The order of convergence in the averaging principle for slow-fast systems of SDE’s in Hilbert spaces

Filippo de Feo
filippo.defeo@polimi.it

Department of Mathematics, Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy

June 7, 2022

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 SDE’s in Hilbert spaces when the stochastic perturbations are general Wiener processes, i.e their covariance operators are allowed to be not trace class. In particular 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 fully coupled 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

Declarations of interest: none

1 Introduction
Consider the following slow-fast system of abstract SDE’s

\[
\begin{cases}
    dU^\varepsilon_t = A_1 U^\varepsilon_t + F(U^\varepsilon_t, V^\varepsilon_t)dt + dW^Q_1 \\
    U^\varepsilon_0 = u \in H \\
    dV^\varepsilon_t = \varepsilon^{-1} A_2 V^\varepsilon_t + \varepsilon^{-1} G(U^\varepsilon_t, V^\varepsilon_t)dt + \varepsilon^{-1/2} dW^Q_2 \\
    V^\varepsilon_0 = v \in K
\end{cases}
\]  

(1)

where $\varepsilon > 0$ is a small parameter representing the ratio of time scales between the slow component of the system $U^\varepsilon$ and the fast one $V^\varepsilon$. Here $H, K$ are Hilbert spaces, $A_1, A_2$ 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, [30] to chemistry, [38] 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 $\varepsilon \to 0$. In particular under
certain hypotheses it is known that the slow component $U^\varepsilon$ converges to the solution $U$ of the so-called averaged equation

$$\begin{align*}
dU_t &= A_1 U_t + \tilde{F}(U_t)dt + dW^Q_t, \\
U_0 &= u
\end{align*}$$

where

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

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

$$\begin{align*}
dv_s^{u,v} &= A_2 vv_s^{u,v} + G(u, v_s^{u,v})ds + dW_s^{Q_2} \\
v_0^{u,v} &= v \in K
\end{align*}$$

Note that the equation for $U$ is uncoupled from $V^\varepsilon$. This fact is known as averaging principle and it is fundamental in applications since $U$ captures the effective dynamic of $U^\varepsilon$ (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 [26]. For generalizations and improvements see [13], [17], [18], [23], [32], [37], [39] and the references therein. It is important to mention that 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], [36] and the references therein. Also for infinite dimensional systems the previous comment about the dependence of the diffusion coefficients holds.

See also [19], [20], [35] 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 $U^\varepsilon \to U$, i.e. see [3], [28]. For the study of the order of convergence for finite dimensional systems we refer to [16], [27], [29], [31], [40] 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 is being addressed by researchers: in [4] the author (generalizing his previous work [2]) considers a slow-fast stochastic reaction diffusion system where the fast component is uncoupled from the slow one (i.e. $G = G(V_t^\varepsilon)$) and the noise is additive. Both the weak and the 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 [38] a 1-dimensional 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. $\text{Tr}(\Lambda^{1/2}Q_1) < \infty$, where $\Lambda$ is the Laplacian.

In [33] the strong order of convergence for (1) is studied. Here it is assumed that the the covariance operators of the noises are trace-class and moreover that $\text{Tr}(-A_1Q_1) < \infty$.
See also [25] 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 $\text{Tr}(Q_1) < \infty$.

Indeed in all these papers the case $\text{Tr}(Q_1) = \infty$, which is very important for applications as it happens very naturally for example when the stochastic perturbation is a white noise i.e. $Q_1 = I$, 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 SDE’s (1) where $W^{Q_1}, W^{Q_2}$ are general Wiener processes on $H, K$ respectively with covariance operators $Q_1, Q_2$ with $\text{Tr}(Q_1) = +\infty, \text{Tr}(Q_2) = +\infty$ 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

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

where $U_t$ is the solution of the averaged equation. Notice that this result is much stronger than [4], [33] where $\sup_{t \in [0,T]}$ is outside the expectation.

Our proof is based on a classical time-discretization argument inspired by Khasminskii [26]. In particular $[0, T]$ is divided into intervals of the form $[K\Delta, (K + 1)\Delta]$ with $\Delta = \Delta(\epsilon)$ and an approximated process $(\tilde{U}^\epsilon, \tilde{V}^\epsilon)$ is built by fixing the slow component on each $[K\Delta, (K + 1)\Delta]$ (but with the fast component free to move): this corresponds to the intuition for a slow-fast system. Using the discretized process it is then possible to prove that the strong order of convergence is $1/2$.

The key tool in the discretization argument is proposition 3 whose proof is based on a technical result, i.e. lemma 6.2 which is inspired by [7] and it is a consequence of the mixing properties of the fast motion, i.e. lemmas 4.6, 4.7, 4.8. We recall that [7] studies the normal deviations, i.e. the weak convergence of $Z^\epsilon := (U^\epsilon - U)/\sqrt{\epsilon}$ but it has a simpler system, i.e. the equation for the slow component has no stochastic perturbation ($Q_1 = 0$) and the equation for the fast component is uncoupled from the slow-component ($G = G(v)$).

Finally we discuss an application of our theory to a fully coupled 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 section 2 we introduce the problem in a formal way and we state the assumptions that we will use.

in section 3 we prove some a-priori estimates.

in section 4 we prove some results related to the fast motion.

in section 5 we study the well posedness of the averaged equation.

in section 6 we prove some preliminary results.

in section 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.

Call $H, K$ 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 $\phi: H \to \mathbb{R}$ with the sup norm $| \cdot |_{H, \infty}$.

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

$$
[\phi]_{H, \text{Lip}} = \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{|x - y|_H}
$$
\( \mathcal{L}(H) \) will denote the space of linear bounded operators from \( H \) to \( H \), endowed with the operator norm
\[
\|L\|_{H^1} = \sup_{|x|=1} \|Lx\|_H
\]
Next denote by \( \mathcal{L}_2(H) \) the space of Hilbert-Schmidt operators endowed with the norm
\[
\|L\|_{\mathcal{L}_2(H)} = (\text{Tr}(L^*L))^{1/2}
\]
The analogous spaces \( \mathcal{B}(K), \text{Lip}(K), \mathcal{L}(K), \mathcal{L}_2(K) \) are defined for the Hilbert space \( K \) with the corresponding norms \( | \cdot |_{K,\infty}, | \cdot |_{\text{Lip},K}, \| \cdot \|_K, \| \cdot \|_{\mathcal{L}_2(K)} \).
In order to simplify the notation we will omit the pedices \( \mu \) 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 \( 0 \leq t \leq T < \infty \)
\[
\begin{align*}
dU_t^\varepsilon &= A_1U_t^\varepsilon + F(U_t^\varepsilon, V_t^\varepsilon)dt + dW_t^{Q_1} \\
U_0^\varepsilon &= u \in H \\
dV_t^\varepsilon &= \varepsilon^{-1}A_2V_t^\varepsilon + \varepsilon^{-1}G(U_t^\varepsilon, V_t^\varepsilon)dt + \varepsilon^{-1/2}dW_t^{Q_2} \\
V_0^\varepsilon &= \nu \in K
\end{align*}
\]
where
- \( \varepsilon > 0 \) is a parameter,
- \( A_1: D(A_1) \subset H \to H, A_2: D(A_2) \subset K \to K \) are linear operators,
- \( F: H \times K \to H, G: H \times K \to K \),
- \( W^{Q_1}, W^{Q_2} \) 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’ll use throughout the work:

**Hypothesis 1.** \( A_1: D(A_1) \subset H \to H \) is a linear operator generator of an analytical semigroup \( e^{A_1t} \) on \( H \), \( t \geq 0 \).
Moreover there exist an orthonormal basis \( \{e_k\}_{k \in \mathbb{N}} \) of \( H \) and \( \{\alpha_k\}_{k \in \mathbb{N}} \) such that \( \alpha_k \geq 0 \),
\[
A_1e_k = -\alpha_ke_k
\]
Moreover we assume that there exist \( \mu > 0, n \geq 2 \) integer and \( 1/(2n) < \beta < 1/3 \) such that
\[
\sum_{k=1}^{\infty} \alpha_k^{-\mu} < +\infty
\]
and
\[
\sum_{k=1}^{\infty} \alpha_k^{n(\mu+2\beta-1)-\mu} < +\infty
\]

**Remark 2.1.** *Hypothesis 1 is necessary in the proof of proposition 3. It holds for example when \( A_1 \) is the Laplacian on \( [0,1] \). Indeed in this case it is well known that \( \alpha_k \sim k^2 \) and then the two series converge for example by choosing \( \mu = \frac{1}{y} \), \( n = 3 \), \( \beta = \frac{1}{y} \).*
From hypothesis 1 it follows that $e^{A_1 t}$ is a semigroup of contractions, i.e.

$$\|e^{A_1 t}\| \leq 1$$  \hspace{1cm} (3)

for every $t \geq 0$.

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

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

Moreover there exists $\lambda > 0$ such that

$$\|e^{A_2 t}\| \leq e^{-\lambda t}$$  \hspace{1cm} (4)

for every $t \geq 0$.

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

$$|F(x_2, y_2) - F(x_1, y_1)| \leq L_f(|x_2 - x_1| + |y_2 - y_1|)$$
$$|G(x_2, y_2) - G(x_1, y_1)| \leq L_G(|x_2 - x_1| + |y_2 - y_1|)

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 $A_2 + G(x, \cdot)$ is strongly dissipative for fixed $x \in H$, i.e. set

$$\delta = \frac{\lambda - L_G}{2} > 0$$  \hspace{1cm} (5)

then it holds:

$$\langle A_2 (y_2 - y_1) + G(x, y_2) - G(x, y_1), y_2 - y_1 \rangle \leq -2\delta |y_2 - y_1|^2$$  \hspace{1cm} (6)

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

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

$$\left\| e^{A_2 t} Q_1^{1/2} \right\|_{L_2} \leq C(t \wedge 1)^{-\gamma}$$
$$\left\| e^{A_2 t} Q_2^{1/2} \right\|_{L_2} \leq C(t \wedge 1)^{-\gamma}$$

for every $t > 0$.

**Hypothesis 5.** Assume that $Q_2$ is invertible with inverse $Q_2^{-1}$.

**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

$$\begin{cases}
U_t^\xi = e^{A_1 t} u + \int_0^t e^{A_1(t-s)} F(U_s^\xi, V_s^\xi) ds + \int_0^t e^{A_1(t-s)} dW_s^{Q_1} \\
V_t^\xi = e^{A_1 t} v + \int_0^t e^{A_2(t-s)} G(U_s^\xi, V_s^\xi) ds + \int_0^t e^{A_2(t-s)} dW_s^{Q_2}
\end{cases}$$  \hspace{1cm} (7)

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

**Proof.** See e.g. [10]. \hfill \square

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.**

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

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

**Proof.** Fix \( 0 < \eta < 1/2 \), by the factorization method we have:

\[
\Gamma_t^1 = C \int_0^t e^{A_1(t-\rho)} Y_\rho \, d\rho
\]

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:

\[
E \left[ \sup_{t \leq T} \left| \Gamma_t^1 \right|^p \right] \leq C \int_0^T E|Y_\rho|^p \, d\rho
\]

for every \( p > 1/\eta > 2 \).

Now we estimate \( E|Y_\rho|^p \); by Burkholder’s inequality and hypothesis 4 we have:

\[
E|Y_\rho|^p \leq C \left( \int_0^\rho \left\| (-A_1)^{\theta} e^{A_1(\rho-s)} (\rho-s)^{-\eta} Q_1^{1/2} \right\|_{L^2}^2 \, ds \right)^{p/2}
\]

\[
\leq C \left( \int_0^\rho (\rho-s)^{-2\eta} \left\| (-A_1)^{\theta} e^{A_1(\rho-s)/2} \right\|_{L^2}^2 \left\| e^{A_1(\rho-s)/2 Q_1^{1/2}} \right\|_{L^2}^2 \, ds \right)^{p/2}
\]

\[
\leq C \left( \int_0^\rho (\rho-s)^{-2(\eta+\theta+\gamma)} \, ds \right)^{p/2} < C
\]

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

Next inserting the last inequality in (8) 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. \( \square \)
Lemma 3.2. Let \( p \geq 2 \), then there exists \( C = C(p) > 0 \) such that
\[
\sup_{t > 0} \mathbb{E} \left[ |t^2 \varepsilon| |p| \right] \leq C
\]
for every \( \varepsilon > 0 \).

Proof. For \( t > 0 \) by Burkholder’s inequality and by our hypotheses we have:
\[
\mathbb{E} \left[ |t^2 \varepsilon| |p| \right] \leq C \left( \int_0^t \frac{1}{\varepsilon} \left\| e^{\Lambda_2(t-s)/\varepsilon} Q^2_{t-s} \right\|_{L^2}^2 ds \right)^{p/2}
\]
\[
= C \left( \int_0^t \left\| e^{\Lambda_2 p Q^2} \right\|_{L^2}^2 dp \right)^{p/2}
\]
\[
\leq C \left( \int_0^t e^{-\lambda p} \left( \frac{p}{2} \wedge 1 \right)^{-2\gamma} dp \right)^{p/2}
\]
\[
\leq C \left( \int_0^T e^{-\lambda p} \left( \frac{p}{2} \wedge 1 \right)^{-2\gamma} dp \right)^{p/2} \leq C
\]
so that the thesis is proved. \( \square \)

Lemma 3.3. Let \( p \geq 2 \) then there exists \( C = C(T, p) > 0 \) such that
\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |U_t^\varepsilon|^p \right] \leq C (1 + |u|^p + |v|^p)
\]
and
\[
\sup_{t \in [0, T]} \mathbb{E} \left[ |V_t^\varepsilon|^p \right] \leq C (1 + |u|^p + |v|^p)
\]
for every \( \varepsilon > 0 \).

Proof. Define
\[
\Lambda_t^1 = U_t^\varepsilon - \Gamma_t^1
\]
so that
\[
d\Lambda_t^1 = A_1 \Lambda_t^1 + f(\Lambda_t^1 + \Gamma_t^1, V_t^\varepsilon)dt
\]
By Young’s inequality and the sublinearity of \( F \) we have:
\[
\frac{1}{p} \frac{d}{dt} |\Lambda_t^1|^p = \langle A_1 \Lambda_t^1, \Lambda_t^1 \rangle |\Lambda_t^1|^{p-2} + \langle f(\Lambda_t^1 + \Gamma_t^1, V_t^\varepsilon), \Lambda_t^1 \rangle |\Lambda_t^1|^{p-2}
\]
\[
\leq C |\Lambda_t^1|^p + C |f(\Lambda_t^1 + \Gamma_t^1, V_t^\varepsilon)|^p
\]
\[
\leq C |\Lambda_t^1|^p + C (1 + |\Gamma_t^1|^p + |V_t^\varepsilon|^p)
\]
for every \( t \leq T \).

Then by the comparison theorem we have
\[
|\Lambda_t^1|^p \leq C |u|^p + C \int_0^t \left( 1 + |\Gamma_s^1|^p + |V_s^\varepsilon|^p \right) ds
\]
(11)
for every $t \leq T$.  
Then by the definition of $\Lambda^{1e}$ and this last inequality it follows
\[ |u_t^e|^p \leq C \left( 1 + |u|^p + \int_0^T \left( |\tau_s^e|^p + |V_s^e|^p \right) ds + |\tau_t^e|^p \right) \]  
(12) for every $t \leq T$.
Now by lemma 3.1 we have:
\[ \mathbb{E} \left[ \sup_{t \leq \tau} |u_t^e|^p \right] \leq C \left( 1 + |u|^p + \int_0^\tau \mathbb{E} |V_s^e|^p \, dr \right) \]  
(13) for every $\tau \leq T$.
Now we proceed in a similar way to [19, proof of lemma 3.10], i.e. set
\[ \Lambda_t^{2e} = e^{\frac{2t}{\epsilon}} (V_t^e - \Gamma_t^{2e}) \]
so that
\[ d\Lambda_t^{2e} = \frac{2s}{\epsilon} \Lambda_t^{2e} \, dt + \frac{1}{\epsilon} A_2 \Lambda_t^{2e} \, dt + \frac{1}{\epsilon} e^{\frac{2s}{\epsilon}} G(U_t^e, e^{-\frac{2s}{\epsilon}} \Lambda_t^{2e} + \Gamma_t^{2e}) \, dt \]
Now by (6) it follows
\[ d|\Lambda_t^{2e}|^2 = \frac{4s}{\epsilon} |\Lambda_t^{2e}|^2 + \frac{2}{\epsilon} \left( A_2 |\Lambda_t^{2e}|^2, \Lambda_t^{2e} \right) \, dt + \frac{2}{\epsilon} e^{\frac{2s}{\epsilon}} \left( G(U_t^e, e^{-\frac{2s}{\epsilon}} \Lambda_t^{2e} + \Gamma_t^{2e}), \Lambda_t^{2e} \right) \, dt \]
\[ - \frac{2}{\epsilon} e^{\frac{2s}{\epsilon}} \left( G(U_t^e, \Gamma_t^{2e}), \Lambda_t^{2e} \right) \, dt \]
\[ \leq \frac{2}{\epsilon} e^{2s/\epsilon} \left( G(U_t^e, \Gamma_t^{2e}), \Lambda_t^{2e} \right) \, dt \]
Now similarly to [19, proof of lemma 3.10] fix $\theta > 0$, differentiate $f(x) = \sqrt{x + \theta}$ and use the previous inequality, then:
\[ |\Lambda_t^{2e}| \leq \sqrt{|\Lambda_t^{2e}|^2 + \theta} \]
\[ \leq \sqrt{|v|^2 + \theta + \int_0^t \frac{1}{\epsilon} e^{\sqrt{\frac{2s}{\epsilon}}} |G(U_s^e, \Gamma_s^{2e})| |\Lambda_s^{2e}| \, ds} \]
Now by dominated convergence for $\theta \to 0$ we have:
\[ |\Lambda_t^{2e}| \leq |v| + \int_0^t \frac{1}{\epsilon} e^{\sqrt{\frac{2s}{\epsilon}}} |G(U_s^e, \Gamma_s^{2e})| \, ds \]
so that by recalling the definition of $\Lambda_t^{2e}$ we have:
\[ |V_t^e| \leq e^{-\frac{2t}{\epsilon} + 1} |v| + \int_0^t \frac{1}{\epsilon} e^{-\frac{2s}{\epsilon}(1-s)} |G(U_s^e, \Gamma_s^{2e})| ds + |\Gamma_t^{2e}| \]
\[ = e^{-\frac{2t}{\epsilon} + 1} |v| + \int_0^t e^{-\frac{2s}{\epsilon}(1-s)} |G(U_s^e, \Gamma_s^{2e})| ds + |\Gamma_t^{2e}| \]
By Holder’s inequality, lemma 3.2 and (13) we then have:

\[
\mathbb{E}[|V_t^\varepsilon|^p] \leq C|v|^p + C\mathbb{E}\left[ \int_0^{t/\varepsilon} e^{-2\delta p/(t/\varepsilon - s)} |G(U_{\varepsilon s}, \Gamma_0^2)|^p \, ds \right] + C\mathbb{E}[|\Gamma_t^2|^p]
\]

\[
\leq C|v|^p + C\left( \int_0^{t/\varepsilon} e^{-2\delta p/(t/\varepsilon - s)} \, ds \right)^{p-1} \int_0^{t/\varepsilon} e^{-\delta p(t/\varepsilon - s)} \mathbb{E}[|G(U_{\varepsilon s}, \Gamma_0^2)|^p] \, ds + C\mathbb{E}[|\Gamma_t^2|^p]
\]

\[
\leq C(1 + |u|^p + |v|^p) + C\int_0^t \frac{1}{e^{\delta p(t-s)/\varepsilon}} \int_s^t \mathbb{E}[|V_r^\varepsilon|^p] \, dr \, ds
\]

\[
= C(1 + |u|^p + |v|^p) + C\int_0^t \left( \int_0^{t-r/\varepsilon} e^{-\delta p \sigma} \, d\sigma \right) \mathbb{E}[|V_r^\varepsilon|^p] \, dr
\]

\[
\leq C(1 + |u|^p + |v|^p) + C\int_0^t \mathbb{E}[|V_r^\varepsilon|^p] \, dr
\]

Now by Gronwall’s lemma we have (10).

Finally inserting (10) into (13) we have (9). \(\Box\)

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

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

for every \(\varepsilon > 0\).

**Proof.** Consider for \(t \leq T\)

\[
U_t^\varepsilon = e^{A_\alpha t} u + \int_0^t e^{A_\alpha(t-s)} f(U_{s}^\varepsilon, V_s^\varepsilon) \, ds + \int_0^t e^{A_\alpha(t-s)} dW_s^\varepsilon
\]

(14)

First as \(u \in D(-A_\alpha^2)\) we have:

\[
\sup_{t \leq T} |e^{A_\alpha t} u|_\alpha^2 \leq C|u|_\alpha^2
\]

Moreover as \(\|A_\alpha e^{A_\alpha t}\| \leq C t^{-\alpha}\) for every \(0 < t \leq T\) and by lemma 3.3 we have:

\[
\mathbb{E}\left[ \sup_{t \in [0, T]} \left| \int_0^t e^{A_\alpha(t-s)} f(U_{s}^\varepsilon, V_s^\varepsilon) \, ds \right|_\alpha^2 \right] \leq \mathbb{E}\left[ \sup_{t \in [0, T]} \left( \int_0^t (t-s)^{-\alpha} |f(U_{s}^\varepsilon, V_s^\varepsilon)| \, ds \right)^2 \right]
\]

\[
\leq C \mathbb{E}\left[ \sup_{t \in [0, T]} \int_0^t (t-s)^{-2\alpha} ds \int_0^t |f(U_{s}^\varepsilon, V_s^\varepsilon)|^2 \, ds \right]
\]

\[
\leq C \int_0^T (1 + \mathbb{E}[|U_s^\varepsilon|^2 + |V_s^\varepsilon|^2]) \, ds
\]

\[
\leq C(1 + |u|^2 + |v|^2)
\]
Finally by lemma 3.1 we have:
\[ \mathbb{E} \left[ \sup_{t \in [0, T]} \| f(t) \|_\alpha^2 \right] < +\infty \]

By considering (14), 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. \( \Box \)

**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

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

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
\[ U_{t+h}^\epsilon - U_t^\epsilon = (e^{A_1^\alpha h} - I)U_t^\epsilon + \int_t^{t+h} A_1^\alpha (t+h-s) f(U_s^\epsilon, V_s^\epsilon) \, ds + \int_t^{t+h} e^{A_1^\alpha (t+h-s)} \, dW^Q_{s,2} \quad (15) \]

Consider the first term on the right-hand-side, then as \( u \in D(-A_1^\alpha) \) and by lemma 3.4 we have:
\[ \mathbb{E}\left[ |(e^{A_1^\alpha h} - I)U_t^\epsilon\|_\alpha^2 \right] \leq C|h|^{2\alpha} \mathbb{E}\| U_t^\epsilon \|_\alpha^2 \leq C|h|^{2\alpha}(1 + |u|_\alpha^2 + |v|^2) \]

Consider now the second term on the right-hand-side, then by lemma 3.3 we have:
\[ \mathbb{E}\left[ \left| \int_t^{t+h} e^{A_1^\alpha (t+h-s)} f(U_s^\epsilon, V_s^\epsilon) \, ds \right|^2 \right] \leq C|h| \int_t^{t+h} (1 + \mathbb{E}|U_s^\epsilon|^2 + \mathbb{E}|V_s^\epsilon|^2) \, ds \]
\[ \leq C|h|^2(1 + |u|_\alpha^2 + |v|^2) \]

Finally for the third term on the right-hand-side by Ito’s isometry and hypothesis 4 we have:
\[ \mathbb{E}\left[ \left| \int_t^{t+h} e^{A_1^\alpha (t+h-s)} \, dW^Q_{s,2} \right|^2 \right] = \int_t^{t+h} \left\| e^{A_1^\alpha (t+h-s)} Q_1^{1/2} \right\|_L^2 \, ds \]
\[ \leq C \int_t^{t+h} |t + h - s|^{-\gamma} \, ds \]
\[ = C|h|^{-2\gamma} \]

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

## 4 Fast motion

In this section we study some classical properties of the fast motion when the slow motion is fixed. Consider for fixed \( u \in H \)

\[ \begin{align*}
  dv_s^{u,v} &= A_2 v_s^{u,v} \, ds + G(u, v_s^{u,v}) \, ds + dw_s^{Q_2} \\
  v_0^{u,v} &= v \in K
\end{align*} \quad (16) \]

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

First define the semigroup related to the fast motion by
\[ P_s^u \phi(v) = \mathbb{E} \left[ \phi(v_s^{u,v}) \right] \quad (17) \]
Lemma 4.1.

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

for every \( s \geq 0 \), \( u \in H, v_1, v_2 \in K \).

Proof. Define \( \rho_s = v_s^{u,v_1} - v_s^{u,v_2} \), then by (6) we have:

\[ \frac{d}{ds} |\rho_s|^2 = 2\langle A_2 \rho_s, \rho_s \rangle + 2\langle G(u, v_s^{u,v_1}) - G(u, v_s^{u,v_2}), \rho_s \rangle \leq -4\delta |\rho_s|^2 \]

for every \( s \geq 0 \), \( u \in H, v_1, v_2 \in K \). Then by taking the expectation and applying the comparison theorem we have the thesis. □

Moreover we have:

Lemma 4.2. There exists \( C > 0 \) such that:

\[ E \left( |v_s^{u_1,v} - v_s^{u_2,v}|^2 \right) \leq C|u_1 - u_2|^2 \]

for every \( s \leq T, u_1, u_2 \in H, v \in K \).

Proof. Define \( \rho_s = v_s^{u_1,v} - v_s^{u_2,v} \), then by Young’s inequality we have:

\[
\frac{d}{ds} |\rho_s|^2 = 2\langle A_2 \rho_s, \rho_s \rangle + 2\langle G(u_1, v_s^{u_1,v}) - G(u_2, v_s^{u_2,v}), \rho_s \rangle \\
\leq C (|u_1 - u_2| + |\rho_s|)|\rho_s| \\
\leq C|\rho_s|^2 + C|u_1 - u_2|^2
\]

for every \( s \leq T, u_1, u_2 \in H, v \in K \). Finally by taking the expectation and applying the comparison theorem we have the thesis. □

Next we can show:

Lemma 4.3. Let \( p \geq 1 \), then there exists \( C = C(p) > 0 \) such that:

\[ E \left( |v_s^{u,v}|^p \right) \leq C(1 + e^{-\delta p s}|v|^p + |u|^p) \]

for every \( s \geq 0 \), \( u \in H, v \in K \).

Proof. Define

\[ \tilde{r}_s^{Q_2} = \int_0^s e^{A_2(s-r)} dw_r^{Q_2} \]

First by Burkholder’s inequality and hypotheses 2 and 4 similarly to what done for lemma 3.2 we have:

\[
E \left( |\tilde{r}_s^{Q_2}|^p \right) \leq C \left( \int_0^s \left| e^{A_2(s-r)} Q_2^{1/2} \right|_{L_2}^2 \ dr \right)^{p/2} \\
\leq C \left( \int_0^s e^{-\Lambda(s-r)} \left( \frac{s-r}{2} \wedge 1 \right)^{-2\gamma} \ dr \right)^{p/2} \leq C \tag{18}
\]
for every \( s \geq 0 \).

Now set \( \rho_s = v_s u^v - i_s Q^2 \). Then for \( p \geq 2 \) we have:

\[
\frac{1}{p} \frac{d}{ds} |\rho_s|^p = \langle A_2 \rho_s, \rho_s \rangle |\rho_s|^{p-2} + \langle G(u, \rho_s + i_s Q^2) - G(u, i_s Q^2), \rho_s \rangle |\rho_s|^{p-2} + \langle G(u, i_s Q^2), \rho_s \rangle |\rho_s|^{p-2} \\
\leq -2\delta |\rho_s|^p + C(1 + |u| + |\tilde{i}_s Q^2|)|\rho_s|^{p-1} \\
\leq -\delta |\rho_s|^p + C(1 + |u|^p + |\tilde{i}_s Q^2|^p)
\]

Then by the comparison theorem and (18) we have:

\[
E|v_s^u u^v|^p \leq C E|\rho_s|^p + CE|i_s Q^2|^p \\
\leq C \left( e^{-\delta s} |v|^p + \int_0^s e^{-\delta t(s-r)} \left( 1 + |u|^p + E|i_r Q^2|^p \right) \, dr + E|i_s Q^2|^p \right) \\
\leq C \left( e^{-\delta s} |v|^p + |u|^p + 1 \right)
\]

Then by [11, theorem 6.3.3] there exists a unique invariant measure \( \mu^u \) for the semigroup \( P^u \).

Moreover we have:

**Lemma 4.4.** There exists \( C > 0 \) such that:

\[
\int_K |z| \mu^u(dz) \leq C(1 + |u|)
\]

for every \( u \in H \).

**Proof.** By definition of invariant measure and lemma 4.3 we have for every \( t > 0 \)

\[
\int_K |z| \mu^u(dz) = \int_K E \left[ |v_s^u u^v|^t \right] \mu^u(dz) \\
\leq C \left( 1 + |u| + e^{-\delta s} \int_K |z| \mu^u(dz) \right)
\]

By choosing \( s > 0 \) big enough we have the thesis. \( \square \)

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

**Lemma 4.5.** There exists \( C > 0 \) such that

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

for every \( s \geq 0, \ u \in H, \ v \in K, \ \phi \in \text{Lip}(K) \).

Moreover there exists \( C > 0 \) such that

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

for every \( s > 0, \ \phi \in B_2(K), \ u \in H, \ v \in K \).
Proof. First for every $\phi \in \text{Lip}(K)$ by lemma 4.1 we have:

$$
|P^u_s \phi(v_2) - P^u_s \phi(v_1)| \leq [\phi]_{\text{Lip}} E \left[ |v^u_s, v_2 - v^u_s, v_1| \right] \leq [\phi]_{\text{Lip}} e^{-2s} |v_2 - v_1| \quad (20)
$$

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

Now let $s > 0$, by definition of invariant measure, (20) and lemma 4.4 we have:

$$
\left| P^u_s \phi(v) - \int_K \phi(z) \mu^u(dz) \right| = \int_K (P^u_s \phi(v) - P^u_s \phi(z)) \mu^u(dz)
\leq [\phi]_{\text{Lip}} e^{-2s} \int_K |v - z| \mu^u(dz)
\leq C [\phi]_{\text{Lip}} e^{-2s} (1 + |u| + |v|)
$$

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

Now by application of [10, theorem 9.32] we have the Bismut-Elworthy formula:

$$
\sup_{u \in H} |DP^u_s \phi|_{\infty} \leq C(s \wedge 1)^{-1/2} |\phi|_{\infty} \quad (22)
$$

for every $s > 0$, $\phi \in B_{\text{B}}(K)$. In fact in our case the constant $K$ appearing in [10, hypothesis 9.1] is independent of $u$ and the function $\phi(x) = G(u, x)$ is Lipschitz uniformly in $u$. This is an important regularizing property for $P^u_s \phi(-)$ for $s > 0$.

Now by the semigroup property, the regularizing property of the semigroup (22) and by (20) we have:

$$
|P^u_s \phi(v_2) - P^u_s \phi(v_1)| = \left| p^u_{s/2} \left( p^u_{s/2} \phi \right)(v_2) - p^u_{s/2} \left( p^u_{s/2} \phi \right)(v_1) \right|
\leq \left| p^u_{s/2} \phi \right|_{\text{Lip}} e^{-\delta s} |v_2 - v_1|
\leq C |\phi|_{\infty} (s \wedge 1)^{-\frac{s}{2}} e^{-\delta s} |v_2 - v_1| \quad (23)
$$

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

Finally similarly to before by (23) for $s > 0$ we have:

$$
\left| P^u_s \phi(v) - \int_K \phi(z) \mu^u(dz) \right| = \left| \int_K (P^u_s \phi(v) - P^u_s \phi(z)) \mu^u(dz) \right|
\leq C |\phi|_{\infty} (s \wedge 1)^{-\frac{s}{2}} e^{-\delta s} \int_K |v - z| \mu^u(dz)
\leq C |\phi|_{\infty} (s \wedge 1)^{-\frac{s}{2}} e^{-\delta s} (1 + |u| + |v|)
$$

for every $u \in H, v \in K$, $\phi \in B_{\text{B}}(K)$. \hfill \Box

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

$$
\mathcal{H}^u_s(u, v) = \sigma(v^u_s, v, 0 \leq s \leq r \leq t)
$$

Then a classical consequence of lemmas 4.3, 4.5 is the following mixing result whose proof is the same of [7, lemma 3.2] for fixed $u$ and is reported in the appendix for completeness.
Moreover there exists $C > 0$ such that
\[
\sup \{ |P(B_1 \cap B_2) - P(B_1)P(B_2)| : B_1 \in \mathcal{H}_t^i(u, v), B_2 \in \mathcal{H}_{t+s}^i(u, v) \} \leq C e^{-\delta s} s^{-1/2}(1 + |u| + |v|)
\]
for every $0 \leq s \leq t, u \in H, v \in K$.

Now lemma 4.6 implies the following classical result, i.e. see [34] (see also [7, proposition 3.3]). The proof can be found in the appendix for completeness.

Lemma 4.7. There exists $C > 0$ such that for every $0 \leq s_1 \leq t_1 < s_2 \leq t_2$ and $\xi_i \mathcal{H}_{t_i}^i(u, v)$-measurable $i = 1, 2$ and $|\xi_i| \leq 1 a.s.$
\[
|E[\xi_1 \xi_2] - E[\xi_1]E[\xi_2]| \leq C e^{-\delta(s_2 - t_1)} (1 + |u| + |v|)
\]

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 report the proof in the appendix.

Lemma 4.8. 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 $\xi_i \mathcal{H}_{t_i}^i(u, v)$-measurable, $i = 1, 2$ satisfying for some $K_i = K_i(\rho) > 0$
\[
\left( E[|\xi_i|^\frac{\rho}{2}] \right)^{\frac{2}{\rho}} = K_i < \infty
\]

then:
\[
|E[\xi_1 \xi_2] - E[\xi_1]E[\xi_2]| \leq C K_1^{\frac{\rho}{2}} K_2^{\frac{\rho}{2}} (K_1 + K_2)^{\frac{\rho}{2}} e^{-\delta(s_2 - t_1)} (1 + |u| + |v|)^{\frac{\rho}{2}}
\]

5 Averaged equation

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

\[
\tilde{F}(u) = \int_K F(u, v) \mu^u(dv)
\]

for every $u \in H$ and consider the so called averaged equation:
\[
\begin{align*}
&dU_t = A_1 U_t + \tilde{F}(U_t)dt + dW_{t}^{Q_1} \\
&U_0 = u
\end{align*}
\]

for every $t \leq T$.

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

Proposition 2. Equation (26) admits a unique mild solution given by:
\[
U_t = e^{A_1 t} u + \int_0^t e^{A_1(t-s)} \tilde{F}(U_s)ds + \int_0^t e^{A_1(t-s)} dW_{s}^{Q_1}
\]

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

Moreover there exists $C = C(T) > 0$ such that
\[
E \left[ \sup_{t \in [0, T]} |U_t|^2 \right] \leq C(1 + |u|^2)
\]
Proof. In order to prove the first part of the proposition it is sufficient to prove that $\overline{F}$ is Lipschitz, i.e. [10]. To this purpose let $h, u \in H$ fixed and consider

$$\phi_{u,h}(v) = \langle F(u, v), h \rangle$$

for every $v \in K$.

Note that $\phi_{u,h}$ is Lipschitz with Lipschitz constant $L_F[h]$.

By lemmas 4.2, 4.5 and by the Lipschitzianity of $F$ we have

$$\left| \langle \overline{F}(u_1) - \overline{F}(u_2), h \rangle \right| \leq \int_K \left| \langle F(u_1, v), h \rangle \mu^{u_1} - \mathbb{E} \left[ \langle F(u_1, v^{u_1,0}), h \rangle \right] \right| dv$$

$$+ \left| \mathbb{E} \left[ \langle F(u_2, v^{u_1,0}), h \rangle \right] - \int_K \langle F(u_2, v), h \rangle \mu^{u_2} dv \right|$$

$$+ \left| \mathbb{E} \left[ \langle F(u_1, v^{u_1,0}) - F(u_2, v^{u_1,0}), h \rangle \right] \right|$$

$$\leq C L_F e^{-2s} (1 + |u|) + C|u_1 - u_2||h|$$

for every $u_1, u_2 \in H$.

Now letting $s \to +\infty$ we obtain the Lipschitzianity of $\overline{F}$.

The proof of the second part of the proposition follows by a standard application of Gronwall lemma.

\[\square\]

6 Preliminary results

In this section we prove a technical result, i.e. lemma 6.2, which is inspired by [7, lemma 4.2] and follows by the mixing properties of the fast motion studied in section 4, in particular lemma 4.8. We remark that the equation for the slow component in [7] has no stochastic perturbation and the equation for the fast component is uncoupled from the slow-component. Nevertheless in lemma 6.2 we can fix the slow-component $u$ so that to show it we can proceed in a similar way to the proof of [7, lemma 4.2]. This is satisfactory because we will use lemma 6.2 in section 7 where we use a time discretization argument inspired by Khasminskii where the slow motion is frozen.

Fix $n \in \mathbb{N}$, $T > 0$ and define for every $0 < r < T/\epsilon$

$$\Phi(r) = F(u, v^{r,u}) - \overline{F}(u)$$

for every $u \in H, v \in K$.

Moreover set:

$$J_j(r_1, \ldots, r_{2n}) = \mathbb{E} \left[ \prod_{i=1}^{2n} \Phi(r_i) - \mathbb{E} \left[ \prod_{i=1}^{2j} \Phi(r_i) \mathbb{E} \left[ \prod_{i=2j+1}^{2n} \Phi(r_i) \right] \right] \right]$$

(27)

for every $1 \leq j \leq n$, $0 \leq r_1 \leq \ldots \leq r_{2n} \leq T/\epsilon$. First we show the following result:

Lemma 6.1. Let $0 < \rho < 1$ then there exists $C = C(T, \rho, n) > 0$ and $\overline{\eta} = \overline{\eta}(\rho, n) > 0$ such that

$$\left| \mathbb{E} \left[ \prod_{i=j_1}^{j_2} \phi(r_i) \right] \right| \leq C \left( \frac{e^{-\delta(r_{j_2} - r_{j_1} - 1)}}{\sqrt{j_2 - j_1}} \right)^{\overline{\eta}} (1 + |u|^\overline{\eta} + |v|^\overline{\eta})$$

(28)

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

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

$$J_j(r_1, \ldots, r_{2n}) \leq C \left( \frac{e^{-\delta r_1}}{\sqrt{r_1}} \right)^{\overline{\eta}} (1 + |u|^\overline{\eta} + |v|^\overline{\eta})$$

(29)
where
\[ F_j = \max (r_{2n} - r_{2n-1}, r_{2j+1} - r_{2j}) \]
for every \( 1 \leq j \leq n, 0 \leq r_1 \leq ... \leq r_{2n} \leq T / \epsilon \).

Proof. By the sublinearity of \( \Phi \) and lemma 4.3 for every \( p \geq 1 \) we have:

\[
E \left| \prod_{i=j_1}^{j_2} \Phi(r_i) \right|^p \leq C \left( 1 + \sum_{i=j_1}^{j_2} |u|^{(j_2-j_1+1)p} + \sum_{i=j_1}^{j_2} E \left| v_{r_i}^{(j_2-j_1+1)p} \right| \right)
\]

(30)

for every \( 1 \leq j_1 \leq j_2 \leq 2n \).

Then by setting \( p = 2/(1 - \rho) \) and applying lemma 4.8 to \( \xi_1 = \prod_{i=1}^{j_1} \Phi(r_i) \), \( \xi_2 = \prod_{i=j_1+1}^{2n} \Phi(r_i) \) with \( K_1 = C(1 + |u| + |v|^j), K_2 = C(1 + |u|^{2n-j} + |v|^{2n-j}) \) for every \( j < 2n \) we have:

\[
\left| E \prod_{i=1}^{2n} \Phi(r_i) - E \prod_{i=1}^{j_1} \Phi(r_i) E \prod_{i=j_1+1}^{2n} \Phi(r_i) \right| \leq C (1 + |u|^n + |v|^n) \left( \frac{e^{-\delta (r_{j+1} - r_j)}}{\sqrt{r_{j+1} - r_j}} \right)^{\frac{p}{\rho p}}
\]

(31)

where \( \eta = 2n + \frac{2(2n-1)}{(p+2)}. \)

Moreover by definition of \( \Phi \) and lemma 4.5 we have:

\[
|E \Phi (r_{j_2})| \leq Ce^{-2\delta r_{j_2}} (1 + |u| + |v|)
\]

(32)

Now by the last three inequalities we have:

\[
\left| E \prod_{i=j_1}^{j_2} \Phi(r_i) \right| \leq E \prod_{i=j_1}^{j_2} \Phi(r_i) - E \prod_{i=j_1}^{j_2-1} \Phi(r_i) E \Phi(r_{j_2}) + E \prod_{i=j_1}^{j_2-1} \Phi(r_i) E \Phi(r_{j_2})
\]

\[
\leq C \left( 1 + |u|^n + |v|^n \right) \left( \frac{e^{-\delta (r_{j_2} - r_{j_2-1})}}{\sqrt{r_{j_2} - r_{j_2-1}}} \right)^{\frac{p}{\rho p}}
\]

where \( \eta = (j_2 - j_1 + 1) + \frac{2(j_2-j_1)}{p+2} \). This implies (28).

Now by (28) we have:

\[
\left| E \prod_{i=1}^{2n} \Phi(r_i) \right| \leq C (1 + |u|^n + |v|^n) \left( \frac{e^{-\delta (r_{2n} - r_{2n-1})}}{\sqrt{r_{2n} - r_{2n-1}}} \right)^{\frac{p}{\rho p}}
\]

(33)

By the last inequality and (30) we have:

\[
\left| E \prod_{i=1}^{2n} \Phi(r_i) - E \prod_{i=1}^{j_1} \Phi(r_i) E \prod_{i=j_1+1}^{2n} \Phi(r_i) \right| \leq C (1 + |u|^n + |v|^n) \left( \frac{e^{-\delta (r_{2n} - r_{2n-1})}}{\sqrt{r_{2n} - r_{2n-1}}} \right)^{\frac{p}{\rho p}}
\]

(34)

for every \( j < 2n. \)

Now by (31) and (34) we have (29) since the function \( f(s) = e^{-\delta s} s^{-1/2} \) is decreasing. \( \square \)
Let $\epsilon > 0$, $\alpha > 0$, $0 \leq \beta < 1/3$, $u \in H$, $v \in K$, $h \in H$ with $|h| = 1$. Define

$$
\Psi_h(r) = \langle F(u, v^u) - \overline{F(u)}, h \rangle
$$

and $\theta_{\alpha, \beta}(r) = e^{-r^\alpha r^{-\beta}}$ for every $r > 0$.

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

**Lemma 6.2.** Let $n \in \mathbb{N}$ and $0 \leq \beta < 1/3$. Then there exists a constant $C = C(\Gamma, n, \beta) > 0$ 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:

$$
\left| \mathbb{E} \int_{[s, t]^2} \prod_{i=1}^{2n} \theta_{\alpha, \beta}(\Psi_h(r_i/\epsilon)) dr_1 \cdots dr_{2n} \right| \leq C (1 + |u|^n + |v|^n) \epsilon^n \left( \frac{1}{\alpha} \right)^{(1-2\beta)n}
$$

*Proof.* First by a change of variable we have:

$$
\mathbb{E} \int_{[s, t]^2} \prod_{i=1}^{2n} \left( \theta_{\alpha, \beta} \left( t - r_i \right) \Psi_h \left( r_i/\epsilon \right) \right) dr_1 \cdots dr_{2n} = \epsilon^{2n} \mathbb{E} H_\epsilon(s, t) \tag{35}
$$

where we have defined:

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

By symmetry we have:

$$
H_\epsilon(s, t) = C \int_{\left[ \frac{s}{\epsilon}, \frac{t}{\epsilon} \right]^2} \prod_{i=1}^{2n} \left( \theta_{\alpha, \beta} \left( t - cr_i \right) \Psi_h \left( r_i \right) \right) dr_1 \cdots dr_{2n} \tag{36}
$$

We proceed by induction on $n$ and to this purpose fix some $\rho \in (0, 1)$. Consider $n = 1$ then by (29) and some changes of variables we have

$$
e^2 |\mathbb{E} H_\epsilon(s, t)| = 2e^2 \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \theta_{\alpha, \beta} \left( t - cr_1 \right) \theta_{\alpha, \beta} \left( t - cr_2 \right) \mathbb{E} \left[ \Psi_h \left( r_1 \right) \Psi_h \left( r_2 \right) \right] dr_1 dr_2
$$

$$
\leq C e^2 \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \theta_{\alpha, \beta} \left( t - cr_1 \right) \theta_{\alpha, \beta} \left( t - cr_2 \right) \left( \frac{e^{-\delta (r_2 - r_1)}}{\sqrt{r_2 - r_1}} \right) dr_1 dr_2 (1 + |u|^n + |v|^n)
$$

$$
= C e^{-2\beta} \int_{0}^{r_2} \int_{0}^{r_2} e^{-\epsilon r_2 \alpha} \int_{\frac{s}{\epsilon}}^{\frac{t}{\epsilon}} \left( \frac{e^{-\delta (r_2 - r_1)}}{\sqrt{r_2 - r_1}} \right) dr_1 dr_2 (1 + |u|^n + |v|^n)
$$

$$
\leq C e^{-2\beta} \int_{0}^{r_2} \int_{0}^{r_2} e^{-2\epsilon r_2 \alpha} \int_{0}^{r_2} \left( \frac{e^{-\delta r_1}}{\sqrt{r_1}} \right) dr_1 dr_2 (1 + |u|^n + |v|^n)
$$

$$
\leq C e^{-2\beta} \int_{0}^{r_2} e^{-2\epsilon r \alpha} dr (1 + |u|^n + |v|^n)
$$

$$
\leq C e^{-2\beta} (\epsilon \alpha)^{-(1-2\beta)} \int_{0}^{\alpha(t-s)} e^{-2r} dr (1 + |u|^n + |v|^n)
$$

where $\alpha = \alpha(t-s)$.
so that by (35) we have the thesis for \( n = 1 \).

Now assume that

\[
\left| \sum_{j=1}^{n-1} v^j \right| \leq C (1 + |u|^n + |v|^n) e^{j/2} \left( \frac{1}{\alpha} \right)^{(1-\beta)j/2}
\]

for every even \( j < 2n \) where \( \eta_1 = j + \frac{\rho(j-1)}{(\rho+2)j} \).

We prove that then it holds for \( j = 2n \).

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

\[
\max_{j=1, \ldots, n-1} (r_{2j} - r_{2j+1}) = r_{2j(r+1)} - r_{2j(r)}
\]

Now consider \( J_j(r_1, \ldots, r_{2n}) \) given by (27) with \( \Phi = \Psi_t \), then by definition of \( j(r) \) we have:

\[
e^{2n} |E H_\varepsilon(s, t)| \leq C e^{2n} \left( \sum_{j=1}^{n-1} v^j \right) \left( \sum_{i=1}^{2n} \max_{r_{2j} \leq \rho < r_{2j+1}} (r_{2j} - \rho) \right) e^{-\delta j \varepsilon^2} \leq C \left( 1 + |u|^n + |v|^n \right) e^{\frac{\delta}{\bar{\rho}}} \left( \frac{1}{\alpha} \right)^{(1-\beta)j/2}
\]

Now by (29) and by definition of \( j(r) \) we have:

\[
J_j(r_1, \ldots, r_{2n}) \leq C \left( 1 + |u|^n + |v|^n \right) e^{-\delta \varepsilon^2} \left( \frac{1}{\alpha} \right)^{(1-\beta)j/2}
\]

where \( \delta_n = \frac{\delta \varepsilon^2}{n(n-1)\rho}, \bar{\rho} = \frac{\rho}{2(n-2)\rho} \).

We can apply this last inequality in order to estimate \( I_{1,\varepsilon}(s, t) \), i.e.
Now consider $k \geq 1$. Applying now (40) and (41) to the inequality for $\|u\|^n + |v|^n$, we have:

$$I_{1,\epsilon}(s, t) \leq C(1 + |u|^n + |v|^n) \int_0^{2n} \int_0^{2n} \cdots \int_0^{2n} e^{-\alpha \delta_n (r_{2n} - r_{2n-2})} \prod_{i=1}^{2n} (r_{2n-i} - r_{2n-2})^\beta \times \prod_{i=1}^{2n} \theta_{\alpha, \beta} (t - \epsilon r_i) \int_0^{r_{2n-i}} e^{-\alpha \delta_n (r_{2n-i} - r_{2n})} dr_1 \cdots dr_{2n} \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i}) r_{i-1}^\beta} dr_i \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

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

$$I_{1,\epsilon}(s, t) \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

(38)

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

$$I_{1,\epsilon}(s, t) \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

(39)

Now by (38) and (39) we have:

$$I_{1,\epsilon}(s, t) \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

(40)

Now consider the remaining term in the inequality for $I_{1,\epsilon}(s, t)$: by (38) and (39) we have

$$I_{1,\epsilon}(s, t) \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

(41)

Applying now (40) and (41) to the inequality for $I_{1,\epsilon}(s, t)$ we have:

$$I_{1,\epsilon}(s, t) \leq C \epsilon^{k \beta - 1} \prod_{i=1}^{2n} e^{-\alpha \delta_n (r_{2n-i} + 1)} (1 + |u|^n + |v|^n)
$$

(42)
Then by the inductive hypothesis we have:

$$I_{2,ε} \leq C_ε e^{-\bar{α}(1-2β)n} (1 + |u|^n + |v|^n)$$  \hspace{1cm} (43)

Finally applying the last two inequalities to (37) and then going back to (35) we have the thesis. \hfill \Box

### 7 Order of convergence

In this section we finally investigate the order of convergence. We use a time-discretization inspired by Khasminskii where the slow motion is frozen.

Consider a discretization parameter $Δ = Δ(ε) > 0$. Now let $k = 0\ldots\lfloor T/Δ \rfloor - 1$ and consider the following discretized processes $U^ε, V^ε$ defined by

$$\begin{align*}
dU^ε_t &= A_1 U^ε_t + f(U^ε_{kΔ}, V^ε_t)dt + dW^Q_t \\
U^ε_0 &= u \\
dV^ε_t &= ε^{-1}A_2 V^ε_t + ε^{-1}G(U^ε_{kΔ}, V^ε_t)dt + ε^{-1/2}dW^Q_t \\
V^ε_{kΔ} &= V^ε_0
\end{align*}$$

for every $t \in [kΔ, (k+1)Δ]$.

Note that the dynamics of $U^ε_t$ can be written with the compact notation on $[0, T]$ as

$$dU^ε_t = A_1 U^ε_t + f(U^ε_{t/Δ}; V^ε_t)dt + dW^Q_t$$

In the next two lemmas we study the difference between the discretized processes and the original ones:

**Lemma 7.1.** Let $0 < α < 1/2 - γ$ and let $u \in D(-A^α T), v \in K$, then there exists $C = C(T, α) > 0$ such that

$$\mathbf{E} |V^ε_t - V^ε_t|^2 \leq CΔ^{2α}(1 + |u|^2 + |v|^2)$$

for every $ε > 0$, $u \in D(-A^α T), v \in K$, $t \in [0, T]$.

**Proof.** Fix $t \in [kΔ, (k+1)Δ)$, then (6), lemma 3.5 and Young’s inequality we have:

$$\begin{align*}
\frac{1}{2} \frac{d}{dt} \mathbf{E} |V^ε_t - V^ε_t|^2 &\leq \frac{2ε}{ε} \mathbf{E} |V^ε_t - V^ε_t|^2 + \frac{C}{ε} \mathbf{E} [||U^ε_{kΔ} - U^ε_t||^2] \\
&\leq \frac{2ε}{ε} \mathbf{E} |V^ε_t - V^ε_t|^2 + \frac{C}{ε} Δ^α (\mathbf{E} |V^ε_t - V^ε_t|^2)^{1/2} (1 + |u|^2 + |v|) \\
&\leq \frac{2}{ε} \mathbf{E} |V^ε_t - V^ε_t|^2 + \frac{C}{ε} Δ^{2α}(1 + |u|^2 + |v|^2)
\end{align*}$$

where $C > 0$ is some constant independent of $ε, k$.

Now by Gronwall’s lemma we have:

$$\mathbf{E} |V^ε_t - V^ε_t|^2 \leq CΔ^{2α} \int_0^t \frac{1}{ε} e^{-2δ(t-s)/ε} ds (1 + |u|^2 + |v|^2) \leq CΔ^{2α}(1 + |u|^2 + |v|^2)$$

where $C > 0$ is some constant independent of $ε, k$. \hfill \Box
Lemma 7.2. Let $0 < \alpha < 1/2 - \gamma$ and let $u \in \mathcal{D}(-A^n_0)$, $v \in \mathcal{K}$, then there exists $C = C(T, \alpha) > 0$ such that

$$E \left[ \sup_{t \in [0,T]} |\bar{U}^\varepsilon_t - U^\varepsilon_t|^2 \right] \leq C\Delta^{2\alpha}(1 + |u|^2 + |v|^2)$$

for every $\varepsilon > 0$, $u \in \mathcal{D}(-A^n_0)$, $v \in \mathcal{K}$.

Proof. By the Lipschitzianity of $F$ and lemmas 3.5, 7.1 we have:

$$E \left[ \sup_{t \in [0,T]} |\bar{U}^\varepsilon_t - U^\varepsilon_t|^2 \right] = E \left[ \sup_{t \in [0,T]} \left| \int_0^t e^{A_1(t-s)} \left( \bar{F}(U^\varepsilon_{[s]}, \bar{V}^\varepsilon_s) - \bar{F}(U^\varepsilon_s, V^\varepsilon_s) \right) ds \right|^2 \right] \leq C \left\| \bar{F}(U^\varepsilon_{[s]}) - \bar{F}(U^\varepsilon_s, V^\varepsilon_s) \right\| ds \leq C\Delta^{2\alpha}(1 + |u|^2 + |v|^2)$$

We now prove proposition 3 which will be crucial in the derivation of the order of convergence.

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

$$E \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_1(t-s)} \left( \bar{F}(U^\varepsilon_{[s]}) - \bar{F}(U^\varepsilon_s, V^\varepsilon_s) \right) ds \right|^2 \right] \leq C\xi(1 + |u|^2 + |v|^2)$$

for every $0 \leq t_0 \leq \tau \leq T$, $\varepsilon > 0$, $u \in \mathcal{H}, v \in \mathcal{K}$.

Proof. We proceed with similar techniques to the ones in the proof of [7, Theorem 4.1]. Fix $0 \leq t_0 \leq T$, consider (16) with

$$W^Q_{s,t_0} = \left( W^Q_{s,t_0} - W^Q_{t_0} \right) / \sqrt{\varepsilon}$$

and define

$$v^{u,v,c}_{s,t_0} = v^{u,v,c}_{s/\varepsilon}$$

for every $s \geq 0$ and $u \in \mathcal{H}, v \in \mathcal{K}$.

Note that the process $(v^{u,v,c}_{s,t_0})_{t \geq t_0}$ solves the following stochastic differential equation

$$\begin{cases}
\frac{dv^{u,v,c}_{s,t_0}}{dr} = \frac{1}{2} A_2 v^{u,v,c}_{s,t_0} dr + \frac{1}{2} G(u, v^{u,v,c}_{s,t_0}) dr + \frac{1}{\sqrt{\varepsilon}} dW^Q_{s,t_0} \\
v^{u,v,c}_{s,t_0} = v \in \mathcal{H}
\end{cases}$$

for every $r \geq t_0$.

Fix $\mu > 0$, $n \geq 2$ integer and $1/(2n) < \beta < 1/3$ as in hypothesis 1 and define

$$Y^c_s = \int_{t_0}^s (s-r)^{-\beta} e^{A_1(s-r)} \left( F(u, v^{u,v,c}_{r}) - \bar{F}(u) \right) dr$$

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

$$E |Y^c_s|^{2n} \leq C\xi^n (1 + |u|^n + |v|^n)$$

(46)
for every $0 \leq t_0 \leq s \leq T$, $\varepsilon > 0$, $u \in H, v \in K$.
Indeed first recall the spectral representation
\[ e^{A_1 t} x = \sum_{j=1}^{\infty} e^{-\alpha_j t} (x, e_j) e_j \]

Then by Parseval’s identity, Hölder’s inequality and hypothesis 1 we have:
\[
\mathbb{E} \left| \mathbf{Y}_s^\varepsilon \right|^{2n} = \mathbb{E} \left( \sum_{k=1}^{\infty} \alpha_k^{(n-1)\mu - n(1-2\beta)} \alpha_k \left( \int_{t_0}^{s} (s-r)^{-\beta} e^{-(s-r)\alpha_k} (\mathbf{F}(u, v_r^{u,v}) - \mathbf{F}(u, e_k)) dr \right)^2 \right)^n \\
\leq \left( \sum_{k=1}^{\infty} \alpha_k^{-u} \right)^{n-1} \sum_{k=1}^{\infty} \alpha_k^{(n-1)\mu} E \left[ \int_{t_0}^{s} (s-r)^{-\beta} e^{-(s-r)\alpha_k} (\mathbf{F}(u, v_r^{u,v}) - \mathbf{F}(u, e_k)) dr \right]^{2n} \\
\leq C \sum_{k=1}^{\infty} \alpha_k^{(n-1)\mu} \int_{t_0}^{s} \int_{t_0}^{s} (s-r_i)^{-\beta} e^{-(s-r_i)\alpha_k} (\mathbf{F}(u, v_r^{u,v}) - \mathbf{F}(u, e_k)) dr_1 \cdots dr_{2n} 
\]

Now recall (45) and apply lemma 6.2 with $h = e_k, \alpha = \alpha_k$, then:
\[ \mathbb{E} \left| \mathbf{Y}_s^\varepsilon \right|^{2n} \leq C \varepsilon^n (1 + |u|^n + |v|^n) \sum_{k=1}^{\infty} \alpha_k^{(n-1)\mu - n(1-2\beta)} \\
= C \varepsilon^n (1 + |u|^n + |v|^n) \sum_{k=1}^{\infty} \alpha_k^{n(\mu + 2\beta - 1) - \mu} 
\]

The series on the right-hand-side is convergent by hypothesis 1 and we have (46).
Now define
\[ Z^\varepsilon := \int_{t_0}^{t} e^{A_1 (t-s)} \left( \mathbf{F}(\mathbf{U}_{[s/(\Delta \alpha)]}^\varepsilon, \mathbf{V}_s^\varepsilon) - \mathbf{F}(\mathbf{U}_{[s/(\Delta \alpha)]}^\varepsilon) \right) ds \]
We proceed using the factorization method:
\[ Z^\varepsilon = \sum_{k=0}^{\lfloor t/\Delta \rfloor - 1} \int_{k \Delta}^{(k+1) \Delta} e^{A_1 (t-s)} \left( \mathbf{F}(\mathbf{U}_{k \Delta}^\varepsilon, \mathbf{V}_s^\varepsilon) - \mathbf{F}(\mathbf{U}_{k \Delta}^\varepsilon) \right) ds \]
\[ = \frac{\sin \beta \pi}{\pi} \sum_{k=0}^{\lfloor t/\Delta \rfloor - 1} \int_{k \Delta}^{(k+1) \Delta} ((k+1) \Delta - s)^{\beta-1} e^{A_1 (t-s)} Y_{s,k}^\varepsilon ds \]
where
\[ Y_{s,k}^\varepsilon = \int_{k \Delta}^{s} (s-r)^{-\beta} e^{A_1 (s-r)} \left( \mathbf{F}(\mathbf{U}_{k \Delta}^\varepsilon, \mathbf{V}_r^\varepsilon) - \mathbf{F}(\mathbf{U}_{k \Delta}^\varepsilon) \right) dr \]
Notice that $\mathbf{V}_s^\varepsilon$ and $\mathbf{U}_{k \Delta}^\varepsilon, \mathbf{V}_{k \Delta}^\varepsilon$ defined by (45) with $t_0 = K \Delta$ are indistinguishable on $[K \Delta, (K+1) \Delta)$.
Then by Freezing lemma, (46) and lemma 3.3 we have:
\[
\mathbb{E} \left| Y_{s,k}^\varepsilon \right|^{2n} = \mathbb{E} \left[ \mathbb{E} \left[ \left| Y_{s,k}^\varepsilon \right|^{2n} | \mathbf{U}_{k \Delta}, \mathbf{V}_{k \Delta}^\varepsilon \right] \right] \\
\leq C \varepsilon^n \left( 1 + \mathbb{E} \left| \mathbf{U}_{k \Delta}^\varepsilon \right|^{n} + \mathbb{E} \left| \mathbf{V}_{k \Delta}^\varepsilon \right|^{n} \right) \\
\leq C \varepsilon^n \left( 1 + |u|^n + |v|^n \right) 
\]
Proof. for every \( \varepsilon > 0 \), we have:

\[
E \left( \sup_{t \in [0,T]} |Z_t^\varepsilon|^{2n} \right) \leq C \sum_{k=0}^{[T/\Delta]} \int_{k\Delta}^{(k+1)\Delta} E|Y_s^\varepsilon|^{2n} \, ds \leq C \varepsilon^n (1 + |u|^n + |v|^n)
\]

Finally by Holder’s inequality we have the thesis:

\[
E \left( \sup_{t \in [0,T]} |Z_t^\varepsilon|^2 \right) \leq \left( E \left( \sup_{t \in [0,T]} |Z_t^\varepsilon|^{2n} \right) \right)^{1/n} \leq C \varepsilon (1 + |u|^\xi + |v|\xi)
\]

for \( \xi = \eta/n \).

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

**Theorem 1.** Let \( 0 < \alpha < 1/2 - \gamma \), \( u \in D(-A_{\alpha}^n) \), \( v \in K \) and assume hypotheses 1, 2, 3, 4, 5. Then there exists \( C = C(T, \alpha, |u|, |v|) > 0 \) such that

\[
E \left( \sup_{t \in [0,T]} |U_t^\varepsilon - U_t|^2 \right) \leq C \varepsilon
\]

for every \( \varepsilon > 0 \).

**Proof.** For \( t \in [0,T] \) we have:

\[
\tilde{U}_t^\varepsilon - U_t = \int_0^t e^{A_{\alpha}(t-s)} \left( f(U_{[s/\Delta]}^{\varepsilon} \tilde{V}_t^\varepsilon) - \overline{f}(U_s) \right) \, ds
\]

so that we have:

\[
|\tilde{U}_t^\varepsilon - U_t| \leq \left| \int_0^t e^{A_{\alpha}(t-s)} \left( f(U_{[s/\Delta]}^{\varepsilon} \tilde{V}_t^\varepsilon) - \overline{f}(U_{[s/\Delta]}^{\varepsilon}) \right) ds \right| + \left| \int_0^t e^{A_{\alpha}(t-s)} \left( \overline{f}(U_{[s/\Delta]}^{\varepsilon}) - \overline{f}(U_s) \right) ds \right|
\]

Now let \( 0 \leq \tau \leq T \) and compute:

\[
E \left( \sup_{0 \leq t \leq \tau} |\tilde{U}_t^\varepsilon - U_t|^2 \right) \leq CE \left( \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_{\alpha}(t-s)} \left( f(U_{[s/\Delta]}^{\varepsilon} \tilde{V}_t^\varepsilon) - \overline{f}(U_{[s/\Delta]}^{\varepsilon}) \right) ds \right|^2 \right)
\]

\[
+ CE \left( \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_{\alpha}(t-s)} \left( \overline{f}(U_{[s/\Delta]}^{\varepsilon}) - \overline{f}(U_s) \right) ds \right|^2 \right)
\]

For the first term on the right-hand-side by lemma 3 we have:

\[
E \left( \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_{\alpha}(t-s)} \left( f(U_{[s/\Delta]}^{\varepsilon} \tilde{V}_t^\varepsilon) - \overline{f}(U_{[s/\Delta]}^{\varepsilon}) \right) ds \right|^2 \right) \leq C \varepsilon
\]

23
For the second term on the right-hand-side by the Lipschitzianity of $\bar{F}$ and lemma 3.5 we have:

$$E \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_s(t-s)} \left( \bar{F}(U_s^{\epsilon}) - \bar{F}(U_s^{c}) \right) ds \right|^2 \right] \leq C \int_0^\tau E \left[ \left| \bar{F}(U_{t}^{\epsilon}) - \bar{F}(U_{t}^{c}) \right|^2 ds \right]$$

$$\leq C \int_0^\tau E \left[ \left| U_{t}^{\epsilon} - U_{t}^{c} \right|^2 ds \right]$$

$$\leq C A^{2\alpha}$$

For the third term on the right-hand-side by lemma 7.2 we have:

$$E \left[ \sup_{0 \leq t \leq \tau} \left| \int_0^t e^{A_s(t-s)} \left( \bar{F}(U_s^{c}) - \bar{F}(U_s) \right) ds \right|^2 \right] \leq C \int_0^\tau E \left[ \left| U_{t}^{c} - U_{t} \right|^2 ds \right]$$

$$\leq C \left( A^{2\alpha} + \int_0^\tau E \left| \bar{U}_{t}^{c} - U_{t} \right|^2 ds \right)$$

Putting everything together we have:

$$E \left[ \sup_{0 \leq t \leq \tau} \left| \bar{U}_{t}^{c} - U_{t} \right|^2 \right] \leq C \left( \epsilon + A^{2\alpha} + \int_0^\tau E \left| \bar{U}_{t}^{c} - U_{t} \right|^2 ds \right)$$

(47)

for every $\tau \leq T$.

Then by Gronwall lemma we have:

$$E \left[ \sup_{0 \leq t \leq \tau} \left| \bar{U}_{t}^{c} - U_{t} \right|^2 \right] \leq C(\epsilon + A^{2\alpha})$$

Finally by this and lemma 7.2 we have:

$$E \left[ \sup_{0 \leq t \leq T} \left| U_{t}^{\epsilon} - U_{t} \right|^2 \right] \leq 2E \left[ \sup_{0 \leq t \leq T} \left| U_{t}^{\epsilon} - \bar{U}_{t}^{c} \right|^2 \right] + 2E \left[ \sup_{0 \leq t \leq T} \left| \bar{U}_{t}^{c} - U_{t} \right|^2 \right]$$

$$\leq C(\epsilon + A^{2\alpha})$$

so that by choosing $\Delta = \Delta(\epsilon) = \epsilon^{1/(2\alpha)}$, i.e. $\Delta^{2\alpha} = \epsilon$ we have the thesis of the theorem. \hfill $\square$

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:

$$\begin{align*}
\frac{du}{dt}(t, \xi) &= \frac{\partial^2}{\partial \xi^2} u_\varepsilon(t, \xi) + f(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi)) + \dot{w}_1(t, \xi) \\
\frac{dv}{dt}(t, \xi) &= \frac{1}{\varepsilon} \left( \frac{\partial^2}{\partial \xi^2} - \lambda \right) v_\varepsilon(t, \xi) + g(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi)) + \frac{1}{\varepsilon} \dot{w}_2(t, \xi) \\
u_\varepsilon(0, \xi) &= u(\xi), \quad v_\varepsilon(0, \xi) = v(\xi), \quad \xi \in [0, L] \\
u_\varepsilon(t, 0) &= \nu_\varepsilon(t, L) = 0, \quad t \geq 0, \quad \xi \in [0, L]
\end{align*}$$

where

- $t \in [0, T], \xi \in [0, L]$
- $\varepsilon \in (0, 1]$ is a small parameter representing the ratio of time-scales between the two variables of the system $u_\varepsilon$ and $v_\varepsilon$. 

24
\begin{itemize}
  \item $u_c$ and $v_c$ are the slow and fast components respectively,
  \item $u, v \in H = L^2[0, T]$ are the initial conditions,
  \item $\lambda > 0$,
  \item $f, g : [0, L] \times \mathbb{R}^2 \to \mathbb{R}$ are Lipschitz functions uniformly wrt $\xi$, with Lipschitz constants $L_f, L_G$ respectively and $L_G < \lambda$,
  \item $\mathcal{W}_1, \mathcal{W}_2$ are independent white noises both in time and space.
\end{itemize}

Then it is well known \cite{6} that (48) can be rewritten in the abstract form (2) where $H = K = L^2[0, T]$, $F, G : H \times H \to H$ are the Nemytskii operators of $f, g$ respectively, i.e.

$$
F(x, y)(\xi) = f(\xi, x(\xi), y(\xi)) \quad G(x, y)(\xi) = g(\xi, x(\xi), y(\xi))
$$

In this setting the hypotheses of theorem 1 are satisfied (recall remark 2.1) so that the result can be applied.

\section{Proof of lemma 4.6}

\begin{proof}
First observe that

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

where

$$
C_s^t(u, v) = \{v_{r_1,1}^u \in A_1, ..., v_{r_k,1}^u \in A_k : k \in \mathbb{N}, s \leq r_1 < ... < r_k \leq t, A_1, ..., A_k \in \mathcal{B}(K)\}
$$

\text{(49)}

is the family of cylindrical sets.

Consider $B_1 \in C_0^1(u, v)$ and $B_2 \in C_{s+t}^0(u, v)$, i.e.

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

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

First by the tower property we have:

$$
P(B_1 \cap B_2) = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}}(v_{r_{1,i}}^u) \prod_{i=1}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^u) \right] = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}}(v_{r_{1,i}}^u) \mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^u) \bigg| F_t \right] \right]
$$

\text{(50)}

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

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

$$
\mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^u) \bigg| F_t \right] = \mathbb{E} \left[ 1_{A_{2,1}}(v_{r_{2,1}}^u) \mathbb{E} \left[ \prod_{i=2}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^u) \bigg| F_{r_{2,1}} \right] \bigg| F_t \right] = \mathbb{E} \left[ 1_{A_{2,1}}(v_{r_{2,1}}^u) \mathbb{E} \left[ \prod_{i=3}^{k_2} 1_{A_{2,i}}(v_{r_{2,i}}^u) \bigg| F_{r_{2,1}} \bigg| F_{r_{2,2}} \right] \bigg| F_t \right]
$$

25
and then by iteration we have:

\[
\mathbb{E} \left[ \prod_{i=1}^{k_2} 1_{A_{2,i}} \left( v_{r_2,i}^{u,v} \right) \right] \mathcal{F}_t = \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^{u,v} \right) \mathbb{E} \left[ 1_{A_{2,2}} \left( v_{r_{2,2}}^{u,v} \right) \cdots \mathbb{E} \left[ 1_{A_{2,k_2}} \left( v_{r_{2,k_2}}^{u,v} \right) \right] \right] \right] \mathcal{F}_{r_{2,k_2}} \cdots \mathcal{F}_{r_{2,2}} \mathcal{F}_t
\]

\[
= p_{r_{2,1,t}}^{u} 1_{A_{2,1}} \left( v_{r_{2,2}-r_{2,1}}^{u} \left( 1_{A_{2,2}} p_{r_{2,2}-r_{2,1}}^{u} \left( 1_{A_{2,3}} \cdots \right) \right) \right) \left( v_{r_{2,1}t}^{u,v} \right)
\]

so that by (50) we have:

\[
P(B_1 \cap B_2) = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^{u,v} \right) \right] \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^{u,v} \right) \right] \mathcal{F}_{r_{2,1}} \cdots \mathcal{F}_{r_{2,1}} \mathcal{F}_t
\]

In a similar way we have:

\[
P(B_1)P(B_2) = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^{u,v} \right) \right] \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^{u,v} \right) \right] \mathcal{F}_{r_{2,1}} \cdots \mathcal{F}_{r_{2,1}} \mathcal{F}_t
\]

and then:

\[
P(B_1 \cap B_2) - P(B_1)P(B_2) = \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^{u,v} \right) \right] \mathbb{E} \left[ 1_{A_{2,1}} \left( v_{r_{2,1}}^{u,v} \right) \right] \mathcal{F}_{r_{2,1}} \cdots \mathcal{F}_{r_{2,1}} \mathcal{F}_t
\]

\[
- p_{r_{2,1,t}}^{u} 1_{A_{2,1}} \left( v_{r_{2,2}-r_{2,1}}^{u} \left( 1_{A_{2,2}} p_{r_{2,2}-r_{2,1}}^{u} \left( 1_{A_{2,3}} \cdots \right) \right) \right) \left( v_{r_{2,1}t}^{u,v} \right)
\]

\[
= \mathbb{E} \left[ \prod_{i=1}^{k_1} 1_{A_{1,i}} \left( v_{r_{1,i}}^{u,v} \right) \right] \mathbb{E} \left[ p_{r_{2,1}-t}^{u} \phi \left( v_{r_{2,1}t}^{u,v} \right) - \int_k \phi(v)u(v)dv \right]
\]

for \( \phi = 1_{A_{2,1}} p_{r_{2,2}-r_{2,1}}^{u} \left( 1_{A_{2,2}} p_{r_{2,3}-r_{2,2}}^{u} \left( 1_{A_{2,3}} \cdots \right) \right) \).

Then by lemmas 4.3 and 4.5 and as \( f(s) = e^{-\delta s}s^{-1/2} \) is decreasing we have:

\[
|P(B_1 \cap B_2) - P(B_1)P(B_2)| \leq C \frac{e^{-\delta s(r_{2,1-t})}}{\sqrt{r_{2,1-t}}} \left( 1 + \mathbb{E} \left| v_{r_{2,1}t}^{u,v} \right| + |v| \right)
\]

\[
\leq C \frac{e^{-\delta s}}{\sqrt{s}} \left( 1 + |u| + |v| \right)
\]

so that the inequality holds when \( B_1 \in C_0^s(v) \) and \( B_2 \in C_1^{\infty}(v) \).

Finally recalling (49) the validity of the inequality can be extended to every \( B_1 \in \mathcal{H}_0(v) \) and \( B_2 \in \mathcal{H}_1^s(v) \). \( \square \)
B Proof of lemma 4.7

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

$$|E[\xi_1 \xi_2] - E\xi_1 E\xi_2| = |E[E[\xi_1 | \mathcal{H}^{t_1}_{s_1}(u,v)] - E\xi_1]| \leq |E[E[\xi_1 | \mathcal{H}^{t_1}_{s_1}(u,v)] - E\xi_1]| = |E[\xi_1 \xi_2] - E\xi_1 E\xi_2|$$

where we have defined

$$\xi_1 = 2\mathbb{I}_{E[\xi_1 | \mathcal{H}^{t_1}_{s_1}(u,v)] - E\xi_1 > 0} - 1$$

Similarly we have:

$$|E[\tilde{\xi}_1 \xi_2] - E\tilde{\xi}_1 E\xi_2| \leq |E[\tilde{\xi}_1 \xi_2] - E\tilde{\xi}_1 E\tilde{\xi}_2|$$

where we have defined

$$\tilde{\xi}_2 = 2\mathbb{I}_{E[\xi_1 | \mathcal{H}^{t_1}_{s_1}(u,v)] - E\tilde{\xi}_1 > 0} - 1$$

Then it follows:

$$|E[\xi_1 \xi_2] - E\xi_1 E\xi_2| \leq |E[\tilde{\xi}_1 \xi_2] - E\tilde{\xi}_1 E\tilde{\xi}_2| = 4|P(A \cap B) - P(A)P(B)|$$

where we have defined:

$$A = \{\omega \in \Omega : E\left(\xi_2 | \mathcal{H}^{t_1}_{s_1}(u,v)\right) - E\xi_2 > 0\}$$

and

$$B = \{\omega \in \Omega : E\left(\tilde{\xi}_2 | \mathcal{H}^{t_1}_{s_1}(u,v)\right) - E\tilde{\xi}_2 > 0\}$$

Now by lemma 4.6 as $A \in \mathcal{H}^{t_1}_{s_1}(u,v), B \in \mathcal{H}^{t_2}_{s_2}(u,v)$ we have:

$$|E[\xi_1 \xi_2] - E\xi_1 E\xi_2| \leq C \frac{e^{-s(t_2 - t_1)}}{\sqrt{s_2 - t_1}} (1 + |u| + |v|)$$

so that we have the thesis of the lemma.

C Proof of lemma 4.8

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:

$$E[\xi_1 \xi_2] - E\xi_1 E\xi_2 = E\left[\xi_1 \xi_2 \left(1_{A_{1,R}} \cap A_{2,R} + 1_{A_{1,R}} \cup A_{2,R}\right)\right] - E\xi_1 E\xi_2$$

$$= \left\{E[\xi_1 1_{A_{1,R}} \xi_2 1_{A_{2,R}}] - E[\xi_1 1_{A_{1,R}} E[\xi_2 1_{A_{2,R}}]] + E[\xi_1 1_{A_{1,R}}] E[\xi_2 1_{A_{2,R}}] + E[\xi_1 1_{A_{1,R}}] E[\xi_2 1_{A_{2,R}}]\right\}$$

$$= T_{1,R} + T_{2,R} + T_{3,R}$$

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

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

27
Now as $|\mathbf{A}_{1,R}| \leq 1$ a.s then by lemma 4.7 we have:

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

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

$$|T_{2,R}|^{\frac{1}{1+\rho}} \leq E[|\xi_1|^{\frac{1}{1+\rho}} E[|\xi_2|^{\frac{1}{1+\rho}} (P(A_{1,R}^c) + P(A_{2,R}^c))]^{\frac{1}{1+\rho}}$$

$$\leq E[|\xi_1|^{\frac{1}{1+\rho}} E[|\xi_2|^{\frac{1}{1+\rho}} R^{\rho} (E[|\xi_1|] + E[|\xi_2|])^{\frac{2\rho}{1+\rho}}$$

and then by (24) 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:

$$|E[\xi_1 A_{1,R}] | \leq (E[|\xi_1|^{\frac{1}{1+\rho}} E[|\xi_2|^{\frac{1}{1+\rho}} R^{\rho} (E[|\xi_1|] + E[|\xi_2|])^{\frac{2\rho}{1+\rho}}$$

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 $T_{1,R}$ into the first equation we have:

$$|E\xi_1 \xi_2 - E\xi_1 E\xi_2| \leq CK_1 K_2 (K_1 + K_2)^\rho R^{-\rho}$$

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

$$|E\xi_1 \xi_2 - E\xi_1 E\xi_2| \leq CK_1 K_2 (K_1 + K_2)^\rho R^{-\rho}$$

Acknowledgements

The author thanks warmly his PhD supervisor Giuseppina Guatteri for her precious and constant support during the preparation of this manuscript. Moreover he thanks Sandra Cerrai and Gianmario Tessitore for some relevant conversations related to the content of the present paper.

References

[1] M. Bardi, A. Cesaroni, L. Manca, Convergence by viscosity methods in multiscale financial models with stochastic volatility, SIAM J. Financ. Math. 1 (1) (2010) 230–265, https://doi.org/10.1137/090748147.

[2] Bréhier, Charles-Edouard. "Strong and weak orders in averaging for SPDEs." Stochastic Processes and their Applications 122.7 (2012): 2553-2593.
[3] Brehier C. E.: Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. SIAM J. Numer. Anal., 51 (2013), 1185–1210.

[4] Bréhier, Charles-Edouard. "Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component." Stochastic Processes and their Applications 130.6 (2020): 3325-3368.

[5] Cerrai S. and Freidlin M.: Averaging principle for stochastic reaction-diffusion equations. Prob. Theory Related Fields, 144 (2009), 137–177.

[6] S. Cerrai. A Khasminskii type averaging principle for stochastic reaction-diffusion equations. Ann. Appl. Probab. 19(3) (2009) 899-948. doi:10.1214/08-AAP560.

[7] Cerrai, Sandra. "Normal deviations from the averaged motion for some reaction–diffusion equations with fast oscillating perturbation." Journal de mathématiques pures et appliquées 91.6 (2009): 614-647.

[8] Cerrai S.: Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise. SIAM J. Math. Anal., 43 (2011) 2482-2518.

[9] S. Cerrai, A. Lunardi. Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: the almost periodic case. SIAM J. Math. Anal. 49 (2017) 2843-2884. https://doi.org/10.1137/16M1063307

[10] Da Prato, Giuseppe, and Jerzy Zabczyk. Stochastic equations in infinite dimensions. Cambridge university press, 2014.

[11] Da Prato, Giuseppe, Jerzy Zabczyk, and J. Zabczyk. Ergodicity for infinite dimensional systems. Vol. 229. Cambridge University Press, 1996.

[12] Filippo de Feo, 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, Filippo. "The averaging principle for non-autonomous slow-fast stochastic differential equations and an application to a local stochastic volatility model." Journal of Differential Equations 302 (2021): 406-443.

[14] Dong Z., Sun X., Xiao H. and Zhai J.: Averaging principle for one dimensional stochastic Burgers equation. J. Diff. Equ., 265 (2018), 4749–4797.

[15] Fu H., Wan L. and Liu J.: Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. Stoch. Process. Appl., 125 (8) (2015), 3255–3279.

[16] Givon, Dror. "Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems." Multiscale Modeling & Simulation 6.2 (2007): 577-594.

[17] J. Golec, Stochastic averaging principle for systems with pathwise uniqueness, Stoch. Anal. Appl. 13 (1995) 307–322.

[18] J. Golec, G. Ladde, Averaging principle and systems of singularly perturbed stochastic differential equations, J.Math. Phys. 31 (1990) 1116–1123.

[19] Guatteri, Giuseppina, and Gianmario Tessitore. "Singular limit of BSDEs and optimal control of two scale stochastic systems in infinite dimensional spaces." Applied Mathematics & Optimization 83.2 (2021): 1025-1051.
[20] Guatteri, Giuseppina, and Gianmario Tessitore. "Singular limit of two-scale stochastic optimal control problems in infinite dimensions by vanishing noise regularization." SIAM Journal on Control and Optimization 60.1 (2022): 575-596.

[21] J. P. Fouque, G. Papanicolaou, R. Sircar. Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press, 2000.

[22] J. P. Fouque, G. Papanicolaou, R. Sircar and K. Solna. Multiscale stochastic volatility for equity, interest rate, and credit derivatives. Cambridge University Press, 2011.

[23] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems. Springer, New York, NY, 1998.

[24] Hongbo Fu, Li Wan, Jicheng Liu, Strong convergence in averaging principle for stochastic hyperbolic–parabolic equations with two time-scales, Stochastic Processes and their Applications, Volume 125, Issue 8, 2015, Pages 3255-3279, ISSN 0304-4149, https://doi.org/10.1016/j.spa.2015.03.004.

[25] Fu, Hongbo, et al. "Weak order in averaging principle for stochastic wave equation with a fast oscillation." Stochastic Processes and their Applications 128.8 (2018): 2557-2580.

[26] Khasminskii R.Z. On an averaging principle for Itô stochastic differential equations. Kibernetika, 4(3) (1968), 260-279. (Russian).

[27] Khasminskii, Rafail Z., and George Yin. "On averaging principles: an asymptotic expansion approach." SIAM Journal on Mathematical Analysis 35.6 (2004): 1534-1560.

[28] D. Liu, E. Weinan, E. Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. Communications on Pure and Applied Mathematics 58.11 (2005): 1544-1585.

[29] D. Liu, Strong convergence of principle of averaging for multiscale stochastic dynamical systems, Commun. Math. Sci. 8 (2010) 999–1020.

[30] W. Fuke, T. Tian, J. B. Rawlings, and G. Yin. Approximate method for stochastic chemical kinetics with two-time scales by chemical Langevin equations. The Journal of Chemical Physics 144.17 (2016): 174112. https://doi.org/10.1063/1.4948407

[31] Röckner, Michael, Xiaobin Sun, and Longjie Xie. "Strong and weak convergence in the averaging principle for SDEs with Hölder coefficients." arXiv preprint arXiv:1907.09256 (2019).

[32] W. Liu, M. Röckner, X. Sun, Y. Xie, Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients, Journal of Differential Equations, Volume 268, Issue 6, 2020, Pages 2910-2948, ISSN 0022-0396, https://doi.org/10.1016/j.jde.2019.09.047.

[33] Röckner, Michael, Longjie Xie, and Li Yang. "Asymptotic behavior of multiscale stochastic partial differential equations." arXiv preprint arXiv:2010.14897 (2020).

[34] Rozanov, I. A. (1967). Stationary random processes. Holden-Day.

[35] Święch, Andrzej. "Singular perturbations and optimal control of stochastic systems in infinite dimension: HJB equations and viscosity solutions." ESAIM: Control, Optimisation and Calculus of Variations 27 (2021): 6.
[36] Kenneth Uda, Averaging principle for stochastic differential equations in the random periodic regime, Stochastic Processes and their Applications, Volume 139, 2021, Pages 1-36, ISSN 0304-4149, https://doi.org/10.1016/j.spa.2021.04.017.

[37] A. Yu. Veretennikov. On the averaging principle for systems of stochastic differential equations. Mathematics of the USSR-Sbornik 69.1 (1991): 271.

[38] Wang, W., and A. J. Roberts. "Average and deviation for slow–fast stochastic partial differential equations." Journal of Differential Equations 253.5 (2012): 1265-1286.

[39] J. Xu, J. Liu, Y. Miao, Strong averaging principle for two-time-scale SDEs with non-Lipschitz coefficients, J.Math. Anal. Appl. 468(1) (2018) 116–140.

[40] Zhang B., Fu H., Wan L. and Liu J.: Weak order in averaging principle for stochastic differential equations with jumps. Adv. Difference Equ., 2018, Paper No. 197, 20 pp.