{-1})^\alpha, & \text{if } 0 < x \le \delta, \\ c\delta(\log\delta^{-1})^\alpha, & \text{if } x > \delta, \end{cases}$$

where $\alpha \le \frac{1}{2}$, $c \ge 0$, and $\delta \in (0,1)$ is sufficiently small. Let

$$\bar{f}(y_{\varepsilon,t}) = \frac{1}{\pi}\int_0^\pi f(t, y_{\varepsilon,t})\,dt = \frac{1}{2}y_{\varepsilon,t}$$

and

$$\bar{h}(y_{\varepsilon,t}) = \frac{1}{\pi}\int_0^\pi h(t, y_{\varepsilon,t})\,dt = \rho(y_{\varepsilon,t}).$$

Hence, we have the corresponding averaged equation

$$d\big[y_\varepsilon(t) - 0.1y_{\varepsilon,t}\big] = \frac{1}{2}\varepsilon y_{\varepsilon,t}\,dt + \sqrt{\varepsilon}\rho(y_{\varepsilon,t})\,dN(t). \tag{40}$$

Clearly, the coefficient $\rho(\cdot)$ does not satisfy the Lipschitz condition. It is a concave nondecreasing continuous function on $[0,\infty]$ with $\rho(0) = 0$ and

$$\begin{aligned}
\int_{0^+} \frac{x}{\rho^2(x)}\,dx &= -\frac{1}{c^2}\int_{0^+} \frac{1}{x\log x}\,dx = -\frac{1}{c^2}\int_{0^+} \frac{1}{\log x} d(\log x) \\
&= -\frac{1}{c^2}\log|\log x|\Big|_{0^+} = \infty, \quad \text{if } \alpha = \frac{1}{2}, \\
\int_{0^+} \frac{x}{\rho^2(x)}\,dx &= \frac{1}{c^2}\int_{0^+} \frac{1}{x(-\log x)^{2\alpha}}\,dx = -\frac{1}{c^2}\int_{0^+} \frac{1}{(-\log x)^{2\alpha}} d(-\log x) \\
&= -\frac{1}{c^2}\frac{1}{-2\alpha+1}(-\log x)^{-2\alpha+1}\Big|_{0^+} = \infty, \quad \text{if } \alpha < \frac{1}{2}.
\end{aligned}$$

Therefore, it follows that Assumption 2.5 is satisfied. Consequently, by Theorem 2.4 and Corollary 2.2, we see that the solutions of averaged equation (40) will converge to that of the standard equation (39) in the sense of L^2 and probability.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the manuscript and typed, read, and approved the final manuscript.

Author details
[1]School of Mathematics and Information Technology, Jiangsu Second Normal University, Nanjing, 210013, China.
[2]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, UK.

Acknowledgements
The authors would like to thank the referees for their valuable comments and suggestions. The authors would also like to thank the National Natural Science Foundation of China under NSFC grant (11401261) and the NSF of Higher Education Institutions of Jiangsu Province (13KJB110005) for their financial support.

Received: 11 January 2016 Accepted: 6 March 2016 Published online: 15 March 2016

References

1. Krylov, NM, Bogolyubov, NN: Les proprietes ergodiques des suites des probabilites en chaine. C. R. Math. Acad. Sci. **204**, 1454-1546 (1937)
2. Besjes, JG: On the asymptotic methods for non-linear differential equations. J. Méc. **8**, 357-373 (1969)
3. Volosov, VM: Averaging in systems of ordinary differential equations. Russ. Math. Surv. **17**, 1-126 (1962)
4. Matthies, K: Time-averaging under fast periodic forcing of parabolic partial differential equations: exponential estimates. J. Differ. Equ. **174**, 133-180 (2001)
5. Mierczynski, J, Shen, W: Time averaging for non-autonomous random linear parabolic equations. Discrete Contin. Dyn. Syst., Ser. B **9**, 661-699 (2008)
6. Khasminskii, RZ: On the principle of averaging the Ito stochastic differential equations. Kibernetika **4**, 260-279 (1968)
7. Freidlin, M, Wentzell, A: Random Perturbations of Dynamical Systems. Springer, New York (1998)
8. Golec, J, Ladde, G: Averaging principle and systems of singularly perturbed stochastic differential equations. J. Math. Phys. **31**, 1116-1123 (1990)
9. Veretennikov, AY: On the averaging principle for systems of stochastic differential equations. Math. USSR Sb. **69**, 271-284 (1991)
10. Khasminskii, RZ, Yin, G: On averaging principles: an asymptotic expansion approach. SIAM J. Math. Anal. **35**, 1534-1560 (2004)
11. Givon, D, Kevrekidis, IG, Kupferman, R: Strong convergence of projective integration schemes for singular perturbed stochastic differential systems. Commun. Math. Sci. **4**, 707-729 (2006)
12. Stoyanov, IM, Bainov, DD: The averaging method for a class of stochastic differential equations. Ukr. Math. J. **26**, 186-194 (1974)
13. Kolomiets, VG, Melnikov, AI: Averaging of stochastic systems of integral-differential equations with Poisson noise. Ukr. Math. J. **43**, 242-246 (1991)
14. Givon, D: Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems. SIAM J. Multiscale Model. Simul. **6**, 577-594 (2007)
15. Xu, Y, Duan, JQ, Xu, W: An averaging principle for stochastic dynamical systems with Lévy noise. Physica D **240**, 1395-1401 (2011)
16. Yin, G, Ramachandran, KM: A differential delay equation with wideband noise perturbation. Stoch. Process. Appl. **35**, 231-249 (1990)
17. Tan, L, Lei, D: The averaging method for stochastic differential delay equations under non-Lipschitz conditions. Adv. Differ. Equ. **2013**, 38 (2013)

18. Xu, Y, Pei, B, Li, Y: Approximation properties for solutions to non-Lipschitz stochastic differential equations with Lévy noise. Math. Methods Appl. Sci. **38**, 2120-2131 (2015)
19. Mao, W, You, S, Wu, X, Mao, X: On the averaging principle for stochastic delay differential equations with jumps. Adv. Differ. Equ. **2015**, 70 (2015)
20. Bao, J, Yuan, C: Large deviations for neutral functional SDEs with jumps. Stoch. Int. J. Probab. Stoch. Process. **87**, 48-70 (2015)
21. Jankovic, S, Randjelovic, J, Jovanovic, M: Razumikhin-type exponential stability criteria of neutral stochastic functional differential equations. J. Math. Anal. Appl. **355**, 811-820 (2009)
22. Mao, X: Exponential stability in mean square of neutral stochastic differential functional equations. Syst. Control Lett. **26**, 245-251 (1995)
23. Mao, X: Razumikhin-type theorems on exponential stability of neutral stochastic differential equations. SIAM J. Math. Anal. **28**, 389-401 (1997)
24. Wu, F, Mao, X: Numerical solutions of neutral stochastic functional differential equations. SIAM J. Numer. Anal. **46**, 1821-1841 (2008)
25. Zong, X, Wu, F, Huang, C: Exponential mean square stability of the theta approximations for neutral stochastic differential delay equations. J. Comput. Appl. Math. **286**, 172-185 (2015)
26. Applebaum, D: Lévy Process and Stochastic Calculus. Cambridge University Press, Cambridge (2009)
27. Gikhman, II, Skorokhod, AV: Stochastic Differential Equations. Springer, Berlin (1972)
28. Kunita, H: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: Real and Stochastic Analysis. New Perspectives, pp. 305-373. Birkhäuser, Basel (2004)
29. Mao, W, Zhu, Q, Mao, X: Existence, uniqueness and almost surely asymptotic estimations of the solutions to neutral stochastic functional differential equations driven by pure jumps. Appl. Math. Comput. **254**, 252-265 (2015)
30. Bao, J, Hou, Z: Existence of mild solutions to stochastic neutral partial functional differential equations with non-Lipschitz coefficients. Comput. Math. Appl. **59**, 207-214 (2010)
31. Ren, Y, Xia, N: Existence, uniqueness and stability of the solutions to neutral stochastic functional differential equations with infinite delay. Appl. Math. Comput. **210**, 72-79 (2009)
32. Wei, F, Cai, Y: Existence, uniqueness and stability of the solution to neutral stochastic functional differential equations with infinite delay under non-Lipschitz conditions. Adv. Differ. Equ. **2013**, 151 (2013)
33. Mao, X: Stochastic Differential Equations and Applications. Ellis Horwood, Chichester (2008)