



# The Order of Convergence in the Averaging Principle for Slow-Fast Systems of Stochastic Evolution Equations in Hilbert Spaces

Filippo de Feo<sup>1</sup>

Accepted: 15 May 2023 / Published online: 31 May 2023  
© The Author(s) 2023

## Abstract

In this work we are concerned with the study of the strong order of convergence in the averaging principle for slow-fast systems of stochastic evolution equations in Hilbert spaces with additive noise. In particular the stochastic perturbations are general Wiener processes, i.e their covariance operators are allowed to be not trace class. We prove that the slow component converges strongly to the averaged one with order of convergence  $1/2$  which is known to be optimal. Moreover we apply this result to a slow-fast stochastic reaction diffusion system where the stochastic perturbation is given by a white noise both in time and space.

**Keywords** Averaging principle · Order of convergence · Slow fast stochastic differential equations · SPDE's · Stochastic reaction diffusion system

## 1 Introduction

Consider the following slow-fast system of abstract stochastic evolution equations with additive noise

$$\begin{cases} dU_t^\epsilon = A_1 U_t^\epsilon dt + F(U_t^\epsilon, V_t^\epsilon) dt + dW_t^{Q_1}, \\ U_0^\epsilon = u \in H, \\ dV_t^\epsilon = \epsilon^{-1} A_2 V_t^\epsilon dt + \epsilon^{-1} G(V_t^\epsilon) dt + \epsilon^{-1/2} dW_t^{Q_2}, \\ V_0^\epsilon = v \in K, \end{cases} \quad (1)$$

where  $\epsilon > 0$  is a small parameter representing the ratio of time scales between the slow component of the system  $U^\epsilon$  and the fast one  $V^\epsilon$ . Here  $H, K$  are Hilbert spaces,

---

✉ Filippo de Feo  
filippo.defeo@polimi.it

<sup>1</sup> Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

$A_1, A_2$  are unbounded linear operators on  $H, K$  respectively and  $W^{Q_1}, W^{Q_2}$  are Wiener processes on  $H, K$  respectively.

Slow-fast systems are very used in applications since it is very natural for real-world systems to present very different time-scales. We refer the reader for example to [23] for applications to physics, [32] to chemistry, [34] to neurophysiology, [1, 13, 21, 22] to mathematical finance (see also [12] for a slightly different financial model) and the references therein.

A natural idea is then to study the behaviour of the system when  $\epsilon \rightarrow 0$ . In particular under certain hypotheses it is known that the slow component  $U^\epsilon$  converges to the solution  $U$  of the so called averaged equation

$$\begin{cases} dU_t = A_1 U_t dt + \bar{F}(U_t) dt + dW_t^{Q_1}, \\ U_0 = u, \end{cases}$$

where

$$\bar{F}(u) = \int_K F(u, v) \mu(dv),$$

and  $\mu$  is the invariant measure related to the fast motion, i.e.

$$\begin{cases} dv_t^v = A_2 v_t^v dt + G(v_t^v) dt + dw_t^{Q_2}, \\ v_0^v = v \in K. \end{cases}$$

Note that the equation for  $U$  is uncoupled from  $V^\epsilon$ . This fact is known as averaging principle and it is fundamental in applications since  $U$  captures the effective dynamic of  $U^\epsilon$  (which is usually the most interesting variable in applications) and it is then a rigorous dimensionality reduction of the original system.

The first general result for the averaging principle for finite-dimensional stochastic differential equations can be found in [27]. For generalizations and improvements see [13, 17, 18, 23, 31, 40, 41] and the references therein. It is important to mention that the drift of the fast equation is allowed to depend also on the slow component (i.e. fully coupled system) and the stochastic perturbations of the slow and fast equations are allowed to be multiplicative, i.e. the diffusion coefficient of the slow equation can depend on both the slow and fast variables. Moreover when the diffusion coefficient of the slow equation is independent of the fast variable then a strong convergence in probability is obtained. Otherwise only a weak convergence can be proved.

The averaging principle for infinite dimensional systems follows more delicate arguments: for this we refer to [5, 6, 8, 9, 14, 15, 24, 39] and the references therein. Also for infinite dimensional systems the previous comment about the dependence of the diffusion coefficients holds. See also [19, 20, 38] for optimal control problems of slow-fast systems in infinite dimension.

However for numerical applications it is very important to know the speed of convergence for which  $U^\epsilon \rightarrow U$ , e.g. see [3, 29]. For the study of the order of convergence for finite dimensional systems we refer to [16, 25, 28, 30, 35] and the references therein.

It is important to mention that the order of convergence can be studied in two ways: in the strong sense and in the weak sense. Moreover the optimal order for the strong and weak convergence are known to be  $1/2$  and  $1$  respectively.

Recently the problem of estimating the order of convergence in the averaging principle for infinite dimensional systems systems is being addressed by researchers:

in [4] the author (generalizing his previous work [2]) considers a slow-fast stochastic reaction diffusion system with additive noise. Both the weak and strong orders of convergence are obtained: in particular under strong regularity of the noise (it is for example assumed that the covariance operator is trace class but for the precise statement see [4]) it is proved that the strong order of convergence is  $1/2$  and the weak order is  $1$  with both orders being optimal. Instead under more general assumptions on the noise only weaker orders of convergence are obtained for both the strong and weak convergence.

In [34] a 1-dimensional fully coupled reaction-diffusion system is considered and the strong order of convergence is proved to be  $1/2$  under very strong assumptions on the covariance operators of the noises, i.e.  $Tr(\Delta^{1/2}Q_i) < \infty$ , where  $\Delta$  is the Laplacian.

In [36] the strong order of convergence for a fully coupled slow-fast stochastic system is studied. Here it is assumed that the the covariance operators of the noises are trace-class and moreover that  $Tr(-A_1Q_1) < \infty$ .

See also [26] where the weak order of convergence for a stochastic wave equation with fast oscillation given by a fast reaction-diffusion stochastic system is proved to be  $1$ . Also here it assumed  $Tr(Q_i) < \infty$ .

Indeed in all these papers the case  $Tr(Q_i) = \infty$ , which is very important for applications as it happens very naturally for example when the stochastic perturbation is a white noise i.e.  $Q_i = I$ , can't be treated.

In this manuscript we are then interested in studying the strong order of convergence for the slow-fast infinite-dimensional system of stochastic evolution equations (1) where  $W^{Q_1}, W^{Q_2}$  are general Wiener processes on  $H, K$  respectively with covariance operators  $Q_1, Q_2$  with  $Tr(Q_1) = +\infty, Tr(Q_2) = +\infty$  possibly. Under some hypotheses, see Hypotheses 1, 2, 3, 4, 5 below, we prove that the strong order of convergence is  $1/2$  which is known to be optimal. In particular we show in Theorem 1 that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t^\epsilon - U_t|^2 \right] \leq C\epsilon$$

where  $U_t$  is the solution of the averaged equation. Notice that this result is much stronger than [4, 36] where  $\sup_{t \in [0, T]}$  is outside the expectation.

The key tool in the proof of Theorem 1 is Proposition 3. The proof of this proposition is based on a technical result, i.e. Lemma 6.3, which is inspired by [7], and it is a consequence of the mixing properties of the fast motion, i.e. Lemmas 4.5, 4.6, 4.7. We recall that [7] studies the normal deviations, i.e. the weak convergence of  $Z^\epsilon := (U^\epsilon - U)/\sqrt{\epsilon}$ , when the equation for the slow component has no stochastic perturbation ( $Q_1 = 0$ ).

Finally we discuss an application of our theory to a 1-dimensional slow-fast stochastic reaction diffusion system where the stochastic perturbation is given by a white noise both in time and space which to the best of our knowledge, as said before, can not be treated by the existing literature.

The paper is organized as follows: in Sect. 2 we introduce the problem in a formal way and we state the assumptions that we will use. In Sect. 3 we prove some a-priori estimates. In Sect. 4 we prove some results related to the fast motion. In Sect. 5 we study the well posedness of the averaged equation. In Sect. 6 we prove some preliminary results. In Sect. 7 we prove that the order of convergence is 1/2 and we give an application of our theory.

## 2 Setup and Assumptions

In this section we define the notation and the assumptions for the rest of the paper.

$H, K$  will be Hilbert spaces with scalar products  $\langle \cdot, \cdot \rangle_H, \langle \cdot, \cdot \rangle_K$  and  $|\cdot|_H, |\cdot|_K$  the induced norms.

$B_B(H)$  will denote the space of bounded functions  $\phi: H \rightarrow \mathbb{R}$  with the sup norm  $|\cdot|_{H,\infty}$ .

$Lip(H)$  will denote the set of Lipschitz functions  $\phi: H \rightarrow \mathbb{R}$  and set

$$[\phi]_{H,Lip} = \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{|x - y|_H}.$$

$\mathcal{L}(H)$  will denote the space of linear bounded operators from  $H$  to  $H$ , endowed with the operator norm

$$\|L\|_H = \sup_{|x|_H=1} |Lx|_H.$$

Next denote by  $\mathcal{L}_2(H)$  the space of Hilbert-Schmidt operators endowed with the norm

$$\|L\|_{\mathcal{L}_2(H)} = (Tr(L^*L))^{1/2}.$$

The analogous spaces  $B_B(K), Lip(K), \mathcal{L}(K), \mathcal{L}_2(K)$  are defined for the Hilbert space  $K$  with the corresponding norms  $|\cdot|_{K,\infty}, [\cdot]_{K,Lip}, \|\cdot\|_K, \|\cdot\|_{\mathcal{L}_2(K)}$ .

In order to simplify the notation we will omit the subscripts  $K$  and  $H$  in the various norms when no confusion is possible.

$\mathcal{B}(H)$  and  $\mathcal{B}(K)$  will denote the Borel sigma-algebra in  $H$  and  $K$  respectively.

Consider now the following infinite dimensional system for  $0 \leq t \leq T < \infty$

$$\begin{cases} dU_t^\epsilon = A_1 U_t^\epsilon dt + F(U_t^\epsilon, V_t^\epsilon) dt + dW_t^{Q_1}, \\ U_0^\epsilon = u \in H, \\ dV_t^\epsilon = \epsilon^{-1} A_2 V_t^\epsilon dt + \epsilon^{-1} G(V_t^\epsilon) dt + \epsilon^{-1/2} dW_t^{Q_2}, \\ V_0^\epsilon = v \in K, \end{cases} \tag{2}$$

where

- $\epsilon > 0$  is a parameter,
- $A_1: D(A_1) \subset H \rightarrow H, A_2: D(A_2) \subset K \rightarrow K$  are linear operators,
- $F: H \times K \rightarrow H, G: K \rightarrow K,$
- $W^{Q_1}, W^{Q_2}$  are independent cylindrical Wiener processes on  $H, K$  respectively with covariance operator  $Q_1, Q_2$  respectively and they are defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a normal filtration  $\mathcal{F}_t, t \geq 0.$

We now state the assumptions that we will use throughout the work:

**Hypothesis 1**  $A_1: D(A_1) \subset H \rightarrow H$  is a linear operator generator of an analytical semigroup  $e^{A_1 t}$  on  $H, t \geq 0.$

Moreover there exist an orthonormal basis  $\{e_k\}_{k \in \mathbb{N}}$  of  $H$  and  $\{\alpha_k\}_{k \in \mathbb{N}} \subset (0, +\infty)$  such that

$$A_1 e_k = -\alpha_k e_k.$$

Moreover we assume that there exist  $\zeta > 0, n \geq 2$  integer and  $1/(2n) < \beta < 1/3$  such that

$$\sum_{k=1}^{\infty} \alpha_k^{-\zeta} < +\infty$$

and

$$\sum_{k=1}^{\infty} \alpha_k^{n(\zeta+2\beta-1)-\zeta} < +\infty.$$

**Remark 2.1** Hypothesis 1 is necessary in the proof of Proposition 3. It holds for example when  $A_1 = \Delta$  is the Laplacian on  $[0, L].$  Indeed in this case it is well known that  $\alpha_k = Ck^2$  and then the two series converge for example by choosing  $\zeta = \frac{3}{5}, n = 3, \beta = \frac{1}{5}.$

From Hypothesis 1 we have the following spectral representation:

$$e^{A_1 t} x = \sum_{k=1}^{\infty} e^{-\alpha_k t} \langle x, e_k \rangle e_k, \quad \forall t > 0. \tag{3}$$

Moreover we can define the fractional powers of  $-A_1$  denoted by  $(-A_1)^\theta$  for  $\theta \geq 0$  with domain  $D((-A_1)^\theta).$  We will denote by  $|\cdot|_\theta$  the norm  $|\cdot|_{D((-A_1)^\theta)} = |(-A_1)^\theta \cdot|.$

The following standard results holds, e.g. see [33, Chapter 2, Theorem 6.13],

$$\|(-A_1)^\theta e^{A_1 t}\| \leq C_\theta t^{-\theta} e^{-\nu t}, \quad \forall \theta \geq 0, t > 0, \tag{4}$$

for some  $\nu > 0$  and

$$|e^{A_1 t} x - x| \leq C_\theta t^\theta |x|_\theta, \quad \forall x \in D((-A_1)^\theta), 0 < \theta \leq 1, t \geq 0. \tag{5}$$

**Hypothesis 2**  $A_2: D(A_2) \subset K \rightarrow K$  is a linear operator generator of a  $C_0$ -semigroup  $e^{A_2t}$  on  $K$ ,  $t \geq 0$ .

Moreover there exists  $\lambda > 0$  such that

$$\|e^{A_2t}\| \leq e^{-\lambda t}, \tag{6}$$

for every  $t \geq 0$ .

**Hypothesis 3** There exist  $L_F, L_G > 0$  such that

$$\begin{aligned} |F(x_2, y_2) - F(x_1, y_1)| &\leq L_F(|x_2 - x_1| + |y_2 - y_1|), \\ |G(y_2) - G(y_1)| &\leq L_G|y_2 - y_1|, \end{aligned}$$

for every  $x_1, x_2 \in H, y_1, y_2 \in K$ .

Moreover we assume that

$$L_G < \lambda.$$

We remark that this implies that  $A_2 + G(\cdot)$  is strongly dissipative, i.e. set

$$\delta = \frac{\lambda - L_G}{2} > 0, \tag{7}$$

then it holds:

$$\langle A_2(y_2 - y_1) + G(y_2) - G(y_1), y_2 - y_1 \rangle \leq -2\delta|y_2 - y_1|^2, \tag{8}$$

for every  $y_1, y_2 \in D(A_2)$ .

**Hypothesis 4** There exist  $C > 0, \gamma \in (0, 1/2)$  such that

$$\begin{aligned} \|e^{A_1t} Q_1^{1/2}\|_{\mathcal{L}_2} &\leq C(t \wedge 1)^{-\gamma}, \\ \|e^{A_2t} Q_2^{1/2}\|_{\mathcal{L}_2} &\leq C(t \wedge 1)^{-\gamma}, \end{aligned}$$

for every  $t > 0$ .

**Hypothesis 5** Assume that  $Q_2$  is invertible (with inverse  $Q_2^{-1} \in \mathcal{L}(K)$ ).

**Remark 2.2** Hypotheses 4, 5 hold for example choosing  $A_1 = A_2 = A$  to be the Laplacian on  $[0, L]$  and  $Q_1 = Q_2 = Q = I$ .

Indeed Hypothesis 5 is immediately satisfied. Moreover by setting  $H = K = L^2([0, L])$ , we have that  $Ae_k = -Ck^2e_k$  for some orthonormal basis  $\{e_k\}$  of eigenvectors of  $A$ . Thus, by the spectral representation (3) with  $\alpha_k = Ck^2$ , we have:

$$\|e^{At} Q^{1/2}\|_{\mathcal{L}_2}^2 = \|e^{At}\|_{\mathcal{L}_2}^2 = \sum_{h=1}^{\infty} |e^{At} e_h|^2 = \sum_{h=1}^{\infty} e^{-2Ch^2t} \leq \int_0^{\infty} e^{-2Ch^2t} dh$$

$$= \frac{1}{\sqrt{2Ct}} \int_0^\infty e^{-y^2} dy = Ct^{-1/2},$$

where the inequality follows since  $\forall t > 0$  the function  $h \mapsto e^{-Ch^2t}$  is non-increasing. This shows that Hypothesis 4 holds with  $\gamma = 1/4$ .

**Proposition 1** *Let  $u \in H, v \in K$ , under Hypotheses 1, 2, 3, 4, 5 there exists a unique mild solution of (2) given by*

$$\begin{cases} U_t^\epsilon = e^{A_1 t} u + \int_0^t e^{A_1(t-s)} F(U_s^\epsilon, V_s^\epsilon) ds + \int_0^t e^{A_1(t-s)} dW_s^{Q_1}, \\ V_t^\epsilon = e^{A_1 t} v + \int_0^t e^{A_2(t-s)} G(V_s^\epsilon) ds + \int_0^t e^{A_2(t-s)} dW_s^{Q_2}, \end{cases} \tag{9}$$

for every  $t \in [0, T]$ .

**Proof** See e.g. [10]. □

In the sequel we will always assume that Hypotheses 1, 2, 3, 4, 5 hold. Moreover  $C > 0$  will denote a generic constant independent of  $\epsilon$  which may change from line to line.

### 3 A Priori Estimates

In this section we prove some classical a-priori estimates for the slow and fast components.

In the following for every  $t \geq 0$  denote by

$$\Gamma_t^1 = \int_0^t e^{A_1(t-s)} dW_s^{Q_1}$$

and

$$\Gamma_t^{2\epsilon} = \frac{1}{\sqrt{\epsilon}} \int_0^t e^{A_2(t-s)/\epsilon} dW_s^{Q_2}$$

the stochastic convolutions.

First we prove some estimates related to  $\Gamma_t^1$  and  $\Gamma_t^{2\epsilon}$ .

**Lemma 3.1**

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \left| \Gamma_t^1 \right|_\theta^p \right] < +\infty,$$

for every  $0 \leq \theta < 1/2 - \gamma, p \geq 1$ .

**Proof** Fix  $0 < \eta < 1/2$ , by the factorization method, e.g. see [10, Chapter 5, Section 3], we have:

$$\Gamma_t^1 = C \int_0^t e^{A_1(t-\rho)}(t-\rho)^{\eta-1} Y_\rho d\rho$$

where

$$Y_\rho = \int_0^\rho e^{A_1(\rho-s)}(\rho-s)^{-\eta} dW_s^{Q_1}.$$

Now fix  $0 \leq \theta < 1/2 - \gamma$ , then by Holder’s inequality:

$$\mathbb{E} \left[ \sup_{t \leq T} \left| \Gamma_t^1 \right|_\theta^p \right] \leq C \int_0^T \mathbb{E} |Y_\rho|_\theta^p d\rho, \tag{10}$$

for every  $p > 1/\eta > 2$ .

Now we estimate  $\mathbb{E}|Y_\rho|_\theta^p$ ; since  $Y_\rho$  is a Gaussian random variable, by Ito’s isometry, (4) and Hypothesis 4 we have:

$$\begin{aligned} \mathbb{E}|Y_\rho|_\theta^p &\leq C \left( \mathbb{E}|Y_\rho|_\theta^2 \right)^{p/2} = C \left( \int_0^\rho \left\| (-A_1)^\theta e^{A_1(\rho-s)}(\rho-s)^{-\eta} Q_1^{1/2} \right\|_{\mathcal{L}_2}^2 ds \right)^{p/2} \\ &\leq C \left( \int_0^\rho (\rho-s)^{-2\eta} \left\| (-A_1)^\theta e^{A_1(\rho-s)/2} \right\|^2 \left\| e^{A_1(\rho-s)/2} Q_1^{1/2} \right\|_{\mathcal{L}_2}^2 ds \right)^{p/2} \\ &\leq C \left( \int_0^\rho (\rho-s)^{-2(\eta+\theta+\gamma)} ds \right)^{p/2} < C, \end{aligned}$$

for every  $\rho \leq T$  and  $\theta, \eta$  such that  $0 \leq \theta + \eta < 1/2 - \gamma$ .

Next inserting the last inequality in (10) and recalling that  $p > 1/\eta$ , which yields  $p > (1/2 - \gamma - \theta)^{-1}$  we obtain the thesis of the Lemma for  $0 \leq \theta < 1/2 - \gamma$ ,  $p > (1/2 - \gamma - \theta)^{-1} > 2$ .

Finally by Holder’s inequality we have the thesis of the Lemma. □

**Lemma 3.2** *Let  $p \geq 2$ , then there exists  $C = C(p) > 0$  such that*

$$\sup_{t>0} \mathbb{E} \left[ \left| \Gamma_t^{2\epsilon} \right|^p \right] \leq C,$$

for every  $\epsilon > 0$ .

**Proof** For  $t > 0$ , by [10, Theorem 4.36] and by our hypotheses we have

$$\begin{aligned} \mathbb{E} \left[ \left| \Gamma_t^{2\epsilon} \right|^p \right] &\leq C \left( \int_0^t \frac{1}{\epsilon} \left\| e^{A_2(t-s)/\epsilon} Q_2^{1/2} \right\|_{\mathcal{L}_2}^2 ds \right)^{p/2} \\ &= C \left( \int_0^{t/\epsilon} \left\| e^{A_2\rho} Q_2^{1/2} \right\|_{\mathcal{L}_2}^2 d\rho \right)^{p/2} \end{aligned}$$



$$\begin{aligned} &\leq C \left( \int_0^{t/\epsilon} \|e^{A_2 \rho/2}\|^2 \|e^{A_2 \rho/2} Q_2^{1/2}\|_{\mathcal{L}_2}^2 d\rho \right)^{p/2} \\ &\leq C \left( \int_0^{t/\epsilon} e^{-\lambda \rho} \left(\frac{\rho}{2} \wedge 1\right)^{-2\gamma} d\rho \right)^{p/2} \\ &\leq C \left( \int_0^\infty e^{-\lambda \rho} \left(\frac{\rho}{2} \wedge 1\right)^{-2\gamma} d\rho \right)^{p/2} \leq C, \end{aligned}$$

so that the thesis is proved. □

**Lemma 3.3** *Let  $p \geq 2$  then there exists  $C = C(T, p) > 0$  such that*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t^\epsilon|^p \right] \leq C(1 + |u|^p + |v|^p) \tag{11}$$

and

$$\sup_{t \in [0, T]} \mathbb{E} [ |V_t^\epsilon|^p ] \leq C(1 + |u|^p + |v|^p), \tag{12}$$

for every  $\epsilon > 0$ .

**Proof** Define

$$\Lambda_t^{1\epsilon} = U_t^\epsilon - \Gamma_t^1,$$

so that

$$d\Lambda_t^{1\epsilon} = A_1 \Lambda_t^{1\epsilon} dt + F(\Lambda_t^{1\epsilon} + \Gamma_t^1, V_t^\epsilon) dt.$$

By Young’s inequality and Hypotheses 1, 3 we have:

$$\begin{aligned} \frac{1}{p} \frac{d}{dt} |\Lambda_t^{1\epsilon}|^p &= \langle A_1 \Lambda_t^{1\epsilon}, \Lambda_t^{1\epsilon} \rangle |\Lambda_t^{1\epsilon}|^{p-2} + \langle F(\Lambda_t^{1\epsilon} + \Gamma_t^1, V_t^\epsilon), \Lambda_t^{1\epsilon} \rangle |\Lambda_t^{1\epsilon}|^{p-2} \\ &\leq C |\Lambda_t^{1\epsilon}|^p + C |F(\Lambda_t^{1\epsilon} + \Gamma_t^1, V_t^\epsilon)|^p \\ &\leq C |\Lambda_t^{1\epsilon}|^p + C \left( 1 + |\Gamma_t^1|^p + |V_t^\epsilon|^p \right), \end{aligned}$$

for every  $t \leq T$ .

Then by the comparison Theorem we have

$$|\Lambda_t^{1\epsilon}|^p \leq C|u|^p + C \int_0^t \left( 1 + |\Gamma_s^1|^p + |V_s^\epsilon|^p \right) ds, \tag{13}$$

for every  $t \leq T$ .

Then by the definition of  $\Lambda^{1\epsilon}$  and this last inequality it follows

$$|U_t^\epsilon|^p \leq C \left( 1 + |u|^p + \int_0^t \left( |\Gamma_s^1|^p + |V_s^\epsilon|^p \right) ds + |\Gamma_t^1|^p \right), \tag{14}$$

for every  $t \leq T$ .

Now by Lemma 3.1 we have:

$$\mathbb{E} \left[ \sup_{t \leq \tau} |U_t^\epsilon|^p \right] \leq C \left( 1 + |u|^p + \int_0^\tau \mathbb{E} |V_r^\epsilon|^p dr \right), \tag{15}$$

for every  $\tau \leq T$ .

Now we proceed in a similar way to [19, proof of Lemma 3.10], i.e. set

$$\Lambda_t^{2\epsilon} = e^{\frac{2\delta}{\epsilon}t} \left( V_t^\epsilon - \Gamma_t^{2\epsilon} \right),$$

so that

$$d\Lambda_t^{2\epsilon} = \frac{2\delta}{\epsilon} \Lambda_t^{2\epsilon} dt + \frac{1}{\epsilon} A_2 \Lambda_t^{2\epsilon} dt + \frac{1}{\epsilon} e^{\frac{2\delta}{\epsilon}t} G(e^{-\frac{2\delta}{\epsilon}t} \Lambda_t^{2\epsilon} + \Gamma_t^{2\epsilon}) dt.$$

Now by (8) it follows

$$\begin{aligned} d|\Lambda_t^{2\epsilon}|^2 &= \frac{4\delta}{\epsilon} |\Lambda_t^{2\epsilon}|^2 + \frac{2}{\epsilon} \langle A_2 \Lambda_t^{2\epsilon}, \Lambda_t^{2\epsilon} \rangle dt + \frac{2}{\epsilon} e^{2\delta t/\epsilon} \langle G(e^{-2\delta t/\epsilon} \Lambda_t^{2\epsilon} + \Gamma_t^{2\epsilon}), \Lambda_t^{2\epsilon} \rangle dt \\ &\quad - \frac{2}{\epsilon} e^{2\delta t/\epsilon} \langle G(\Gamma_t^{2\epsilon}), \Lambda_t^{2\epsilon} \rangle dt + \frac{2}{\epsilon} e^{2\delta t/\epsilon} \langle G(\Gamma_t^{2\epsilon}), \Lambda_t^{2\epsilon} \rangle dt \\ &\leq \frac{2}{\epsilon} e^{2\delta t/\epsilon} \langle G(\Gamma_t^{2\epsilon}), \Lambda_t^{2\epsilon} \rangle dt. \end{aligned}$$

Now similarly to [19, proof of Lemma 3.10] fix  $\theta > 0$ , differentiate  $f(x) = \sqrt{x + \theta}$  and use the previous inequality, then:

$$\begin{aligned} |\Lambda_t^{2\epsilon}| &\leq \sqrt{|\Lambda_t^{2\epsilon}|^2 + \theta} \\ &\leq \sqrt{|v|^2 + \theta} + \int_0^t \frac{1}{\epsilon} \frac{1}{\sqrt{|\Lambda_s^{2\epsilon}|^2 + \theta}} e^{\frac{2\delta}{\epsilon}s} |G(\Gamma_s^{2\epsilon})| |\Lambda_s^{2\epsilon}| ds. \end{aligned}$$

Now by dominated convergence for  $\theta \rightarrow 0$  we have:

$$|\Lambda_t^{2\epsilon}| \leq |v| + \int_0^t \frac{1}{\epsilon} e^{\frac{2\delta}{\epsilon}s} |G(\Gamma_s^{2\epsilon})| ds,$$

so that by recalling the definition of  $\Lambda_t^{2\epsilon}$  we have:

$$|V_t^\epsilon| \leq e^{-\frac{2\delta}{\epsilon}t} |v| + \int_0^t \frac{1}{\epsilon} e^{-\frac{2\delta}{\epsilon}(t-s)} |G(\Gamma_s^{2\epsilon})| ds + |\Gamma_t^{2\epsilon}|$$

$$= e^{-\frac{2\delta}{\epsilon}t}|v| + \int_0^{t/\epsilon} e^{-2\delta(t/\epsilon-s)}|G(\Gamma_{\epsilon s}^{2\epsilon})|ds + |\Gamma_t^{2\epsilon}|.$$

Then, by Holder’s inequality and Lemma 3.2, we have:

$$\begin{aligned} \mathbb{E}|V_t^\epsilon|^p &\leq C|v|^p + C\mathbb{E}\left|\int_0^{t/\epsilon} e^{-2\delta(t/\epsilon-s)}|G(\Gamma_{\epsilon s}^{2\epsilon})|ds\right|^p + C\mathbb{E}\left|\Gamma_t^{2\epsilon}\right|^p \\ &\leq C|v|^p + C\left(\int_0^{t/\epsilon} e^{-\frac{2\delta p}{(p-1)(t/\epsilon-s)}ds}\right)^{p-1} \int_0^{t/\epsilon} e^{-2\delta p(t/\epsilon-s)}\mathbb{E}|G(\Gamma_{\epsilon s}^{2\epsilon})|^p ds \\ &\quad + C\mathbb{E}\left|\Gamma_t^{2\epsilon}\right|^p \\ &= C|v|^p + C\left(\int_0^{t/\epsilon} e^{-\frac{2\delta p}{(p-1)\sigma}d\sigma}\right)^{p-1} \int_0^t \frac{1}{\epsilon}e^{-2\delta p(t-s)/\epsilon}\mathbb{E}|G(\Gamma_s^{2\epsilon})|^p ds \\ &\quad + C\mathbb{E}\left|\Gamma_t^{2\epsilon}\right|^p \\ &\leq C|v|^p + C\left(\int_0^{t/\epsilon} e^{-\frac{2\delta p}{(p-1)\sigma}d\sigma}\right)^{p-1} \int_0^t \frac{1}{\epsilon}e^{-2\delta p(t-s)/\epsilon}(1 + \mathbb{E}|\Gamma_s^{2\epsilon}|^p)ds \\ &\quad + C\mathbb{E}\left|\Gamma_t^{2\epsilon}\right|^p \\ &\leq C|v|^p + C\left(\int_0^{t/\epsilon} e^{-\frac{2\delta p}{(p-1)\sigma}d\sigma}\right)^{p-1} \int_0^t \frac{1}{\epsilon}e^{-2\delta p(t-s)/\epsilon}ds + C \\ &\leq C(1 + |v|^p). \end{aligned}$$

This proves (12).

Finally inserting (12) into (15) we have (11). □

**Lemma 3.4** *Let  $0 < \alpha < 1/2 - \gamma$ ,  $u \in D((-A_1)^\alpha)$ ,  $v \in K$ , then there exists  $C = C(T, \alpha) > 0$  such that*

$$\mathbb{E}\left[\sup_{t \in [0, T]} |U_t^\epsilon|_\alpha^2\right] \leq C(1 + |u|_\alpha^2 + |v|^2),$$

for every  $\epsilon > 0$ .

**Proof** Consider for  $t \leq T$

$$U_t^\epsilon = e^{A_1 t}u + \int_0^t e^{A_1(t-s)}F(U_s^\epsilon, V_s^\epsilon)ds + \int_0^t e^{A_1(t-s)}dW_s^{Q_1}. \tag{16}$$

First as  $u \in D((-A_1)^\alpha)$  we have:

$$\sup_{t \leq T} |e^{A_1 t}u|_\alpha^2 = \sup_{t \leq T} |e^{A_1 t}(-A_1)^\alpha u|^2 \leq C|u|_\alpha^2.$$

Moreover by (4) and by Lemma 3.3 we have:

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \int_0^t e^{A_1(t-s)} F(U_s^\epsilon, V_s^\epsilon) ds \right|_\alpha^2 \right] \\ & \leq \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \int_0^t (t-s)^{-\alpha} |F(U_s^\epsilon, V_s^\epsilon)| ds \right|_\alpha^2 \right] \\ & \leq C \mathbb{E} \left[ \sup_{t \in [0, T]} \int_0^t (t-s)^{-2\alpha} ds \int_0^t |F(U_s^\epsilon, V_s^\epsilon)|^2 ds \right] \\ & \leq C \int_0^T (1 + \mathbb{E}|U_s^\epsilon|^2 + \mathbb{E}|V_s^\epsilon|^2) ds \\ & \leq C(1 + |u|^2 + |v|^2). \end{aligned}$$

Finally by Lemma 3.1 we have:

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \left| \Gamma_t^1 \right|_\alpha^2 \right] < +\infty.$$

By considering (16), calculating  $\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t^\epsilon|_\alpha^2 \right]$  and using the last three inequalities we have the thesis. □

**Lemma 3.5** *Let  $0 < \alpha < 1/2 - \gamma$ ,  $u \in D((-A_1)^\alpha)$ ,  $v \in K$ , then there exists  $C = C(T, \alpha) > 0$  such that*

$$\mathbb{E} |U_{t+h}^\epsilon - U_t^\epsilon|^2 \leq C|h|^{2\alpha} (1 + |u|_\alpha^2 + |v|^2),$$

for every  $\epsilon > 0$ ,  $0 \leq t \leq T$ ,  $h \geq 0$  such that  $t + h \leq T$ .

**Proof** For  $0 \leq t \leq T$ ,  $h \geq 0$  such that  $t + h \leq T$  we have

$$\begin{aligned} U_{t+h}^\epsilon - U_t^\epsilon &= (e^{A_1 h} - I)U_t^\epsilon + \int_t^{t+h} e^{A_1(t+h-s)} F(U_s^\epsilon, V_s^\epsilon) ds \\ &\quad + \int_t^{t+h} e^{A_1(t+h-s)} dW_s^{\mathcal{Q}_1}. \end{aligned} \tag{17}$$

Consider the first term on the right-hand-side, as  $u \in D((-A_1)^\alpha)$  by (5) and by Lemma 3.4 we have:

$$\begin{aligned} \mathbb{E} \left| (e^{A_1 h} - I)U_t^\epsilon \right|^2 &\leq C|h|^{2\alpha} \mathbb{E} |U_t^\epsilon|_\alpha^2 \\ &\leq C|h|^{2\alpha} (1 + |u|_\alpha^2 + |v|^2). \end{aligned}$$

Consider now the second term on the right-hand-side, then by Lemma 3.3 we have:

$$\begin{aligned} \mathbb{E} \left| \int_t^{t+h} e^{A_1(t+h-s)} F(U_s^\epsilon, V_s^\epsilon) ds \right|^2 &\leq C|h| \int_t^{t+h} \left( 1 + \mathbb{E}|U_s^\epsilon|^2 + \mathbb{E}|V_s^\epsilon|^2 \right) ds \\ &\leq C|h|^2(1 + |u|^2 + |v|^2). \end{aligned}$$

Finally for the third term on the right-hand-side by Ito’s isometry and Hypothesis 4 we have:

$$\begin{aligned} \mathbb{E} \left| \int_t^{t+h} e^{A_1(t+h-s)} dW_s^{Q_1} \right|^2 &= \int_t^{t+h} \left\| e^{A_1(t+h-s)} Q_1^{1/2} \right\|_{\mathcal{L}_2}^2 ds \\ &\leq C \int_t^{t+h} |t+h-s|^{-2\gamma} ds \\ &= C|h|^{1-2\gamma}. \end{aligned}$$

As by assumption  $2\alpha \wedge 2 \wedge (1 - 2\gamma) = 2\alpha$  then we have the thesis. □

### 4 Fast Motion

In this section we study some classical properties of the fast motion. Consider

$$\begin{cases} dv_s^v = A_2 v_s^v ds + G(v_s^v) ds + dw_s^{Q_2}, \\ v_0^v = v \in K, \end{cases} \tag{18}$$

for every  $s \geq 0$  and for some  $Q_2$ -Wiener process  $w_s^{Q_2}$ .

First define the semigroup related to the fast motion by

$$P_s \phi(v) = \mathbb{E} [\phi(v_s^v)], \tag{19}$$

for every  $\phi \in B_B(K)$ ,  $s \geq 0$ .

Next recall that  $\delta$  was defined by (7), then we have:

#### Lemma 4.1

$$\mathbb{E} \left[ |v_s^{v_1} - v_s^{v_2}|^2 \right] \leq e^{-4\delta s} |v_1 - v_2|^2,$$

for every  $s \geq 0$ ,  $v_1, v_2 \in K$ .

**Proof** Define  $\rho_s = v_s^{v_1} - v_s^{v_2}$ , then by (8) we have:

$$\begin{aligned} \frac{d}{ds} |\rho_s|^2 &= 2\langle A_2 \rho_s, \rho_s \rangle + 2\langle G(v_s^{v_1}) - G(v_s^{v_2}), \rho_s \rangle \\ &\leq -4\delta |\rho_s|^2, \end{aligned}$$

for every  $s \geq 0, v_1, v_2 \in K$ .

Then by taking the expectation and applying the comparison Theorem we have the thesis. □

Next we can show:

**Lemma 4.2** *Let  $p \geq 1$ , then there exists  $C = C(p) > 0$  such that:*

$$\mathbb{E} [|v_s^v|^p] \leq C(1 + e^{-\delta ps} |v|^p),$$

for every  $s \geq 0, v \in K$ .

**Proof** Define

$$\tilde{\Gamma}_s^{Q_2} = \int_0^s e^{A_2(s-r)} dw_r^{Q_2}.$$

First by Burkholder’s inequality and Hypotheses 2 and 4 similarly to what is done for Lemma 3.2 we have:

$$\begin{aligned} \mathbb{E} [| \tilde{\Gamma}_s^{Q_2} |^p] &\leq C \left( \int_0^s \| e^{A_2(s-r)} Q_2^{1/2} \|^2_{\mathcal{L}_2} dr \right)^{p/2} \\ &\leq C \left( \int_0^s e^{-\lambda(s-r)} \left( \frac{s-r}{2} \wedge 1 \right)^{-2\gamma} dr \right)^{p/2} \leq C, \end{aligned} \tag{20}$$

for every  $s \geq 0$ .

Now set  $\rho_s = v_s^v - \tilde{\Gamma}_s^{Q_2}$ . Then for  $p \geq 2$  we have:

$$\begin{aligned} \frac{1}{p} \frac{d}{ds} |\rho_s|^p &= \langle A_2 \rho_s, \rho_s \rangle |\rho_s|^{p-2} + \langle G(\rho_s + \tilde{\Gamma}_s^{Q_2}) - G(\tilde{\Gamma}_s^{Q_2}), \rho_s \rangle |\rho_s|^{p-2} \\ &\quad + \langle G(\tilde{\Gamma}_s^{Q_2}), \rho_s \rangle |\rho_s|^{p-2} \\ &\leq -2\delta |\rho_s|^p + C(1 + |\tilde{\Gamma}_s^{Q_2}|) |\rho_s|^{p-1} \\ &\leq -\delta |\rho_s|^p + C(1 + |\tilde{\Gamma}_s^{Q_2}|^p). \end{aligned}$$

Then by the comparison Theorem and (20) we have:

$$\begin{aligned} \mathbb{E} |v_s^{u,v}|^p &\leq C \mathbb{E} |\rho_s|^p + C \mathbb{E} |\tilde{\Gamma}_s^{Q_2}|^p \\ &\leq C \left[ e^{-\delta ps} |v|^p + \int_0^s e^{-\delta p(s-r)} \left( 1 + \mathbb{E} |\tilde{\Gamma}_r^{Q_2}|^p \right) dr + \mathbb{E} |\tilde{\Gamma}_s^{Q_2}|^p \right] \\ &\leq C (e^{-\delta ps} |v|^p + 1). \end{aligned}$$

□

Now by [11, Theorem 6.3.3] there exists a unique invariant measure  $\mu$  for the semigroup  $P_t$ . Moreover we have:

**Lemma 4.3** *We have:*

$$\int_K |z| \mu(dz) < \infty. \tag{21}$$

**Proof** Fix  $N > 0$ , then by definition of invariant measure and Lemma 4.2 we have for every  $s > 0$

$$\begin{aligned} \int_K (|z| \wedge N) \mu(dz) &= \int_K \mathbb{E} [ |v_s^z| \wedge N ] \mu(dz) \leq \int_K \{ [C(1 + e^{-\delta s}|z|)] \wedge N \} \mu(dz) \\ &\leq \int_K C(1 \wedge N + e^{-\delta s}|z| \wedge N) \mu(dz) \\ &\leq C \left( 1 + e^{-\delta s} \int_K (|z| \wedge N) \mu(dz) \right), \end{aligned}$$

where we have used the fact that  $(a + b) \wedge c \leq a \wedge c + b \wedge c$  for every  $a, b, c \geq 0$ . By choosing  $s > 0$  large enough we have

$$\int_K (|z| \wedge N) \mu(dz) \leq C,$$

for some  $C > 0$  independent of  $N$ . Letting  $N \rightarrow \infty$  by the monotone convergence Theorem we have the result. □

Next we study the convergence to equilibrium of the semigroup of the fast motion, i.e. we prove:

**Lemma 4.4** *There exists  $C > 0$  such that*

$$\left| P_s \phi(v) - \int_K \phi(z) \mu(dz) \right| \leq C [\phi]_{Lip} e^{-2\delta s} (1 + |v|),$$

for every  $s \geq 0, v \in K, \phi \in Lip(K)$ .

Moreover there exists  $C > 0$  such that

$$\left| P_s \phi(v) - \int_K \phi(z) \mu(dz) \right| \leq C |\phi|_{\infty} e^{-\delta s} (s \wedge 1)^{-1/2} (1 + |v|),$$

for every  $s > 0, \phi \in B_B(K), v \in K$ .

**Proof** First for every  $\phi \in Lip(K)$  by Lemma 4.1 we have:

$$\begin{aligned} |P_s \phi(v_2) - P_s \phi(v_1)| &\leq [\phi]_{Lip} \mathbb{E} [ |v_s^{v_2} - v_s^{v_1}| ] \\ &\leq [\phi]_{Lip} e^{-2\delta s} |v_2 - v_1|, \end{aligned} \tag{22}$$

for every  $s \geq 0, v_1, v_2 \in K$ .

Now let  $s > 0$ , by definition of invariant measure, (22) and Lemma 4.3 we have:

$$\begin{aligned} \left| P_s \phi(v) - \int_K \phi(z) \mu(dz) \right| &= \left| \int_K (P_s \phi(v) - P_s \phi(z)) \mu(dz) \right| \\ &\leq [\phi]_{Lip} e^{-2\delta s} \int_K |v - z| \mu(dz) \\ &\leq C [\phi]_{Lip} e^{-2\delta s} (1 + |v|), \end{aligned} \tag{23}$$

for every  $v \in K$ ,  $\phi \in Lip(K)$  so that we have the first inequality.

Now thanks to Hypothesis 5 we can apply [10, Theorem 9.32] to have the Bismut-Elworthy formula:

$$\sup_{u \in H} |DP_s \phi|_\infty \leq C(s \wedge 1)^{-1/2} |\phi|_\infty, \tag{24}$$

for every  $s > 0$ ,  $\phi \in B_B(K)$ .

Now by the semigroup property, the regularizing property of the semigroup (24) and by (22) we have:

$$\begin{aligned} |P_s \phi(v_2) - P_s \phi(v_1)| &= |P_{s/2} (P_{s/2} \phi)(v_2) - P_{s/2} (P_{s/2} \phi)(v_1)| \\ &\leq [P_{s/2} \phi]_{Lip} e^{-2\delta \frac{s}{2}} |v_2 - v_1| \\ &\leq C |\phi|_\infty (s \wedge 1)^{-\frac{1}{2}} e^{-\delta s} |v_2 - v_1|, \end{aligned} \tag{25}$$

for every  $s > 0$ ,  $v_1, v_2 \in K$ .

Finally similarly to before by (25) for  $s > 0$  we have:

$$\begin{aligned} \left| P_s \phi(v) - \int_K \phi(z) \mu(dz) \right| &= \left| \int_K (P_s \phi(v) - P_s \phi(z)) \mu(dz) \right| \\ &\leq C |\phi|_\infty (s \wedge 1)^{-\frac{1}{2}} e^{-\delta s} \int_K |v - z| \mu(dz) \\ &\leq C |\phi|_\infty (s \wedge 1)^{-\frac{1}{2}} e^{-\delta s} (1 + |v|), \end{aligned}$$

for every  $v \in K$ ,  $\phi \in B_B(K)$ . □

Now we study the mixing properties of the semigroup of the fast motion. To this purpose define for  $0 \leq s \leq t \leq \infty$ ,  $v \in K$

$$\mathcal{H}_s^t(v) = \sigma(v_r^v, 0 \leq s \leq r \leq t).$$

Then a classical consequence of Lemmas 4.2, 4.4 is the following mixing result whose proof is the same of [7, Lemma 3.2] and is reported in the appendix for completeness.

**Lemma 4.5** *There exists  $C > 0$  such that*

$$\sup \{ |\mathbb{P}(B_1 \cap B_2) - \mathbb{P}(B_1)\mathbb{P}(B_2)| : B_1 \in \mathcal{H}_0^t(v), B_2 \in \mathcal{H}_{t+s}^\infty(v) \}$$



$$\leq C e^{-\delta s} s^{-1/2} (1 + |v|),$$

for every  $0 \leq s \leq t, v \in K$ .

Now Lemma 4.5 implies the following classical result, e.g. see [37] (see also [7, Proposition 3.3]). The proof can be found in the appendix for completeness.

**Lemma 4.6** *There exists  $C > 0$  such that for every  $0 \leq s_1 \leq t_1 < s_2 \leq t_2$  and  $\xi_i \mathcal{H}_{s_i}^{t_i}(v)$ -measurable  $i = 1, 2$  and  $|\xi_i| \leq 1$  a.s  $i = 1, 2$*

$$|\mathbb{E} [\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq C \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|).$$

Since in our case  $|\xi_i|$  will not be bounded by 1 we need the following result which is similar to [7, Proposition 3.3]. Also in this case we postpone the proof in the appendix.

**Lemma 4.7** *Let  $\rho \in (0, 1)$ . Then there exists  $C = C(\rho) > 0$  such that for every  $0 \leq s_1 \leq t_1 < s_2 \leq t_2$  and  $\xi_i \mathcal{H}_{s_i}^{t_i}(v)$ -measurable,  $i = 1, 2$  satisfying for some  $K_i = K_i(\rho) > 0$*

$$\left( \mathbb{E} |\xi_i|^{\frac{2}{1-\rho}} \right)^{\frac{1-\rho}{2}} = K_i < \infty \tag{26}$$

then:

$$|\mathbb{E} [\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq C K_1^{\frac{2}{2+\rho}} K_2^{\frac{2}{2+\rho}} (K_1 + K_2)^{\frac{2\rho}{2+\rho}} \left( \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|) \right)^{\frac{\rho}{2+\rho}}.$$

### 5 Averaged Equation

In this section we introduce the averaged equation and we prove its well-posedness.

$$\bar{F}(u) = \int_K F(u, v) \mu(dv), \tag{27}$$

for every  $u \in H$  and consider the so called averaged equation:

$$\begin{cases} dU_t = A_1 U_t dt + \bar{F}(U_t) dt + dW_t^{Q_1}, \\ U_0 = u, \end{cases} \tag{28}$$

for every  $t \leq T$ .

Now we prove the well-posedness of the averaged equation:

**Proposition 2** *Equation (28) admits a unique mild solution given by:*

$$U_t = e^{A_1 t} u + \int_0^t e^{A_1(t-s)} \bar{F}(U_s) ds + \int_0^t e^{A_1(t-s)} dW_s^{Q_1},$$

for every  $t \in [0, T]$ .

Moreover for every  $p > 0$  there exists  $C = C(T, p) > 0$  such that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t|^p \right] \leq C(1 + |u|^p).$$

**Proof** In order to prove the first part of the Proposition it is sufficient to prove that  $\bar{F}$  is Lipschitz (e.g. see [10]). But this follows from the Lipschitz continuity of  $F$ , indeed:

$$|\bar{F}(u_1) - \bar{F}(u_2)| \leq \int_K |F(u_1, v) - F(u_2, v)| \mu(dv) \leq L_F |u_1 - u_2|, \quad \forall u_1, u_2 \in H.$$

Hence we obtain the Lipschitzianity of  $\bar{F}$  and the first claim of the Proposition.

The proof of the second claim of the Proposition follows by a standard application of Gronwall’s Lemma. □

### 6 Preliminary Results

In this section we prove a technical result, i.e. Lemma 6.3, which is inspired by [7, Lemma 4.2] and follows by the mixing properties of the fast motion studied in Sect. 4, in particular Lemma 4.7. In order to prove it we proceed with similar techniques to the ones of the proof of [7, Lemma 4.2].

Fix  $\bar{\xi} \in \mathcal{C}([0, T]; H)$ ,  $v \in K$ ,  $h \in H$  with  $|h| = 1$ , and define

$$\Phi_h(r) := \langle F(\bar{\xi}_r, V_r^\epsilon) - \bar{F}(\bar{\xi}_r), h \rangle \quad \forall r \leq T$$

and

$$\Psi_h(r) := \Phi_h(\epsilon r) = \langle F(\bar{\xi}_{\epsilon r}, V_{\epsilon r}^\epsilon) - \bar{F}(\bar{\xi}_{\epsilon r}), h \rangle.$$

Moreover let  $n \in \mathbb{N}$  and set:

$$J_j(r_1, \dots, r_{2n}) := \left| \mathbb{E} \prod_{i=1}^{2n} \Psi_h(r_i) - \mathbb{E} \prod_{i=1}^{2j} \Psi_h(r_i) \mathbb{E} \prod_{i=2j+1}^{2n} \Psi_h(r_i) \right|, \quad (29)$$

for every  $1 \leq j \leq n$ ,  $0 \leq r_1 \leq \dots \leq r_{2n} \leq T/\epsilon$ .

First we show the following result:

**Lemma 6.1** *Let  $0 < \rho < 1$  then there exists  $C = C(T, \rho, n) > 0$  and  $\eta = \eta(\rho, j_2 - j_1) > 0$  such that*

$$\left| \mathbb{E} \prod_{i=j_1}^{j_2} \Psi_h(r_i) \right| \leq C \left( \frac{e^{-\delta(r_{j_2} - r_{j_2-1})}}{\sqrt{r_{j_2} - r_{j_2-1}}} \right)^{\frac{\rho}{2+\rho}} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right), \quad (30)$$

for every  $u \in H, v \in K, 1 \leq j_1 \leq j_2 \leq n, 0 \leq r_1 \leq \dots \leq r_{2n} \leq T/\epsilon$ .

Moreover there exists  $C = C(T, \rho, n) > 0$  and  $\eta = \eta(\rho, n) > 0$  such that

$$J_j(r_1, \dots, r_{2n}) \leq C \left( \frac{e^{-\delta \hat{r}_j}}{\sqrt{\hat{r}_j}} \right)^{\frac{\rho}{2+\rho}} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right), \tag{31}$$

where

$$\hat{r}_j = \max(r_{2n} - r_{2n-1}, r_{2j+1} - r_{2j}),$$

for every  $1 \leq j \leq n, 0 \leq r_1 \leq \dots \leq r_{2n} \leq T/\epsilon$ .

**Remark 6.2** Note that the dependence of the exponents with respect to  $j_2 - j_1$  and  $n$  is not restrictive: once  $n$  has been fixed (together with  $\rho$  and  $T$ ) we are allowed to take any  $0 \leq j_1 \leq j_2 \leq n$ . In this sense  $\eta = \eta(\rho, j_2 - j_1)$ . Of course since  $j_2 - j_1 \leq n$  we could choose  $\eta' = \eta'(\rho, n) \geq \eta$  and replace  $\eta$  with  $\eta'$  in the estimates of the Lemma. However the estimate with  $\eta$  is more precise.

**Proof** By the sublinearity of  $\Psi_h$  and Lemma 4.2 for every  $p \geq 1$  we have:

$$\begin{aligned} \mathbb{E} \left| \prod_{i=j_1}^{j_2} \Psi_h(r_i) \right|^p &\leq C \left( 1 + \sum_{i=j_1}^{j_2} |\bar{\xi}_{\epsilon r_i}|^{(j_2-j_1+1)p} + \sum_{i=j_1}^{j_2} \mathbb{E} |V_{\epsilon r_i}^\epsilon|^{(j_2-j_1+1)p} \right) \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{(j_2-j_1+1)p} + |v|^{(j_2-j_1+1)p} \right), \end{aligned} \tag{32}$$

for every  $1 \leq j_1 \leq j_2 \leq 2n$ .

Notice that  $V_{\epsilon r}^\epsilon$  and  $v_r^v$ , defined by (18) for  $w_r^{Q_2} = W_{\epsilon r}^{Q_2}/\sqrt{\epsilon}$ , are indistinguishable for  $r \in [0, T/\epsilon]$ . Then by setting  $p = 2/(1 - \rho)$  and applying Lemma 4.7 to  $\xi_1 = \prod_{i=1}^j \Psi_h(r_i), \xi_2 = \prod_{i=j+1}^{2n} \Psi_h(r_i)$  with  $K_1 = C(1 + \sup_{r \leq T} |\bar{\xi}_r|^j + |v|^j), K_2 = C(1 + \sup_{r \leq T} |\bar{\xi}_r|^{2n-j} + |v|^{2n-j})$  for every  $0 \leq j_1 < j < j_2 \leq 2n$  we have:

$$\begin{aligned} &\left| \mathbb{E} \prod_{i=j_1}^{j_2} \Psi_h(r_i) - \mathbb{E} \prod_{i=j_1}^j \Psi_h(r_i) \mathbb{E} \prod_{i=j+1}^{j_2} \Psi_h(r_i) \right| \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta(r_{j+1}-r_j)}}{\sqrt{r_{j+1}-r_j}} \right)^{\frac{\rho}{2+\rho}}, \end{aligned} \tag{33}$$

where  $\eta = j_2 - j_1 + \frac{\rho(j_2-j_1-1)}{(\rho+2)}$ .

By definitions of  $\Psi_h$ , the indistinguishability of  $V_{\epsilon r}^\epsilon, v_r^v$  and by Lemma 4.4 we have:

$$|\mathbb{E} \Psi_h(r_{j_2})| \leq C e^{-2\delta r_{j_2}} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r| + |v| \right). \tag{34}$$

Now by the last three inequalities we have:

$$\begin{aligned} \left| \mathbb{E} \prod_{i=j_1}^{j_2} \Psi_h(r_i) \right| &\leq \left| \mathbb{E} \prod_{i=j_1}^{j_2} \Psi_h(r_i) - \mathbb{E} \prod_{i=j_1}^{j_2-1} \Psi_h(r_i) \mathbb{E} \Psi_h(r_{j_2}) \right| + \mathbb{E} \left| \prod_{i=j_1}^{j_2-1} \Psi_h(r_i) \right| \left| \mathbb{E} \Psi_h(r_{j_2}) \right| \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta(r_{j_2} - r_{j_2-1})}}{\sqrt{r_{j_2} - r_{j_2-1}}} \right)^{\frac{\rho}{2+\rho}} \\ &\quad + C(1 + \sup_{r \leq T} |\bar{\xi}_r|^{j_2-j_1} + |v|^{j_2-j_1}) \left( 1 + \sup_{r \leq T} |\bar{\xi}_r| + |v| \right) e^{-2\delta r_{j_2}} \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta(r_{j_2} - r_{j_2-1})}}{\sqrt{r_{j_2} - r_{j_2-1}}} \right)^{\frac{\rho}{2+\rho}}, \end{aligned}$$

where  $\eta = (j_2 - j_1 + 1) + \frac{\rho(j_2-j_1)}{(\rho+2)}$ . This implies (30).

Now by (30) we have:

$$\left| \mathbb{E} \prod_{i=j_1}^{2n} \Psi_h(r_i) \right| \leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta(r_{2n} - r_{2n-1})}}{\sqrt{r_{2n} - r_{2n-1}}} \right)^{\frac{\rho}{2+\rho}}. \tag{35}$$

for every  $j_1 < 2n$ .

Now fix any  $j < 2n$ . Then by the last inequality and (32) we have:

$$\begin{aligned} &\left| \mathbb{E} \prod_{i=1}^{2n} \Psi_h(r_i) - \mathbb{E} \prod_{i=1}^j \Psi_h(r_i) \mathbb{E} \prod_{i=j+1}^{2n} \Psi_h(r_i) \right| \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta(r_{2n} - r_{2n-1})}}{\sqrt{r_{2n} - r_{2n-1}}} \right)^{\frac{\rho}{2+\rho}}. \end{aligned} \tag{36}$$

Finally consider (33) with  $j_1 = 1, j_2 = 2n$  and (36). Since the function  $f(s) = e^{-\delta s} s^{-1/2}$  is decreasing we have

$$\begin{aligned} &\left| \mathbb{E} \prod_{i=1}^{2n} \Psi_h(r_i) - \mathbb{E} \prod_{i=1}^j \Psi_h(r_i) \mathbb{E} \prod_{i=j+1}^{2n} \Psi_h(r_i) \right| \\ &\leq C \left( \frac{e^{-\delta \max(r_{2n} - r_{2n-1}, r_{2j+1} - r_{2j})}}{\sqrt{\max(r_{2n} - r_{2n-1}, r_{2j+1} - r_{2j})}} \right)^{\frac{\rho}{2+\rho}} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\ &= C \left( \frac{e^{-\delta \hat{r}_j}}{\sqrt{\hat{r}_j}} \right)^{\frac{\rho}{2+\rho}} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right). \end{aligned}$$

This implies (31). □

Let  $0 < \alpha, 0 \leq \beta < 1/3$  and set

$$\theta_{\alpha,\beta}(r) := e^{-r\alpha}r^{-\beta} \quad \forall r > 0.$$

Now we can state and prove the main result of this section.

**Lemma 6.3** *Let  $n \in \mathbb{N}$  and  $0 \leq \beta < 1/3$ . Then there exists a constant  $C = C(T, n, \beta) > 0$  and  $\eta = \eta(n) > 0$  such that for every  $\epsilon > 0, \alpha > 0, u \in H, v \in K, h \in H$  with  $|h| = 1$  we have:*

$$\begin{aligned} & \int_{[s,t]^{2n}} \left| \mathbb{E} \prod_{i=1}^{2n} \theta_{\alpha,\beta}(t - r_i) \langle F(\bar{\xi}_{r_i}, V_{r_i}^\epsilon) - \bar{F}(\bar{\xi}_{r_i}), h \rangle \right| dr_1 \dots dr_{2n} \\ & \leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \epsilon^n \left( \frac{1}{\alpha} \right)^{(1-2\beta)n}. \end{aligned}$$

**Proof** Recall the definition of  $\Psi_h(r)$ , then by a change of variable we have:

$$\int_{[s,t]^{2n}} \left| \mathbb{E} \prod_{i=1}^{2n} \theta_{\alpha,\beta}(t - r_i) \langle F(\bar{\xi}_{r_i}, V_{r_i}^\epsilon) - \bar{F}(\bar{\xi}_{r_i}), h \rangle \right| dr_1 \dots dr_{2n} = \epsilon^{2n} H_\epsilon(s, t), \tag{37}$$

where we have defined:

$$H_\epsilon(s, t) := \int_{[\frac{s}{\epsilon}, \frac{t}{\epsilon}]^{2n}} \left| \mathbb{E} \prod_{i=1}^{2n} (\theta_{\alpha,\beta}(t - \epsilon r_i) \Psi_h(r_i)) \right| dr_1 \dots dr_{2n}.$$

By simmetry we have:

$$H_\epsilon(s, t) = C \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_2} \dots \int_{\frac{s}{\epsilon}}^{r_2} \left| \mathbb{E} \prod_{i=1}^{2n} (\theta_{\alpha,\beta}(t - \epsilon r_i) \Psi_h(r_i)) \right| dr_1 \dots dr_{2n}. \tag{38}$$

We proceed by induction on  $n$  and to this purpose fix some  $\rho \in (0, 1)$ . Consider  $n = 1$  then by the definition of  $\theta_{\alpha,\beta} = e^{-r\alpha}r^{-\beta}$ , (30) and some changes of variables we have

$$\begin{aligned} \epsilon^2 H_\epsilon(s, t) &= 2\epsilon^2 \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_2} \theta_{\alpha,\beta}(t - \epsilon r_1) \theta_{\alpha,\beta}(t - \epsilon r_2) |\mathbb{E}[\Psi_h(r_1) \Psi_h(r_2)]| dr_1 dr_2 \\ &\leq C\epsilon^2 \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_2} \theta_{\alpha,\beta}(t - \epsilon r_1) \theta_{\alpha,\beta}(t - \epsilon r_2) \left( \frac{e^{-\delta(r_2-r_1)}}{\sqrt{r_2-r_1}} \right)^{\frac{\rho}{2+\rho}} dr_1 dr_2 \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\ &= C\epsilon^{2-2\beta} \int_0^{\frac{t-\epsilon}{\epsilon}} y_2^{-\beta} e^{-\epsilon y_2 \alpha} \int_{y_2}^{\frac{t-\epsilon}{\epsilon}} y_1^{-\beta} e^{-\epsilon y_1 \alpha} \left( \frac{e^{-\delta(y_1-y_2)}}{\sqrt{y_1-y_2}} \right)^{\frac{\rho}{2+\rho}} dy_1 dy_2 \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\ &\leq C\epsilon^{2-2\beta} \int_0^{\frac{t-\epsilon}{\epsilon}} y_2^{-2\beta} e^{-2\epsilon y_2 \alpha} \int_0^{+\infty} \left( \frac{e^{-\delta y_1}}{\sqrt{y_1}} \right)^{\frac{\rho}{2+\rho}} dy_1 dy_2 \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \end{aligned}$$

$$\begin{aligned}
 &\leq C\epsilon^{2-2\beta} \int_0^{\frac{t-s}{\epsilon}} y_2^{-2\beta} e^{-2\epsilon y_2 \alpha} dy_2 \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{\eta} + |v|^{\eta} \right) \\
 &= C\epsilon^{2-2\beta} (\epsilon\alpha)^{-(1-2\beta)} \int_0^{\alpha(t-s)} r^{-2\beta} e^{-2r} dr \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{\eta} + |v|^{\eta} \right) \\
 &\leq C\epsilon^{2-2\beta} (\epsilon\alpha)^{-(1-2\beta)} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{\eta} + |v|^{\eta} \right) \\
 &= C\epsilon\alpha^{-(1-2\beta)} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{\eta} + |v|^{\eta} \right),
 \end{aligned}$$

so that by (37) we have the thesis for  $n = 1$ .

Now assume that

$$\begin{aligned}
 &\int_{[s,t]^j} \left| \mathbb{E} \prod_{i=1}^j \theta_{\alpha,\beta}(t - \epsilon r_i) \Psi_h(r_i) \right| dr_1 \dots dr_j \\
 &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^{\eta_j} + |v|^{\eta_j} \right) \epsilon^{j/2} \left( \frac{1}{\alpha} \right)^{(1-2\beta)j/2},
 \end{aligned}$$

for every even  $j < 2n$  where  $\eta_j = j + \frac{\rho(j-1)}{(\rho+2)}$ .

We prove that then it holds for  $j = 2n$ .

Set for  $r = (r_1, \dots, r_{2n}) \in (s, t)^{2n}$  with  $s \leq r_1 \leq \dots \leq r_{2n} \leq t$  the integer  $j(r)$  such that

$$\max_{j=1, \dots, n-1} (r_{2j+1} - r_{2j}) = r_{2j(r)+1} - r_{2j(r)}.$$

and consider  $H_\epsilon(s, t)$  given by (38). Recalling the definition of  $J_j(r_1, \dots, r_{2n})$  given by (29) we have:

$$\begin{aligned}
 \epsilon^{2n} H_\epsilon(s, t) &= C\epsilon^{2n} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \dots \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \prod_{i=1}^{2n} \theta_{\alpha,\beta}(t - \epsilon r_i) \left| \mathbb{E} \prod_{i=1}^{2n} \Psi_h(r_i) \right| dr_1 \dots dr_{2n} \\
 &\leq C\epsilon^{2n} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \dots \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \prod_{i=1}^{2n} \theta_{\alpha,\beta}(t - \epsilon r_i) J_{j(r)}(r_1, \dots, r_{2n}) dr_1 \dots dr_{2n} \\
 &\quad + C\epsilon^{2n} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \dots \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \prod_{i=1}^{2n} \theta_{\alpha,\beta}(t - \epsilon r_i) \left| \mathbb{E} \prod_{i=1}^{2j(r)} \Psi_h(r_i) \mathbb{E} \prod_{i=2j(r)+1}^{2n} \Psi_h(r_i) \right| dr_1 \dots dr_{2n}.
 \end{aligned}$$

Note that by definition of  $j(r)$ , for every  $s \leq r_1 \leq \dots \leq r_{2n} \leq t$ , we have

$$\left| \mathbb{E} \prod_{i=1}^{2j(r)} \Psi_h(r_i) \mathbb{E} \prod_{i=2j(r)+1}^{2n} \Psi_h(r_i) \right| \leq \sum_{j=1}^{n-1} \left| \mathbb{E} \prod_{i=1}^{2j} \Psi_h(r_i) \mathbb{E} \prod_{i=2j+1}^{2n} \Psi_h(r_i) \right|.$$

It follows:

$$\begin{aligned}
 & \epsilon^{2n} H_\epsilon(s, t) \\
 & \leq C \epsilon^{2n} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_{2n}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \prod_{i=1}^{2n} \theta_{\alpha, \beta}(t - \epsilon r_i) J_{j(r)}(r_1, \dots, r_{2n}) dr_1 \dots dr_{2n} \\
 & \quad + C \epsilon^{2n} \sum_{j=1}^{n-1} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_{2n}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \prod_{i=1}^{2n} \theta_{\alpha, \beta}(t - \epsilon r_i) \left| \mathbb{E} \prod_{i=1}^{2j} \Psi_h(r_i) \mathbb{E} \prod_{i=2j+1}^{2n} \Psi_h(r_i) \right| dr_1 \dots dr_{2n} \\
 & = C \epsilon^{2n} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_{2n}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \prod_{i=1}^{2n} \theta_{\alpha, \beta}(t - \epsilon r_i) J_{j(r)}(r_1, \dots, r_{2n}) dr_1 \dots dr_{2n} \tag{39} \\
 & \quad + C \epsilon^{2n} \sum_{j=1}^{n-1} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_{2j}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \prod_{i=1}^{2j} \theta_{\alpha, \beta}(t - \epsilon r_i) \left| \mathbb{E} \prod_{i=1}^{2j} \Psi_h(r_i) \right| dr_1 \dots r_{2j} \\
 & \quad \times \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{r_{2(n-j)}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \prod_{i=1}^{2(n-j)} \theta_{\alpha, \beta}(t - \epsilon r_i) \left| \mathbb{E} \prod_{i=1}^{2(n-j)} \Psi_h(r_i) \right| dr_1 \dots r_{2(n-j)} \\
 & =: \epsilon^{2n} I_{1, \epsilon}(s, t) + \epsilon^{2n} I_{2, \epsilon}(s, t).
 \end{aligned}$$

Now by (31) and the definition of  $j(r)$  we have:

$$\begin{aligned}
 J_{j(r)}(r_1, \dots, r_{2n}) & \leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \left( \frac{e^{-\delta \hat{r}_{j(r)}}}{\sqrt{\hat{r}_{j(r)}}} \right)^{\frac{\rho}{2+\rho}} \\
 & = C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\
 & \quad \left( \frac{e^{-\delta \max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2)}}{\sqrt{\max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2)}} \right)^{\frac{\rho}{2+\rho}} \\
 & = C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\
 & \quad \frac{e^{-\frac{\delta \rho}{2+\rho} \max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2)}}{\left[ \max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2) \right]^{\bar{\rho}}},
 \end{aligned}$$

where  $\bar{\rho} = \frac{\rho}{2(2+\rho)}$ .

Recall that  $r_{i+1} \geq r_i$  for every  $i$  and note that  $\max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}) \geq 1/2(r_{2n} - r_{2n-2})$  and  $\max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2) \geq \frac{1}{n}(r_{2n} - r_{2n-1} + \sum_{i=1}^{n-1} r_{2i+1} - r_{2i})$ . Hence, since  $g(t) = 1/t^{\bar{\rho}}$ ,  $f(t) = e^{-\alpha t}$  for  $\alpha > 0$  are decreasing, we have (recall that  $C$  may change from line to line):

$$J_{j(r)}(r_1, \dots, r_{2n}) \leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \frac{e^{-\frac{\delta \rho}{2+\rho} \max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}, \dots, r_3 - r_2)}}{\left[ \max(r_{2n} - r_{2n-1}, r_{2n-1} - r_{2n-2}) \right]^{\bar{\rho}}}$$

$$\begin{aligned} &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \frac{e^{-\frac{\delta\rho}{2+\rho} \max(r_{2n}-r_{2n-1}, r_{2n-1}-r_{2n-2}, \dots, r_3-r_2)}}{(r_{2n} - r_{2n-2})^{\bar{\rho}}} \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \frac{e^{-\frac{\delta\rho}{2+\rho} \frac{1}{n} (r_{2n}-r_{2n-1} + \sum_{i=1}^{n-1} r_{2i+1}-r_{2i})}}{(r_{2n} - r_{2n-2})^{\bar{\rho}}} \\ &= C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \frac{e^{-\delta_n (r_{2n}-r_{2n-2} + \sum_{i=1}^{n-2} r_{2i+1}-r_{2i})}}{(r_{2n} - r_{2n-2})^{\bar{\rho}}}, \end{aligned}$$

where in the last equality we have defined  $\delta_n = \frac{\delta\rho}{n(2+\rho)}$  and we have used the following identity:  $\sum_{i=1}^{n-1} r_{2i+1} - r_{2i} = r_{2n-1} - r_{2n-2} + \sum_{i=1}^{n-2} r_{2i+1} - r_{2i}$ .

We can apply this last inequality in order to estimate  $I_{1,\epsilon}(s, t)$ , i.e.

$$\begin{aligned} I_{1,\epsilon}(s, t) &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \int_{\frac{s}{\epsilon}}^t \int_{\frac{s}{\epsilon}}^{r_{2n}} \dots \int_{\frac{s}{\epsilon}}^{r_2} \frac{e^{-\delta_n(r_{2n}-r_{2n-2})}}{(r_{2n} - r_{2n-2})^{\bar{\rho}}} \\ &\quad \times \prod_{i=1}^{2n} \vartheta_{\alpha,\beta}(t - \epsilon r_i) \prod_{i=1}^{n-2} e^{-\delta_n(r_{2i+1}-r_{2i})} dr_1 \dots dr_{2n} \\ &= C \epsilon^{-2n\beta} \int_0^{\frac{t-s}{\epsilon}} e^{-(\epsilon\alpha-\delta_n)y_{2n}} y_{2n}^{-\beta} \int_{y_{2n}}^{\frac{t-s}{\epsilon}} e^{-\epsilon\alpha y_{2n-1}} y_{2n-1}^{-\beta} \int_{y_{2n-1}}^{\frac{t-s}{\epsilon}} \frac{e^{-(\epsilon\alpha+\delta_n)y_{2n-2}}}{(y_{2n-2} - y_{2n})^{\bar{\rho}}} y_{2n-2}^{-\beta} \\ &\quad \times \int_{y_{2n-2}}^{\frac{t-s}{\epsilon}} \dots \int_{y_3}^{\frac{t-s}{\epsilon}} \prod_{i=2}^{2n-2} e^{-(\epsilon\alpha+(-1)^i \delta_n)y_i} r_i^{-\beta} \\ &\quad \int_{y_2}^{\frac{t-s}{\epsilon}} e^{-\epsilon\alpha y_1} y_1^{-\beta} dy_1 \dots dy_{2n} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right). \end{aligned}$$

Now for  $k = 1, 2, 3, i = 1, 3, \dots, 2n - 1$  we obtain:

$$\begin{aligned} \int_{y_{i+1}}^{\frac{t-s}{\epsilon}} e^{-k\epsilon\alpha y_i} y_i^{-k\beta} dy_i &= (\epsilon\alpha)^{k\beta-1} \int_{\epsilon\alpha y_{i+1}}^{(t-s)\alpha} e^{-ky_i} y_i^{-k\beta} dy_i \\ &\leq C(\epsilon\alpha)^{k\beta-1}. \end{aligned} \tag{40}$$

Now consider  $i = 2, 4, \dots, 2n$ , we have:

$$\begin{aligned} \int_{y_{i+1}}^{\frac{t-s}{\epsilon}} e^{-(\epsilon\alpha+\delta_n)y_i} y_i^{-\beta} dy_i &\leq y_{i+1}^{-\beta} e^{-(\epsilon\alpha+\delta_n)y_{i+1}} (\epsilon\alpha + \delta_n)^{-1} \\ &\leq C y_{i+1}^{-\beta} e^{-(\epsilon\alpha+\delta_n)y_{i+1}}. \end{aligned} \tag{41}$$

Now by (40) and (41) we have:

$$\begin{aligned} \int_{y_{2n-2}}^{\frac{t-s}{\epsilon}} \dots \int_{y_3}^{\frac{t-s}{\epsilon}} \prod_{i=2}^{2n-2} e^{-(\epsilon\alpha+(-1)^i \delta_n)y_i} y_i^{-\beta} \int_{y_2}^{\frac{t-s}{\epsilon}} e^{-\epsilon\alpha y_1} y_1^{-\beta} dy_1 \dots dy_{2n-3} \\ \leq (\epsilon\alpha)^{\beta-1} (\epsilon\alpha)^{(2\beta-1)(n-2)} \\ = (\epsilon\alpha)^{n(2\beta-1)-(3\beta-1)}. \end{aligned} \tag{42}$$



Now consider the remaining term in the inequality for  $I_{1,\epsilon}(s, t)$ : by (40) and (41) we have

$$\begin{aligned} & \int_0^{\frac{t-s}{\epsilon}} e^{-(\epsilon\alpha-\delta_n)y_{2n}} y_{2n}^{-\beta} \int_{y_{2n}}^{\frac{t-s}{\epsilon}} e^{-\epsilon\alpha y_{2n-1}} y_{2n-1}^{-\beta} \int_{y_{2n-1}}^{\frac{t-s}{\epsilon}} \frac{e^{-(\epsilon\alpha+\delta_n)y_{2n-2}}}{(y_{2n-2}-y_{2n})^{\bar{\rho}}} y_{2n-2}^{-\beta} dy_{2n-2} dy_{2n-1} dy_{2n} \\ & \leq C \int_0^{\frac{t-s}{\epsilon}} e^{-3\epsilon\alpha y_{2n}} y_{2n}^{-3\beta} \int_{y_{2n}}^{\frac{t-s}{\epsilon}} \frac{e^{-(2\epsilon\alpha+\delta_n)(y_{2n-1}-y_{2n})}}{(y_{2n-1}-y_{2n})^{\bar{\rho}}} dy_{2n-1} dy_{2n} \tag{43} \\ & \leq C \int_0^{\frac{t-s}{\epsilon}} e^{-3\epsilon\alpha y_{2n}} y_{2n}^{-3\beta} \int_0^\infty \frac{e^{-\delta_n y_{2n-1}}}{y_{2n-1}^{\bar{\rho}}} dy_{2n-1} dy_{2n} \\ & \leq C(\epsilon\alpha)^{3\beta-1}. \end{aligned}$$

Applying now (42) and (43) to the inequality for  $I_{1,\epsilon}(s, t)$  we have:

$$\begin{aligned} I_{1,\epsilon} & \leq C\epsilon^{-2n\beta} (\epsilon\alpha)^{n(2\beta-1)-(3\beta-1)} (\epsilon\alpha)^{3\beta-1} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \\ & = C\epsilon^{-n} \alpha^{-(1-2\beta)n} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right). \tag{44} \end{aligned}$$

In addition by the inductive hypothesis we have:

$$I_{2,\epsilon} \leq C\epsilon^{-n} \alpha^{-(1-2\beta)n} \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right). \tag{45}$$

Finally applying the last two inequalities to (39) and then going back to (37) we have the thesis. □

### 7 Order of Convergence

In this section we finally investigate the order of convergence. We first prove the following proposition which will be crucial in the derivation of the order of convergence.

**Proposition 3** *There exist  $C = C(T) > 0, \iota > 0$  such that*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} (F(U_s, V_s^\epsilon) - \bar{F}(U_s)) ds \right|^2 \right] \leq C\epsilon(1 + |u|^\iota + |v|^\iota),$$

for every  $0 \leq \tau \leq T, \epsilon > 0, u \in H, v \in K$ .

**Proof** We proceed with similar techniques to the ones in the proof of [7, Theorem 4.1].

Define

$$Z_t^\epsilon := \int_0^t e^{A_1(t-s)} (F(U_s, V_s^\epsilon) - \bar{F}(U_s)) ds.$$

We proceed using the factorization method, e.g. see [10, Chapter 5, Section 3]: fix  $\zeta > 0, n \geq 2$  integer and  $1/(2n) < \beta < 1/3$  as in Hypothesis 1, then:

$$Z_t^\epsilon = \frac{\sin \beta \pi}{\pi} \int_0^t (t-s)^{\beta-1} e^{A_1(t-s)} Y_s^\epsilon ds,$$

where

$$Y_s^\epsilon = \int_0^s (s-r)^{-\beta} e^{A_1(s-r)} (F(U_r, V_r^\epsilon) - \bar{F}(U_r)) dr.$$

By Holder’s inequality we have:

$$\mathbb{E} \left[ \sup_{t \in [0, \tau]} |Z_t^\epsilon|^{2n} \right] \leq C \int_0^\tau \mathbb{E} |Y_s^\epsilon|^{2n} ds. \tag{46}$$

We now claim that there exist  $C = C(T) > 0$  and  $\eta > 0$  such that

$$\mathbb{E} |Y_s^\epsilon|^{2n} \leq C \epsilon^n (1 + |u|^\eta + |v|^\eta), \tag{47}$$

for every  $0 \leq t_0 \leq s \leq T, \epsilon > 0, u \in H, v \in K$ .

Indeed first recall the spectral representation (3). Then by Parseval’s identity, Holder’s inequality, Hypothesis 1 and the properties of conditional expectations we have:

$$\begin{aligned} \mathbb{E} |Y_s^\epsilon|^{2n} &= \mathbb{E} \left( \sum_{k=1}^\infty \alpha_k^{-\frac{(n-1)\zeta}{n}} \alpha_k^{\frac{(n-1)\zeta}{n}} \left| \int_0^s (s-r)^{-\beta} e^{-(s-r)\alpha_k} \langle F(U_r, V_r^\epsilon) - \bar{F}(U_r), e_k \rangle dr \right|^2 \right)^n \\ &\leq \left( \sum_{k=1}^\infty \alpha_k^{-\zeta} \right)^{n-1} \sum_{k=1}^\infty \left[ \alpha_k^{(n-1)\zeta} \mathbb{E} \left| \int_0^s (s-r)^{-\beta} e^{-(s-r)\alpha_k} \langle F(U_r, V_r^\epsilon) - \bar{F}(U_r), e_k \rangle dr \right|^{2n} \right] \\ &\leq C \sum_{k=1}^\infty \left[ \alpha_k^{(n-1)\zeta} \mathbb{E} \left| \int_0^s (s-r)^{-\beta} e^{-(s-r)\alpha_k} \langle F(U_r, V_r^\epsilon) - \bar{F}(U_r), e_k \rangle dr \right|^{2n} \right] \\ &= C \sum_{k=1}^\infty \left[ \alpha_k^{(n-1)\zeta} \mathbb{E} \int_{[0,s]^{2n}} \prod_{i=1}^{2n} (s-r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(U_{r_i}, V_{r_i}^\epsilon) - \bar{F}(U_{r_i}), e_k \rangle dr_1 \cdots dr_{2n} \right] \\ &= C \sum_{k=1}^\infty \left[ \alpha_k^{(n-1)\zeta} \mathbb{E} \left[ \int_{[0,s]^{2n}} \prod_{i=1}^{2n} (s-r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(U_{r_i}, V_{r_i}^\epsilon) - \bar{F}(U_{r_i}), e_k \rangle \right. \right. \\ &\quad \left. \left. dr_1 \cdots dr_{2n} \Big| \{U_r, \forall r \leq s\} \right] \right]. \end{aligned}$$

Now consider the conditional expectation on the right-hand-side. Fixing  $s \geq 0$ , due to the independence of the averaged component  $\{U_r, \forall r \leq s\}$  (which is  $\sigma\{W_r^Q, \forall r \leq s\}$ -measurable) and the fast component  $\{V_r^\epsilon, \forall r \leq s\}$  (which is

$\sigma\{W_r^{Q_2}, \forall r \leq s\}$ -measurable being independent of  $\sigma\{W_r^{Q_1}, \forall r \leq s\}$ , for every  $\bar{\xi} \in C([0, s]; H)$  we have

$$\begin{aligned} & \mathbb{E} \left[ \int_{[0,s]^{2n}} \prod_{i=1}^{2n} (s - r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(U_{r_i}, V_{r_i}^\epsilon) - \bar{F}(U_{r_i}), e_k \rangle dr_1 \cdots dr_{2n} \Big| \{U_r, \forall r \leq s\} = \{\bar{\xi}_r, \forall r \leq s\} \right] \\ &= \mathbb{E} \left[ \int_{[0,s]^{2n}} \prod_{i=1}^{2n} (s - r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(\bar{\xi}_{r_i}, V_{r_i}^\epsilon) - \bar{F}(\bar{\xi}_{r_i}), e_k \rangle dr_1 \cdots dr_{2n} \right] \\ &\leq \int_{[0,s]^{2n}} \left| \mathbb{E} \left[ \prod_{i=1}^{2n} (s - r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(\bar{\xi}_{r_i}, V_{r_i}^\epsilon) - \bar{F}(\bar{\xi}_{r_i}), e_k \rangle \right] \right| dr_1 \cdots dr_{2n} \\ &\leq C \left( 1 + \sup_{r \leq T} |\bar{\xi}_r|^\eta + |v|^\eta \right) \epsilon^n \left( \frac{1}{\alpha_k} \right)^{(1-2\beta)n}, \end{aligned}$$

where the last inequality follows by Lemma 6.3.

Hence we have:

$$\begin{aligned} & \mathbb{E} \left[ \int_{[0,s]^{2n}} \prod_{i=1}^{2n} (s - r_i)^{-\beta} e^{-(s-r_i)\alpha_k} \langle F(U_{r_i}, V_{r_i}^\epsilon) - \bar{F}(U_{r_i}), e_k \rangle dr_1 \cdots dr_{2n} \Big| \{U_r, \forall r \leq s\} \right] \\ &\leq C \left( 1 + \sup_{r \leq T} |U_r|^\eta + |v|^\eta \right) \epsilon^n \left( \frac{1}{\alpha_k} \right)^{(1-2\beta)n}. \end{aligned}$$

Inserting this inequality in the one for  $\mathbb{E} |Y_s^\epsilon|^{2n}$  we have:

$$\begin{aligned} \mathbb{E} |Y_s^\epsilon|^{2n} &\leq C \epsilon^n \left( 1 + \mathbb{E} \left[ \sup_{r \leq T} |U_r|^\eta \right] + |v|^\eta \right) \sum_{k=1}^\infty \alpha_k^{(n-1)\zeta - (1-2\beta)n} \\ &\leq C \epsilon^n (1 + |u|^\eta + |v|^\eta) \sum_{k=1}^\infty \alpha_k^{n(\zeta + 2\beta - 1) - \zeta}. \end{aligned}$$

The series on the right-hand-side is convergent by Hypothesis 1 and we have (47), so that the claim is proved.

Inserting (47) into (46) we have:

$$\mathbb{E} \left[ \sup_{t \in [0, \tau]} |Z_t^\epsilon|^{2n} \right] \leq C \epsilon^n (1 + |u|^\eta + |v|^\eta).$$

Finally by Holder’s inequality we have the thesis:

$$\mathbb{E} \left[ \sup_{t \in [0, \tau]} |Z_t^\epsilon|^2 \right] \leq \left( \mathbb{E} \left[ \sup_{t \in [0, \tau]} |Z_t^\epsilon|^{2n} \right] \right)^{1/n} \leq C \epsilon (1 + |u|^t + |v|^t),$$

for  $t = \eta/n$ .

□

We can now state and prove the main Theorem of this work:

**Theorem 1** *Let  $u \in H, v \in K$  and assume Hypotheses 1, 2, 3, 4, 5. Then there exists  $C = C(T, |u|, |v|) > 0$  such that*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t^\epsilon - U_t|^2 \right] \leq C\epsilon,$$

for every  $\epsilon > 0$ .

**Proof** For  $t \in [0, T]$  we have:

$$U_t^\epsilon - U_t = \int_0^t e^{A_1(t-s)} (F(U_s^\epsilon, V_s^\epsilon) - \bar{F}(U_s)) ds,$$

so that:

$$\begin{aligned} |U_t^\epsilon - U_t| &\leq \left| \int_0^t e^{A_1(t-s)} (F(U_s^\epsilon, V_s^\epsilon) - F(U_s, V_s^\epsilon)) ds \right| \\ &\quad + \left| \int_0^t e^{A_1(t-s)} (F(U_s, V_s^\epsilon) - \bar{F}(U_s)) ds \right|. \end{aligned}$$

Now let  $0 \leq \tau \leq T$  and compute

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq t \leq \tau} |\tilde{U}_t^\epsilon - U_t|^2 \right] &\leq 2\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} (F(U_s^\epsilon, V_s^\epsilon) - F(U_s, V_s^\epsilon)) ds \right|^2 \right] \\ &\quad + 2\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} (F(U_s, V_s^\epsilon) - \bar{F}(U_s)) ds \right|^2 \right]. \end{aligned}$$

For the first term on the right-hand-side by the Lipschitz continuity of  $F$  we have:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} (F(U_s^\epsilon, V_s^\epsilon) - F(U_s, V_s^\epsilon)) ds \right|^2 \right] \leq C \int_0^\tau \mathbb{E} |U_s^\epsilon - U_s|^2 ds.$$

For the second term on the right-hand-side by Proposition 3 we have:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} (F(U_s, V_s^\epsilon) - \bar{F}(U_s)) ds \right|^2 \right] \leq C\epsilon.$$

Putting everything together we have:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau} |U_t^\epsilon - U_t|^2 \right] \leq C \left( \epsilon + \int_0^\tau \mathbb{E} |U_s^\epsilon - U_s|^2 ds \right), \tag{48}$$

for every  $\tau \leq T$ .

Then by Gronwall’s Lemma we have the thesis of the Theorem:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |U_t^\epsilon - U_t|^2 \right] \leq C\epsilon.$$

□

Finally we can provide an application to which our theory can be applied and which is not covered by the existing literature.

**Example 1** Consider the following fully coupled slow-fast stochastic reaction-diffusion system:

$$\begin{cases} \frac{\partial u_\epsilon}{\partial t}(t, \xi) = \frac{\partial^2}{\partial \xi^2} u_\epsilon(t, \xi) + f(\xi, u_\epsilon(t, \xi), v_\epsilon(t, \xi)) + \dot{w}_1(t, \xi), \\ \frac{\partial v_\epsilon}{\partial t}(t, \xi) = \frac{1}{\epsilon} \left[ \left( \frac{\partial^2}{\partial \xi^2} - \lambda \right) v_\epsilon(t, \xi) + g(\xi, v_\epsilon(t, \xi)) \right] + \frac{1}{\sqrt{\epsilon}} \dot{w}_2(t, \xi), \\ u_\epsilon(0, \xi) = u(\xi), \quad v_\epsilon(0, \xi) = v(\xi), \quad \xi \in [0, L], \\ u_\epsilon(t, \xi) = v_\epsilon(t, \xi) = 0, \quad t \geq 0, \quad \xi = 0, L, \end{cases} \quad (49)$$

where

- $t \in [0, T], \xi \in [0, L]$ ,
- $\epsilon \in (0, 1]$  is a small parameter representing the ratio of time-scales between the two variables of the system  $u_\epsilon$  and  $v_\epsilon$ ,
- $u_\epsilon$  and  $v_\epsilon$  are the slow and fast components respectively,
- $u, v \in H = L^2[0, T]$  are the initial conditions,
- $\lambda > 0$ ,
- $f, g: [0, L] \times \mathbb{R} \rightarrow \mathbb{R}$  are Lipschitz functions uniformly wrt  $\xi$  with Lipschitz constants  $L_f, L_G$  respectively and  $L_G < \lambda$ ,
- $\dot{w}_1, \dot{w}_2$  are independent white noises both in time and space.

Then it is well known [6] that (49) can be rewritten in the abstract form (2) where  $H = K = L^2[0, T], F: H \times H \rightarrow H, G: H \rightarrow H$  are the Nemytskii operators of  $f, g$  respectively, i.e.

$$\begin{aligned} F(x, y)(\xi) &= f(\xi, x(\xi), y(\xi)), \\ G(y)(\xi) &= g(\xi, y(\xi)). \end{aligned}$$

In this setting the hypotheses of Theorem 1 are satisfied (recall Remarks 2.1, 2.2) so that the result can be applied.

**Acknowledgements** The author thanks warmly his PhD supervisor Giuseppina Guatteri for her precious and constant support during the preparation of this manuscript. He also thanks Sandra Cerrai and Gianmario Tessitore for queryA funding declaration is mandatory for publication in this journal. Please confirm that this declaration is accurate, or provide an alternative.some relevant conversations related to the content of the present paper. Finally he thanks the two anonymous referees for their very careful reading of the manuscript and their precious remarks and suggestions.

**Funding** This research was partially financed by the INdAM - GNAMPA Project ‘‘Riduzione del modello in sistemi dinamici stocastici in dimensione infinita a due scale temporali’’. Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

### Declarations

**Conflict of interest** Author declared that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### A Proof of Lemma 4.5

**Proof** First observe that

$$\mathcal{H}_s^t(v) = \sigma(C_s^t(v), 0 \leq s \leq t),$$

where

$$C_s^t(v) = \{v_{r_1}^v \in A_1, \dots, v_{r_k}^v \in A_k : k \in \mathbb{N}, s \leq r_1 < \dots < r_k \leq t, A_1, \dots, A_k \in \mathcal{B}(K)\} \tag{50}$$

is the family of cylindrical sets.

Consider  $B_1 \in \mathcal{C}_0^t(v)$  and  $B_2 \in \mathcal{C}_{s+t}^\infty(v)$ , i.e.

$$B_1 = \bigcap_{i=1}^{k_1} \{v_{r_{1,i}}^v \in A_{1,i}\}, \quad B_2 = \bigcap_{i=1}^{k_2} \{v_{r_{2,i}}^v \in A_{2,i}\},$$

for  $0 \leq r_{1,1} < \dots < r_{1,k_1} \leq t$  and  $s + t \leq r_{2,1} < \dots < r_{2,k_2} < \infty$  and  $A_{j,i} \in \mathcal{B}(K)$ , for  $j = 1, 2$  and  $i = 1, \dots, k_j$ .

First by the tower property of conditional expectations we have:

$$\begin{aligned} \mathbb{P}(B_1 \cap B_2) &= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}}(v_{r_{1,i}}^v) \prod_{i=1}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^v) \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}}(v_{r_{1,i}}^v) \mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^v) \middle| \mathcal{F}_t \right] \right], \end{aligned} \tag{51}$$

where  $1_{A_{i,j}}(\cdot)$  is the indicator function.

Now, as  $t + s \leq r_{2,1} < \dots < r_{2,k_2}$ , we have:

$$\begin{aligned} & \mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}} \left( v_{r_{2,i}}^v \right) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^v \right) \mathbb{E} \left[ \prod_{i=2}^{k_2} 1_{A_{2,i}} \left( v_{r_{2,i}}^v \right) \middle| \mathcal{F}_{r_{2,1}} \right] \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^v \right) \mathbb{E} \left[ 1_{A_{2,2}} \left( v_{r_{2,2}}^v \right) \mathbb{E} \left[ \prod_{i=3}^{k_2} 1_{A_{2,i}} \left( v_{r_{2,i}}^v \right) \middle| \mathcal{F}_{r_{2,2}} \right] \middle| \mathcal{F}_{r_{2,1}} \right] \middle| \mathcal{F}_t \right]. \end{aligned}$$

By iteration we have:

$$\begin{aligned} & \mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}} \left( v_{r_{2,i}}^v \right) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^v \right) \mathbb{E} \left[ 1_{A_{2,2}} \left( v_{r_{2,2}}^v \right) \mathbb{E} \left[ 1_{A_{2,3}} \left( v_{r_{2,3}}^v \right) \dots \mathbb{E} \left[ 1_{A_{2,k_2}} \left( v_{r_{2,k_2}}^v \right) \middle| \mathcal{F}_{r_{2,k_2-1}} \right] \middle| \dots \middle| \mathcal{F}_{r_{2,2}} \right] \middle| \mathcal{F}_{r_{2,1}} \right] \middle| \mathcal{F}_t \right] \\ &= P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v_t^v \right), \end{aligned}$$

so that by (51) we have:

$$P(B_1 \cap B_2) = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^v \right) P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v_t^v \right) \right].$$

In a similar way we have:

$$\mathbb{E} \prod_{i=1}^{k_2} 1_{A_{2,i}} \left( v^v \left( r_{2,i} \right) \right) = P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v \right),$$

so that

$$\begin{aligned} & P(B_1)P(B_2) \\ &= \mathbb{E} \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^v \right) \mathbb{E} \prod_{i=1}^{k_2} 1_{A_{2,i}} \left( v^v \left( r_{2,i} \right) \right) \\ &= \mathbb{E} \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^v \right) P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v \right). \end{aligned}$$

It follows

$$\begin{aligned} & \mathbb{P}(B_1 \cap B_2) - \mathbb{P}(B_1) \mathbb{P}(B_2) \\ &= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v^{u,v} \left( r_{1,i} \right) \right) \left\{ P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v_t^v \right) \right. \right. \\ & \quad \left. \left. - P_{r_{2,1}-t} \left[ 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right) \right] \left( v \right) \right\} \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v^v \left( r_{1,i} \right) \right) \left\{ P_{r_{2,1}-t} \phi \left( v_t^v \right) - P_{r_{2,1}-t} \phi \left( v \right) \right\} \right] \end{aligned}$$

$$= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}}(v^v(r_{1,i})) \left\{ P_{r_{2,1}-t} \phi(v_i^v) - \int_K \phi(v) \mu(dv) \right\} \right],$$

for  $\phi := 1_{A_{2,1}} P_{r_{2,2}-r_{2,1}} \left( 1_{A_{2,2}} P_{r_{2,3}-r_{2,2}} \left( \dots I_{A_{2,k_2-1}} P_{r_{2,k_2}-r_{2,k_2-1}} \left( I_{A_{2,k_2}} \right) \dots \right) \right)$ .

Then by Lemmas 4.2 and 4.4 and as  $f(s) = e^{-\delta s} s^{-1/2}$  is decreasing we have:

$$\begin{aligned} |\mathbb{P}(B_1 \cap B_2) - \mathbb{P}(B_1) \mathbb{P}(B_2)| &\leq C \frac{e^{-\delta(r_{2,1}-t)}}{\sqrt{r_{2,1}-t}} (1 + \mathbb{E}|v_i^v| + |v|) \\ &\leq C \frac{e^{-\delta s}}{\sqrt{s}} (1 + |v|), \end{aligned}$$

so that the inequality holds when  $B_1 \in \mathcal{C}_0^t(v)$  and  $B_2 \in \mathcal{C}_{t+s}^\infty(v)$ .

Finally recalling (50) the validity of the inequality can be extended to every  $B_1 \in \mathcal{H}_0^t(v)$  and  $B_2 \in \mathcal{H}_{s+t}^\infty(v)$ . □

### B Proof of Lemma 4.6

**Proof** As  $|\xi_1| \leq 1$  a.s. we have:

$$\begin{aligned} |\mathbb{E}[\xi_1 \xi_2] - \mathbb{E}\xi_1 \mathbb{E}\xi_2| &= |\mathbb{E}[\xi_1 \mathbb{E}[\xi_2 | \mathcal{H}_{s_1}^{t_1}(v)]] - \mathbb{E}\xi_1 \mathbb{E}\xi_2| \\ &= |\mathbb{E}[\xi_1 (\mathbb{E}[\xi_2 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\xi_2)]| \\ &\leq |\mathbb{E}[\tilde{\xi}_1 (\mathbb{E}[\xi_2 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\xi_2)]| \\ &= |\mathbb{E}[\tilde{\xi}_1 \xi_2] - \mathbb{E}\tilde{\xi}_1 \mathbb{E}\xi_2|, \end{aligned}$$

where we have defined

$$\tilde{\xi}_1 := \begin{cases} 1, & \text{if } \mathbb{E}[\xi_2 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\xi_2 > 0, \\ -1, & \text{if } \mathbb{E}[\xi_2 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\xi_2 \leq 0. \end{cases}$$

Similarly, as  $|\xi_2| \leq 1$  a.s., we have:

$$\begin{aligned} |\mathbb{E}[\tilde{\xi}_1 \xi_2] - \mathbb{E}\tilde{\xi}_1 \mathbb{E}\xi_2| &= |\mathbb{E}[\xi_2 \mathbb{E}[\tilde{\xi}_1 | \mathcal{H}_{s_2}^{t_2}(v)]] - \mathbb{E}\tilde{\xi}_1 \mathbb{E}\xi_2| \\ &= |\mathbb{E}[\xi_2 (\mathbb{E}[\tilde{\xi}_1 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\tilde{\xi}_1)]| \\ &\leq |\mathbb{E}[\tilde{\xi}_2 (\mathbb{E}[\tilde{\xi}_1 | \mathcal{H}_{s_1}^{t_1}(v)] - \mathbb{E}\tilde{\xi}_1)]| \\ &= |\mathbb{E}[\tilde{\xi}_1 \tilde{\xi}_2] - \mathbb{E}\tilde{\xi}_1 \mathbb{E}\tilde{\xi}_2|, \end{aligned}$$



where we have defined

$$\tilde{\xi}_2 := \begin{cases} 1, & \text{if } \mathbb{E} \left[ \tilde{\xi}_1 \mid \mathcal{H}_{s_2}^{t_2}(v) \right] - \mathbb{E} \tilde{\xi}_1 > 0, \\ -1, & \text{if } \mathbb{E} \left[ \tilde{\xi}_1 \mid \mathcal{H}_{s_2}^{t_2}(v) \right] - \mathbb{E} \tilde{\xi}_1 \leq 0. \end{cases}$$

Then it follows:

$$|\mathbb{E}[\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq \left| \mathbb{E}[\tilde{\xi}_1 \tilde{\xi}_2] - \mathbb{E} \tilde{\xi}_1 \mathbb{E} \tilde{\xi}_2 \right|.$$

Notice that

$$\tilde{\xi}_1 = 21_A - 1, \quad \tilde{\xi}_2 = 21_B - 1$$

with

$$A = \left\{ \omega \in \Omega : \mathbb{E} \left[ \xi_2 \mid \mathcal{H}_{s_1}^{t_1}(v) \right] - \mathbb{E} \xi_2 > 0 \right\}, \\ B = \left\{ \omega \in \Omega : \mathbb{E} \left[ \tilde{\xi}_1 \mid \mathcal{H}_{s_2}^{t_2}(v) \right] - \mathbb{E} \tilde{\xi}_1 > 0 \right\}.$$

It follows:

$$\begin{aligned} |\mathbb{E}[\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2| &\leq \left| \mathbb{E}[\tilde{\xi}_1 \tilde{\xi}_2] - \mathbb{E} \tilde{\xi}_1 \mathbb{E} \tilde{\xi}_2 \right| = |\mathbb{E}[(21_A - 1)(21_B - 1)] \\ &\quad - \mathbb{E}[21_A - 1] \mathbb{E}[21_B - 1]| \\ &= 4|\mathbb{E}[1_A 1_B] - \mathbb{E}[1_A] \mathbb{E}[1_B]| = 4|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|. \end{aligned}$$

Now by Lemma 4.5 as  $A \in \mathcal{H}_{s_1}^{t_1}(v)$ ,  $B \in \mathcal{H}_{s_2}^{t_2}(v)$  we have:

$$|\mathbb{E}[\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq C \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|)$$

so that we have the thesis of the Lemma. □

### C Proof of Lemma 4.7

**Proof** We proceed in a similar way to the proof of [7, Proposition 3.3].

Indeed fix  $R > 0$  and set  $A_{1,R} = \{\omega \in \Omega : |\xi_1| \leq R\}$ ,  $A_{2,R} = \{\omega \in \Omega : |\xi_2| \leq R\}$ . Then we have:

$$\begin{aligned} &\mathbb{E}[\xi_1 \xi_2] - \mathbb{E} \xi_1 \mathbb{E} \xi_2 \\ &= \mathbb{E} \left[ \xi_1 \xi_2 \left( 1_{A_{1,R} \cap A_{2,R}} + 1_{A_{1,R}^c \cup A_{2,R}^c} \right) \right] - \mathbb{E} \xi_1 \mathbb{E} \xi_2 \\ &= \left\{ \mathbb{E}[\xi_1 1_{A_{1,R}} \xi_2 1_{A_{2,R}}] - \mathbb{E}[\xi_1 1_{A_{1,R}}] \mathbb{E}[\xi_2 1_{A_{2,R}}] \right\} + \mathbb{E}[\xi_1 \xi_2 1_{A_{1,R}^c \cup A_{2,R}^c}] \end{aligned}$$

$$\begin{aligned}
 & - \left\{ \mathbb{E}[\xi_1 1_{A_{1,R}}] \mathbb{E}[\xi_2 1_{A_{2,R}^c}] + \mathbb{E}[\xi_1 1_{A_{1,R}^c}] \mathbb{E}[\xi_2 1_{A_{2,R}}] + \mathbb{E}[\xi_1 1_{A_{1,R}^c}] \mathbb{E}[\xi_2 1_{A_{2,R}^c}] \right\} \\
 & = T_{1,R} + T_{2,R} + T_{3,R}.
 \end{aligned}$$

Consider  $T_{1,R}$ , then we have:

$$T_{1,R} = R^2 \left( \mathbb{E} \left[ \frac{\xi_1}{R} 1_{A_{1,R}} \frac{\xi_2}{R} 1_{A_{2,R}} \right] - \mathbb{E} \left[ \frac{\xi_1}{R} 1_{A_{1,R}} \right] \mathbb{E} \left[ \frac{\xi_2}{R} 1_{A_{2,R}} \right] \right).$$

Now as  $\left| \frac{\xi_i}{R} 1_{A_{i,R}} \right| \leq 1$  a.s then by Lemma 4.6 we have:

$$|T_{1,R}| \leq CR^2 \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|).$$

For  $T_{2,R}$  by Holder’s and Markov’s inequalities we have:

$$\begin{aligned}
 |T_{2,R}|^{\frac{2}{1-\rho}} & \leq \mathbb{E} |\xi_1|^{\frac{2}{1-\rho}} \mathbb{E} |\xi_2|^{\frac{2}{1-\rho}} (\mathbb{P}(A_{1,R}^c) + \mathbb{P}(A_{2,R}^c))^{\frac{2\rho}{1-\rho}} \\
 & \leq \mathbb{E} |\xi_1|^{\frac{2}{1-\rho}} \mathbb{E} |\xi_2|^{\frac{2}{1-\rho}} R^{-\frac{2\rho}{1-\rho}} (\mathbb{E} |\xi_1| + \mathbb{E} |\xi_2|)^{\frac{2\rho}{1-\rho}}
 \end{aligned}$$

and then by (26) we have:

$$|T_{2,R}| \leq CK_1 K_2 (K_1 + K_2)^\rho R^{-\rho}.$$

For the first term of  $T_{3,R}$  we have:

$$|\mathbb{E}[\xi_1 1_{A_{1,R}}] \mathbb{E}[\xi_2 1_{A_{2,R}^c}]| \leq \left( \mathbb{E} |\xi_1|^{\frac{2}{1-\rho}} \right)^{\frac{1-\rho}{2}} \left( \mathbb{E} |\xi_2|^{\frac{2}{1-\rho}} \right)^{\frac{1-\rho}{2}} \mathbb{P}(A_{2,R}^c)^\rho.$$

Also the other terms can be treated in an analogous way and then similarly to before we have:

$$|T_{3,R}| \leq CK_1 K_2 (K_1 + K_2)^\rho R^{-\rho}.$$

Now by inserting the inequalities for  $T_{i,R}$  into the first equation we have:

$$|\mathbb{E} \xi_1 \xi_2 - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq C \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|) R^2 + CK_1 K_2 (K_1 + K_2)^\rho R^{-\rho}.$$

By minimizing over  $R > 0$  the right-hand-side of the previous inequality we have:

$$|\mathbb{E} \xi_1 \xi_2 - \mathbb{E} \xi_1 \mathbb{E} \xi_2| \leq CK_1^{\frac{2}{2+\rho}} K_2^{\frac{2}{2+\rho}} (K_1 + K_2)^{\frac{2\rho}{2+\rho}} \left( \frac{e^{-\delta(s_2-t_1)}}{\sqrt{s_2-t_1}} (1 + |v|) \right)^{\frac{\rho}{2+\rho}}.$$

□

## References

1. Bardi, M., Cesaroni, A., Manca, L.: Convergence by viscosity methods in multiscale financial models with stochastic volatility. *SIAM J. Financ. Math.* **1**(1), 230–265 (2010). <https://doi.org/10.1137/090748147>
2. Bréhier, C.E.: Strong and weak orders in averaging for SPDEs. *Stoch. Process. Appl.* **122**(7), 2553–2593 (2012)
3. Bréhier, C.E.: Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. *SIAM J. Numer. Anal.* **51**, 1185–1210 (2013)
4. Bréhier, C.E.: Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component. *Stoch. Process. Appl.* **130**(6), 3325–3368 (2020)
5. Cerrai, S., Freidlin, M.: Averaging principle for stochastic reaction-diffusion equations. *Prob. Theory Relat. Fields* **144**, 137–177 (2009)
6. Cerrai, S.: A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *Ann. Appl. Probab.* **19**(3), 899–948 (2009). <https://doi.org/10.1214/08-AAP560>
7. Cerrai, S.: Normal deviations from the averaged motion for some reaction-diffusion equations with fast oscillating perturbation. *J. Math. Pures Appl.* **91**(6), 614–647 (2009)
8. Cerrai, S.: Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise. *SIAM J. Math. Anal.* **43**, 2482–2518 (2011)
9. Cerrai, S., Lunardi, A.: Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: the almost periodic case. *SIAM J. Math. Anal.* **49**, 2843–2884 (2017). <https://doi.org/10.1137/16M1063307>
10. Da Prato, G., Zabczyk, J.: *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge (2014)
11. Da Prato, G., Zabczyk, J.: *Ergodicity for Infinite Dimensional Systems*, vol. 229. Cambridge University Press, Cambridge (1996)
12. de Feo, F.: The averaging principle for stochastic differential equations and a financial application. Master's Thesis, Politecnico di Milano (2020). <https://www.politesi.polimi.it/handle/10589/165293>
13. de Feo, F.: The averaging principle for non-autonomous slow-fast stochastic differential equations and an application to a local stochastic volatility model. *J. Differ. Equ.* **302**, 406–443 (2021)
14. Dong, Z., Sun, X., Xiao, H., Zhai, J.: Averaging principle for one dimensional stochastic Burgers equation. *J. Differ. Equ.* **265**, 4749–4797 (2018)
15. Fu, H., Wan, L., Liu, J.: Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. *Stoch. Process. Appl.* **125**(8), 3255–3279 (2015)
16. Givon, D.: Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems. *Multiscale Modeling Simul.* **6**(2), 577–594 (2007)
17. Golec, J.: Stochastic averaging principle for systems with pathwise uniqueness. *Stoch. Anal. Appl.* **13**, 307–322 (1995)
18. Golec, J., Ladde, G.: Averaging principle and systems of singularly perturbed stochastic differential equations. *J. Math. Phys.* **31**, 1116–1123 (1990)
19. Guatteri, G., Tessitore, G.: Singular limit of BSDEs and optimal control of two scale stochastic systems in infinite dimensional spaces. *Appl. Math. Optim.* **83**(2), 1025–1051 (2021)
20. Guatteri, G., Tessitore, G.: Singular limit of two-scale stochastic optimal control problems in infinite dimensions by vanishing noise regularization. *SIAM J. Control Optim.* **60**(1), 575–596 (2022)
21. Fouque, J.P., Papanicolaou, G., Sircar, R.: *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, Cambridge (2000)
22. Fouque, J.P., Papanicolaou, G., Sircar, R., Solna, K.: *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Cambridge University Press, Cambridge (2011)
23. Freidlin, M.I., Wentzell, A.D.: *Random Perturbations of Dynamical Systems*. Springer, New York (1998)
24. Fu, H., Wan, L., Liu, J.: Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales, *Stochastic Processes and their Applications*, **125**(8), 3255–3279. ISSN 0304-4149 (2015) <https://doi.org/10.1016/j.spa.2015.03.004>
25. Fu, H., Wan, L., Liu, J., Zhang, B.: Weak order in averaging principle for stochastic differential equations with jumps. *Adv. Differ. Equ. Paper No.* 197, 20 pp (2018)
26. Fu, H., Wan, L., Liu, J., Liu, X.: Weak order in averaging principle for stochastic wave equation with a fast oscillation. *Stoch. Process. Appl.* **128**(8), 2557–2580 (2018)

27. Khasminskii, R.Z.: On an averaging principle for Itô stochastic differential equations. *Kibernetika* **4**(3), 260–279 (1968). ((Russian))
28. Khasminskii, R.Z., Yin, G.: On averaging principles: an asymptotic expansion approach. *SIAM J. Math. Anal.* **35**(6), 1534–1560 (2004)
29. Liu, D., Weinan, E., Vanden-Eijnden, E.: Analysis of multiscale methods for stochastic differential equations. *Commun. Pure Appl. Math.* **58**(11), 1544–1585 (2005)
30. Liu, D.: Strong convergence of principle of averaging for multiscale stochastic dynamical systems. *Commun. Math. Sci.* **8**, 999–1020 (2010)
31. Liu, W., Rockner, M., Sun, X., Xie, Y.: Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients, *Journal of Differential Equations*, **268**(6), 2910–2948, ISSN 0022-0396 (2020). <https://doi.org/10.1016/j.jde.2019.09.047>
32. Fuke, W., Tian, T., Rawlings, J.B., Yin, G.: Approximate method for stochastic chemical kinetics with two-time scales by chemical Langevin equations. *J. Chem. Phys.* **144**(17), 174112 (2016). <https://doi.org/10.1063/1.4948407>
33. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York (1983)
34. Roberts, A.J., Wang, W.: Average and deviation for slow-fast stochastic partial differential equations. *J. Differ. Equ.* **253**(5), 1265–1286 (2012)
35. Röckner, M., Sun, X., Xie, L.: Strong and weak convergence in the averaging principle for SDEs with Hölder coefficients. arXiv preprint [arXiv:1907.09256](https://arxiv.org/abs/1907.09256) (2019)
36. Röckner, M., Xie, L., Yang, L.: Asymptotic behavior of multiscale stochastic partial differential equations. arXiv preprint [arXiv:2010.14897](https://arxiv.org/abs/2010.14897) (2020)
37. Rozanov, I.A.: *Stationary random processes*. Holden-Day, (1967)
38. Świąch, A.: Singular perturbations and optimal control of stochastic systems in infinite dimension: HJB equations and viscosity solutions. *ESAIM Control Optim. Calculus Variations* **27**, 6 (2021)
39. Uda, K.: Averaging principle for stochastic differential equations in the random periodic regime. *Stoch. Process. Appl.* **139**, 1–36. ISSN 0304-4149, (2021) <https://doi.org/10.1016/j.spa.2021.04.017>
40. Veretennikov, AYu.: On the averaging principle for systems of stochastic differential equations. *Math. USSR-Sbornik* **69**(1), 271 (1991)
41. Xu, J., Liu, J., Miao, Y.: Strong averaging principle for two-time-scale SDEs with non-Lipschitz coefficients. *J. Math. Anal. Appl.* **468**(1), 116–140 (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.