Guo *Advances in Difference Equations* (2017) 2017:386
https://doi.org/10.1186/s13662-017-1442-5

Advances in Difference Equations
a SpringerOpen Journal

RESEARCH Open Access

L^p ($p \geq 2$)-strong convergence in averaging principle for multivalued stochastic differential equation with non-Lipschitz coefficients

Zhongkai Guo*

*Correspondence:
zhongkaiguo@hust.edu.cn
School of Mathematics and Statistics, South-Central University for Nationalities, Wuhan, 430074, China

Abstract

We investigate the averaging principle for multivalued stochastic differential equations (MSDEs) driven by a random process under non-Lipschitz conditions. We consider the convergence of solutions in L^p ($p \geq 2$) and in probability between the MSDEs and the corresponding averaged MSDEs.

Keywords: multivalued stochastic differential equations; non-Lipschitz; averaging principle; L^p ($p \geq 2$)-strong convergence

1 Introduction

Most systems in science and industry are perturbed by some random environmental effects described by stochastic differential equations with (fractional) Brownian motion, Lévy process, Poisson process, and so on. A series of useful theories and methods have been proposed to explore stochastic differential equations, such as invariant manifolds [1–3], averaging principle [3–12], homogenization principle, and so on. All these theories and methods develop to extract an effective dynamics from these stochastic differential equations, which is more effective for analysis and simulation. Averaging principle is often used to approximate dynamical systems with random fluctuations and provides a powerful tool for simplifying nonlinear dynamical systems. The essence of averaging principle is to establish an approximation theorem for a simplified stochastic differential equation that replaces the original one in some sense and the corresponding optimal order convergence. The theory of stochastic averaging principle has a long and rich history. It was first introduced by Khasminskii [13] in 1968, and since then, the principle for stochastic differential equations was intensively and extensively studied. Stoyanov and Bainov [11] investigated the averaging method for a class of stochastic differential equations with Poisson noise, proving that under some conditions the solutions of averaged systems converge to the solutions of the original systems in mean square and in probability. Xu, Duan, and Xu [4] established an averaging principle for stochastic differential equations with general non-Gaussian Lévy noise. Quite recently, L^2 (mean square) strong averaging principle for multivalued stochastic differential equations with Brownian motion was established by

Xu and Liu [14]. Note that all the works mentioned are under the Lipschitz condition; however, in the real world, the Lipschitz condition seems to be exceedingly harsh when discussing various applications. So it is necessary and significant to consider some non-Lipschitz cases; see [15].

In [16], the author discussed the existence and uniqueness of a solution in L^p (pth moment) sense for some multivalued stochastic differential equations under a non-Lipschitz condition. From the dynamic view, we concern the p-moment averaging principle for a multivalued stochastic differential equation under a non-Lipschitz condition, which is different from [14] under the Lipschitz condition. Although the method used in our paper is similar to that of [14], compared with the results of article [14], our conclusion is more general, as it is well known that the L^2-strong convergence does not imply $L^p(p \geq 2)$ strong convergence in general. Meanwhile, results for higher-order moments are needed that possess a good robustness and can be applied in computations in statistic, finance, and other aspects.

Recently, many authors considered multivalued and set-valued stochastic differential equations; see, for example, [17–20]. In this article, we study the averaging principle for MSDEs of the form

$$dX_t + \mathcal{A}(X_t)\,dt \ni f(t,X_t)\,dt + g(t,X_t)\,dB_t, \qquad X_0 = x \in \overline{D(\mathcal{A})}, \tag{1.1}$$

with $t \in [0,T]$, where $\mathcal{A}$ is a multivalued maximal monotone operator, which we introduce in the next section, $f : [0,T] \times R^d \to R^d$ and $g : [0,T] \times R^d \to R^d$ are measurable functions and satisfy non-Lipschitz conditions with respect to x. To derive the averaging principle for non-Lipschitz multivalued stochastic differential equation, we need some assumptions given in the next section.

This paper is organized as follows. In Section 2, we give some assumptions for our theory and then introduce the definition of a multivalued maximal monotone operator and related results. The convergence of solutions in L^p ($p \geq 2$) and in probability between the MSDEs and the corresponding averaged MSDEs are considered in Section 3.

Throughout this paper, the letter C will denote positive constants with values changing in different occasions. When necessary, we will explicitly write the dependence of constants on parameters.

2 Framework and preliminaries

2.1 Basic hypothesis

In this paper, we impose the following assumptions.

H1 Non-Lipschitz condition: Suppose that f and b are bounded and satisfy the following conditions:

For any $x, y \in \mathbb{R}^d$ and $t \in [0,T]$,

$$\bigl\|g(t,x) - g(t,y)\bigr\|^2 \leq \rho_{2,\eta}^2\bigl(\|x-y\|\bigr) \quad \text{and} \quad \bigl\|f(t,x) - f(t,y)\bigr\| \leq \rho_{1,\eta}\bigl(\|x-y\|\bigr). \tag{2.1}$$

For $0 < \eta < \frac{1}{e}$, let $\rho_{1,\eta}, \rho_{2,\eta}$ be two concave functions defined by

$$\rho_{j,\eta}(x) := \begin{cases} x[\log x^{-1}]^{\frac{1}{j}}, & x \leq \eta, \\ ([\log \eta^{-1}]^{\frac{1}{j}} - \frac{1}{j}[\log \eta^{-1}]^{\frac{1}{j}-1})x + \frac{1}{j}[\log \eta^{-1}]^{\frac{1}{j}-1}\eta, & x > \eta. \end{cases}$$

Let $\bar{f}: R^d \to R^d$ and $\bar{g}: R^d \to R^d$ be measurable functions satisfying the Lipschitz conditions with respect to x as $f(t,x)$ and $g(x,t)$. Moreover, we assume that $f(t,x), \bar{f}(x), g(x,t)$, and $\bar{g}(x)$ satisfied the following conditions:

H2

$$\frac{1}{T}\int_0^T \|f(s,x)-\bar{f}(x)\|^2\,ds \le \varphi_1(T)\bigl(1+\|x\|^2\bigr), \tag{2.2}$$

H3

$$\frac{1}{T}\int_0^T \|g(s,x)-\bar{g}(x)\|^2\,ds \le \varphi_2(T)\bigl(1+\|x\|^2\bigr), \tag{2.3}$$

where $\varphi_i(T), i = 1,2$, are positive bounded functions; moreover, if T is fixed, then $\varphi_i(T)$ is a constant, which means that $\varphi_i(\cdot)$ only depends on time.

H4 The operator $\mathcal{A}$ is a maximal monotone operator with $D(\mathcal{A}) = \mathbb{R}^d$.

2.2 Multivalued operators and MSDEs

A map $\mathcal{A}: \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is called a multivalued operator. Define the domain and image of $\mathcal{A}$ as

$$D(\mathcal{A}) := \bigl\{x \in \mathbb{R}^d : \mathcal{A}(x) \neq \varnothing\bigr\}, \qquad \operatorname{Im}(\mathcal{A}) := \bigcup_{x\in D(\mathcal{A})} \mathcal{A}(x),$$

and the graph of $\mathcal{A}$ is

$$\operatorname{Gr}(\mathcal{A}) := \bigl\{(x,y) \in \mathbb{R}^{2d} : x \in \mathbb{R}^d, y \in \mathcal{A}(x)\bigr\}.$$

Definition 2.1 (1) A multivalued operator $\mathcal{A}$ is called monotone if

$$\langle y_1 - y_2, x_1 - x_2\rangle \ge 0 \quad \text{for all } (x_1,y_1),(x_2,y_2) \in \operatorname{Gr}(\mathcal{A}).$$

(2) A monotone operator $\mathcal{A}$ is called maximal monotone if and only if

$$(x_1,y_1) \in \operatorname{Gr}(\mathcal{A}) \quad \Leftrightarrow \quad \bigl\{\langle y_1 - y_2, x_1 - x_2\rangle \ge 0 \text{ for all } (x_2,y_2) \in \operatorname{Gr}(\mathcal{A})\bigr\}.$$

Now, we give a precise definition of the solution to equation (1.1).

Definition 2.2 A pair of continuous and $\mathcal{F}_t$-adapted processes (X,K) is called a strong solution of equation (1.1) if:

- $X_0 = x$, and $X(t) \in \overline{D(\mathcal{A})}$ a.s.;
- $K = \{K(t), \mathcal{F}_t; t \in R^+\}$ is of finite variation, and $K(0) = 0$ a.s.;
- $dX_t = f(t,X_t)\,dt + g(t,X_t)\,dZ_t - dK_t, t \in R^+$, a.s.;
- for any continuous processes $(\alpha(t),\beta(t))$ satisfying

  $$\bigl(\alpha(t),\beta(t)\bigr) \in \operatorname{Gr}(\mathcal{A}), \quad t \in R^+,$$

  the measure

  $$\bigl\langle X(t) - \alpha(t), dK(t) - \beta(t)\,dt\bigr\rangle \ge 0.$$

Also, we need the following lemma from [21].

Lemma 2.1 *Let $\mathcal{A}$ be a multivalued maximal monotone operator, let $t \mapsto (X(t), K(t))$ and $t \mapsto (\tilde{X}(t), \tilde{K}(t))$ be continuous functions with $X(t), \tilde{X}(t) \in \overline{D(\mathcal{A})}$, and let $t \mapsto K(t), \tilde{K}(t)$ be of finite variation. Let $(\alpha(t), \beta(t))$ be continuous functions satisfying*

$$\bigl(\alpha(t), \beta(t)\bigr) \in \mathrm{Gr}(\mathcal{A}), \quad t \in R^+.$$

If

$$\bigl\langle X(t) - \alpha(t), dK(t) - \beta(t)\,dt \bigr\rangle \geq 0$$

and

$$\bigl\langle \tilde{X}(t) - \alpha(t), d\tilde{K}(t) - \beta(t)\,dt \bigr\rangle \geq 0,$$

then

$$\bigl\langle X(t) - \tilde{X}(t), dK(t) - d\tilde{K}(t) \bigr\rangle \geq 0.$$

Lemma 2.2 ([22]) *Under* **H1** *and* **H4***, let the initial condition satisfy $E\|x\|^{2p} < +\infty$. For any $p \geq 1$ and $0 \leq t \leq T$, equation* (1.1) *has a unique solution satisfying*

$$E\Bigl[\sup_{0\leq t\leq T} \|X_t\|^{2p}\Bigr] \leq C_T^{(p,\|x\|)} < +\infty.$$

The following example and two lemmas are taken from [22].

Lemma 2.3 *Let $\rho : R^+ \to R^+$ be a continuous nondecreasing function. If $g(s)$ and $q(s)$ are two strictly positive functions on R^+ such that*

$$g(t) \leq g(0) + \int_0^t q(s)\rho\bigl(g(s)\bigr)\,ds, \quad t \geq 0,$$

then

$$g(t) \leq f^{-1}\biggl(f\bigl(g(0)\bigr) + \int_0^t q(s)\,ds\biggr), \tag{2.4}$$

where $f(x) := \int_{x_0}^x \frac{1}{\rho(y)}\,dy$ is well-defined for some $x_0 > 0$.

Example 2.1 For $0 < \eta < \frac{1}{e}$, define a concave function as

$$\rho_\eta(x) := \begin{cases} x \log x^{-1}, & x \leq \eta, \\ \eta \log \eta^{-1} + (\log \eta^{-1} - 1)(x - \eta), & x > \eta. \end{cases}$$

Choosing $x_0 = \eta$, we have

$$f(x) = \log\biggl(\frac{\log \eta}{\log x}\biggr), \quad 0 < x < \eta,$$

$$f^{-1}(x) = \exp\{\log\eta \cdot \exp\{-x\}\}, \quad x < 0.$$

If $g(0) < \eta$, then substituting these into (2.4), we obtain

$$g(t) \le (g(0))^{\exp\{-\int_0^t q(s)\,ds\}}. \tag{2.5}$$

Lemma 2.4

- *For $j = 1, 2$, $\rho_{j,\eta}$ is decreasing in η, that is, $\rho_{j,\eta_1} \le \rho_{j,\eta_2}$ if $1 > \eta_1 > \eta_2$.*
- *For any $p \ge 0$ and η sufficiently small, we have*

$$x^p \rho^j_{j,\eta}(x) \le \frac{1}{j+p}\rho_{1,\eta^{j+p}}(x^{j+p}), \quad j = 1, 2.$$

3 Averaging principle for MSDEs

In this section, we prove an averaging principle for multivalued stochastic differential equations (MSDEs) driven by a random process under non-Lipschitz conditions. We consider the convergence of solutions in L^p ($p \ge 2$) and in probability between the MSDEs and the corresponding averaged MSDEs.

For $t \in [0, T]$, consider

$$dX^\epsilon_t + \epsilon\mathcal{A}(X^\epsilon_t)\,dt \ni \epsilon f(t, X^\epsilon_t)\,dt + \sqrt{\epsilon}g(t, X^\epsilon_t)\,dB_t, \qquad X^\epsilon_0 = x \in \overline{D(\mathcal{A})}. \tag{3.1}$$

The standard form of (3.1) is defined as

$$X^\epsilon_t = X^\epsilon(0) + \epsilon\int_0^t f(s, X^\epsilon(s))\,ds + \sqrt{\epsilon}\int_0^t g(s, X^\epsilon_s)\,dB_s - \epsilon K(t), \quad t \in [0, T], \tag{3.2}$$

and the corresponding averaged MSDEs of (3.2) are defined as

$$Y^\epsilon_t = Y^\epsilon(0) + \epsilon\int_0^t \overline{f}(Y^\epsilon(s))\,ds + \sqrt{\epsilon}\int_0^t \overline{g}(Y^\epsilon_s)\,dB_s - \epsilon\overline{K}(t), \quad t \in [0, T]. \tag{3.3}$$

Here $\overline{f} : R^d \to R^d$ and $\overline{g} : R^d \to R^d$ are measurable functions satisfying the non-Lipschitz conditions with respect to x as $f(t, x)$ and $g(t, x)$, $Y^\epsilon(0) = X^\epsilon(0) = x$, and $f, \overline{f}, g, \overline{g}$ satisfy **H2** and **H3**.

Now, we are in the position to investigate the relationship between the processes X^ϵ_t and Y^ϵ_t.

Theorem 3.1 *Suppose that conditions* **H1**-**H4** *hold. Then, for a given arbitrarily small number $\delta > 0$ and for $\alpha \in (0, \frac{1}{2})$, there exists a number $\tilde{\epsilon} \in (0, \epsilon_0]$ ($\epsilon_0 = \frac{1}{16p^2}$) such that, for all $\epsilon \in (0, \tilde{\epsilon})$ and $p \ge 1$, we have*

$$E\Big(\sup_{t\in[0,\epsilon^{\alpha-\frac{1}{2}}(1-4p\sqrt{\epsilon})]}\|X^\epsilon_t - Y^\epsilon_t\|^{2p}\Big) \le \delta.$$

Proof Consider the difference $X_t^\epsilon - Y_t^\epsilon$. From (3.2) and (3.3) we have

$$\begin{aligned} X_t^\epsilon - Y_t^\epsilon &= \epsilon \int_0^t \bigl[f\bigl(s, X^\epsilon(s)\bigr) - \overline{f}\bigl(Y^\epsilon(s)\bigr)\bigr]\,ds \\ &\quad + \sqrt{\epsilon}\int_0^t \bigl[g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr]\,dB_s - \epsilon\bigl[K(t) - \overline{K}(t)\bigr]. \end{aligned}$$

By Itô's formula [23],

$$\begin{aligned} \bigl\|X_t^\epsilon - Y_t^\epsilon\bigr\|^{2p} &= -\epsilon 2p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle X_s^\epsilon - Y_s^\epsilon, dK(s) - d\overline{K}(s)\bigr\rangle \\ &\quad + \epsilon 2p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle f\bigl(s, X^\epsilon(s)\bigr) - \overline{f}\bigl(Y^\epsilon(s)\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,ds \\ &\quad + \sqrt{\epsilon}2p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,dB_s \\ &\quad + \epsilon p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\|g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\|^2\,ds \\ &\quad + 2\epsilon p(p-1)\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-4}\bigl\langle X_s^\epsilon - Y_s^\epsilon, g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\rangle^2\,ds. \end{aligned}$$

By Definition 2.2 and Lemma 2.1 we get

$$\begin{aligned} \bigl\|X_t^\epsilon - Y_t^\epsilon\bigr\|^{2p} &\le \epsilon 2p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle f\bigl(s, X^\epsilon(s)\bigr) - \overline{f}\bigl(Y^\epsilon(s)\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,ds \\ &\quad + \sqrt{\epsilon}2p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,dB_s \\ &\quad + \epsilon p \int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\|g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\|^2\,ds \\ &\quad + 2\epsilon p(p-1)\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-4}\bigl\langle X_s^\epsilon - Y_s^\epsilon, g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\rangle^2\,ds. \end{aligned}$$

Then

$$\begin{aligned} &E\sup_{0\le t\le T}\bigl\|X_t^\epsilon - Y_t^\epsilon\bigr\|^{2p} \\ &\quad \le \epsilon 2pE\sup_{0\le t\le T}\biggl|\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle f\bigl(s, X^\epsilon(s)\bigr) - \overline{f}\bigl(Y^\epsilon(s)\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,ds\biggr| \\ &\qquad + \sqrt{\epsilon}2pE\sup_{0\le t\le T}\biggl|\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\langle g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr), X_s^\epsilon - Y_s^\epsilon\bigr\rangle\,dB_s\biggr| \\ &\qquad + \epsilon pE\sup_{0\le t\le T}\biggl|\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-2}\bigl\|g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\|^2\,ds\biggr| \\ &\qquad + 2\epsilon p(p-1)E\sup_{0\le t\le T}\int_0^t \bigl\|X_s^\epsilon - Y_s^\epsilon\bigr\|^{2p-4}\bigl\langle X_s^\epsilon - Y_s^\epsilon, g\bigl(s, X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\bigr\rangle^2\,ds \\ &\quad = I_1 + I_2 + I_3 + I_4. \end{aligned}$$

We now estimate I_1, I_2, I_3, I_4 separately.

Estimate of I_1. Using the trigonometric inequality, we have

$$\begin{aligned} I_1 &= 2\epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-2}\langle f(s,X^\epsilon(s)) - \bar{f}(Y^\epsilon(s)), X_s^\epsilon - Y_s^\epsilon\rangle\, ds\right| \\ &\le 2\epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-2}\langle f(s,X^\epsilon(s)) - f(s,Y^\epsilon(s)), X_s^\epsilon - Y_s^\epsilon\rangle\, ds\right| \\ &\quad + 2\epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-2}\langle f(s,Y^\epsilon(s)) - \bar{f}(Y^\epsilon(s)), X_s^\epsilon - Y_s^\epsilon\rangle\, ds\right| \\ &= I_{11} + I_{12}. \end{aligned}$$

For I_{11}, using the non-Lipschitz condition of f and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} I_{11} &\le 2\epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-1}\rho_{1,\eta}(\|X_s^\epsilon - Y_s^\epsilon\|)\, ds\right| \\ &\le 2\epsilon p\int_0^T E\|X_s^\epsilon - Y_s^\epsilon\|^{2p-1}\rho_{1,\eta}(\|X_s^\epsilon - Y_s^\epsilon\|)\, ds. \end{aligned}$$

For I_{12}, using the Hölder and Young inequalities, we deduce

$$\begin{aligned} I_{12} &= 2\epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-2}\langle f(s,Y^\epsilon(s)) - \bar{f}(Y^\epsilon(s)), X_s^\epsilon - Y_s^\epsilon\rangle\, ds\right| \\ &\le \epsilon pE \sup_{0\le t\le T}\left|\int_0^t \|X_s^\epsilon - Y_s^\epsilon\|^{2p-2}\left(\|f(s,Y^\epsilon(s)) - \bar{f}(Y^\epsilon(s))\|^2 + \|X_s^\epsilon - Y_s^\epsilon\|^2\right) ds\right| \\ &\le \epsilon pE \sup_{0\le t\le T}\int_0^t \left(\frac{2p-2}{2p}\|X_s^\epsilon - Y_s^\epsilon\|^{2p} + \frac{1}{p}\right)\left(\|f(s,Y^\epsilon(s)) - \bar{f}(Y^\epsilon(s))\|^2\right) ds \\ &\quad + \epsilon p\int_0^t E\sup_{0\le s\le t}\|X_s^\epsilon - Y_s^\epsilon\|^{2p}\, ds \\ &\le \epsilon\left[(p-1)C_T^{(p,\|x\|)} + 1\right]E\left(\sup_{0\le t\le T} t\frac{1}{t}\int_0^t \|f(s,Y^\epsilon(s)) - \bar{f}(Y^\epsilon(s))\|^2\, ds\right) \\ &\quad + \epsilon p\int_0^t E\sup_{0\le s\le t}\|X_s^\epsilon - Y_s^\epsilon\|^{2p}\, ds. \end{aligned}$$

Taking condition **H2**, Lemma 2.2, and the Young inequality into account, we have

$$\begin{aligned} I_{12} &\le \epsilon\left[(p-1)C_T^{(p,\|x\|)} + 1\right]\sup_{0\le t\le T}\left\{t\varphi_1(t)\left[1 + E\left(\sup_{0\le s\le T}\|Y_s^\epsilon\|^{2p}\right)\right]\right\} \\ &\quad + \epsilon p\int_0^t E\sup_{0\le s\le t}\|X_s^\epsilon - Y_s^\epsilon\|^{2p}\, ds \\ &\le \epsilon p\int_0^t E\sup_{0\le s\le t}\|X_s^\epsilon - Y_s^\epsilon\|^{2p}\, ds + \epsilon C(p,T,\|x\|)T. \end{aligned}$$

Finally, we have

$$
\begin{aligned}
I_1 &\le \epsilon TC\big(p,T,\|x\|\big) + \epsilon p \int_0^t E \sup_{0\le s\le t} \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p}\,ds \\
&\quad + 2\epsilon p \int_0^T E \sup_{0\le u\le s} \big\|X_u^\epsilon - Y_u^\epsilon\big\|^{2p-1} \rho_{1,\eta}\big(\big\|X_u^\epsilon - Y_u^\epsilon\big\|\big)\,ds.
\end{aligned}
$$

Estimate of I_2. Using the Burkholder-Davis-Gundy and Young inequalities, we have

$$
\begin{aligned}
I_2 &= \sqrt{\epsilon}2pE \sup_{0\le t\le T} \left| \int_0^t \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big\langle g\big(s,X_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big), X_s^\epsilon - Y_s^\epsilon \big\rangle\,dB_s \right| \\
&\le \sqrt{\epsilon}8pE \left\{ \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{4p-4} \big|\big\langle g\big(s,X_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big), X_s^\epsilon - Y_s^\epsilon\big\rangle\big|^2\,ds \right\}^{\frac{1}{2}} \\
&\le \sqrt{\epsilon}8pE \left\{ \int_0^T \sup_{0\le s\le T} \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p} \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big\|g\big(s,X_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\,ds \right\}^{\frac{1}{2}} \\
&\le \sqrt{\epsilon}4pE \sup_{0\le t\le T} \big\|X_t^\epsilon - Y_t^\epsilon\big\|^{2p} + \sqrt{\epsilon}4pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big\|g\big(s,X_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\,ds \\
&= I_{21} + I_{22}.
\end{aligned}
$$

In the following, we estimate

$$
I_{22} = \sqrt{\epsilon}4pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big\|g\big(s,X_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\,ds.
$$

Using conditions **H1** and **H3** and the Young inequality, we get:

$$
\begin{aligned}
I_{22} &\le \sqrt{\epsilon}8pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big(\big\|g\big(s,X_s^\epsilon\big) - g\big(s,Y_s^\epsilon\big)\big\|^2 + \big\|g\big(s,Y_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\big)\,ds \\
&\le \sqrt{\epsilon}8pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \big(\rho_{2,\eta}^2\big(\big\|X_s^\epsilon - Y_s^\epsilon\big\|\big) + \big\|g\big(s,Y_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\big)\,ds \\
&\le \sqrt{\epsilon}8pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \rho_{2,\eta}^2\big(\big\|X_s^\epsilon - Y_s^\epsilon\big\|\big)\,ds \\
&\quad + \sqrt{\epsilon}8pE \int_0^T \left[\frac{2p-2}{2p} \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p} + \frac{1}{p} \right] \big(\big\|g\big(s,Y_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\big)\,ds \\
&\le \sqrt{\epsilon}8pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \rho_{2,\eta}^2\big(\big\|X_s^\epsilon - Y_s^\epsilon\big\|\big)\,ds \\
&\quad + \sqrt{\epsilon}C\big(p,T,\|x\|\big)E\left(\sup_{0\le t\le T} t\frac{1}{t} \int_0^t \big\|g\big(s,Y_s^\epsilon\big) - \overline{g}\big(Y_s^\epsilon\big)\big\|^2\,ds \right) \\
&\le \sqrt{\epsilon}TC_2\big(p,T,\|x\|\big) + \sqrt{\epsilon}8pE \int_0^T \big\|X_s^\epsilon - Y_s^\epsilon\big\|^{2p-2} \rho_{2,\eta}^2\big(\big\|X_s^\epsilon - Y_s^\epsilon\big\|\big)\,ds.
\end{aligned}
$$

Combing the estimates of I_{21} and I_{22}, we conclude that

$$
\begin{aligned}
I_2 &\leq \sqrt{\epsilon}4pE\sup_{0\leq t\leq T}\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} + \sqrt{\epsilon}TC_2\bigl(p,T,\|x\|\bigr) \\
&\quad + \sqrt{\epsilon}8pE\int_0^T \left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds.
\end{aligned}
$$

Estimate of I_3. Note that

$$
I_3 = \epsilon pE\sup_{0\leq t\leq T}\left|\int_0^t \left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\left\|g\bigl(s,X_s^\epsilon\bigr) - \overline{g}\bigl(Y_s^\epsilon\bigr)\right\|^2\,ds\right|.
$$

Using the same estimate as for I_{22}, we have

$$
I_3 \leq \epsilon TC_3\bigl(p,T,\|x\|\bigr) + \epsilon 2pE\int_0^T \left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds.
$$

Estimate of I_4. Using the Cauchy-Schwarz inequality, the term I_4 has the same form with I_3 with a different constant:

$$
I_4 \leq \epsilon TC_4\bigl(p,T,\|x\|\bigr) + \epsilon 4p(p-1)E\int_0^T \left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds.
$$

Combing the estimates of I_1, I_2 and I_3, I_4, we have

$$
\begin{aligned}
&E\sup_{0\leq t\leq T}\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} \\
&\quad\leq \epsilon TC_1\bigl(p,T,\|x\|\bigr) + \epsilon C_1(p)\int_0^T E\sup_{0\leq s\leq t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p}\,dt \\
&\qquad + 2\epsilon p\int_0^T E\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p-1}\rho_{1,\eta}\bigl(\left\|X_t^\epsilon - Y_t^\epsilon\right\|\bigr)\,dt \\
&\qquad + \sqrt{\epsilon}4pE\sup_{0\leq t\leq T}\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} + \sqrt{\epsilon}C_2\bigl(p,T,\|x\|\bigr) \\
&\qquad + \sqrt{\epsilon}8p\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds \\
&\qquad + \epsilon TC_3\bigl(p,T,\|x\|\bigr) + \epsilon 2p\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds \\
&\qquad + \epsilon TC_4\bigl(p,T,\|x\|\bigr) + \epsilon 4p(p-1)\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\bigl(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\bigr)\,ds.
\end{aligned}
$$

Taking $\sqrt{\epsilon}4p < 1$, that is, $\epsilon < \frac{1}{16p^2}$, we have

$$\begin{aligned}
&E \sup_{0\le t\le T} \left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} \\
&\quad\le \frac{\sqrt{\epsilon}TC_5(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p} + \frac{\epsilon C_1(p)}{1-\sqrt{\epsilon}4p}\int_0^T E\sup_{0\le s\le t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p}\,dt \\
&\qquad + \frac{2\epsilon p}{1-\sqrt{\epsilon}4p}\int_0^T E\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p-1}\rho_{1,\eta}\left(\left\|X_t^\epsilon - Y_t^\epsilon\right\|\right)dt \\
&\qquad + \frac{\sqrt{\epsilon}8p}{1-\sqrt{\epsilon}4p}\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\left(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\right)ds \\
&\qquad + \frac{\epsilon 2p}{1-\sqrt{\epsilon}4p}\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\left(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\right)ds \\
&\qquad + \frac{\epsilon 4p(p-1)}{1-\sqrt{\epsilon}4p}\int_0^T E\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p-2}\rho_{2,\eta}^2\left(\left\|X_s^\epsilon - Y_s^\epsilon\right\|\right)ds.
\end{aligned}$$

By Lemma 2.4 and the concavity of the function $\rho_{1,\eta}$ we have

$$\begin{aligned}
&E \sup_{0\le t\le T} \left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} \\
&\quad\le \frac{\sqrt{\epsilon}TC_5(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p} + \frac{\epsilon C_1(p)}{1-\sqrt{\epsilon}4p}\int_0^T E\sup_{0\le s\le t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p}\,dt \\
&\qquad + \frac{\sqrt{\epsilon}C_6(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p}\int_0^T \rho_{1,\eta}\Bigl(E\sup_{0\le s\le t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p}\Bigr)dt \\
&\quad\le \frac{\sqrt{\epsilon}TC_5(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p} \\
&\qquad + \frac{\sqrt{\epsilon}C_7(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p}\int_0^T E\sup_{0\le s\le t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p} + \rho_{1,\eta}\Bigl(E\sup_{0\le s\le t}\left\|X_s^\epsilon - Y_s^\epsilon\right\|^{2p}\Bigr)dt.
\end{aligned}$$

Note that, for sufficiently small ϵ, we have $g(0) = \frac{\sqrt{\epsilon}C_5(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p} \le \eta < \frac{1}{e}$, and from Lemma 2.3 and Example 2.1 we get the following estimate:

$$E\sup_{0\le t\le T}\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p} \le \frac{\sqrt{\epsilon}TC_5(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p}\exp^{(1-\ln\eta)\exp\{-\frac{\sqrt{\epsilon}TC_7(p,\|x\|,T,\epsilon)}{1-\sqrt{\epsilon}4p}\}}.$$

Choose $\alpha \in (0,\frac{1}{2})$ such that, for every $t \in [0,\epsilon^{\alpha-\frac{1}{2}}(1-4p\sqrt{\epsilon})] \subseteq [0,T]$, we have

$$E\Bigl(\sup_{t\in[0,\epsilon^{\alpha-\frac{1}{2}}(1-4p\sqrt{\epsilon})]}\left\|X_t^\epsilon - Y_t^\epsilon\right\|^{2p}\Bigr) \le C\epsilon^\alpha,$$

where

$$C = C_5\bigl(p,\|x\|,T,\epsilon\bigr)\exp^{\{(1-\ln\eta)\exp\{-\epsilon^\alpha(C_7(p,\|x\|,T,\epsilon)\}\}}.$$

Consequently, given any number $\delta > 0$, we can choose $\tilde{\epsilon} \in (0, \epsilon_0]$ ($\epsilon_0 = \frac{1}{16p^2}$) such that, for each $\epsilon \in (0, \tilde{\epsilon})$ and every $t \in [0, \epsilon^{\alpha - \frac{1}{2}}(1 - 4p\sqrt{\epsilon})]$,

$$E\Big(\sup_{t\in[0,\epsilon^{\alpha-\frac{1}{2}}(1-4p\sqrt{\epsilon})]} \big\|X_t^\epsilon - Y_t^\epsilon\big\|^{2p}\Big) \le \delta,$$

which completes the proof of the theorem. □

Using the Chebyshev-Markov inequality, we can also get the convergence in probability.

Theorem 3.2 *Suppose that conditions* **H1**-**H4** *hold. Then, for a given arbitrarily small number* $\theta > 0$ *and for* $\alpha \in (0, \frac{1}{2})$, *there exists a number* $\tilde{\epsilon} \in (0, \epsilon_0]$ ($\epsilon_0 = \frac{1}{16p^2}$) *such that, for all* $\epsilon \in (0, \tilde{\epsilon})$ *and* $p \ge 1$, *we have*

$$\lim_{\epsilon\to 0} \mathbb{P}\Big(\sup_{t\in[0,\epsilon^{\alpha-\frac{1}{2}}(1-4p\sqrt{\epsilon})]} \big\|X_t^\epsilon - Y_t^\epsilon\big\|^{2p} > \theta\Big) = 0.$$

Acknowledgements
We would like to thank the anonymous referee for careful reading of the manuscript and valuable comments and suggestions.
This work was supported by NSFs of China (No. 11526196) and the Fundamental Research Funds for the Central Universities (SCUEC: CZQ17005; CZW15124).

Competing interests
The author declares that they have no competing interests.

Author's contributions
All authors read and approved the final manuscript.

Publisher's Note

Received: 1 May 2017 Accepted: 11 December 2017 Published online: 19 December 2017

References

1. Duan, J, Lu, K, Schmalfuss, B: Smooth stable and unstable manifolds for stochastic evolutionary equations. J. Dyn. Differ. Equ. **16**(4), 949-972 (2004)
2. Duan, J, Lu, K, Schmalfuss, B: Invariant manifolds for stochastic partial differential equations. Ann. Probab. **31**(4), 2109-2135 (2003)
3. Duan, J, Wang, W: Effective Dynamics of Stochastic Partial Differential Equations. Elsevier, Amsterdam (2014)
4. Xu, Y, Duan, J, Xu, W: An averaging principle for stochastic dynamical systems with Lévy noise. Physica D **240**, 1395-1401 (2011)
5. Wang, W, Roberts, AJ: Average and deviation for slow-fast stochastic differential equations. J. Differ. Equ. **253**, 1265-1286 (2012)
6. Cerrai, S: Stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term. Probab. Theory Relat. Fields **125**, 271-304 (2003)
7. Cerrai, S: A Khasminkii type averaging principle for stochastic reaction-diffusion equations. Ann. Appl. Probab. **19**(3), 899-948 (2009)
8. Cerrai, S, Freidlin, MI: Averaging principle for a class of stochastic reaction-diffusion equations. Probab. Theory Relat. Fields **144**, 137-177 (2009)
9. Fu, H, Liu, J: Strong convergence in stochastic averaging for two time-scales stochastic partial differential equations. J. Math. Anal. Appl. **384**, 70-86 (2011)
10. Liu, D: Strong convergence of principle of averaging for multiscale stochastic dynamical systems. Commun. Math. Sci. **8**(4), 999-1020 (2010)
11. Stoyanov, IM, Bainov, DD: The averaging method for a class of stochastic differential equations. Ukr. Math. J. **26**(2), 186-194 (1974)
12. Pei, B, Xu, Y: L^p($p > 2$)-strong convergence in averaging principle for two time-scales stochastic evolution equations driven by Lévy process. arXiv:1511.03438v3
13. Khasminskii, RZ: On the principle of averaging the Itô stochastic differential equations. Kibernetika **4**, 260-279 (1968)
14. Xu, J, Liu, J: An averaging principle for multivalued stochastic differential equations. Stoch. Anal. Appl. **32**(6), 962-974 (2014)

15. Tian, L, Lei, D: The averaging method for stochastic differential delay equations under non-Lipschitz conditions. Adv. Differ. Equ. **2013**, 38 (2013)
16. Xu, S: Multivalued stochastic differential equations with non-Lipschitz coefficients. Chin. Ann. Math. **30B**(3), 321-332 (2009)
17. Malinowski, MT: The narrowing set-valued stochastic integral equations. Dyn. Syst. Appl. **24**(4), 399-419 (2015)
18. Malinowski, MT: Fuzzy and set-valued stochastic differential equations with local Lipschitz condition. IEEE Trans. Fuzzy Syst. **23**(5), 1891-1898 (2015)
19. Malinowski, MT: Set-valued and fuzzy stochastic differential equations in M-type 2 Banach spaces. Tohoku Math. J. **67**(3), 349-381 (2015)
20. Malinowski, MT, Agarwal, RP: On solutions set of a multivalued stochastic differential equation. Czechoslov. Math. J. **67**(1), 11-28 (2017)
21. Cépa, E: Équations différentielles stochastiques multivoques. In: Séminaire de Probabilités XXIX. Lecture Notes in Math., vol. 1631, pp. 86-107. Springer, Berlin (1995)
22. Ren, J, Zhang, X: Stochastic flows for SDEs with non-Lipschitz coefficients. Bull. Sci. Math. **127**(8), 739-754 (2003)
23. Chow, PL: Stochastic Partial Differential Equations. Chapman Hall/CRC, New York (2007)