Zhang et al. *Advances in Difference Equations* (2018) 2018:197
https://doi.org/10.1186/s13662-018-1638-3

Advances in Difference Equations
a SpringerOpen Journal

RESEARCH Open Access

Weak order in averaging principle for stochastic differential equations with jumps

Bengong Zhang[1], Hongbo Fu[1*], Li Wan[1] and Jicheng Liu[2]

*Correspondence: hbfu@wtu.edu.cn
[1]College of Mathematics and Computer Science, Wuhan Textile University, Wuhan, P.R. China
Full list of author information is available at the end of the article

Abstract

In this paper, we deal with the averaging principle for a two-time-scale system of jump-diffusion stochastic differential equations. Under suitable conditions, we expand the weak error in powers of timescale parameter. We prove that the rate of weak convergence to the averaged dynamics is of order 1. This reveals that the rate of weak convergence is essentially twice that of strong convergence.

MSC: 60H10; 70K70

Keywords: Jump-diffusion; Averaging principle; Invariant measure; Weak convergence; Asymptotic expansion

1 Introduction

We consider a two-time-scale system of jump-diffusion stochastic differential equation of the form

$$dX_t^\epsilon = a(X_t^\epsilon, Y_t^\epsilon)\,dt + b(X_t^\epsilon)\,dB_t + c(X_{t-}^\epsilon)\,dP_t, X_0^\epsilon = x, \tag{1.1}$$

$$dY_t^\epsilon = \frac{1}{\epsilon} f(X_t^\epsilon, Y_t^\epsilon)\,dt + \frac{1}{\sqrt{\epsilon}} g(X_t^\epsilon, Y_t^\epsilon)\,dW_t + h(X_{t-}^\epsilon, Y_{t-}^\epsilon)\,dN_t^\epsilon, \qquad Y_0^\epsilon = y, \tag{1.2}$$

where $X_t^\epsilon \in \mathbb{R}^n$, $Y_t^\epsilon \in \mathbb{R}^m$, the drift functions $a(x,y) \in \mathbb{R}^n$, $f(x,y) \in \mathbb{R}^m$, the diffusion functions $b(x) \in \mathbb{R}^{n\times d_1}$, $c(x) \in \mathbb{R}^n$, $g(x,y) \in \mathbb{R}^{m\times d_2}$, and $h(x,y) \in \mathbb{R}^m$, B_t and W_t are d_1- and d_2-dimensional independent Brownian motions on a complete stochastic base $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, respectively, P_t is a scalar Poisson process with intensity λ_1, and N_t^ϵ is a scalar Poisson process with intensity $\frac{\lambda_2}{\epsilon}$. The positive parameter ϵ is small and describes the ratio of time scales between X_t^ϵ and Y_t^ϵ. Systems (1.1)–(1.2) with two time scales occur frequently in applications, including chemical kinetics, signal processing, complex fluids, and financial engineering.

With the separation of time scale, we can view the state variable of the system as being divided into two parts, the "slow" variable X_t^ϵ and the "fast" variable Y_t^ϵ. It is often the case that we are interested only in the dynamics of the slow component. Then a simplified equation, which is independent of the fast variable and possesses the essential features of the system, is highly desirable. Such a simplified equation is often constructed by averaging procedure as in [2, 20] for deterministic ordinary differential equations and in the further development [7, 8, 13–16, 18, 19, 25] for stochastic differential equations with continuous Gaussian processes. As far as averaging for stochastic dynamical systems in infinite-dimensional space is concerned, it is worth quoting the important works [4–6, 26] and also the works [9, 10, 21]. For related works on averaging for multivalued stochastic differential equations, we refer the reader to [12, 22].

To derive the averaged dynamics of system (1.1)–(1.2), we introduce the fast motion equation with a frozen slow component $x \in \mathbb{R}^n$ of the form

$$dY_t^x = f\bigl(x, Y_t^x\bigr)\,dt + g\bigl(x, Y_t^x\bigr)\,dW_t + h\bigl(x, Y_{t-}^x\bigr)\,dN_t, \qquad Y_0^x = y, \tag{1.3}$$

and denote its solution by $Y_t^\epsilon(y)$. Under suitable conditions on f, g, and h, $Y_t^\epsilon(y)$ induces a unique invariant measure $\mu^x(dy)$ on $\mathbb{R}^m$, which is ergodic and ensures the averaged equation

$$d\bar{X}_t = \bar{a}(\bar{X}_t)\,dt + b(\bar{X}_t)\,dB_t + c(\bar{X}_{t-})\,dP_t, \qquad \bar{X}_0 = x,$$

where the averaging nonlinearity is defined by setting

$$\begin{aligned}\bar{a}(x) &= \int_{\mathbb{R}^m} a(x,y)\mu^x(dy)\\ &= \lim_{t\to+\infty} \mathbb{E}a\bigl(x, Y_t^x(y)\bigr).\end{aligned}$$

In [11], it was shown that, under the stated conditions, the slow motion X_t^ϵ converges strongly to the solution $\bar{X}_t$ of the averaged equation with jumps. The order of convergence $\frac{1}{2}$ in the strong sense was provided in [17]. To our best knowledge, there is no literature addressing the weak order in averaging principle for jump-diffusion stochastic differential systems. In fact, it is fair to say that the weak convergence in stochastic averaging theory of systems driven by jump noise is not fully developed yet, although some strong approximation results on the rate of strong convergence were obtained [1, 23, 24].

Therefore, in this paper, we aim to study this problem. Here we are interested in the rate of weak convergence of the averaging dynamics to the true solution of slow motion X_t^ϵ. In other word, we will determine the order, with respect to timescale parameter ϵ, of weak deviation between original solution to slow equation and the solution of the corresponding averaged equation. The main technique we adapted is finding an expansion with respect to ϵ of the solutions of the Kolmogorov equations associated with the jump diffusion system. The solvability of the Poisson equation associated with the generator of frozen equation provides an expression for the coefficients of the expansion. As a result, the boundedness for the coefficients of expansion can be proved by smoothing effect of the corresponding transition semigroup in the space of bounded and uniformly continuous functions, where some regular conditions on the drift and diffusion terms are needed.

Our result shows that the weak convergence rate is 1 even when there are jump components in the system. It is the main contribution of this work. We would like to stress that an asymptotic method was first applied by Bréhier [3] to an averaging result for stochastic reaction–diffusion equations in the case of Gaussian noise of additive type, which was included only in the fast motion. However, the extension of this argument is not straightforward. The method used in the proof of weak order in [3] is strictly related to the differentiability in time of averaged process. Therefore, once the noise is introduced in the slow equation, difficulties arise, and the procedure becomes more complicated. Our result in this paper bridges such a gap, in which the slow and fast motions are both perturbed by noise with jumps.

The rest of the paper is structured as follows. Section 2 is devoted to notations and assumptions and summarizes preliminary results. The ergodicity of a fast process and the averaged dynamics of system with jumps is introduced in Sect. 3. Then the main result of this article, which is derived via the asymptotic expansions and uniform error estimates, is presented in Sect. 4. Finally, we give the Appendix.

It should be pointed out that in the whole paper the letter C with or without subscripts denotes generic positive constants independent of ϵ.

2 Assumptions and preliminary results

For any integer d, the scalar product and norm on the d-dimensional Euclidean space $\mathbb{R}^d$ are denoted by $(\cdot,\cdot)_{\mathbb{R}^d}$ and $\|\cdot\|_{\mathbb{R}^d}$, respectively. For any integer k, we denote by $C_b^k(\mathbb{R}^d,\mathbb{R})$ the space of all k-times differentiable functions on $\mathbb{R}^d$ with bounded uniformly continuous derivatives up to the kth order.

In what follows, we assume that the drift and diffusion coefficients arising in the system fulfill the following conditions.

(A1) The mappings $a(x,y)$, $b(x)$, $c(x)$, $f(x,y)$, $g(x,y)$, and $h(x,y)$ are of class C^2 and have bounded first and second derivatives. Moreover, we assume that $a(x,y)$, $b(x)$, and $c(x)$ are bounded.

(A2) There exists a constant $\alpha > 0$ such that, for any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$,

$$y^T g(x,y) g^T(x,y) y \geq \alpha \|y\|_{\mathbb{R}^m}.$$

(A3) There exists a constant $\beta > 0$ such that, for any $y_1, y_2 \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}&\big(y_1 - y_2, f(x,y_1) - f(x,y_2) + \lambda_2\big(h(x,y_1) - h(x,y_2)\big)\big)_{\mathbb{R}^m}\\&\qquad + \big\|g(x,y_1) - g(x,y_2)\big\|_{\mathbb{R}^m}^2 + \lambda_2\big|h(x,y_1) - h(x,y_2)\big|^2\\&\quad\leq -\beta\|y_1 - y_2\|_{\mathbb{R}^m}^2.\end{aligned}$$

Remark 2.1 Notice that from (A1) it immediately follows that the following directional derivatives exist and are controlled:

$$\begin{aligned}&\big\|D_x a(x,y)\cdot k_1\big\|_{\mathbb{R}^n} \leq L\|k_1\|_{\mathbb{R}^n},\\&\big\|D_y a(x,y)\cdot l_1\big\|_{\mathbb{R}^n} \leq L\|l_1\|_{\mathbb{R}^m},\\&\big\|D_{xx}^2 a(x,y)\cdot(k_1,k_2)\big\|_{\mathbb{R}^n} \leq L\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n},\\&\big\|D_{yy}^2 a(x,y)\cdot(l_1,l_2)\big\|_{\mathbb{R}^n} \leq L\|l_1\|_{\mathbb{R}^m}\|l_2\|_{\mathbb{R}^m},\end{aligned}$$

where L is a constant independent of x, y, k_1, k_2, l_1, and l_2. For the differentiability of mappings b, c, f, g, and h, we possess similar results. For example, we have

$$\begin{aligned}&\big\|D_{xx}^2 b(x)\cdot(k_1,k_2)\big\|_{\mathbb{R}^n} \leq L\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n},\quad k_1,k_2\in\mathbb{R}^n,\\&\big\|D_{yy}^2 f(x,y)\cdot(l_1,l_2)\big\|_{\mathbb{R}^m} \leq L\|l_1\|_{\mathbb{R}^m}\|l_2\|_{\mathbb{R}^m}\|,\quad l_1,l_2\in\mathbb{R}^m.\end{aligned}$$

As far as assumption (A2) is concerned, it is a sort of nondegeneracy condition, which we assume in order to have the regularizing effect of the Markov transition semigroup associated with the fast dynamics. Assumption (A3) is the dissipative condition, which determines how the fast equation converges to its equilibrium state.

As assumption (A1) holds, for any $\epsilon > 0$ and any initial conditions $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, system (1.1)–(1.2) admits a unique solution, which, to emphasize the dependence on the initial data, is denoted by $(X_t^\epsilon(x,y), Y_t^\epsilon(x,y))$. Moreover, we have the following lemma (for a proof, see, e.g., [17]).

Lemma 2.1 *Under assumptions* (A1), (A2), *and* (A3), *for any* $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, *and* $\epsilon > 0$, *we have*

$$\mathbb{E}\|X_t^\epsilon(x,y)\|_{\mathbb{R}^n}^2 \le C_T\big(1 + \|x\|_{\mathbb{R}^n}^2 + \|y\|_{\mathbb{R}^m}^2\big), \quad t \in [0,T], \tag{2.1}$$

and

$$\mathbb{E}\|Y_t^\epsilon(x,y)\|_{\mathbb{R}^n}^2 \le C_T\big(1 + \|x\|_{\mathbb{R}^m}^2 + \|y\|_{\mathbb{R}^m}^2\big), \quad t \in [0,T]. \tag{2.2}$$

3 Frozen equation and averaged equation

Fixing $\epsilon = 1$, we consider the fast equation with frozen slow component $x \in \mathbb{R}^n$,

$$\begin{cases} dY_t^x(y) = f(x, Y_t^x(y))\,dt + g(x, Y_t^x(y))\,dW_t + h(x, Y_{t-}^x(y))\,dN_t, \\ Y_0^x = y. \end{cases} \tag{3.1}$$

Under assumptions (A1)–(A3), such a problem has a unique solution, which satisfies [17]

$$\mathbb{E}\|Y_t^x(y)\|_{\mathbb{R}^m}^2 \le C\big(1 + \|x\|_{\mathbb{R}^n}^2 + e^{-\beta t}\|y\|_{\mathbb{R}^m}^2\big), \quad t \ge 0. \tag{3.2}$$

Let $Y_t^x(y')$ be the solution of problem (3.1) with initial value $Y_0^x = y'$, the Itô formula implies that, for any $t \ge 0$,

$$\mathbb{E}\|Y_t^x(y) - Y_t^x(y')\|_{\mathbb{R}^m}^2 \le \|y - y'\|_{\mathbb{R}^m}^2 e^{-\beta t}. \tag{3.3}$$

Moreover, as discussed in [17] and [11], equation (3.1) admits a unique ergodic invariant measure μ^x satisfying

$$\int_{\mathbb{R}^m} \|y\|_{\mathbb{R}^m}^2 \mu^x(dy) \le C\big(1 + \|x\|_{\mathbb{R}^n}^2\big). \tag{3.4}$$

Then, by averaging the coefficient a with respect to the invariant measure μ^x we can define the $\mathbb{R}^n$-valued mapping

$$\bar{a}(x) := \int_{\mathbb{R}^m} a(x,y)\mu^x(dy), \quad x \in \mathbb{R}^n.$$

Due to assumption (A1), it is easily to check that $\bar{a}(x)$ is twice differentiable with bounded derivatives, and hence it is Lipschitz-continuous:

$$\left\|\bar{a}(x_1)-\bar{a}(x_2)\right\|_{\mathbb{R}^n}\le C\|x_1-x_2\|_{\mathbb{R}^n},\quad x_1,x_2\in\mathbb{R}^n.$$

According to the invariance property of μ^x, (3.4), and assumption (A1), we have

$$\begin{aligned}\left\|\mathbb{E}a\big(x,Y_t^x(y)\big)-\bar{a}(x)\right\|_{\mathbb{R}^n}^2&=\left\|\int_{\mathbb{R}^m}\mathbb{E}\big(a\big(x,Y_t^x(y)\big)-a\big(x,Y_t^x(z)\big)\big)\mu^x(dz)\right\|_{\mathbb{R}^n}^2\\&\le\int_{\mathbb{R}^m}\mathbb{E}\left\|Y_t^x(y)-Y_t^x(z)\right\|_{\mathbb{R}^m}^2\mu^x(dz)\\&\le e^{-\beta t}\int_{\mathbb{R}^m}\|y-z\|_{\mathbb{R}^m}^2\mu^x(dz)\\&\le Ce^{-\beta t}\big(1+\|x\|_{\mathbb{R}^n}^2+\|y\|_{\mathbb{R}^m}^2\big).\end{aligned}\tag{3.5}$$

Now we can introduce the effective dynamical system

$$\begin{cases}d\bar{X}_t(x)=\bar{a}(\bar{X}_t(x))\,dt+b(\bar{X}_t(x))\,dB_t+c(\bar{X}_{t-}(x))\,dP_t,\\\bar{X}_0=x.\end{cases}\tag{3.6}$$

As the coefficients $\bar{a}$, b, and c are Lipschitz-continuous, this equation admits a unique solution such that

$$\mathbb{E}\left\|\bar{X}_t(x)\right\|_{\mathbb{R}^n}^2\le C_T\big(1+\|x\|_{\mathbb{R}^n}^2\big),\quad t\in[0,T].\tag{3.7}$$

With these assumptions and notation, we have the following result, which is a direct consequence of Lemmas 4.1, 4.2, and 4.5.

Theorem 3.1 *Assume that $x\in\mathbb{R}^n$ and $y\in\mathbb{R}^m$. Then, under assumptions* (A1), (A2), *and* (A3), *for any $T>0$ and $\phi\in C_b^3(\mathbb{R}^n,\mathbb{R})$, there exists a constant $C_{T,\phi,x,y}$ such that*

$$\left|\mathbb{E}\phi\big(X_T^\epsilon(x,y)\big)-\mathbb{E}\phi\big(\bar{X}_T(x)\big)\right|\le C_{T,\phi,x,y}\epsilon.$$

As a consequence, the weak order in the averaging principle for jump-diffusion stochastic systems is 1.

4 Asymptotic expansion

Let $\phi\in C_b^3(\mathbb{R}^n,\mathbb{R})$ and define the function $u^\epsilon(t,x,y):[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ by

$$u^\epsilon(t,x,y)=\mathbb{E}\phi\big(X_t^\epsilon(x,y)\big).$$

We are now ready to seek an expansion formula for $u^\epsilon(t,x,y)$ with respect to ϵ of the form

$$u^\epsilon(t,x,y)=u_0(t,x,y)+\epsilon u_1(t,x,y)+r^\epsilon(t,x,y),\tag{4.1}$$

where u_0 and u_1 are smooth functions, which will be constructed further, and r^ϵ is the remainder term. To this end, let us recall the Kolmogorov operator corresponding to the slow motion equation, with a frozen fast component $y \in \mathbb{R}^m$, which is a second-order operator of the form

$$\begin{aligned}\mathcal{L}_1\Phi(x) &= \big(a(x,y), D_x\Phi(x)\big)_{\mathbb{R}^n} + \frac{1}{2}\operatorname{Tr}\big[D^2_{xx}\Phi(x)\cdot b(x)b^T(x)\big] \\ &\quad + \lambda_1\big(\Phi\big(x+c(x)\big)-\Phi(x)\big), \quad \Phi \in C_b^2\big(\mathbb{R}^n,\mathbb{R}\big).\end{aligned}$$

For any frozen slow component $x \in \mathbb{R}^m$, the Kolmogorov operator for equation (3.1) is given by

$$\begin{aligned}\mathcal{L}_2\Psi(y) &= \big(f(x,y), D_y\Psi(y)\big)_{\mathbb{R}^m} + \frac{1}{2}\operatorname{Tr}\big[D^2_{yy}\Psi(y)\cdot g(x,y)g^T(x,y)\big] \\ &\quad + \lambda_2\big(\Psi\big(y+h(x,y)\big)-\Psi(y)\big), \quad \Psi \in C_b^2\big(\mathbb{R}^m,\mathbb{R}\big).\end{aligned}$$

We set

$$\mathcal{L}^\epsilon := \mathcal{L}_1 + \frac{1}{\epsilon}\mathcal{L}_2.$$

It is known that $u^\epsilon(t,x,y)$ solves the equation

$$\begin{cases}\frac{\partial}{\partial t}u^\epsilon(t,x,y) = \mathcal{L}^\epsilon u^\epsilon(t,x,y), \\ u^\epsilon(0,x,y) = \phi(x).\end{cases} \tag{4.2}$$

Also, recall the Kolmogorov operator associated with the averaged equation (3.6) is defined as

$$\begin{aligned}\bar{\mathcal{L}}\Phi(x) &= \big(\bar{a}(x), D_x\Phi(x)\big)_{\mathbb{R}^n} + \frac{1}{2}\operatorname{Tr}\big[D^2_{xx}\Phi(x)\cdot b(x)b^T(x)\big] \\ &\quad + \lambda_1\big(\Phi\big(x+c(x)\big)-\Phi(x)\big), \quad \Phi \in C_b^2\big(\mathbb{R}^n,\mathbb{R}\big).\end{aligned}$$

Setting

$$\bar{u}(t,x) = \mathbb{E}\phi\big(\bar{X}_t(x)\big),$$

we have

$$\begin{cases}\frac{\partial}{\partial t}\bar{u}(t,x) = \bar{\mathcal{L}}\bar{u}(t,x), \\ \bar{u}(0,x) = \phi(x).\end{cases} \tag{4.3}$$

4.1 The leading term

Let us begin with constructing the leading term. By substituting expansion (4.1) into (4.2) we see that

$$\begin{aligned}\frac{\partial u_0}{\partial t} + \epsilon\frac{\partial u_1}{\partial t} + \frac{\partial r^\epsilon}{\partial t} &= \mathcal{L}_1 u_0 + \epsilon\mathcal{L}_1 u_1 + \mathcal{L}_1 r^\epsilon \\ &\quad + \frac{1}{\epsilon}\mathcal{L}_2 u_0 + \mathcal{L}_2 u_1 + \frac{1}{\epsilon}\mathcal{L}_2 r^\epsilon.\end{aligned}$$

By equating the powers of ϵ, we obtain the following system of equations:

$$\mathcal{L}_2 u_0 = 0, \tag{4.4}$$

$$\frac{\partial u_0}{\partial t} = \mathcal{L}_1 u_0 + \mathcal{L}_2 u_1. \tag{4.5}$$

According to (4.4), we can conclude that u_0 does not depend on y, that is,

$$u_0(t,x,y) = u_0(t,x).$$

We also impose the initial condition $u_0(0,x) = \phi(x)$. Noting that $\mathcal{L}_2$ is the generator of a Markov process defined by equation (3.1), which admits a unique invariant measure μ^x, we have

$$\int_{\mathbb{R}^m} \mathcal{L}_2 u_1(t,x,y)\mu^x(dy) = 0. \tag{4.6}$$

Thanks to (4.5), this yields

$$\begin{aligned}
\frac{\partial u_0}{\partial t}(t,x) &= \int_{\mathbb{R}^m} \frac{\partial u_0}{\partial t}(t,x)\mu^x(dy) \\
&= \int_{\mathbb{R}^m} \mathcal{L}_1 u_0(t,x)\mu^x(dy) \\
&= \int_{\mathbb{R}^m} \bigl(a(x,y), D_x u_0(t,x)\bigr)_{\mathbb{R}^n} \mu^x(dy) \\
&\quad + \frac{1}{2}\operatorname{Tr}\bigl[D^2_{xx} u_0(t,x)\cdot b(x)b^T(x)\bigr] \\
&\quad + \lambda_1\bigl(u_0\bigl(x + c(x)\bigr) - u_0(x)\bigr) \\
&= \bar{\mathcal{L}} u_0(t,x),
\end{aligned}$$

so that u_0 and $\bar{u}$ are described by the same evolutionary equation. By a uniqueness argument, we easily have the following lemma.

Lemma 4.1 *Under assumptions* (A1), (A2), *and* (A3), *for any* $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, *and* $T > 0$, *we have* $u_0(T,x,y) = \bar{u}(T,x)$.

4.2 Construction of u_1

According to Lemma 4.1, (4.3), and (4.5), we get

$$\bar{\mathcal{L}}\bar{u} = \mathcal{L}_1 \bar{u} + \mathcal{L}_2 u_1,$$

which means that

$$\begin{aligned}
\mathcal{L}_2 u_1(t,x,y) &= \bigl(\bar{a}(x) - a(x,y), D_x \bar{u}(t,x)\bigr)_{\mathbb{R}^n} \\
&:= -\rho(t,x,y),
\end{aligned} \tag{4.7}$$

where ρ is of class C^2 with respect to y, with uniformly bounded derivatives. Moreover, for any $t \geq 0$ and $x \in \mathbb{R}^n$, equality (4.6) guarantees that

$$\int_{\mathbb{R}^m} \rho(t,x,y)\mu^x(dy) = 0.$$

For any $y \in \mathbb{R}^m$ and $s > 0$, we have

$$\begin{aligned}\frac{\partial}{\partial s}\mathcal{P}_s\rho(t,x,y) &= \big(f(x,y), D_y\big[\mathcal{P}_s\rho(t,x,y)\big]\big)_{\mathbb{R}^m} + \frac{1}{2}\operatorname{Tr}\big[D^2_{yy}\big[\mathcal{P}_s\rho(t,x,y)\big]\cdot g(x,y)g^T(x,y)\big] \\ &\quad + \lambda_2\big(\mathcal{P}_s\big[\rho\big(t,x,y+h(x,y)\big)\big] - \mathcal{P}_s\big[\rho(t,x,y)\big]\big),\end{aligned} \tag{4.8}$$

where

$$\mathcal{P}_s\big[\rho(t,x,y)\big] := \mathbb{E}\rho\big(t,x,Y^x_s(y)\big).$$

Recalling that μ^x is the unique invariant measure corresponding to the Markov process $Y^x_t(y)$ defined by equation (3.1), from Lemma A.1 we infer that

$$\begin{aligned}&\left|\mathbb{E}\rho\big(t,x,Y^x_s(y)\big) - \int_{\mathbb{R}^m}\rho(t,x,z)\mu^x(dz)\right| \\ &\quad = \left|\int_{\mathbb{R}^m}\mathbb{E}\big[\rho\big(t,x,Y^x_s(y)\big) - \rho\big(t,x,Y^x_s(z)\big)\big]\mu^x(dz)\right| \\ &\quad \leq \int_{\mathbb{R}^m}\big|\mathbb{E}\big(a\big(x,Y^x_s(z)\big) - a\big(x,Y^x_s(y)\big), D_x\bar{u}(t,x)\big)_{\mathbb{R}^n}\big|\mu^x(dz) \\ &\quad \leq C\int_{\mathbb{R}^m}\mathbb{E}\big\|Y^x_s(z) - Y^x_s(y)\big\|_{\mathbb{R}^n}\mu^x(dz).\end{aligned}$$

Now it follows from (3.3) and (3.4) that

$$\left|\mathbb{E}\rho\big(t,x,Y^x_s(y)\big) - \int_{\mathbb{R}^m}\rho(t,x,z)\mu^x(dz)\right| \leq C\big(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\big)e^{-\frac{\beta}{2}s},$$

which implies

$$\lim_{s\to+\infty}\mathbb{E}\rho\big(t,x,Y^x_s(y)\big) = \int_{\mathbb{R}^m}\rho(t,x,z)\mu^x(dz) = 0.$$

With the aid of the last limit, we can deduce from (4.8) that

$$\begin{aligned}&\left(f(x,y), D_y\int_0^{+\infty}\big[\mathcal{P}_s\rho(t,x,y)\big]\,ds\right)_{\mathbb{R}^m} \\ &\quad + \frac{1}{2}\operatorname{Tr}\left[D^2_{yy}\int_0^{+\infty}\big[\mathcal{P}_s\rho(t,x,y)\big]\cdot g(x,y)g^T(x,y)\,ds\right] \\ &\quad + \lambda_2\left(\int_0^{+\infty}\mathcal{P}_s\big[\rho\big(t,x,y+h(x,y)\big)\big]\,ds - \int_0^{+\infty}\mathcal{P}_s\big[\rho(t,x,y)\big]\,ds\right) \\ &= \int_0^{+\infty}\frac{\partial}{\partial s}\mathcal{P}_s\big[\rho(t,x,y)\big]\,ds\end{aligned}$$

$$= \lim_{s\to+\infty} \mathbb{E}\rho\big(t,x,Y_s^x(y)\big) - \rho(t,x,y)$$
$$= \int_{\mathbb{R}^m} \rho(t,x,z)\mu^x(dz) - \rho(t,x,y)$$
$$= -\rho(t,x,y),$$

which implies

$$\mathcal{L}_2\left(\int_0^{+\infty} \mathcal{P}_s\rho(t,x,y)\,ds\right) = -\rho(t,x,y).$$

Therefore,

$$u_1(t,x,y) := \int_0^{+\infty} \mathbb{E}\rho\big(t,x,Y_s^x(y)\big)\,ds \tag{4.9}$$

is the solution to equation (4.7).

Lemma 4.2 *Under assumptions* (A1), (A2), *and* (A3), *for any* $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$, *and* $T>0$, *we have*

$$\big|u_1(t,x,y)\big| \le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big), \quad t\in[0,T]. \tag{4.10}$$

Proof By (4.9) we have

$$u_1(t,x,y) = \int_0^{+\infty} \mathbb{E}\big(\bar{a}(x) - a\big(x,Y_s^x(y)\big), D_x\bar{u}(t,x)\big)_{\mathbb{R}^n}\,ds,$$

so that

$$\big|u_1(t,x,y)\big| \le \int_0^{+\infty} \big\|\bar{a}(x) - \mathbb{E}\big[a\big(x,Y_s^x(y)\big)\big]\big\|_{\mathbb{R}^n} \cdot \big\|D_x\bar{u}(t,x)\big\|_{\mathbb{R}^n}\,ds.$$

Therefore, from Lemma A.1 and (3.5) we get

$$\big|u_1(t,x,y)\big| \le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\int_0^{+\infty} e^{-\frac{\beta}{2}s}\,ds \le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).$$ □

4.3 Determination of remainder r^ϵ

We now turn to the construction for the remainder term r^ϵ. It is known that

$$\big(\partial_t - \mathcal{L}^\epsilon\big)u^\epsilon = 0,$$

which, together with (4.4) and (4.5), implies

$$\begin{aligned}
\big(\partial_t - \mathcal{L}^\epsilon\big)r^\epsilon &= -\big(\partial_t - \mathcal{L}^\epsilon\big)u_0 - \epsilon\big(\partial_t - \mathcal{L}^\epsilon\big)u_1 \\
&= -\left(\partial_t - \frac{1}{\epsilon}\mathcal{L}_2 - \mathcal{L}_1\right)u_0 - \epsilon\left(\partial_t - \frac{1}{\epsilon}\mathcal{L}_2 - \mathcal{L}_1\right)u_1 \\
&= \epsilon(\mathcal{L}_1 u_1 - \partial_t u_1).
\end{aligned} \tag{4.11}$$

To estimate the remainder term r^ϵ, we need the following two lemmas.

Lemma 4.3 *Under assumptions* (A1), (A2), *and* (A3), *for any* $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, *and* $T > 0$, *we have*

$$\left|\frac{\partial u_1}{\partial t}(t,x,y)\right| \le C_T\big(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\big).$$

Proof In view of (4.9), we get

$$\frac{\partial u_1}{\partial t}(t,x,y) = \int_0^{+\infty} \mathbb{E}\bigg(\bar{a}(x) - a\big(x, Y_s^x(y)\big), \frac{\partial}{\partial t} D_x \bar{u}(t,x)\bigg)_{\mathbb{R}^n} ds.$$

By Lemma A.6 introduced in Sect. 4.3 we have

$$\begin{aligned}\left|\frac{\partial u_1}{\partial t}(t,x,y)\right| &\le \int_0^{+\infty} \mathbb{E}\bigg(\big\|\bar{a}(x) - a\big(x, Y_s^x(y)\big)\big\|_{\mathbb{R}^n} \cdot \bigg\|\frac{\partial}{\partial t} D_x \bar{u}(t,x)\bigg\|_{\mathbb{R}^n}\bigg) ds \\ &\le C_T \int_0^{+\infty} \mathbb{E}\big\|\bar{a}(x) - a\big(x, Y_s^x(y)\big)\big\|_{\mathbb{R}^n} ds,\end{aligned}$$

so that from (3.5) we have

$$\left|\frac{\partial u_1}{\partial t}(t,x,y)\right| \le C_T\big(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\big). \qquad \square$$

Lemma 4.4 *Under assumptions* (A1), (A2), *and* (A3), *for any* $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, *and* $T > 0$, *we have*

$$\big|\mathcal{L}_1 u_1(t,x,y)\big| \le C_T\big(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\big), \quad t \in [0,T].$$

Proof Recalling that $u_1(t,x,y)$ is the solution of equation (4.7) and equality (4.9) holds, we have

$$\begin{aligned}\mathcal{L}_1 u_1(t,x,y) &= \big(a(x,y), D_x u_1(t,x,y)\big)_{\mathbb{R}^n} + \frac{1}{2}\operatorname{Tr}\big[D_{xx}^2 u_1(t,x,y) \cdot b(x)b^T(x)\big] \\ &\quad + \lambda_1\big[u_1\big(t, x + c(x), y\big) - u_1(t,x,y)\big],\end{aligned} \tag{4.12}$$

and then, in order to prove the boundedness of $\mathcal{L}_1 u_1$, we have to estimate the three terms arising in the right-hand side of (4.12).

Step 1: Estimate of $(a(x,y), D_x u_1(t,x,y))_{\mathbb{R}^n}$.

For any $k \in \mathbb{R}^n$, we have

$$\begin{aligned}D_x u_1(t,x,y) \cdot k &= \int_0^{+\infty} \Big(D_x\big(\bar{a}(x) - \mathbb{E}a\big(x, Y_s^x(y)\big)\big) \cdot k, D_x \bar{u}(t,x)\Big)_{\mathbb{R}^n} ds \\ &\quad + \int_0^{+\infty} \big(\bar{a}(x) - \mathbb{E}a\big(x, Y_s^x(y)\big), D_{xx}^2 \bar{u}(t,x) \cdot k\big)_{\mathbb{R}^n} ds \\ &=: I_1(t,x,y,k) + I_2(t,x,y,k).\end{aligned}$$

By Lemma A.1 and A.4 we infer that

$$\begin{aligned}&\big|I_1(t,x,y,k)\big| \\ &\quad \le \big\|D_x \bar{u}(t,x)\big\|_{\mathbb{R}^n} \int_0^{+\infty} \big\|D_x\big(\bar{a}(x) - \mathbb{E}a\big(x, Y_s^x(y)\big)\big) \cdot k\big\|_{\mathbb{R}^n} ds\end{aligned}$$

$$
\begin{aligned}
&\le C_T\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\int_0^{+\infty}e^{-\frac{\beta}{2}s}\,ds\\
&\le C_T\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).
\end{aligned}
\tag{4.13}
$$

By Lemma A.2 and inequality (3.5) we obtain

$$
\begin{aligned}
\big|I_2(t,x,y,k)\big| &\le C_T\|k\|_{\mathbb{R}^n}\int_0^{+\infty}\big\|\bar{a}(x)-\mathbb{E}a\big(x,Y_s^x(y)\big)\big\|_{\mathbb{R}^n}\,ds\\
&\le C_T\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\int_0^{+\infty}e^{-\frac{\beta}{2}s}\,ds\\
&\le C_T\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).
\end{aligned}
$$

This, together with (4.13), implies

$$
\big\|D_xu_1(t,x,y)\cdot k\big\|\le C_T\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big),
$$

and then, as $a(x,y)$ is bounded, it follows that

$$
\begin{aligned}
\big|\big(a(x,y),D_xu_1(t,x,y)\big)_{\mathbb{R}^n}\big| &\le C_T\big\|a(x,y)\big\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\\
&\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).
\end{aligned}
$$

Step 2: Estimate of $\mathrm{Tr}[D_{xx}^2u_1(t,x,y)\cdot b(x)b^T(x)]$.

Since $u_1(t,x,y)$ is given by the representation formula (4.9), for any $k_1,k_2\in\mathbb{R}^n$, we have

$$
\begin{aligned}
&D_{xx}^2u_1(t,x,y)\cdot(k_1,k_2)\\
&\quad=\int_0^{+\infty}\mathbb{E}\big(D_{xx}^2\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big)\big)\cdot(k_1,k_2),D_x\bar{u}(t,x)\big)_{\mathbb{R}^n}\,ds\\
&\qquad+\int_0^{+\infty}\mathbb{E}\big(D_x\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big)\big)\cdot k_1,D_{xx}^2\bar{u}(t,x)\cdot k_2\big)_{\mathbb{R}^n}\,ds\\
&\qquad+\int_0^{+\infty}\mathbb{E}\big(D_x\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big)\big)\cdot k_2,D_{xx}^2\bar{u}(t,x)\cdot k_1\big)_{\mathbb{R}^n}\,ds\\
&\qquad+\int_0^{+\infty}\mathbb{E}\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big),D_{xxx}^3\bar{u}(t,x)\cdot(k_1,k_2)\big)_{\mathbb{R}^n}\,ds\\
&\quad:=\sum_{i=1}^4J_i(t,x,y,k_1,k_2).
\end{aligned}
$$

Thanks to Lemma A.1 and Lemma A.5, we get

$$
\begin{aligned}
&\big|J_1(t,x,y,k_1,k_2)\big|\\
&\quad\le\int_0^{+\infty}\big|\mathbb{E}\big(D_{xx}^2\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big)\big)\cdot(k_1,k_2),D_x\bar{u}(t,x)\big)_{\mathbb{R}^n}\big|\,ds\\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}\int_0^{+\infty}e^{-\frac{\beta}{2}s}\,ds\\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}.
\end{aligned}
\tag{4.14}
$$

By Lemma A.4 and (3.5) we infer that

$$\begin{aligned}
&\left|J_2(t,x,y,k_1,k_2)\right| \\
&\quad\le \int_0^{+\infty} \left|\mathbb{E}\big(D_x\big(\bar{a}(x)-a\big(x,Y_s^x(y)\big)\big)\cdot k_1, D_{xx}^2\bar{u}(t,x)\cdot k_2\big)_{\mathbb{R}^n}\right| ds \\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}\int_0^{+\infty} e^{-\frac{\beta}{2}s}\,ds \\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}.
\end{aligned} \tag{4.15}$$

With a similar argument, we can also show that

$$\begin{aligned}
&\left|J_3(t,x,y,k_1,k_2)\right| \\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}.
\end{aligned} \tag{4.16}$$

Using Lemma A.3 and (3.5), we get

$$\begin{aligned}
&\left|J_4(t,x,y,k_1,k_2)\right| \\
&\quad\le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}\cdot\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big)\int_0^{+\infty} e^{-\frac{\beta}{2}s}\,ds \\
&\quad\le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).
\end{aligned} \tag{4.17}$$

In view of estimates (4.14), (4.15), (4.16), and (4.17), we can conclude that there exists a constant C_T such that

$$\left|D_{xx}^2 u_1(t,x,y)\cdot(k_1,k_2)\right| \le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big), \quad t\in[0,T],$$

which means that, for fixed $y\in\mathbb{R}^m$ and $t\in[0,T]$,

$$\left\|D_{xx}^2 u_1(t,x,y)\right\|_{L(\mathbb{R}^n,\mathbb{R})} \le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big),$$

where $\|\cdot\|_{L(\mathbb{R}^n,\mathbb{R})}$ denotes the usual operator norm on the Banach space consisting of bounded and linear operators from $\mathbb{R}^n$ to $\mathbb{R}$. As the diffusion function g is bounded, we get

$$\begin{aligned}
\operatorname{Tr}\big(D_{xx}^2 u_1(t,x,y)gg^T\big) &\le C_T\left\|D_{xx}^2 u_1(t,x,y)\right\|_{L(\mathbb{R}^n,\mathbb{R})} \\
&\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).
\end{aligned}$$

Step 3: Estimate of $\lambda_1[u_1(t,x+c(x),y)-u_1(t,x,y)]$.

By Lemma 4.2 and the boundedness condition of $c(x)$, we directly have

$$\begin{aligned}
&\left|\lambda_1\big[u_1\big(t,x+c(x),y\big)-u_1(t,x,y)\big]\right| \\
&\lambda_1\big[\left|u_1\big(t,x+c(x),y\big)\right|+\left|u_1(t,x,y)\right|\big] \\
&\quad\le C_T\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big), \quad t\in[0,T].
\end{aligned}$$

Finally, it is now easy to gather all previous estimates for terms in (4.12) and conclude

$$\left|\mathcal{L}_1 u_1(t,x,y)\right| \le C_T\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right), \quad t \in [0,T]. \qquad \square$$

Lemma 4.5 *Under the conditions of Lemma* 4.3, *for any* $T > 0$, $x \in \mathbb{R}^n$, *and* $y \in \mathbb{R}^m$, *we have*

$$\left|r^\epsilon(T,x,y)\right| \le C_T\epsilon\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right).$$

Proof By the variation-of-constant formula we write equation (4.11) in thes integral form

$$\begin{aligned} r^\epsilon(T,x,y) &= \mathbb{E}\left[r^\epsilon\left(0, X_T^\epsilon(x,y), Y_T^\epsilon(x,y)\right)\right] \\ &\quad + \epsilon\left[\int_0^T \mathbb{E}\left(\mathcal{L}_1 u_1 - \frac{\partial u_1}{\partial s}\right)\left(s, X_{T-s}^\epsilon(x,y), Y_{T-s}^\epsilon(x,y)\right) ds\right]. \end{aligned}$$

Since u^ϵ and $\bar{u}$ satisfy the same initial condition, we have

$$\begin{aligned} \left|r^\epsilon(0,x,y)\right| &= \left|u^\epsilon(0,x,y) - \bar{u}(0,x) - \epsilon u_1(0,x,y)\right| \\ &= \epsilon\left|u_1(0,x,y)\right|, \end{aligned}$$

so that, thanks to (4.10), (2.1), and (2.2), we have

$$\mathbb{E}\left[r^\epsilon(0, X_T^\epsilon(x,y), Y_T^\epsilon(x,y)\right] \le C\epsilon\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right). \tag{4.18}$$

Using Lemmas 4.3 and 4.4 yields

$$\begin{aligned} &\mathbb{E}\left[\left(\mathcal{L}_1 u_1 - \frac{\partial u_1}{\partial s}\right)\left(s, X_{T-s}^\epsilon(x,y), Y_{T-s}^\epsilon(x,y)\right)\right] \\ &\quad \le C\mathbb{E}\left(1 + \left\|X_{T-s}^\epsilon(x,y)\right\| + \left\|Y_{T-s}^\epsilon(x,y)\right\|\right), \end{aligned}$$

and, according to (2.1) and (2.2), this implies that

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \left(\mathcal{L}_1 u_1 - \frac{\partial u_1}{\partial s}\right)\left(s, X_{T-s}^\epsilon(x,y), Y_{T-s}^\epsilon(x,y)\right) ds\right] \\ &\quad \le C_T\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right). \end{aligned}$$

The last inequality, together with (4.18), yields

$$\left|r^\epsilon(T,x,y)\right| \le \epsilon C_T\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right). \qquad \square$$

Appendix

In this appendix, we collect some technical results to which we appeal in the proofs of the main results in Sect. 4.

Lemma A.1 *For any $T > 0$, there exists a constant $C_T > 0$ such that, for any $x, k \in \mathbb{R}^n$ and $t \in [0, T]$, we have*

$$\left|D_x \bar{u}(t,x) \cdot k\right| \le C_T \|k\|_{\mathbb{R}^n}.$$

Proof Observe that, for any $k \in \mathbb{R}^n$,

$$\begin{aligned} D_x \bar{u}(t,x) \cdot k &= \mathbb{E}\bigl[D\phi\bigl(\bar{X}_t(x)\bigr) \cdot \eta_t^{k,x}\bigr] \\ &= \mathbb{E}\bigl(\phi'\bigl(\bar{X}_t(x)\bigr), \eta_t^{k,x}\bigr)_{\mathbb{R}^n}, \end{aligned}$$

where $\eta_t^{k,x}$ denotes the first mean-square derivative of $\bar{X}_t(x)$ with respect to $x \in \mathbb{R}^n$ along the direction $k \in \mathbb{R}^n$. Then we have

$$\begin{cases} d\eta_t^{k,x} = D_x \bar{a}(\bar{X}_t(x)) \cdot \eta_t^{k,x}\, dt + D_x b(\bar{X}_t(x)) \cdot \eta_t^{k,x}\, dB_t + D_x c(\bar{X}_{t-}(x)) \cdot \eta_{t-}^{k,x}\, dP_t, \\ \eta_0^{k,x} = k. \end{cases}$$

This means that $\eta_t^{k,x}$ is the solution of the integral equation

$$\begin{aligned} \eta_t^{k,x} = {} & k + \int_0^t D_x \bar{a}\bigl(\bar{X}_s(x)\bigr) \cdot \eta_s^{k,x}\, ds + \int_0^t D_x b\bigl(\bar{X}_s(x)\bigr) \cdot \eta_s^{k,x}\, dB_s \\ & + \int_0^t D_x c\bigl(\bar{X}_{s-}(x)\bigr) \cdot \eta_{s-}^{k,x}\, dP_s, \end{aligned}$$

and then, thanks to assumption (A1), we get

$$\mathbb{E}\bigl\|\eta_t^{k,x}\bigr\|_{\mathbb{R}^n}^2 \le C_T \|k\|_{\mathbb{R}^n}^2 + C_T \int_0^t \mathbb{E}\bigl\|\eta_s^{k,x}\bigr\|_{\mathbb{R}^n}^2\, ds.$$

Then by the Gronwall lemma it follows that

$$\mathbb{E}\bigl\|\eta_t^{k,x}\bigr\|_{\mathbb{R}^n}^2 \le C_T \|k\|_{\mathbb{R}^n}^2, \quad t \in [0, T], \tag{A.1}$$

so that

$$\left|D_x \bar{u}(t,x) \cdot k\right| \le C_T \|k\|_{\mathbb{R}^n}. \qquad \square$$

Next, we introduce a similar result for the second derivative of $\bar{u}(t,x)$.

Lemma A.2 *For any $T > 0$, there exists a constant $C_T > 0$ such that, for any $x, k_1, k_2 \in \mathbb{R}^n$ and $t \in [0, T]$, we have*

$$\left|D_{xx}^2 \bar{u}(t,x) \cdot (k_1, k_2)\right| \le C_T \|k_1\|_{\mathbb{R}^n} \cdot \|k_2\|_{\mathbb{R}^n}.$$

Proof For any $k_1, k_2 \in \mathbb{R}^n$, we have

$$\begin{aligned} D_{xx}^2 \bar{u}(t,x) \cdot (k_1, k_2) = {} & \mathbb{E}\bigl[\phi''\bigl(\bar{X}_t(x)\bigr) \cdot \bigl(\eta_t^{k_1,x}, \eta_t^{k_2,x}\bigr) \\ & + \phi'\bigl(\bar{X}_t(x)\bigr) \cdot \xi_t^{k_1,k_2,x}\bigr], \end{aligned} \tag{A.2}$$

where $\xi_t^{k_1,k_2,x}$ is the solution of the second variation equation corresponding to the averaged equation, which may be rewritten in the following form:

$$\begin{aligned}\xi_t^{k_1,k_2,x} = &\int_0^t \big[D_x\bar{a}\big(\bar{X}_s(x)\big)\cdot\xi_s^{k_1,k_2,x} + D_{xx}^2\bar{a}\big(\bar{X}_s(x)\big)\cdot\big(\eta_s^{k_1,x},\eta_s^{k_2,x}\big)\big]\,ds \\ &+ \int_0^t \big[D_{xx}^2 b\big(\bar{X}_s(x)\big)\cdot\big(\eta_s^{k_1,x},\eta_s^{k_2,x}\big) + D_x b\big(\bar{X}_s(x)\big)\cdot\xi_s^{k_1,k_2,x}\big]\,dB_s \\ &+ \int_0^t \big[D_{xx}^2 c\big(\bar{X}_{s-}(x)\big)\cdot\big(\eta_{s-}^{k_1,x},\eta_{s-}^{k_2,x}\big) + D_x c\big(\bar{X}_{s-}(x)\big)\cdot\xi_{s-}^{k_1,k_2,x}\big]\,dP_s.\end{aligned}$$

Thus, by assumption (A1) and (A.1) we have

$$\begin{aligned}\mathbb{E}\big\|\xi_t^{k_1,k_2,x}\big\|_{\mathbb{R}^n}^2 &\le C_T\int_0^t \Big(\big\{\mathbb{E}\big\|\eta_s^{k_1,x}\big\|_{\mathbb{R}^n}^2\big\}^{\frac{1}{2}}\big\{\mathbb{E}\big\|\eta_s^{k_2,x}\big\|_{\mathbb{R}^n}^2\big\}^{\frac{1}{2}} + \mathbb{E}\big\|\xi_s^{k_1,k_2,x}\big\|_{\mathbb{R}^n}^2\Big)\,ds \\ &\le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n} + C_T\int_0^t \mathbb{E}\big\|\xi_s^{k_1,k_2,x}\big\|_{\mathbb{R}^n}^2\,ds.\end{aligned}$$

By the Gronwall lemma we have

$$\mathbb{E}\big\|\zeta_t^{k_1,k_2,x}\big\|_{\mathbb{R}^n}^2 \le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}.$$

Returning to (A.2), we get

$$\big|D_{xx}^2\bar{u}(t,x)\cdot(k_1,k_2)\big| \le C_T\|h_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}. \qquad \square$$

Using analogous arguments, we can prove the following estimate for the third-order derivative of $\bar{u}(t,x)$ with respect to x.

Lemma A.3 *For any $T>0$, there exists a constant $C_T>0$ such that, for any $x,k_1,k_2,k_3\in\mathbb{R}^n$ and $t\in[0,T]$, we have*

$$\big|D_{xxx}^3\bar{u}(t,x)\cdot(k_1,k_2,k_3)\big| \le C_T\|k_1\|_{\mathbb{R}^n}\cdot\|k_2\|_{\mathbb{R}^n}\cdot\|k_3\|_{\mathbb{R}^n}.$$

The following lemma states the boundedness for the first derivative of $\bar{a}(x)-\mathbb{E}a(x,Y_t^x(y))$ with respect to x.

Lemma A.4 *There exists a constant $C>0$ such that, for any $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$, $k\in\mathbb{R}^n$, and $t>0$,*

$$\big\|D_x\big(\bar{a}(x)-\mathbb{E}a\big(x,Y_t^x(y)\big)\big)\cdot k\big\|_{\mathbb{R}^n} \le Ce^{-\frac{\beta}{2}t}\|k\|_{\mathbb{R}^n}\big(1+\|x\|_{\mathbb{R}^n}+\|y\|_{\mathbb{R}^m}\big).$$

Proof The proof is a modification of the proof of [3, Prop. C.2]. For any $t_0>0$, we set

$$\tilde{a}_{t_0}(x,y,t) = \hat{a}(x,y,t) - \hat{a}(x,y,t+t_0),$$

where

$$\hat{a}(x,y,t) := \mathbb{E}a\big(x,Y_t^x(y)\big).$$

Then we have

$$\lim_{t_0\to+\infty}\tilde{a}_{t_0}(x,y,t)=\mathbb{E}a\big(x,Y_t^x(y)\big)-\bar{a}(x).$$

By the Markov property we have

$$\begin{aligned}\tilde{a}_{t_0}(x,y,t)&=\hat{a}(x,y,t)-\mathbb{E}a\big(x,Y_{t+t_0}^x(y)\big)\\&=\hat{a}(x,y,t)-\mathbb{E}\hat{a}\big(x,Y_{t_0}^x(y),t\big).\end{aligned}$$

Due to assumption (A1), for any $k\in\mathbb{R}^n$, we have

$$\begin{aligned}D_x\tilde{a}_{t_0}(x,y,t)\cdot k&=D_x\hat{a}(x,y,t)\cdot k-\mathbb{E}D_x\big(\hat{a}\big(x,Y_{t_0}^x(y),t\big)\big)\cdot k\\&=\hat{a}'_x(x,y,t)\cdot k-\mathbb{E}\hat{a}'_x\big(x,Y_{t_0}^x(y),t\big)\cdot k\\&\quad-\mathbb{E}\hat{a}'_y\big(x,Y_{t_0}^x(y),t\big)\cdot\big(D_xY_{t_0}^x(y)\cdot k\big),\end{aligned}\tag{A.3}$$

where the symbols $\hat{a}'_x$ and $\hat{a}'_y$ denote the directional derivatives with respect to x and y, respectively. Note that the first derivative $\zeta_t^{x,y,k}=D_xY_t^x(y)\cdot k$, at the point x and along the direction $k\in\mathbb{R}^n$, is the solution of the equation

$$\begin{aligned}d\zeta_t^{x,y,k}&=\big(f'_x\big(x,Y_t^x(y)\big)\cdot k+f'_y\big(x,Y_t^x(y)\big)\cdot\zeta_t^{x,y,k}\big)\,dt\\&\quad+\big(g'_x\big(x,Y_t^x(y)\big)\cdot k+g'_y\big(x,Y_t^x(y)\big)\cdot\zeta_t^{x,y,k}\big)\,dW_t\\&\quad+\big(h'_x\big(x,Y_{t-}^x(y)\big)\cdot k+h'_y\big(x,Y_{t-}^x(y)\big)\cdot\zeta_{t-}^{x,y,k}\big)\,dN_t\end{aligned}$$

with initial data $\zeta_0^{x,y,k}=0$. Hence, by assumption (A1), it is straightforward to check

$$\mathbb{E}\big\|\zeta_t^{x,y,k}\big\|_{\mathbb{R}^m}\le C\|k\|_{\mathbb{R}^n}\tag{A.4}$$

for any $t\ge 0$. Note that, for any $y_1,y_2\in\mathbb{R}^m$, we have

$$\begin{aligned}\big\|\hat{a}(x,y_1,t)-\hat{a}(x,y_2,t)\big\|_{\mathbb{R}^n}&=\big\|\mathbb{E}a\big(x,Y_t^x(y_1)\big)-\mathbb{E}a\big(x,Y_t^x(y_2)\big)\big\|_{\mathbb{R}^n}\\&\le C\mathbb{E}\big\|Y_t^x(y_1)-Y_t^x(y_2)\big\|_{\mathbb{R}^m}\\&\le Ce^{-\frac{\beta}{2}t}\|y_1-y_2\|_{\mathbb{R}^m},\end{aligned}$$

where (3.3) was used to obtain the last inequality. This means that

$$\big\|\hat{a}'_y(x,y,t)\cdot l\big\|_{\mathbb{R}^m}\le Ce^{-\frac{\beta}{2}t}\|l\|_{\mathbb{R}^m},\quad l\in\mathbb{R}^m.\tag{A.5}$$

From (A.4) and (A.5) we obtain

$$\begin{aligned}&\big\|\mathbb{E}\big[\hat{a}'_y\big(x,Y_{t_0}^x(y),t\big)\cdot\big(D_xY_{t_0}^x(y)\cdot k\big)\big]\big\|_{\mathbb{R}^m}\\&\quad=\big\|\mathbb{E}\big[\hat{a}'_y\big(x,Y_{t_0}^x(y),t\big)\cdot\big(\zeta_{t_0}^{x,y,k}\big)\big]\big\|_{\mathbb{R}^m}\\&\quad\le Ce^{-\frac{\beta}{2}t}\|k\|_{\mathbb{R}^n}.\end{aligned}\tag{A.6}$$

Then by easy calculations we have

$$\begin{aligned}
&\hat{a}'_x(x,y_1,t)\cdot k-\hat{a}'_x(x,y_2,t)\cdot k\\
&\quad=\mathbb{E}\bigl(a'_x\bigl(x,Y^x_t(y_1)\bigr)\bigr)\cdot k-\mathbb{E}\bigl(a'_x\bigl(x,Y^x_t(y_2)\bigr)\bigr)\cdot k\\
&\qquad+\mathbb{E}\bigl(a'_y\bigl(x,Y^x_t(y_1)\bigr)\cdot\zeta^{x,y_1,k}_t-a'_y\bigl(x,Y^x_t(y_2)\bigr)\cdot\zeta^{x,y_2,k}_t\bigr)\\
&\quad=\mathbb{E}\bigl(a'_x\bigl(x,Y^x_t(y_1)\bigr)\bigr)\cdot k-\mathbb{E}\bigl(a'_x\bigl(x,Y^x_t(y_2)\bigr)\bigr)\cdot k\\
&\qquad+\mathbb{E}\bigl(\bigl[a'_y\bigl(x,Y^x_t(y_1)\bigr)-a'_y\bigl(x,Y^x_t(y_2)\bigr)\bigr]\cdot\zeta^{x,y_1,k}_t\bigr)\\
&\qquad+\mathbb{E}\bigl(a'_y\bigl(x,Y^x_t(y_2)\bigr)\cdot\bigl(\zeta^{x,y_1,k}_t-\zeta^{x,y_2,k}_t\bigr)\bigr)\\
&\quad:=\sum_{i=1}^{3}\mathcal{N}_i(t,x,y_1,y_2,k).
\end{aligned}\tag{A.7}$$

Now, we estimate the three terms in the right-hand side of the equality. Concerning $\mathcal{N}_1(t,x,y_1,y_2,k)$, we have

$$\begin{aligned}
&\bigl\|\mathcal{N}_1(t,x,y_1,y_2,k)\bigr\|_{\mathbb{R}^n}\\
&\quad\le\mathbb{E}\bigl\|\bigl(a'_x\bigl(x,Y^x_t(y_1)\bigr)\bigr)\cdot k-\bigl(a'_x\bigl(x,Y^x_t(y_2)\bigr)\bigr)\cdot k\bigr\|_{\mathbb{R}^n}\\
&\quad\le C\mathbb{E}\bigl\|Y^x_t(y_1)-Y^x_t(y_2)\bigr\|_{\mathbb{R}^m}\cdot\|k\|_{\mathbb{R}^n}\\
&\quad\le Ce^{-\frac{\beta}{2}t}\|y_1-y_2\|_{\mathbb{R}^m}\cdot\|k\|_{\mathbb{R}^n}.
\end{aligned}\tag{A.8}$$

Next, by assumption (A1) we get

$$\begin{aligned}
&\bigl\|\mathcal{N}_2(t,x,y_1,y_2,k)\bigr\|_{\mathbb{R}^n}\\
&\quad\le\mathbb{E}\bigl\|\bigl[a'_y\bigl(x,Y^x_t(y_1)\bigr)-a'_y\bigl(x,Y^x_t(y_2)\bigr)\bigr]\cdot\zeta^{x,y_1,k}_t\bigr\|_{\mathbb{R}^n}\\
&\quad\le C\bigl\{\mathbb{E}\bigl\|\zeta^{x,y_1,k}_t\bigr\|^2_{\mathbb{R}^m}\bigr\}^{\frac{1}{2}}\cdot\bigl\{\mathbb{E}\bigl\|Y^x_t(y_1)-Y^x_t(y_2)\bigr\|^2_{\mathbb{R}^m}\bigr\}^{\frac{1}{2}}\\
&\quad\le Ce^{-\frac{\beta}{2}t}\|k\|_{\mathbb{R}^n}\cdot\|y_1-y_2\|_{\mathbb{R}^m}.
\end{aligned}\tag{A.9}$$

For the third term, using assumption (A1) again, we can infer that

$$\begin{aligned}
&\bigl\|\mathcal{N}_3(t,x,y_1,y_2,k)\bigr\|_{\mathbb{R}^n}\\
&\quad\le\mathbb{E}\bigl\|a'_y\bigl(x,Y^x_t(y_2)\bigr)\cdot\bigl(\zeta^{x,y_1,k}_t-\zeta^{x,y_2,k}_t\bigr)\bigr\|_{\mathbb{R}^n}\\
&\quad\le C\mathbb{E}\bigl\|\zeta^{x,y_1,k}_t-\zeta^{x,y_2,k}_t\bigr\|_{\mathbb{R}^m}\\
&\quad\le Ce^{-\frac{\beta}{2}t}\|y_1-y_2\|_{\mathbb{R}^m}\cdot\|k\|_{\mathbb{R}^n}.
\end{aligned}\tag{A.10}$$

Now, returning to (A.7) and taking into account (A.8), (A.9), and (A.10), we get

$$\begin{aligned}
&\bigl\|\hat{a}'_x(x,y_1,t)\cdot k-\hat{a}'_x(x,y_2,t)\cdot k\bigr\|\\
&\quad\le Ce^{-\frac{\beta}{2}t}\|y_1-y_2\|_{\mathbb{R}^m}\cdot\|k\|_{\mathbb{R}^n},
\end{aligned}$$

which leads to

$$\begin{aligned}
&\left\|\hat{a}'_x(x,y,t)\cdot h - \mathbb{E}\hat{a}'_x\left(x, Y^x_{t_0}(y), t\right)\cdot k\right\|_{\mathbb{R}^n} \\
&\quad \le Ce^{-\frac{\beta}{2}t}\left(1 + \|y\|_{\mathbb{R}^m} + \left\|Y^x_{t_0}(y)\right\|_{\mathbb{R}^m}\right)\cdot \|k\|_{\mathbb{R}^n} \\
&\quad \le e^{-\frac{\beta}{2}t}\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right)\cdot \|k\|_{\mathbb{R}^n},
\end{aligned} \tag{A.11}$$

where we used inequality (3.2). Returning to (A.3), by (A.6) and (A.11) we conclude that

$$\left\|D_x\tilde{a}_{t_0}(x,y,t)\cdot k\right\|_{\mathbb{R}^n} \le Ce^{-\frac{\beta}{2}t}\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right)\|k\|_{\mathbb{R}^n}.$$

Taking the limit as $t_0 \to +\infty$, we obtain

$$\left\|D_x\left(\bar{a}(x) - \mathbb{E}a\left(x, Y^x_t(y)\right)\right)\right\|_{\mathbb{R}^n} \le Ce^{-\frac{\beta}{2}t}\|k\|_{\mathbb{R}^n}\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right). \qquad \square$$

Proceeding with similar arguments, we obtain the following higher-order differentiability.

Lemma A.5 *There exists a constant $C > 0$ such that, for any $x, k_1, k_2 \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, and $t > 0$,*

$$\begin{aligned}
&\left\|D^2_{xx}\left(\bar{a}(x) - \mathbb{E}a\left(x, Y^x_t(y)\right)\right)(k_1, k_2)\right\|_{\mathbb{R}^n} \\
&\quad \le Ce^{-\frac{\beta}{2}t}\|k_1\|_{\mathbb{R}^n}\|k_2\|_{\mathbb{R}^n}\left(1 + \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^m}\right).
\end{aligned}$$

Finally, we introduce the following auxiliary result.

Lemma A.6 *There exists a constant $C > 0$ such that, for any $x, k \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, and $t > 0$,*

$$\left\|\frac{\partial}{\partial t}D_x\bar{u}(t,x)\cdot k\right\|_{\mathbb{R}^n} \le C\|k\|_{\mathbb{R}^n}.$$

Proof For simplicity of presentation, we will prove it for the one-dimensional case. The multidimensional situation can be treated similarly, and only notations are somewhat involved. In this case, we only need to show that

$$\left|\frac{\partial}{\partial t}\frac{\partial}{\partial x}\bar{u}(t,x)\right| \le C. \tag{A.12}$$

In fact, for any $\phi \in C^3_b(\mathbb{R}, \mathbb{R})$, we have

$$\frac{\partial}{\partial x}\bar{u}(t,x) = \frac{\partial}{\partial x}\mathbb{E}\phi\left(\bar{X}_t(x)\right) = \mathbb{E}\left(\phi'\left(\bar{X}_t(x)\right)\cdot\frac{\partial}{\partial x}\bar{X}_t(x)\right).$$

If we define

$$\varsigma^x_t := \frac{\partial}{\partial x}\bar{X}_t(x),$$

then we have

$$\begin{aligned}\varsigma_t^x &= 1 + \int_0^t \bar{a}'\big(\bar{X}_s(x)\big)\cdot\varsigma_s^x\,ds + \int_0^t b'\big(\bar{X}_s(x)\big)\cdot\varsigma_s^x\,dB_s\\&\quad + \int_0^t c'\big(\bar{X}_{s-}(x)\big)\cdot\varsigma_{s-}^x\,dP_s.\end{aligned}$$

The boundedness of $\bar{a}'$, b', and c' guarantees that

$$\mathbb{E}\big|\varsigma_t^x\big|^2 \le C_T, \quad t\in[0,T]. \tag{A.13}$$

By Itô's formula we have

$$\begin{aligned}&\mathbb{E}\big[\phi'\big(\bar{X}_t(x)\big)\cdot\varsigma_t^x\big]\\&\quad = \phi'(x) + \mathbb{E}\int_0^t \big[\phi'\big(\bar{X}_s(x)\big)\bar{a}'\big(\bar{X}_s(x)\big)\varsigma_s^x + \varsigma_s^x\phi''\big(\bar{X}_s(x)\big)\bar{a}\big(\bar{X}_s(x)\big)\big]\,ds\\&\qquad + \mathbb{E}\int_0^t b'\big(\bar{X}_s(x)\big)\big)\varsigma_s^x\phi''\big(\bar{X}_s(x)\big)b\big(\bar{X}_s(x)\big)\,ds\\&\qquad + \frac{1}{2}\mathbb{E}\int_0^t \varsigma_s^x\phi'''\big(\bar{X}_s(x)\big)b^2\big(\bar{X}_s(x)\big)\,ds\\&\qquad + \lambda_1\mathbb{E}\int_0^t \phi'\big(\bar{X}_s(x)\big)c'\big(\bar{X}_{s-}(x)\big)\varsigma_s^x\,ds\\&\qquad + \lambda_1\mathbb{E}\int_0^t \varsigma_s^x\big[\phi'\big(\bar{X}_{s-}(x) + c\big(\bar{X}_{s-}(x)\big)\big) - \phi'\big(\bar{X}_{s-}(x)\big)\big]\,ds\\&\qquad + \lambda_1\mathbb{E}\int_0^t c'\big(\bar{X}_{s-}(x)\big)\varsigma_s^x\big[\phi'\big(\bar{X}_{s-}(x) + c\big(\bar{X}_{s-}(x)\big)\big) - \phi'\big(\bar{X}_{s-}(x)\big)\big]\,ds.\end{aligned}$$

Since ϕ belongs to $C_b^3(\mathbb{R},\mathbb{R})$, from assumption (A1) it follows that, for any $t\in[0,T]$,

$$\begin{aligned}\left|\frac{\partial}{\partial t}\left[\frac{\partial}{\partial x}\bar{u}(t,x)\right]\right| &= \left|\frac{\partial}{\partial t}\mathbb{E}\big[\phi'\big(\bar{X}_t(x)\big)\cdot\varsigma_t^x\big]\right|\\&\le C\big|\mathbb{E}\varsigma_t^x\big|.\end{aligned}$$

Then, taking (A.13) into account, we easily arrive at (A.12). □

Acknowledgements
We would like to thank Professor Jinqiao Duan for helpful discussions and comments. Hongbo Fu is supported by Natural Science Foundation of Hubei Province (No. 2018CFB688), NSF of China (No. 11301403), and Chinese Scholarship Council (No. [2015]5104). Bengong Zhang is supported by NSF of China (No. 11401448). Li Wan is supported by NSF of China (No. 61573011). Jicheng Liu is supported by NSF of China (No. 11271013).

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

Authors' contributions
The authors declare that the work was realized in collaboration with the same responsibility. All authors read and approved the final manuscript.

Author details
[1] College of Mathematics and Computer Science, Wuhan Textile University, Wuhan, P.R. China. [2] School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, P.R. China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 6 February 2018 Accepted: 8 May 2018 Published online: 22 May 2018

References

1. Bao, J., Yin, G., Yuan, C.: Two-time-scale stochastic partial differential equations driven by α-stable noises: averaging principles. Bernoulli **23**(1), 645–669 (2017)
2. Bogoliubov, N.N., Mitropolsky, Y.A.: Asymptotic Methods in the Theory of Non-linear Oscillations. Gordon & Breach, New York (1961)
3. Bréhier, C.E.: Strong and weak orders in averaging for SPDEs. Stoch. Process. Appl. **122**, 2553–2593 (2012)
4. Cerrai, S., Freidlin, M.I.: Averaging principle for a class of stochastic reaction–diffusion equations. Probab. Theory Relat. Fields **144**, 137–177 (2009)
5. Cerrai, S.: A Khasminskii type averaging principle for stochastic reaction–diffusion equations. Ann. Appl. Probab. **19**(3), 899–948 (2009)
6. Cerrai, S.: Averaging principle for systems of reaction–diffusion equations with polynomial nonlinearities perturbed by multiplicative type noise. SIAM J. Math. Anal. **43**(6), 2482–2518 (2011)
7. Freidlin, M.I., Wentzell, A.D.: Random Perturbation of Dynamical Systems, 2nd edn. Springer, New York (1998)
8. Freidlin, M.I., Wentzell, A.D.: Long-time behavior of weakly coupled oscillators. J. Stat. Phys. **123**, 1311–1337 (2006)
9. Fu, H., Wan, L., Wang, Y., Liu, J.: Strong convergence rate in averaging principle for stochastic FitzHug–Nagumo system with two time-scales. J. Math. Anal. Appl. **416**, 609–628 (2014)
10. Fu, H., Wan, L., Liu, J.: Strong convergence in averaging principle for stochastic hyperbolic–parabolic equations with two time-scales. Stoch. Process. Appl. **125**, 3255–3279 (2015)
11. Givon, D.: Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems. SIAM J. Multiscale Model. Simul. **6**, 577–594 (2007)
12. Guo, Z.: L^p ($p \geq 2$)-strong convergence in averaging principle for multivalued stochastic differential equation with non-Lipschitz coefficients. Adv. Differ. Equ. **2017**, 386 (2017)
13. Khasminskii, R.Z.: On the principle of averaging the Itô stochastic differential equations. Kibernetika **4**, 260–279 (1968) (Russian)
14. Kifer, Y.: Some recent advance in averaging. In: Modern Dynamical Systems and Applications, pp. 385–403. Cambridge University Press, Cambridge (2004)
15. Kifer, Y.: Diffusion approximation for slow motion in fully coupled averaging. Probab. Theory Relat. Fields **129**, 157–181 (2004)
16. Kifer, Y.: Another proof of the averaging principle for fully coupled dynamical systems with hyperbolic fast motions. Discrete Contin. Dyn. Syst. **13**, 1187–1201 (2005)
17. Liu, D.: Strong convergence rate of principle of averaging for jump-diffusion processes. Front. Math. China **7**(2), 305–320 (2012)
18. Veretennikov, A.Y.: On the averaging principle for systems of stochastic differential equations. Math. USSR Sb. **69**, 271–284 (1991)
19. Veretennikov, A.Y.: On large deviations in the averaging principle for SDEs with full dependence. Ann. Probab. **27**, 284–296 (1999)
20. Volosov, V.M.: Averaging in systems of ordinary differential equations. Russ. Math. Surv. **17**, 1–126 (1962)
21. Xu, J.: L^p-strong convergence of the averaging principle for slow–fast SPDEs with jumps. J. Math. Anal. Appl. **445**, 342–373 (2017)
22. Xu, J., Liu, J.: An averaging principle for multivalued stochastic differential equations. Stoch. Anal. Appl. **32**, 962–974 (2014)
23. Xu, Y., Duan, J., Xu, W.: An averaging principle for stochastic dynamical systems with Lévy noise. Physica D **240**(17), 1395–1401 (2011)
24. Xu, Y., Pei, B., Li, Y.: Approximation properties for solutions to non-Lipschitz stochastic differential equations with Lévy noise. Math. Methods Appl. Sci. **11**(38), 2120–2131 (2014)
25. Wainrib, G.: Double averaging principle for periodically forced slow–fast stochastic systems. Electron. Commun. Probab. **18**(51), 1–12 (2013)
26. Wang, W., Roberts, A.J.: Average and deviation for slow–fast stochastic partial differential equations. J. Differ. Equ. **253**, 1265–1286 (2012)