Research Article

Tao Jiang and Yancai Liu*

Averaging principle for two-time-scale stochastic differential equations with correlated noise

https://doi.org/10.1515/math-2022-0538
received September 18, 2022; accepted December 1, 2022

Abstract: This article is devoted to studying the averaging principle for two-time-scale stochastic differential equations with correlated noise. By the technique of multiscale expansion of the solution to the backward Kolmogorov equation and consequent elimination of variables, we obtain the Kolmogorov equation corresponding to the reduced simplified system. The approximation of the slow component of the original system to the solution of the corresponding averaged equation is in the weak sense. An example is also provided to illustrate our result.

Keywords: averaging principle, stochastic differential equation, correlated noise, multiscale expansion

MSC 2020: Primary: 34F05, 60H10, Secondary: 34C29

1 Introduction

Almost all dynamical systems in sciences, such as materials science, fluid dynamics, climate dynamics, celestial mechanics, and radiophysics, have a time hierarchy where the components evolve at quite different rates. In other words, some components evolve rapidly, while others vary slowly. Moreover, all important and interesting systems are essentially nonlinear and disturbed by inner and outer noises. Thus, to explore the interactions between multiscale, nonlinearity and uncertainty, many authors propose a lot of different methods in which the theory of averaging principle stands out. The averaging principle provides an effective tool to explore the asymptotic behavior of the multiscale dynamical systems.

The primary target of the averaging principle is to find the *averaged equation*, when the scale parameter ε tends to 0, for the slow component x of original multiscale system, and then to demonstrate the approximation between x and the solution of this *averaged equation*.

The averaging principle was initiated by Bogoliubov and Mitropolsky [1] for ordinary differential equations and then this theory is developed to deal with partial differential equations (PDEs). It is well known that realistic models incorporate uncertainty as an indispensable ingredient. The first result in this direction for stochastic differential equations (SDEs) was obtained by Khasminskii [2], which demonstrated the correctness of averaging principle in a weak sense. Now, the averaging principle for stochastic systems is attracting more and more interest and has a wide range of applications, see, e.g., [3–9] and references therein for the SDEs and [10–22] and references therein for stochastic partial differential equations (SPDEs).

* **Corresponding author: Yancai Liu,** School of Science, Henan University of Technology, Zhengzhou, 450001, China, e-mail: yancailiu@haut.edu.cn
Tao Jiang: Hubei Center for Data and Analysis, Hubei University of Economics, Wuhan, 430205, China, e-mail: jiangtao2017@hbue.edu.cn

However, as far as we know, all these works assume that the noises driving fast and slow components are independent. No result has been obtained for stochastic dynamical systems with two timescales as the noises are correlated. In order to fill the gap in this field of research, we consider a stochastic slow-fast dynamical system where the noise driving the fast components is correlated with the noise driving the slow components. More exactly, we consider the following SDEs driven by correlated noises on $\mathcal{Z}$:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} f_1(x,y) + f_2(x,y) + \alpha_1(x,y)\dfrac{\mathrm{d}B_1}{\mathrm{d}t} + \alpha_2(x,y)\dfrac{\mathrm{d}B_2}{\mathrm{d}t}, x(0) = x_0, & (1.1) \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2} g_1(x,y) + \dfrac{1}{\varepsilon} g_2(x,y) + \dfrac{1}{\varepsilon}\beta(x,y)\dfrac{\mathrm{d}B_2}{\mathrm{d}t}, y(0) = y_0, & (1.2) \end{cases}$$

where the small parameter $\varepsilon > 0$ is the ratio between the fast component y and the slow component x. B_1 and B_2 are independent standard Brownian motions. We take $\mathcal{Z} = \mathbb{R}^d$, thus $x \in \mathbb{R}^l$, $y \in \mathbb{R}^{d-l}$, or $x \in \mathbb{T}^l$, $y \in \mathbb{T}^{d-l}$, then we have $\mathcal{Z} = \mathbb{T}^d$. $\mathbb{T}^d$ denotes the d-dimensional unit torus.

This article is organized as follows. In Section 2, some notations, hypotheses, and useful results we need in later sections are introduced. We present the main result of this article in Section 3, and the proof of this article is given in Section 4. An example is given to illustrate our result in Section 5.

2 Preliminaries

For the convenience of the following description, we first introduce some notations. For two $d \times d$ matrices $A = (a_{ij})$, $B = (b_{ij})$ denote the *inner product* by $A : B = tr(A^T B) = \sum_{i,j} a_{ij} b_{ij}$. Note that $S : T = S^T : T = \frac{1}{2}(S + S^T)$: T, if matrix T is symmetric. For vectors $a, b, c \in \mathbb{R}^d$, define the *outer product*, which is a matrix, of a and b by $(a \otimes b)c = (b \cdot c)a$.

The drift and diffusion coefficients should satisfy some appropriate conditions to derive the existence and uniqueness of the solution to the SDEs (1.1) and (1.2), as is stated in the following result (Theorem 6.1 in [23]):

Lemma 2.1. *Assume that f_1, f_2, α_1, α_2, g_1, and g_2 are all globally Lipschitz on $\mathcal{Z}$. $z_0 = (x_0, y_0)$ is independent of the Brownian motions B_1 and B_2, and*

$$\mathbb{E}|z_0|^2 < \infty.$$

Then the SDEs (1.1) *and* (1.2) *have a unique solution $z = (x, y) \in C(\mathbb{R}^+; \mathcal{Z})$.*

The backward Kolmogorov equation for equations (1.1) and (1.2) is

$$\frac{\partial v}{\partial t} = \frac{1}{\varepsilon^2}\mathcal{L}_0 v + \frac{1}{\varepsilon}\mathcal{L}_1 v + \mathcal{L}_2 v, \quad \text{for } (x, y, t) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^+, \tag{2.1}$$

$$v = \phi(x), \quad \text{for } (x, y, t) \in \mathcal{X} \times \mathcal{Y} \times \{0\}, \tag{2.2}$$

where

$$\mathcal{L}_0 := g_1 \cdot \nabla_y + \frac{1}{2} B(x,y) : \nabla_y \nabla_y, \tag{2.3}$$

$$\mathcal{L}_1 := f_1 \cdot \nabla_x + g_2 \cdot \nabla_y + \frac{1}{2}\alpha_2 \beta^T : \nabla_y \nabla_x + \frac{1}{2}\beta \alpha_2^T : \nabla_x \nabla_y, \tag{2.4}$$

$$\mathcal{L}_2 := f_2 \cdot \nabla_x + \frac{1}{2} A(x,y) : \nabla_x \nabla_x, \tag{2.5}$$

with

$$A(x, y) \coloneqq \alpha_1\alpha_1^T + \alpha_2\alpha_2^T, \tag{2.6}$$

$$B(x, y) \coloneqq \beta\beta^T. \tag{2.7}$$

The significance of the backward Kolmogorov equation above is displayed in the proof of the main result of this work in Section 4. Roughly speaking, we first write the solution to the backward Kolmogorov equation above in multiple-scale expansion. Then we identify the PDE satisfied by the main term of the expansion. Through the correspondence between the SDE and its backward Kolmogorov equation, we can finally derive a simplified equation to approximate the dynamics of x.

The adjoint operator of $\mathcal{L}_0$ is denoted by $\mathcal{L}_0^*$. We impose the following assumptions:

A1: $B(x, y)$ is a strictly and uniformly positive-definite matrix;

A2: $\mathcal{L}_0$ satisfies the Fredholm alternative. To be exact, let $\mathcal{L}_0 : H \to H$ be an operator in H. Then

(i) Either

$$\begin{cases} \mathcal{L}_0 u = \gamma, & (2.8) \\ \mathcal{L}_0^* U = \Gamma, & (2.9) \end{cases}$$

have unique solutions for every $\gamma, \Gamma \in H$;

or

(ii) the corresponding homogeneous equations

$$\begin{cases} \mathcal{L}_0 V_0 = 0, & (2.10) \\ \mathcal{L}_0^* v_0 = 0, & (2.11) \end{cases}$$

satisfy

$$1 \le \dim(\mathcal{N}(\mathcal{L}_0)) = \dim(\mathcal{N}(\mathcal{L}_0^*)) < \infty,$$

where $\mathcal{N}(\mathcal{L}_0)$ and $\mathcal{N}(\mathcal{L}_0^*)$ denote the null spaces of $\mathcal{L}_0$ and $\mathcal{L}_0^*$, respectively. In the latter case, then equations (2.8) and (2.9) have a solution iff

$$\forall v_0 \in \mathcal{N}(\mathcal{L}_0^*), (\gamma, v_0) = 0$$

and

$$\forall V_0 \in \mathcal{N}(\mathcal{L}_0)(\Gamma, V_0) = 0;$$

A3 (Centering condition): $\int_y f_1(x, y)\rho^\infty(y; x)\mathrm{d}y = 0, \quad \forall x \in \mathcal{X}$.

With the help of the assumption (A1), the following result (Theorem 6.16 in [23]) holds:

Lemma 2.2. *Suppose $\mathcal{L}_0$, $\mathcal{L}_0^*$ on $\mathbb{T}^d$ are equipped with periodic boundary conditions and $B(x, y)$ is a strictly positive-definite matrix, uniformly in $z = (x, y) \in \mathbb{T}^d$. Then the following hold*

- $\mathcal{N}(\mathcal{L}_0) = \mathrm{span}\{\mathbf{1}\}$;
- $\mathcal{N}(\mathcal{L}_0^*) = \mathrm{span}\{\rho^\infty\}$, $inf_{z\in\mathbb{T}^d}\rho^\infty(z) > 0$.

1 *stands for all functions that are independent of y.*

Remark 2.1. Here we explain the significance of the result above. $\mathcal{L}_0$ is viewed as a differential operator in y, with x being a parameter. For variable elimination in Section 4, the natural *ergodicity assumption* to make is that

$$\mathcal{L}_0\mathbf{1}(y) = 0, \tag{2.12}$$

$$\mathcal{L}_0^*\rho^\infty(y; x) = 0. \tag{2.13}$$

Thus, Lemma 2.2 validates this ergodicity assumption.

3 Main result

As preparation for the main result of this article, we define the *cell problem* as follows:

$$-\mathcal{L}_0\Phi(x,y) = f_1(x,y), \quad \int_{\mathcal{Y}} \Phi(x,y)\rho^{\infty}(y;x)\mathrm{d}y = 0. \tag{3.1}$$

This is a PDE in y, with x as a parameter. The solution $\Phi(x,y)$ will be crucial to the derivation in the next section. By assumptions (A2) and (A3), the existence and uniqueness of the solution to the cell problem (3.1) is guaranteed.

Remark 3.1. The existence and uniqueness for solutions to equation (3.1) is more complicated, in the case the matrix $B(x,y)$ is degenerate or $\mathcal{Y} = \mathbb{R}^d$. Analogous results are still possible, however, in function space settings with appropriate decay properties. See [24–26].

Then define ρ by

$$\begin{aligned}\rho(x) &= \int_{\mathcal{Y}} \rho^{\infty}(y;x)(f_2(x,y) + \nabla_x\Phi(x,y)\cdot f_1(x,y) + g_2(x,y)\cdot\nabla_y\Phi(x,y) + \alpha_2(x,y)\beta^T(x,y) : \nabla_x\nabla_y\Phi(x,y))\mathrm{d}y \\ &=: F_2(x) + F_1(x)\end{aligned} \tag{3.2}$$

and $\sigma(x)$ by

$$\sigma(x)\sigma(x)^T = A_2(x) + \frac{1}{2}(A_1(x) + A_1(x)^T), \tag{3.3}$$

where

$$A_1(x) := 2\int_{\mathcal{Y}} \rho^{\infty}(y;x)[f_1(x,y)\otimes\Phi(x,y) + \beta\alpha_2^T(\nabla_y\Phi)^T]\mathrm{d}y, \tag{3.4}$$

$$A_2(x) := \int_{\mathcal{Y}} \rho^{\infty}(y;x)A(x,y)\mathrm{d}y. \tag{3.5}$$

Now we state our main result, the derivation of which is in the next section.

Theorem 3.1. *For t in scale of $O(1)$, the solution $x(t)$ of (1.1) can be approximated by the solution $X(t)$ of the following SDE,*

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \rho(X) + \sigma(X)\frac{\mathrm{d}W}{\mathrm{d}t}, \quad X(0) = x_0, \tag{3.6}$$

as $\varepsilon \to 0$, where $\rho(X), \sigma(X)$ are defined before, and $W(t)$ stands for a standard Wiener process, i.e., a standard Brownian motion which is independent of U and V.

Note that ρ and σ in equation (3.6) for X both depend on f_0 in the x equation.

Remark 3.2. The big-oh notation O in the theorem above is defined as follows: $f(x) = O(g(x))$ as $x \to x_0$ if there exists a constant C such that $|f(x)| \le C|g(x)|$ for all x sufficiently close to x_0. Thus, the phrase "t in scale of $O(1)$" means that the aforementioned theorem holds in finite time interval.

Remark 3.3. Note that merely from the knowledge of $\sigma(X)\sigma^T(X)$, we cannot determine the diffusion matrix $\sigma(X)$ uniquely. Therefore, the limiting equation (3.6) can neither be uniquely determined by equation (3.3), a consequence of the well-known fact that different SDEs can have the same generator. As a result,

the approximation of the solution of (1.1) by the solution of (3.6) is in the sense of *weak convergence of probability measures*.

4 Derivation

In this section, we give the derivation of Theorem 3.1. By multiscale expansion, we seek the solution of (2.1) with the form

$$v(x, y, t) = v_0(x, y, t) + \varepsilon v_1(x, y, t) + \varepsilon^2 v_2(x, y, t) + \cdots . \tag{4.1}$$

Substitute this expansion into (2.1) and equate the coefficients of equal powers in ε. The first three of the hierarchy of equations are

$$O\left(\frac{1}{\varepsilon^2}\right) : -\mathcal{L}_0 v_0 = 0, \tag{4.2}$$

$$O\left(\frac{1}{\varepsilon}\right) : -\mathcal{L}_0 v_1 = \mathcal{L}_1 v_0, \tag{4.3}$$

$$O(1) : -\mathcal{L}_0 v_2 = -\frac{\partial v_0}{\partial t} + \mathcal{L}_1 v_1 + \mathcal{L}_2 v_0. \tag{4.4}$$

By (2.12), we can deduce, from equation (4.2), $v_0(x, y, t)$ is independent of y, i.e., $v_0 = v_0(x, t)$. As for equation (4.3), the solvability condition is satisfied by assumption (A3). From the expression of (2.4), we have

$$\mathcal{L}_1 v_0 = f_1(x, y)\cdot\nabla_x v_0(x, t).$$

Equation (4.3) becomes

$$-\mathcal{L}_0 v_1 = f_1(x, y)\cdot\nabla_x v_0(x, t). \tag{4.5}$$

The general solution of (4.5) has the form

$$v_1(x, y, t) = \Phi(x, y)\cdot\nabla_x v_0(x, t) + \Phi_1(x, t), \tag{4.6}$$

since $\mathcal{L}_0$ can be viewed as a differential operator in y with x being a parameter.

We set the function Φ_1 to 0 because it plays no role in what follows. Thus, solution v_1 is represented as a linear operator acting on v_0. This form for v_1 is a useful representation since we aim to find a closed equation for v_0. Substituting for v_1 in (4.5) indicates that Φ is the solution to the cell problem (3.1). The assumptions (A2) and (A3) ensure the existence of the solution to the cell problem and the normalization condition, the second equation in (3.1), makes it unique. The right-hand side equation (4.4) becomes

$$-\left(\frac{\partial v_0}{\partial t} - \mathcal{L}_2 v_0 - \mathcal{L}_1(\Phi\cdot\nabla_x v_0)\right).$$

Hence for each fixed x, solvability of (4.4) leads to

$$\frac{\partial v_0}{\partial t} = \int_{\mathcal{Y}} \mathcal{L}_2 v_0(x, t)\rho^{\infty}\mathrm{d}y + \int_{\mathcal{Y}} (y; x)\mathcal{L}_1(\Phi\cdot\nabla_x v_0(x, t))\rho^{\infty}(y; x)\mathrm{d}y =: I_1 + I_2. \tag{4.7}$$

Using the symbols we introduce, the first term on the right-hand side of the equation above is

$$\begin{aligned} I_1 &= \int_{\mathcal{Y}} (f_2\nabla_x + \frac{1}{2}A(x, y) : \nabla_x\nabla_x)v_0(x, t)\rho^{\infty}(y; x)\mathrm{d}y \\ &= F_2(x)\cdot\nabla_x v_0(x, t) + \frac{1}{2}A_2(x) : \nabla_x\nabla_x v_0(x, t). \end{aligned} \tag{4.8}$$

As for the term I_2, note that

$$\begin{aligned}\mathcal{L}_1 &= f_1 \cdot \nabla_x + g_2 \cdot \nabla_y + \frac{1}{2}\alpha_2\beta^T : \nabla_y\nabla_x + \frac{1}{2}\beta\alpha_2^T : \nabla_x\nabla_y \\ &= f_1 \cdot \nabla_x + g_2 \cdot \nabla_y + \beta\alpha_2^T : \nabla_x\nabla_y.\end{aligned} \tag{4.9}$$

The last equality is from the definition of the inner product between matrices.

By direct calculation, we have

$$f_1 \cdot \nabla_x(\Phi \cdot \nabla_x v_0) = f_1 \otimes \Phi : \nabla_x\nabla_x v_0 + (\nabla_x \Phi f_1)\cdot \nabla_x v_0, \tag{4.10}$$

$$g_2 \cdot \nabla_y(\Phi \cdot \nabla_x v_0) = (\nabla_y \Phi g_2)\cdot \nabla_x v_0, \tag{4.11}$$

and

$$\begin{aligned}\beta\alpha_2^T : \nabla_x\nabla_y(\Phi \cdot \nabla_x v_0) &= \beta\alpha_2^T : [(\nabla_x\nabla_y\Phi)\cdot \nabla_x v_0] + \beta\alpha_2^T : [(\nabla_y\Phi)^T\nabla_x\nabla_x v_0] \\ &= (\beta\alpha_2^T : \nabla_x\nabla_y\Phi)\cdot \nabla_x v_0 + [\beta\alpha_2^T(\nabla_y\Phi)^T] : \nabla_x\nabla_x v_0.\end{aligned} \tag{4.12}$$

The last equality holds when $\beta\alpha_1^T$ is symmetric. The first term of the right-hand side of the last equality is from the property $A : (B \cdot \mathbf{c}) = (A : B)\cdot \mathbf{c}$. The second term of the right-hand side of the last equality is from the property $A : (BC) = (AB) : C$ while A, C are both symmetric matrices.

Hence, $I_2 = I_3 + I_4$ where

$$I_3 := \int_{\mathcal{Y}} [\nabla_x \Phi f_1 + \nabla_y \Phi g_2 + (\beta\alpha_2^T : \nabla_x\nabla_y\Phi)]\cdot \nabla_x v_0(x, t)\rho^\infty(y; x)\mathrm{d}y \tag{4.13}$$

and

$$I_4 := \int_{\mathcal{Y}} [f_1 \otimes \Phi + \beta\alpha_2^T(\nabla_y\Phi)^T] : \nabla_x\nabla_x v_0(x, t)\rho^\infty(y; x)\mathrm{d}y. \tag{4.14}$$

Thus,

$$I_2 = F_1(x)\cdot \nabla_x v_0(x, t) + \frac{1}{2}A_1(x) : \nabla_x\nabla_x v_0(x, t). \tag{4.15}$$

By equations (4.7), (4.8), and (4.15), we obtain the following equation:

$$\frac{\partial v_0}{\partial t} = \rho(x)\cdot \nabla_x v_0 + \frac{1}{2}\sigma(x)\sigma(x)^T : \nabla_x\nabla_x v_0. \tag{4.16}$$

This is exactly the backward Kolmogorov equation corresponding to the reduced dynamics given in (3.5). This completes the proof.

Remark 4.1. We note that equation (1.2) for the fast component y contains only one noise term, but the method we exploit still applies in the case that the fast component is driven by two independent Gaussian white noises. To be more exact, consider the following SDEs:

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon}f_1(x, y) + f_2(x, y) + \alpha_1(x, y)\dfrac{\mathrm{d}B_1}{\mathrm{d}t} + \alpha_2(x, y)\dfrac{\mathrm{d}B_2}{\mathrm{d}t}, & (4.17)\\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2}g_1(x, y) + \dfrac{1}{\varepsilon}g_2(x, y) + \dfrac{1}{\varepsilon}\beta_1(x, y)\dfrac{\mathrm{d}B_1}{\mathrm{d}t} + \dfrac{1}{\varepsilon}\beta_2(x, y)\dfrac{\mathrm{d}B_2}{\mathrm{d}t}. & (4.18)\end{cases}$$

The operators $\mathcal{L}_1$ in (2.4) should be replaced by

$$\mathcal{L}_1 = f_1 \cdot \nabla_x + g_2 \cdot \nabla_y + \frac{1}{2}\alpha_2\beta_2^T : \nabla_y\nabla_x + \frac{1}{2}\beta_2\alpha_2^T : \nabla_x\nabla_y. \tag{4.19}$$

Meanwhile, $B(x, y)$ in (2.3) is replaced by

$$B(x, y) = \beta_1\beta_1^T + \beta_2\beta_2^T. \tag{4.20}$$

Consequently, (3.2) and (3.4) are replaced by

$$\begin{aligned}\rho(x) = &\int_{\mathcal{Y}} (f_2(x,y) + \nabla_x\Phi(x,y)\cdot f_1(x,y) + g_2(x,y)\cdot\nabla_y\Phi(x,y) + (\alpha_1(x,y)\beta_1^T(x,y) + \alpha_2(x,y)\beta_2^T(x,y)) \\ &: \nabla_x\nabla_y\Phi(x,y))\rho^\infty(y;x)\mathrm{d}y \\ =&: F_2(x) + F_1(x)\end{aligned} \tag{4.21}$$

and

$$A_1(x) := 2\int_{\mathcal{Y}} [f_1(x,y)\otimes\Phi(x,y) + (\beta_1\alpha_1^T + \beta_2\alpha_2^T)(\nabla_y\Phi)^T]\rho^\infty(y;x)\mathrm{d}y. \tag{4.22}$$

After all these changes, the simplified equation has the same form as (3.6).

5 Example

To illustrate our result, consider the following system in $\mathbb{T}^2$:

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(1-y^2)x + xy + x\dfrac{\mathrm{d}B_1}{\mathrm{d}t} + y\dfrac{\mathrm{d}B_2}{\mathrm{d}t}, \quad x(0) = x_0, & (5.1)\\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2}\alpha y + \dfrac{1}{\varepsilon}\sqrt{2\alpha}\dfrac{\mathrm{d}B_2}{\mathrm{d}t}, \quad y(0) = y_0. & (5.2)\end{cases}$$

Hence, we have the corresponding functions $f_1(x,y) = (1-y^2)x, f_2(x,y) = xy, \alpha_1(x,y) = x, \alpha_2(x,y) = y, g_1(x,y) = \alpha y, g_2(x,y) = 0, \beta(x,y) = \sqrt{2\alpha}$.

In this case, $\mathcal{L}_0 = -\alpha y\frac{\partial}{\partial y} + \alpha\frac{\partial^2}{\partial y^2}$ and the cell problem becomes

$$\alpha y\frac{\partial\Phi}{\partial y} - \alpha\frac{\partial^2\Phi}{\partial y^2} = (1-y^2)x.$$

This equation has a unique centered solution

$$\Phi(y,x) = \frac{1}{2\alpha}(1-y^2)x.$$

It is well known that the second and fourth moments take values 1 and 3, respectively, under the standard normal distribution, by which we obtain

$$\begin{aligned}\frac{1}{2}A_1 &= \int_{\mathcal{Y}}\rho^\infty(y;x)\left[\frac{x^2(1-y^2)^2}{2\alpha} - \sqrt{\frac{2}{\alpha}}xy^2\right]\mathrm{d}y = \frac{1}{\alpha}x^2 - \sqrt{\frac{2}{\alpha}}x,\\ A_2 &= \int_{\mathcal{Y}}\rho^\infty(y;x)(x^2+y^2)\mathrm{d}y = x^2+1,\\ F_1 &= \int_{\mathcal{Y}}\rho^\infty(y;x)\left[\frac{x(1-y^2)^2}{2\alpha} - \sqrt{\frac{2}{\alpha}}y^2\right]\mathrm{d}y = \frac{1}{\alpha}x - \sqrt{\frac{2}{\alpha}},\\ F_2 &= \int_{\mathcal{Y}}xy\rho^\infty(y;x)\mathrm{d}y = 0.\end{aligned}$$

Thus, we derive the simplified equation as follows to approximate the slow component $x(t)$:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{X}{\alpha} - \sqrt{\frac{2}{\alpha}} + \sqrt{\left(1+\frac{2}{\alpha}\right)X^2 - \sqrt{\frac{8}{\alpha}}X + 1}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{5.3}$$

where $W(t)$ is a standard Brownian motion.

6 Conclusion

In this article, we study the averaging principle for systems of SDEs with two widely separated time scales, which is driven by correlated noises. By means of perturbation expansion of the solution to the backward Kolmogorov equation and consequent elimination of variables, we obtain the backward equation corresponding to the reduced simplified system. The solution to reduced SDE (3.6) approximates the slow variable $x(t)$ of the initial systems (1.1) and (1.2) in the weak sense.

Funding information: Tao Jiang is supported by the National Natural Science Foundation of China (11901178). Yancai Liu is supported by the High-Level Personal Foundation of Henan University of Technology (2020BS049).

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations*, Gordon and Breach Science Publishers, New York, 1961.
[2] R. Z. Khasminskii, *On the principle of averaging the Itô stochastic differential equations*, Kibernetika **4** (1968), 260–279, (in Russian).
[3] M. I. Freidlin and A. D. Wentzell, *Long-time behavior of weakly coupled oscillators*, J. Stat. Phys. **123** (2006), 1311–1337.
[4] D. Givon, *Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems*, Multiscale Model. Simul. **6** (2007), no. 2, 577–594.
[5] D. Givon, I. G. Kevrekidis, and R. Kupferman, *Strong convergence of projective integration schemes for singular perturbed stochastic differential systems*, Commun. Math. Sci. **4** (2006), 707–729.
[6] J. Golec, *Stochastic averaging principle for systems with pathwise uniqueness*, Stoch. Anal. Appl. **13** (1995), no. 3, 307–322.
[7] J. Golec and G. Ladde, *Averaging principle and systems of singularly perturbed stochastic differential equations*, J. Math. Phys. **31** (1990), 1116–1123.
[8] A. Y. Veretennikov, *On the averaging principle for systems of stochastic differential equations*, Mat. Sb. **69** (1991), no. 1, 271–284.
[9] A. Y. Veretennikov, *On large deviations in the averaging principle for SDEs with full dependence*, Ann. Probab. **27** (1999), no. 1, 284–296.
[10] J. Bao, G. Yin, and C. Yuan, *Two-time-scale stochastic partial differential equations driven by α-stable noises: averaging principles*, Bernoulli **23** (2017), no. 1, 645–669.
[11] C.-E. Bréhier, *Strong and weak orders in averaging for SPDEs*, Stochastic Process. Appl. **122** (2012), no. 7, 2553–2593.
[12] S. Cerrai and M. I. Freidlin, *Averaging principle for a class of stochastic reaction-diffusion equations*, Probab. Theory Related Fields **144** (2009), 137–177.
[13] S. Cerrai and A Khasminkii *type averaging principle for stochastic reaction-diffusion equations*, Ann. Appl. Probab. **19** (2009), no. 3, 899–948.
[14] S. Cerrai, *Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise*, SIAM J. Math. Anal. **43** (2011), no. 6, 2482–2518.
[15] H. Fu and J. Duan, *An averaging principle for two-scale stochastic partial differential equations*, Stoch. Dyn. **11** (2011), no. 2–3, 353–367.
[16] H. Fu and J. Liu, *Strong convergence in stochastic averaging principle for two time-scales stochastic partial differential equations*, J. Math. Anal. Appl. **384** (2011), no. 1, 70–86.

[17] H. Fu, L. Wan, and J. Liu, *Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with twotime-scales*, Stochastic Process Appl. **125** (2015), no. 8, 3255–3279.
[18] B. Pei, Y. Xu, and J. L. Wu, *Two-time-scales hyperbolic-parabolic equations driven by Poisson random measures: existence, uniqueness and averaging principles*, J. Math. Anal. Appl. **447** (2017), no. 1, 243–268.
[19] W. Wang and A. J. Roberts, *Average and deviation for slow-fast stochastic partial differential equations*, J. Differential Equations **253** (2012), no. 5, 1265–1286.
[20] J. Xu, *L^p– strong convergence of the averaging principle for slow-fast SPDEs with jumps*, J. Math. Anal. Appl. **445** (2017), no. 1, 342–373.
[21] J. Xu, Y. Miao, and J. Liu, *Strong averaging principle for two-time-scale non-autonomous stochastic FitzHugh-Nagumo system with jumps*, J. Math. Phys. **57** (2016), no. 9, 092704.
[22] J. Xu, Y. Miao, and J. Liu, *Strong averaging principle for slow-fast SPDEs with Poisson random measures*, Discrete Contin. Dyn. Syst. Ser. B **20** (2015), no. 7, 2233–2256.
[23] G. A. Pavliotis and A. M. Stuart, *Multiscale Methods: Averaging and Homogenization*, Springer, New York, 2008.
[24] E. Pardoux and Yu. Verrtennikov, *On the Poisson equation and diffusion approximation. I*, Ann. Probab. **29** (2001), no. 3, 1061–1085.
[25] E. Pardoux and Yu. Verrtennikov, *On the Poisson equation and diffusion approximation. II*, Ann. Probab. **31** (2003), no. 3, 1166–1192.
[26] E. Pardoux and Yu. Verrtennikov, *On the Poisson equation and diffusion approximation. III*, Ann. Probab. **33** (2005), no. 3, 1111–1133.