

## A KHASMINSKII TYPE AVERAGING PRINCIPLE FOR STOCHASTIC REACTION–DIFFUSION EQUATIONS

BY SANDRA CERRAI

*Università di Firenze*

We prove that an averaging principle holds for a general class of stochastic reaction–diffusion systems, having unbounded multiplicative noise, in any space dimension. We show that the classical Khasminskii approach for systems with a finite number of degrees of freedom can be extended to infinite-dimensional systems.

**1. Introduction.** Consider the deterministic system with a finite number of degrees of freedom

$$(1.1) \quad \begin{cases} \frac{d\hat{X}_\varepsilon}{dt}(t) = \varepsilon b(\hat{X}_\varepsilon(t), \hat{Y}_\varepsilon(t)), & \hat{X}_\varepsilon(0) = x \in \mathbb{R}^n, \\ \frac{d\hat{Y}_\varepsilon}{dt}(t) = g(\hat{X}_\varepsilon(t), \hat{Y}_\varepsilon(t)), & \hat{Y}_\varepsilon(0) = y \in \mathbb{R}^k \end{cases}$$

for some parameter  $0 < \varepsilon \ll 1$  and some mappings  $b: \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$  and  $g: \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ . Under reasonable conditions on  $b$  and  $g$ , it is clear that as the parameter  $\varepsilon$  goes to zero, the first component  $\hat{X}_\varepsilon(t)$  of the perturbed system (1.1) converges to the constant first component  $x$  of the unperturbed system, uniformly with respect to  $t$  in any bounded interval  $[0, T]$ , with  $T > 0$ .

But in applications that is more interesting is the behavior of  $\hat{X}_\varepsilon(t)$  for  $t$  in intervals of order  $\varepsilon^{-1}$  or even larger. Actually, it is indeed on those time scales that the most significant changes happen, such as exit from the neighborhood of an equilibrium point or of a periodic trajectory. With the natural time scaling  $t \mapsto t/\varepsilon$ , if we set  $X_\varepsilon(t) := \hat{X}_\varepsilon(t/\varepsilon)$  and  $Y_\varepsilon(t) := \hat{Y}_\varepsilon(t/\varepsilon)$ , (1.1) can be rewritten as

$$(1.2) \quad \begin{cases} \frac{dX_\varepsilon}{dt}(t) = b(X_\varepsilon(t), Y_\varepsilon(t)), & X_\varepsilon(0) = x \in \mathbb{R}^n, \\ \frac{dY_\varepsilon}{dt}(t) = \frac{1}{\varepsilon} g(X_\varepsilon(t), Y_\varepsilon(t)), & Y_\varepsilon(0) = y \in \mathbb{R}^k \end{cases}$$

and with this time scale the variable  $X_\varepsilon$  is always referred as the *slow* component and  $Y_\varepsilon$  as the *fast* component. In particular, the study of system (1.1) in time intervals of order  $\varepsilon^{-1}$  is equivalent to the study of system (1.2) on finite time intervals.

---

Received August 2007; revised July 2008.

*AMS 2000 subject classifications.* 60H15, 34C29, 37L40.

*Key words and phrases.* Stochastic reaction diffusion equations, invariant measures, ergodic and strongly mixing processes, averaging principle.

Now, assume that for any  $x \in \mathbb{R}^n$  there exists the limit

$$(1.3) \quad \bar{b}(x) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T b(x, Y^x(t)) dt,$$

where  $Y^x(t)$  is the fast motion with frozen slow component  $x \in \mathbb{R}^n$

$$\frac{dY^x}{dt}(t) = g(x, Y^x(t)), \quad Y^x(0) = y.$$

Such a limit exists, for example, in the case the function  $Y^x(t)$  is periodic. Moreover, assume that the mapping  $\bar{b}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  satisfies some reasonable assumption, for example, it is Lipschitz continuous. In this setting, the *averaging principle* says that the trajectory of  $X_\varepsilon$  can be approximated by the solution  $\bar{X}$  of the so-called *averaged equation*

$$\frac{d\bar{X}}{dt}(t) = \bar{b}(\bar{X}(t)), \quad \bar{X}(0) = x,$$

uniformly in  $t \in [0, T]$ , for any fixed  $T > 0$ . This means that by averaging principle a good approximation of the slow motion can be obtained by averaging its parameters in the fast variables.

The theory of averaging, originated by Laplace and Lagrange, has been applied in its long history in many fields as, for example, celestial mechanics, oscillation theory and radiophysics, and for a long period it has been used without a rigorous mathematical justification. The first rigorous results are due to Bogoliubov (cfr. [3]) and concern both the case of uncoupled systems and the case of  $g(x, y) = g(x)$ . Further developments of the theory, for more general systems, were obtained by Volosov, Anosov and Neishtadt (to this purpose, we refer to [23] and [28]) and a good understanding of the involved phenomena was obtained by Arnold et al. (cfr. [1]).

A further development in the theory of averaging, which is of great interest in applications, concerns the case of random perturbations of dynamical systems. For example, in system (1.1), the coefficient  $g$  may be assumed to depend also on a parameter  $\omega \in \Omega$ , for some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , so that the fast variable is a random process, or even the perturbing coefficient  $b$  may be taken random. Of course, in these cases, one has to reinterpret condition (1.3) and the type of convergence of the stochastic process  $X_\varepsilon$  to  $\bar{X}$ . One possible way is to require (1.3) with probability 1, but in most cases this assumption turns out to be too restrictive. More reasonable is to have (1.3) either in probability or in the mean, and in this case one expects to have convergence in probability of  $X_\varepsilon$  to  $\bar{X}$ . As far as averaging for randomly perturbed systems is concerned, it is worthwhile to quote the important work of Brin, Freidlin and Wentzell (see [4, 12–14]) and also the work of Kifer and Veretennikov (see, e.g., [2, 17–19] and [27]).

An important contribution in this direction has been given by Khasminskii with his paper [16] which appeared in 1968. In this paper, he has considered the following system of stochastic differential equations:

$$(1.4) \quad \begin{cases} dX^\varepsilon(t) = A(X^\varepsilon(t), Y^\varepsilon(t)) dt + \sum_{r=1}^l \sigma^r(X_i^\varepsilon(t), Y^\varepsilon(t)) dw_r(t), \\ X^\varepsilon(0) = x_0, \\ dY^\varepsilon(t) = \frac{1}{\varepsilon} B(X^\varepsilon(t), Y^\varepsilon(t)) dt + \frac{1}{\sqrt{\varepsilon}} \sum_{r=1}^l \varphi^r(X^\varepsilon(t), Y^\varepsilon(t)) dw_r(t), \\ Y^\varepsilon(0) = y_0 \end{cases}$$

for some  $l$ -dimensional Brownian motion  $w(t) = (w_1(t), \dots, w_l(t))$ . In this case, the perturbation in the slow motion is given by the sum of a deterministic part and a stochastic part

$$\varepsilon b(x, y) dt = \varepsilon A(x, y) dt + \sqrt{\varepsilon} \sigma(x, y) dw(t)$$

and the fast motion is described by a stochastic differential equation.

In [16], the coefficients  $A : \mathbb{R}^{l_1} \times \mathbb{R}^{l_2} \rightarrow \mathbb{R}^{l_1}$  and  $\sigma : \mathbb{R}^{l_1} \times \mathbb{R}^{l_2} \rightarrow M(l \times l_1)$  in the slow motion equation are assumed to be Lipschitz continuous and uniformly bounded in  $y \in \mathbb{R}^{l_2}$ . The coefficients  $B : \mathbb{R}^{l_1} \times \mathbb{R}^{l_2} \rightarrow \mathbb{R}^{l_2}$  and  $\varphi : \mathbb{R}^{l_1} \times \mathbb{R}^{l_2} \rightarrow M(l \times l_2)$  in the fast motion equation are assumed to be Lipschitz continuous, so that in particular the fast equation with frozen slow component  $x$ ,

$$dY^{x,y}(t) = B(x, Y^{x,y}(t)) dt + \sum_{r=1}^l \varphi^r(x, Y^{x,y}(t)) dw_r(t), \quad Y^{x,y}(0) = y,$$

admits a unique solution  $Y^{x,y}$ , for any  $x \in \mathbb{R}^{l_1}$  and  $y \in \mathbb{R}^{l_2}$ . Moreover, it is assumed that there exist two mappings  $\bar{A} : \mathbb{R}^{l_1} \rightarrow \mathbb{R}^{l_1}$  and  $\{a_{ij}\} : \mathbb{R}^{l_1} \rightarrow M(l \times l_1)$  such that

$$(1.5) \quad \left| \frac{1}{T} \int_0^T \mathbb{E} A(x, Y^{x,y}(t)) dt - \bar{A}(x) \right| \leq \alpha(T)(1 + |x|^2)$$

and for any  $i = 1, \dots, l_1$  and  $j = 1, \dots, l_2$ ,

$$\left| \frac{1}{T} \int_0^T \mathbb{E} \sum_{r=1}^l \sigma_i^r \sigma_j^r(x, Y^{x,y}(t)) dt - a_{ij}(x) \right| \leq \alpha(T)(1 + |x|^2)$$

for some function  $\alpha(T)$  vanishing as  $T$  goes to infinity.

In his paper, Khasminskii shows that an averaging principle holds for system (1.4). Namely, the slow motion  $X^\varepsilon(t)$  converges in weak sense, as  $\varepsilon$  goes to zero, to the solution  $\bar{X}$  of the averaged equation

$$dX(t) = \bar{A}(X(t)) dt + \bar{\sigma}(X(t)) dw(t), \quad X(0) = x_0,$$

where  $\bar{\sigma}$  is the square root of the matrix  $\{a_{ij}\}$ .

The behavior of solutions of infinite-dimensional systems on time intervals of order  $\varepsilon^{-1}$  is at present not very well understood, even if applied mathematicians do believe that the averaging principle holds and usually approximate the slow motion by the averaged motion, also with  $n = \infty$ . As far as we know, the literature on averaging for systems with an infinite number of degrees of freedom is extremely poor (to this purpose it is worth mentioning the papers [25] by Seidler–Vrkoč and [21] by Maslowskii–Seidler–Vrkoč, concerning with averaging for Hilbert-space valued solutions of stochastic evolution equations depending on a small parameter, and the paper [20] by Kuksin and Piatnitski concerning with averaging for a randomly perturbed KdV equation) and almost all has still to be done.

In the present paper, we are trying to extend the Khasminskii argument to a system with an infinite number of degrees of freedom. We are dealing with the following system of stochastic reaction–diffusion equations on a bounded domain  $D \subset \mathbb{R}^d$ , with  $d \geq 1$ ,

$$(1.6) \quad \begin{cases} \frac{\partial u_\varepsilon}{\partial t}(t, \xi) = \mathcal{A}_1 u_\varepsilon(t, \xi) + b_1(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi)) \\ \quad \quad \quad + g_1(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi)) \frac{\partial w^{Q_1}}{\partial t}(t, \xi), \\ \frac{\partial v_\varepsilon}{\partial t}(t, \xi) = \frac{1}{\varepsilon} [\mathcal{A}_2 v_\varepsilon(t, \xi) + b_2(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi))] \\ \quad \quad \quad + \frac{1}{\sqrt{\varepsilon}} g_2(\xi, u_\varepsilon(t, \xi), v_\varepsilon(t, \xi)) \frac{\partial w^{Q_2}}{\partial t}(t, \xi), \\ u_\varepsilon(0, \xi) = x(\xi), \quad v_\varepsilon(0, \xi) = y(\xi), \quad \xi \in D, \\ \mathcal{N}_1 u_\varepsilon(t, \xi) = \mathcal{N}_2 v_\varepsilon(t, \xi) = 0, \quad t \geq 0, \quad \xi \in \partial D \end{cases}$$

for a positive parameter  $\varepsilon \ll 1$ . The stochastic perturbations are given by Gaussian noises which are white in time and colored in space, in the case of space dimension  $d > 1$ , with covariances operators  $Q_1$  and  $Q_2$ . The operators  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are second order uniformly elliptic operators, having continuous coefficients on  $D$ , and the boundary operators  $\mathcal{N}_1$  and  $\mathcal{N}_2$  can be either the identity operator (Dirichlet boundary condition) or a first order operator satisfying a uniform nontangentiality condition.

In our previous paper [8], written in collaboration with Mark Freidlin, we have considered the simpler case of  $g_1 \equiv 0$  and  $g_2 \equiv 1$ , and we have proved that an averaging principle is satisfied by using a completely different approach based on Kolmogorov equations and martingale solutions of stochastic equations, which is more in the spirit of the general method introduced by Papanicolaou, Strook and Varadhan in their paper [24] of 1977. Here, we are considering the case of general reaction coefficients  $b_1$  and  $b_2$  and diffusion coefficients  $g_1$  and  $g_2$ , and the method based on the martingale approach seems to be very complicated to be applied.

We would like to stress that both here and in our previous paper [8] we are considering averaging for randomly perturbed reaction–diffusion systems, which are of interest in the description of diffusive phenomena in reactive media, such

as combustion, epidemic propagation and diffusive transport of chemical species through cells and dynamics of populations. However, the arguments we are using adapt easily to more general models of semilinear stochastic partial differential equations.

Together with system (1.6), for any  $x, y \in H := L^2(D)$ , we introduce the fast motion equation

$$\begin{cases} \frac{\partial v}{\partial t}(t, \xi) = [\mathcal{A}_2 v(t, \xi) + b_2(\xi, x(\xi), v(t, \xi))] + g_2(\xi, x(\xi), v(t, \xi)) \frac{\partial w^{Q_2}}{\partial t}(t, \xi), \\ v(0, \xi) = y(\xi), \quad \xi \in D, \quad \mathcal{N}_2 v(t, \xi) = 0, \quad t \geq 0, \quad \xi \in \partial D \end{cases}$$

with initial datum  $y$  and frozen slow component  $x$ , whose solution is denoted by  $v^{x,y}(t)$ . The previous equation has been widely studied, as far as existence and uniqueness of solutions are concerned. In Section 3, we introduce the transition semigroup  $P_t^x$  associated with it and, by using methods and results from our previous paper [7], we study its asymptotic properties and its dependence on the parameters  $x$  and  $y$  (cfr. also [5] and [6]).

Under this respect, in addition to suitable conditions on the operators  $\mathcal{A}_i$  and  $Q_i$  and on the coefficients  $b_i$  and  $g_i$ , for  $i = 1, 2$  (see Section 2 for all hypotheses), in the spirit of Khasminskii’s work, we assume that there exist a mapping  $\alpha(T)$ , which vanishes as  $T$  goes to infinity, and two Lipschitz-continuous mappings  $\bar{B}_1 : H \rightarrow H$  and  $\bar{G} : H \rightarrow \mathcal{L}(L^\infty(D), H)$  such that for any choice of  $T > 0$ ,  $t \geq 0$  and  $x, y \in H$

$$(1.7) \quad \mathbb{E} \left| \frac{1}{T} \int_t^{t+T} \langle B_1(x, v^{x,y}(s)), h \rangle_H ds - \langle \bar{B}_1(x), h \rangle_H \right| \leq \alpha(T)(1 + |x|_H + |y|_H)|h|_H$$

for any  $h \in H$ , and

$$(1.8) \quad \left| \frac{1}{T} \int_t^{t+T} \mathbb{E} \langle G_1(x, v^{x,y}(s))h, G_1(x, v^{x,y}(s))k \rangle_H ds - \langle \bar{G}(x)h, \bar{G}(x)k \rangle_H \right| \leq \alpha(T)(1 + |x|_H^2 + |y|_H^2)|h|_{L^\infty(D)}|k|_{L^\infty(D)}$$

for any  $h, k \in L^\infty(D)$ . Here,  $B_1$  and  $G_1$  are the Nemytskii operators associated with  $b_1$  and  $g_1$ , respectively. Notice that unlike  $B_1$  and  $G_1$  which are local operators, the coefficients  $\bar{B}$  and  $\bar{G}$  are not local. Actually, they are defined as general mappings on  $H$ , and also in applications, there is no reason why they should be composition operators.

In Section 3, we describe some remarkable situations in which conditions (1.7) and (1.8) are fulfilled: for example, when the fast motion admits a strongly mixing invariant measure  $\mu^x$ , for any fixed frozen slow component  $x \in H$ , and the diffusion coefficient  $g_1$  of the slow motion equation is bounded and nondegenerate.

Our purpose is showing that under the above conditions the slow motion  $u_\varepsilon$  converges weakly to the solution  $\bar{u}$  of the averaged equation

$$(1.9) \quad \begin{cases} \frac{\partial u}{\partial t}(t, \xi) = \mathcal{A}_1 u(t, \xi) + \bar{B}(u)(t, \xi) + \bar{G}(u)(t, \xi) \frac{\partial w^{Q_1}}{\partial t}(t, \xi), \\ u(0, \xi) = x(\xi), \quad \xi \in D, \quad \mathcal{N}_1 u(t, \xi) = 0, \quad t \geq 0, \quad \xi \in \partial D. \end{cases}$$

More precisely, we prove that for any  $T > 0$

$$(1.10) \quad \mathcal{L}(u_\varepsilon) \rightharpoonup \mathcal{L}(\bar{u}) \text{ in } C([0, T]; H) \quad \text{as } \varepsilon \rightarrow 0$$

(see Theorem 6.2). Moreover, in the case the diffusion coefficient  $g_1$  in the slow equation does not depend on the fast oscillating variable  $v_\varepsilon$ , we show that the convergence of  $u_\varepsilon$  to  $\bar{u}$  is in probability, that is, for any  $\eta > 0$

$$(1.11) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{P}(|u_\varepsilon - \bar{u}|_{C([0, T]; H)} > \eta) = 0$$

(see Theorem 6.4).

In order to prove (1.10), we have to proceed in several steps. First of all, we show that the family  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0, 1]}$  is tight in  $\mathcal{P}(C([0, T]; H))$  and this is obtained by a priori bounds for processes  $u_\varepsilon$  in a suitable Hölder norm with respect to time and in a suitable Sobolev norm with respect to space. We would like to stress that as we are only assuming (1.7) and (1.8) and not a law of large numbers, we also need to prove a priori bounds for the conditioned momenta of  $u_\varepsilon$ .

Once we have the tightness of the family  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0, 1]}$ , we have the weak convergence of the sequence  $\{\mathcal{L}(u_{\varepsilon_n})\}_{n \in \mathbb{N}}$ , for some  $\varepsilon_n \downarrow 0$ , to some probability measure  $\mathbb{Q}$  on  $C_x([0, T]; H)$ . The next steps consist in identifying  $\mathbb{Q}$  with  $\mathcal{L}(\bar{u})$  and proving that limit (1.10) holds. To this purpose, we introduce the martingale problem with parameters  $(x, \mathcal{A}_1, \bar{B}, \bar{G}, Q_1)$  and we show that  $\mathbb{Q}$  is a solution to such martingale problem. As the coefficients  $\bar{B}$  and  $\bar{Q}$  are Lipschitz-continuous, we have uniqueness, and hence we can conclude that  $\mathbb{Q} = \mathcal{L}(\bar{u})$ . This in particular implies that for any  $\varepsilon_n \downarrow 0$  the sequence  $\{\mathcal{L}(u_{\varepsilon_n})\}_{n \in \mathbb{N}}$  converges weakly to  $\mathcal{L}(\bar{u})$ , and hence (1.10) holds. Moreover, in the case  $g_1$  does not depend on  $v_\varepsilon$ , by a uniqueness argument, this implies convergence in probability.

In the general case, the key point in the identification of  $\mathbb{Q}$  with the solution of the martingale problem associated with the averaged (1.9) is the following limit

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_{t_1}^{t_2} \mathbb{E}(\mathcal{L}_{sl} \varphi(u_\varepsilon(r), v_\varepsilon(r)) - \mathcal{L}_{av} \varphi(u_\varepsilon(r)) | \mathcal{F}_{t_1}) dr \right| = 0,$$

where  $\mathcal{L}_{sl}$  and  $\mathcal{L}_{av}$  are the Kolmogorov operators associated, respectively, with the slow motion equation, with frozen fast component, and with the averaged equation, and  $\{\mathcal{F}_t\}_{t \geq 0}$  is the filtration associated with the noise. Notice that it is sufficient to check the validity of such a limit for any cylindrical function  $\varphi$  and any  $0 \leq t_1 \leq t_2 \leq T$ . The proof of the limit above is based on the Khasminskii

argument introduced in [16], but it is clearly more delicate than in [16], as it concerns a system with an infinite number of degrees of freedom (with all well-known problems arising from that).

In the particular case of  $g_1$  not depending on  $v_\varepsilon$ , in order to prove (1.11), we do not need to pass through the martingale formulation. For any  $h \in D(A_1)$ , we write

$$\begin{aligned} \langle u_\varepsilon(t), h \rangle_H &= \langle x, h \rangle_H + \int_0^t \langle u_\varepsilon(s), A_1 h \rangle_H ds + \int_0^t \langle \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \\ &\quad + \int_0^t \langle G_1(u_\varepsilon(s))h, dw_1^{Q_1}(s) \rangle_H + R_\varepsilon(t), \end{aligned}$$

where

$$R_\varepsilon(t) := \int_0^t \langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds$$

and we show that for any  $T > 0$

$$(1.12) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |R_\varepsilon(t)| = 0.$$

Thanks to the Skorokhod theorem and to a general argument due to Gyöngy and Krylov (see [15]), this allows us to obtain (1.11).

**2. Assumptions and preliminaries.** Let  $D$  be a smooth bounded domain of  $\mathbb{R}^d$ , with  $d \geq 1$ . Throughout the paper, we shall denote by  $H$  the Hilbert space  $L^2(D)$ , endowed with the usual scalar product  $\langle \cdot, \cdot \rangle_H$  and with the corresponding norm  $|\cdot|_H$ . The norm in  $L^\infty(D)$  will be denoted by  $|\cdot|_0$ .

We shall denote by  $B_b(H)$  the Banach space of bounded Borel functions  $\varphi : H \rightarrow \mathbb{R}$ , endowed with the sup-norm

$$\|\varphi\|_0 := \sup_{x \in H} |\varphi(x)|.$$

$C_b(H)$  is the subspace of uniformly continuous mappings and  $C_b^k(H)$  is the subspace of all  $k$ -times differentiable mappings, having bounded and uniformly continuous derivatives, up to the  $k$ th order, for  $k \in \mathbb{N}$ .  $C_b^k(H)$  is a Banach space endowed with the norm

$$|\varphi|_k := |\varphi|_0 + \sum_{i=1}^k \sup_{x \in H} |D^i \varphi(x)|_{\mathcal{L}^i(H)} =: |\varphi|_0 + \sum_{i=1}^k [\varphi]_i,$$

where  $\mathcal{L}^1(H) := H$  and, by recurrence,  $\mathcal{L}^i(H) := \mathcal{L}(H, \mathcal{L}^{i-1}(H))$ , for any  $i > 1$ . Finally, we denote by  $\text{Lip}(H)$  the set of functions  $\varphi : H \rightarrow \mathbb{R}$  such that

$$[\varphi]_{\text{Lip}(H)} := \sup_{\substack{x, y \in H \\ x \neq y}} \frac{|\varphi(x) - \varphi(y)|}{|x - y|_H} < \infty.$$

We shall denote by  $\mathcal{L}(H)$  the space of bounded linear operators in  $H$  and we shall denote by  $\mathcal{L}_2(H)$  the subspace of Hilbert–Schmidt operators, endowed with the norm

$$\|Q\|_2 = \sqrt{\text{Tr}[Q^*Q]}.$$

The stochastic perturbations in the slow and in the fast motion equations (1.6) are given, respectively, by the Gaussian noises  $\partial w^{Q_1}/\partial t(t, \xi)$  and  $\partial w^{Q_2}/\partial t(t, \xi)$ , for  $t \geq 0$  and  $\xi \in D$ , which are assumed to be white in time and colored in space, in the case of space dimension  $d > 1$ . Formally, the cylindrical Wiener processes  $w^{Q_i}(t, \xi)$  are defined as the infinite sums

$$w^{Q_i}(t, \xi) = \sum_{k=1}^{\infty} Q_i e_k(\xi) \beta_k(t), \quad i = 1, 2,$$

where  $\{e_k\}_{k \in \mathbb{N}}$  is a complete orthonormal basis in  $H$ ,  $\{\beta_k(t)\}_{k \in \mathbb{N}}$  is a sequence of mutually independent standard Brownian motions defined on the same complete stochastic basis  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  and  $Q_i$  is a compact linear operator on  $H$ .

The operators  $\mathcal{A}_1$  and  $\mathcal{A}_2$  appearing, respectively, in the slow and in the fast motion equation, are second-order uniformly elliptic operators, having continuous coefficients on  $D$ , and the boundary operators  $\mathcal{N}_1$  and  $\mathcal{N}_2$  can be either the identity operator (Dirichlet boundary condition) or a first-order operator of the following type

$$\sum_{j=1}^d \beta_j(\xi) D_j + \gamma(\xi) I, \quad \xi \in \partial D$$

for some  $\beta_j, \gamma \in C^1(\bar{D})$  such that

$$\inf_{\xi \in \partial D} |\langle \beta(\xi), \nu(\xi) \rangle| > 0,$$

where  $\nu(\xi)$  is the unit normal at  $\xi \in \partial D$  (uniform nontangentiality condition).

The realizations  $A_1$  and  $A_2$  in  $H$  of the differential operators  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , endowed, respectively, with the boundary conditions  $\mathcal{N}_1$  and  $\mathcal{N}_2$ , generate two analytic semigroups  $e^{tA_1}$  and  $e^{tA_2}$ ,  $t \geq 0$ . In what follows, we shall assume that  $A_1$ ,  $A_2$  and  $Q_1, Q_2$  satisfy the following conditions.

**HYPOTHESIS 1.** For  $i = 1, 2$ , there exist a complete orthonormal system  $\{e_{i,k}\}_{k \in \mathbb{N}}$  in  $H$  and two sequences of nonnegative real numbers  $\{\alpha_{i,k}\}_{k \in \mathbb{N}}$  and  $\{\lambda_{i,k}\}_{k \in \mathbb{N}}$ , such that

$$A_i e_{i,k} = -\alpha_{i,k} e_{i,k}, \quad Q_i e_{i,k} = \lambda_{i,k} e_{i,k}, \quad k \geq 1.$$

If  $d = 1$ , we have

$$(2.1) \quad \kappa_i := \sup_{k \in \mathbb{N}} \lambda_{i,k} < \infty, \quad \zeta_i := \sum_{k=1}^{\infty} \alpha_{i,k}^{-\beta_i} |e_{i,k}|_0^2 < \infty$$

for some constant  $\beta_i \in (0, 1)$ , and if  $d \geq 2$ , we have

$$(2.2) \quad \kappa_i := \sum_{k=1}^{\infty} \lambda_{i,k}^{\rho_i} |e_{i,k}|_0^2 < \infty, \quad \zeta_i := \sum_{k=1}^{\infty} \alpha_{i,k}^{-\beta_i} |e_{i,k}|_0^2 < \infty$$

for some constants  $\beta_i \in (0, +\infty)$  and  $\rho_i \in (2, +\infty)$  such that

$$(2.3) \quad \frac{\beta_i(\rho_i - 2)}{\rho_i} < 1.$$

Moreover,

$$(2.4) \quad \inf_{k \in \mathbb{N}} \alpha_{2,k} =: \lambda > 0.$$

REMARK 2.1. 1. In several cases as, for example, in the case of space dimension  $d = 1$ , and in the case of the Laplace operator on a hypercube, endowed with Dirichlet boundary conditions, the eigenfunctions  $e_k$  are equibounded in the sup-norm and then conditions (2.1) and (2.2) become

$$\kappa_i = \sum_{k=1}^{\infty} \lambda_{i,k}^{\rho_i} < \infty, \quad \zeta_i = \sum_{k=1}^{\infty} \alpha_{i,k}^{-\beta_i} < \infty$$

for positive constants  $\beta_i, \rho_i$  fulfilling (2.3). In general,

$$|e_{i,k}|_0 \sim k^{a_i}, \quad k \in \mathbb{N}$$

for some  $a_i \geq 0$ . Thus, the two conditions in (2.2) become

$$\kappa_i := \sum_{k=1}^{\infty} \lambda_{i,k}^{\rho_i} k^{2a_i} < \infty, \quad \zeta_i := \sum_{k=1}^{\infty} \alpha_{i,k}^{-\beta_i} k^{2a_i} < \infty.$$

2. For any reasonable domain  $D \subset \mathbb{R}^d$ , one has

$$\alpha_{i,k} \sim k^{2/d}, \quad k \in \mathbb{N}.$$

Thus, if the eigenfunctions  $e_k$  are equibounded in the sup-norm, we have

$$\zeta_i \leq c \sum_{k=1}^{\infty} \alpha_{i,k}^{-\beta_i} \sim \sum_{k=1}^{\infty} k^{-2\beta_i/d}.$$

This means that in order to have  $\zeta_i < \infty$ , we need

$$\beta_i > \frac{d}{2}.$$

In particular, in order to have also  $\kappa_i < \infty$  and condition (2.3) satisfied, in space dimension  $d = 1$  we can take  $\rho_i = +\infty$ , so that we can deal with white noise, both in time and in space. In space dimension  $d = 2$ , we can take any  $\rho_i < \infty$  and in space dimension  $d \geq 3$ , we need

$$\rho_i < \frac{2d}{d-2}.$$

In any case, notice that it is never required to take  $\rho_i = 2$ , which means to have a noise with trace-class covariance. To this purpose, it can be useful to compare these conditions with Hypotheses 2 and 3 in [6].

As far as the coefficients  $b_1, b_2$  and  $g_1, g_2$  are concerned, we assume the following conditions.

**HYPOTHESIS 2.** 1. The mappings  $b_i : D \times \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $g_i : D \times \mathbb{R}^2 \rightarrow \mathbb{R}$  are measurable, both for  $i = 1$  and  $i = 2$ , and for almost all  $\xi \in D$  the mappings  $b_i(\xi, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $g_i(\xi, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  are Lipschitz-continuous, uniformly with respect to  $\xi \in D$ . Moreover,

$$\sup_{\xi \in D} |b_2(\xi, 0, 0)| < \infty.$$

2. It holds

$$(2.5) \quad \sup_{\substack{\xi \in D \\ \sigma_1 \in \mathbb{R}}} \sup_{\substack{\sigma_2, \rho_2 \in \mathbb{R} \\ \sigma_2 \neq \rho_2}} \frac{|b_2(\xi, \sigma_1, \sigma_2) - b_2(\xi, \sigma_1, \rho_2)|}{|\sigma_2 - \rho_2|} =: L_{b_2} < \lambda,$$

where  $\lambda$  is the constant introduced in (2.4).

3. There exists  $\gamma < 1$  such that

$$(2.6) \quad \sup_{\xi \in D} |g_2(\xi, \sigma)| \leq c(1 + |\sigma_1| + |\sigma_2|^\gamma), \quad \sigma = (\sigma_1, \sigma_2) \in \mathbb{R}^2.$$

**REMARK 2.2.** Notice that condition (2.6) on the growth of  $g_2(\xi, \sigma_1, \cdot)$  could be replaced with the condition

$$\sup_{\substack{\xi \in D \\ \sigma_1 \in \mathbb{R}}} [g_2(\xi, \sigma_1, \cdot)]_{\text{Lip}} \leq \eta$$

for some  $\eta$  sufficiently small.

In what follows we shall set

$$L_{g_2} := \sup_{\substack{\xi \in D \\ \sigma_1 \in \mathbb{R}}} \sup_{\substack{\sigma_2, \rho_2 \in \mathbb{R} \\ \sigma_2 \neq \rho_2}} \frac{|g_2(\xi, \sigma_1, \sigma_2) - g_2(\xi, \sigma_1, \rho_2)|}{|\sigma_2 - \rho_2|}.$$

Moreover, we shall set

$$B_i(x, y)(\xi) := b_i(\xi, x(\xi), y(\xi))$$

and

$$[G_i(x, y)z](\xi) := g_i(\xi, x(\xi), y(\xi))z(\xi)$$

for any  $\xi \in D$ ,  $x, y, z \in H$  and  $i = 1, 2$ . Due to Hypothesis 2, the mappings

$$(x, y) \in H \times H \mapsto B_i(x, y) \in H,$$

are Lipschitz-continuous, as well as the mappings

$$(x, y) \in H \times H \mapsto G_i(x, y) \in \mathcal{L}(H; L^1(D))$$

and

$$(x, y) \in H \times H \mapsto G_i(x, y) \in \mathcal{L}(L^\infty(D); H).$$

Now, for any fixed  $T > 0$  and  $p \geq 1$ , we denote by  $\mathcal{H}_{T,p}$  the space of processes in  $C([0, T]; L^p(\Omega; H))$ , which are adapted to the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  associated with the noise.  $\mathcal{H}_{T,p}$  is a Banach space, endowed with the norm

$$\|u\|_{\mathcal{H}_{T,p}} = \left( \sup_{t \in [0, T]} \mathbb{E}|u(t)|_H^p \right)^{1/p}.$$

Moreover, we denote by  $\mathcal{C}_{T,p}$  the subspace of processes  $u \in L^p(\Omega; C([0, T]; H))$ , endowed with the norm

$$\|u\|_{\mathcal{C}_{T,p}} = \left( \mathbb{E} \sup_{t \in [0, T]} |u(t)|_H^p \right)^{1/p}.$$

With all notation we have introduced, system (1.6) can be rewritten as the following abstract evolution system

$$(2.7) \quad \begin{cases} du_\varepsilon(t) = [A_1 u_\varepsilon(t) + B_1(u_\varepsilon(t), v_\varepsilon(t))] ds \\ \quad + G_1(u_\varepsilon(t), v_\varepsilon(t)) dw^{\mathcal{Q}_1}(t), & u_\varepsilon(0) = x \\ dv_\varepsilon(t) = \frac{1}{\varepsilon} [A_2 v_\varepsilon(t) + B_2(u_\varepsilon(t), v_\varepsilon(t))] ds \\ \quad + \frac{1}{\sqrt{\varepsilon}} G_2(u_\varepsilon(t), v_\varepsilon(t)) dw^{\mathcal{Q}_2}(t), & v_\varepsilon(0) = y. \end{cases}$$

As known from the existing literature (see, e.g., [9]), according to Hypotheses 1 and 2 for any  $\varepsilon > 0$  and  $x, y \in H$  and for any  $p \geq 1$  and  $T > 0$  there exists a unique mild solution  $(u_\varepsilon, v_\varepsilon) \in \mathcal{C}_{T,p} \times \mathcal{C}_{T,p}$  to system (1.6). This means that there exist two processes  $u_\varepsilon$  and  $v_\varepsilon$  in  $\mathcal{C}_{T,p}$ , which are unique, such that

$$u_\varepsilon(t) = e^{tA_1} x + \int_0^t e^{(t-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \\ + \int_0^t e^{(t-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s)$$

and

$$v_\varepsilon(t) = e^{tA_2/\varepsilon} y + \frac{1}{\varepsilon} \int_0^t e^{(t-s)A_2/\varepsilon} B_2(u_\varepsilon(s), v_\varepsilon(s)) ds \\ + \frac{1}{\sqrt{\varepsilon}} \int_0^t e^{(t-s)A_2/\varepsilon} G_2(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_2}(s).$$

2.1. *The fast motion equation.* For any fixed  $x \in H$ , we consider the problem

$$(2.8) \quad \begin{cases} \frac{\partial v}{\partial t}(t, \xi) = \mathcal{A}_2 v(t, \xi) + b_2(\xi, x(\xi), v(t, \xi)) \\ \quad + g_2(\xi, x(\xi), v(t, \xi)) \frac{\partial w^{Q_2}}{\partial t}(t, \xi), \\ v(0, \xi) = y(\xi), \quad \xi \in D, \\ \mathcal{N}_2 v(t, \xi) = 0, \quad t \geq 0, \quad \xi \in \partial D. \end{cases}$$

Under Hypotheses 1 and 2, such a problem admits a unique mild solution  $v^{x,y} \in \mathcal{C}_{T,p}$ , for any  $T > 0$  and  $p \geq 1$ , and for any fixed frozen slow variable  $x \in H$  and any initial condition  $y \in H$  (for a proof, see, e.g., [10], Theorem 5.3.1).

By arguing as in the proof of [7], Theorem 7.3, is it possible to show that there exists some  $\delta_1 > 0$  such that for any  $p \geq 1$

$$(2.9) \quad \mathbb{E}|v^{x,y}(t)|_H^p \leq c_p(1 + |x|_H^p + e^{-\delta_1 pt} |y|_H^p), \quad t \geq 0.$$

In particular, as shown in [7], this implies that there exists some  $\theta > 0$  such that for any  $a > 0$

$$(2.10) \quad \sup_{t \geq a} \mathbb{E}|v^{x,y}(t)|_{D((-A_2)^\theta)} \leq c_a(1 + |x|_H + |y|_H).$$

Now, for any  $x \in H$ , we denote by  $P_t^x$  the transition semigroup associated with problem (2.8), which is defined by

$$P_t^x \varphi(y) = \mathbb{E}\varphi(v^{x,y}(t)), \quad t \geq 0, \quad y \in H$$

for any  $\varphi \in B_b(H)$ . Due to (2.10), the family  $\{\mathcal{L}(v^{x,y}(t))\}_{t \geq a}$  is tight in  $\mathcal{P}(H, \mathcal{B}(H))$  and then by the Krylov–Bogoliubov theorem, there exists an invariant measure  $\mu^x$  for the semigroup  $P_t^x$ . Moreover, due to (2.9) for any  $p \geq 1$ , we have

$$(2.11) \quad \int_H |z|_H^p \mu^x(dz) \leq c_p(1 + |x|_H^p)$$

(for a proof see [8], Lemma 3.4).

As in [7], Theorem 7.4, it is possible to show that if  $\lambda$  is sufficiently large and/or  $L_{b_2}, L_{g_2}, \zeta_2$  and  $\kappa_2$  are sufficiently small, then there exist some  $c, \delta_2 > 0$  such that

$$(2.12) \quad \sup_{x \in H} \mathbb{E}|v^{x,y_1}(t) - v^{x,y_2}(t)|_H \leq ce^{-\delta_2 t} |y_1 - y_2|_H, \quad t \geq 0$$

for any  $y_1, y_2 \in H$ . In particular, this implies that  $\mu^x$  is the unique invariant measure for  $P_t^x$  and is strongly mixing. Moreover, by arguing as in [8], Theorem 3.5 and Remark 3.6, from (2.11) and (2.12), we have

$$(2.13) \quad \left| P_t^x \varphi(y) - \int_H \varphi(z) \mu^x(dz) \right| \leq c(1 + |x|_H + |y|_H) e^{-\delta_2 t} [\varphi]_{\text{Lip}(H)}$$

for any  $x, y \in H$  and  $\varphi \in \text{Lip}(H)$ . In particular, this implies the following fact.

LEMMA 2.3. *Under the above conditions, for any  $\varphi \in \text{Lip}(H)$ ,  $T > 0$ ,  $x, y \in H$  and  $t \geq 0$*

$$(2.14) \quad \mathbb{E} \left| \frac{1}{T} \int_t^{t+T} \varphi(v^{x,y}(s)) ds - \int_H \varphi(z) \mu^x(dz) \right| \leq \frac{c}{\sqrt{T}} (H_\varphi(x, y) + |\varphi(0)|),$$

where

$$(2.15) \quad H_\varphi(x, y) := [\varphi]_{\text{Lip}(H)}(1 + |x|_H + |y|_H).$$

PROOF. We have

$$\begin{aligned} & \mathbb{E} \left( \frac{1}{T} \int_t^{t+T} \varphi(v^{x,y}(s)) ds - \bar{\varphi}^x \right)^2 \\ &= \frac{1}{T^2} \int_t^{t+T} \int_t^{t+T} \mathbb{E}(\varphi(v^{x,y}(s)) - \bar{\varphi}^x)(\varphi(v^{x,y}(r)) - \bar{\varphi}^x) ds dr \\ &= \frac{2}{T^2} \int_t^{t+T} \int_r^{t+T} \mathbb{E}(\varphi(v^{x,y}(s)) - \bar{\varphi}^x)(\varphi(v^{x,y}(r)) - \bar{\varphi}^x) ds dr, \end{aligned}$$

where

$$\bar{\varphi}^x := \int_H \varphi(z) \mu^x(dz).$$

From the Markovianity of  $v^{x,y}(t)$ , for  $r \leq s$ , we have

$$\begin{aligned} & \mathbb{E}(\varphi(v^{x,y}(s)) - \bar{\varphi}^x)(\varphi(v^{x,y}(r)) - \bar{\varphi}^x) \\ &= \mathbb{E}[(\varphi(v^{x,y}(r)) - \bar{\varphi}^x) P_{s-r}(\varphi(v^{x,y}(r)) - \bar{\varphi}^x)], \end{aligned}$$

so that in view of (2.9) and (2.13),

$$\begin{aligned} & \mathbb{E} \left( \frac{1}{T} \int_t^{t+T} \varphi(v^{x,y}(s)) ds - \bar{\varphi}^x \right)^2 \\ & \leq \frac{c}{T^2} \int_t^{t+T} \int_r^{t+T} ([\varphi]_{\text{Lip}(H)} (\mathbb{E}|v^{x,y}(r)|_H^2)^{1/2} + |\varphi(0)| + |\bar{\varphi}^x|) \\ & \quad \times (\mathbb{E}[P_{s-r}\varphi(v^{x,y}(r)) - \bar{\varphi}^x]^2)^{1/2} ds dr \\ & \leq \frac{c}{T^2} (H_\varphi(x, y) + |\varphi(0)| + |\bar{\varphi}^x|) H_\varphi(x, y) \int_t^{t+T} \int_r^{t+T} e^{-\delta_2(s-r)} ds dr \\ & \leq \frac{c}{T} (H_\varphi(x, y) + |\varphi(0)| + |\bar{\varphi}^x|) H_\varphi(x, y) \end{aligned}$$

with  $H_\varphi(x, y)$  defined as in (2.15). As from (2.11), we have

$$|\bar{\varphi}^x| \leq [\varphi]_{\text{Lip}(H)}(1 + |x|_H) + |\varphi(0)|,$$

we can conclude that (2.14) holds.  $\square$

2.2. *The averaged coefficients.* In the next hypotheses, we introduce the coefficients of the *averaged* equation, and we give conditions which assure the convergence of the slow motion component  $u_\varepsilon$  to its solution. For the reaction coefficient, we assume the following condition.

HYPOTHESIS 3. There exists a Lipschitz-continuous mapping  $\bar{B}: H \rightarrow H$  such that for any  $T > 0, t \geq 0$  and  $x, y, h \in H$

$$(2.16) \quad \left| \frac{1}{T} \int_t^{t+T} \mathbb{E} \langle B_1(x, v^{x,y}(s)), h \rangle_H ds - \langle \bar{B}(x), h \rangle_H \right| \leq \alpha(T)(1 + |x|_H + |y|_H)|h|_H$$

for some function  $\alpha(T)$  such that

$$\lim_{T \rightarrow \infty} \alpha(T) = 0.$$

Concerning the diffusion coefficient, we assume the following condition.

HYPOTHESIS 4. There exists a Lipschitz-continuous mapping  $\bar{G}: H \rightarrow \mathcal{L}(L^\infty(D); H)$  such that for any  $T > 0, t \geq 0, x, y \in H$  and  $h, k \in L^\infty(D)$

$$(2.17) \quad \left| \frac{1}{T} \int_t^{t+T} \mathbb{E} \langle G_1(x, v^{x,y}(s))h, G_1(x, v^{x,y}(s))k \rangle_H ds - \langle \bar{G}(x)h, \bar{G}(x)k \rangle_H \right| \leq \alpha(T)(1 + |x|_H^2 + |y|_H^2)|h|_\infty |k|_\infty$$

for some  $\alpha(T)$  such that

$$\lim_{T \rightarrow \infty} \alpha(T) = 0.$$

**3. The averaged equation.** In this section, we describe some relevant situations in which Hypotheses 3 and 4 are verified and we give some notation and some results about the martingale problem and the mild solution for the averaged equation.

3.1. *The reaction coefficient  $\bar{B}$ .* For any fixed  $x, h \in H$ , the mapping

$$y \in H \mapsto \langle B_1(x, y), h \rangle_H \in \mathbb{R}$$

is Lipschitz-continuous. Then if we define

$$\bar{B}(x) := \int_H B_1(x, z) \mu^x(dz), \quad x \in H,$$

thanks to (2.13) we have that limit (2.16) holds, with  $\alpha(T) = c/\sqrt{T}$ .

Due to (2.16), for any  $x_1, x_2, y, h \in H$ , we have

$$\begin{aligned} & \langle \bar{B}_1(x_1) - \bar{B}_1(x_2), h \rangle_H \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T \langle B_1(x_1, v^{x_1, y}(s)) - B_1(x_2, v^{x_2, y}(s)), h \rangle_H ds. \end{aligned}$$

Then as the mapping  $B_1 : H \times H \rightarrow H$  is Lipschitz continuous, we have

$$\begin{aligned} & |\langle \bar{B}_1(x_1) - \bar{B}_1(x_2), h \rangle_H| \\ & \leq c|h|_H \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T (|x_1 - x_2|_H + \mathbb{E}|v^{x_1, y}(s) - v^{x_2, y}(s)|_H) ds \\ & = c|h|_H \left( |x_1 - x_2|_H + \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbb{E}|\rho(s)|_H ds \right), \end{aligned}$$

where  $\rho(t) := v^{x_1, y}(t) - v^{x_2, y}(t)$ , for any  $t \geq 0$ . In the next lemma, we show that under suitable conditions on the coefficients there exists some constant  $c > 0$  such that for any  $T > 0$

$$\frac{1}{T} \int_0^T \mathbb{E}|\rho(s)|_H ds \leq c|x_1 - x_2|_H.$$

Clearly, this implies the Lipschitz continuity of  $\bar{B}_1$ .

LEMMA 3.1. *Assume that*

$$\begin{aligned} (3.1) \quad & \frac{L_{b_2}}{\lambda} + L_{g_2} \left( \frac{\beta_2}{e} \right)^{\beta_2(\rho_2 - 2)/(2\rho_2)} \zeta_2^{(\rho_2 - 2)/(2\rho_2)} \kappa_2^{2/(2\rho_2)} \\ & \times \left( \frac{\rho_2}{\lambda(\rho_2 + 2)} \right)^{1/2 - \beta_2(\rho_2 - 2)/(2\rho_2)} =: M_0 < 1. \end{aligned}$$

Then under Hypotheses 1 and 2, there exists  $c > 0$  such that for any  $x_1, x_2, y \in H$  and  $t > 0$

$$\frac{1}{t} \int_0^t \mathbb{E}|v^{x_1, y}(s) - v^{x_2, y}(s)|_H ds \leq c|x_1 - x_2|_H.$$

PROOF. We set  $\rho(t) := v^{x_1, y}(t) - v^{x_2, y}(t)$  and define

$$\Gamma(t) := \int_0^t e^{(t-s)A_2} [G_2(x_1, v^{x_1, y}(s)) - G_2(x_2, v^{x_2, y}(s))] dw^{Q_2}(s)$$

and set  $\Lambda(t) := \rho(t) - \Gamma(t)$ . For any  $\eta \in (0, \lambda/2)$ , we can fix  $c_{1, \eta} > 0$  such that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |\Lambda(t)|_H^2 & \leq -\lambda |\Lambda(t)|_H^2 + (c|x_1 - x_2|_H + L_{b_2}|\rho(t)|_H) |\Lambda(t)|_H \\ & \leq -\left(\frac{\lambda}{2} - \eta\right) |\Lambda(t)|_H^2 + \frac{L_{b_2}^2}{2\lambda} |\rho(t)|_H^2 + c_{1, \eta} |x_1 - x_2|^2. \end{aligned}$$

This implies

$$|\Lambda(t)|_H^2 \leq \left(1 + \frac{2c_{1,\eta}}{\lambda - 2\eta}\right) |x_1 - x_2|^2 + \frac{L_{b_2}^2}{\lambda} \int_0^t e^{-(\lambda-2\eta)(t-s)} |\rho(s)|_H^2 ds,$$

so that for any  $\varepsilon > 0$  and  $\eta < \lambda/2$

$$\begin{aligned} \mathbb{E}|\rho(t)|_H^2 &\leq (1 + \varepsilon) \left(1 + \frac{2c_{1,\eta}}{\lambda - 2\eta}\right) |x_1 - x_2|^2 \\ (3.2) \quad &+ \frac{(1 + \varepsilon)L_{b_2}^2}{\lambda} \int_0^t e^{-(\lambda-2\eta)(t-s)} \mathbb{E}|\rho(s)|_H^2 ds \\ &+ \frac{(1 + \varepsilon)}{\varepsilon} \mathbb{E}|\Gamma(t)|_H^2. \end{aligned}$$

Thanks to Hypothesis 1, for any  $J \in \mathcal{L}(L^\infty(D), H) \cap \mathcal{L}(H, L^1(D))$ , with  $J = J^*$  and for any  $s \geq 0$ , we have

$$\begin{aligned} &\|e^{sA_2} J Q_2\|_2^2 \\ &= \sum_{k=1}^\infty \lambda_{2,k}^2 |e^{sA_2} J e_{2,k}|_H^2 \\ &\leq \left(\sum_{k=1}^\infty \lambda_{2,k}^{\rho_2} |e_{2,k}|_0^2\right)^{2/\rho_2} \left(\sum_{k=1}^\infty |e^{sA_2} J e_{2,k}|_H^{2\rho_2/(\rho_2-2)} |e_{2,k}|_H^{-4/(\rho_2-2)}\right)^{(\rho_2-2)/\rho_2} \\ &\leq \kappa_2^{2/\rho_2} \sup_{k \in \mathbb{N}} |e^{sA_2} J e_{2,k}|_H^{4/\rho_2} |e_{2,k}|_0^{-4/\rho_2} \left(\sum_{k=1}^\infty |e^{sA_2} J e_{2,k}|_H^2\right)^{(\rho_2-2)/\rho_2}. \end{aligned}$$

Hence, thanks to (2.4), we obtain

$$\begin{aligned} (3.3) \quad &\|e^{sA_2} J Q_2\|_2^2 \\ &\leq \kappa_2^{2/\rho_2} \|J\|_{\mathcal{L}(L^\infty(D), H)}^{4/\rho_2} e^{-4\lambda/\rho_2 s} \left(\sum_{k=1}^\infty |e^{sA_2} J e_{2,k}|_H^2\right)^{(\rho_2-2)/\rho_2}. \end{aligned}$$

We have

$$\begin{aligned} \sum_{k=1}^\infty |e^{sA_2} J e_{2,k}|_H^2 &= \sum_{k=1}^\infty \sum_{h=1}^\infty |\langle e^{sA_2} J e_{2,k}, e_{2,h} \rangle_H|^2 = \sum_{h=1}^\infty \sum_{k=1}^\infty |\langle e_{2,k}, J e^{sA_2} e_{2,h} \rangle_H|^2 \\ &= \sum_{h=1}^\infty |J e_{2,h}|_H^2 e^{-2\alpha_2 h s} \leq e^{-\lambda s} \|J\|_{\mathcal{L}(L^\infty(D), H)}^2 \sum_{h=1}^\infty |e_{2,h}|_0^2 e^{-\alpha_2 h s}. \end{aligned}$$

Then as for any  $\beta > 0$ ,

$$(3.4) \quad e^{-\alpha t} \leq \left(\frac{\beta}{e}\right)^\beta t^{-\beta} \alpha^{-\beta}, \quad \alpha, t > 0,$$

if we take  $\beta_2$  as in condition (2.2), we get

$$\begin{aligned} \sum_{k=1}^{\infty} |e^{sA_2} J e_{2,k}|_H^2 &\leq \left(\frac{\beta_2}{e}\right)^{\beta_2} s^{-\beta_2} e^{-\lambda s} \|J\|_{\mathcal{L}(L^\infty(D), H)}^2 \sum_{h=1}^{\infty} |e_{2,h}|_0^2 \alpha_{2,h}^{-\beta_2} \\ &\leq \left(\frac{\beta_2}{e}\right)^{\beta_2} \zeta_2 s^{-\beta_2} e^{-\lambda s} \|J\|_{\mathcal{L}(L^\infty(D), H)}^2 \end{aligned}$$

and from (3.3) we can conclude

$$\begin{aligned} \|e^{sA_2} J Q_2\|_2^2 &\leq \left(\frac{\beta_2}{e}\right)^{\beta_2(\rho_2-2)/\rho_2} \zeta_2^{(\rho_2-2)/\rho_2} \kappa_2^{2/\rho_2} s^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/\rho_2 s} \\ (3.5) \quad &\times \|J\|_{\mathcal{L}(L^\infty(D), H)}^2. \end{aligned}$$

This means that if we set

$$K_2 := \left(\frac{\beta_2}{e}\right)^{\beta_2(\rho_2-2)/\rho_2} \zeta_2^{(\rho_2-2)/\rho_2} \kappa_2^{2/\rho_2}$$

and if we take

$$J := G_2(x_1, v^{x_1, y}(s)) - G_2(x_2, v^{x_2, y}(s))$$

for any  $0 < \eta < \lambda/2$ ,

$$\begin{aligned} \mathbb{E}|\Gamma(t)|_H^2 &= \int_0^t \mathbb{E}\|e^{(t-s)A_2} [G_2(x_1, v^{x_1, y}(s)) - G_2(x_2, v^{x_2, y}(s))] Q_2\|_2^2 ds \\ &\leq K_2 \int_0^t (t-s)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/\rho_2(t-s)} \\ &\quad \times \mathbb{E}(c|x_1 - x_2|_H + L_{g_2}|\rho(s)|_H)^2 ds \\ &\leq \frac{c(1+\eta)}{\eta} |x_1 - x_2|_H^2 \\ &\quad + (1+\eta)L_{g_2}^2 K_2 \int_0^t (t-s)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/\rho_2(t-s)} \mathbb{E}|\rho(s)|_H^2 ds, \end{aligned}$$

last inequality following from the fact that, according to (2.3),  $\beta_2(\rho_2 - 2)/\rho_2 < 1$ .

Now, if we plug the inequality above into (3.2), for any  $\varepsilon > 0$  and  $0 < \eta < \lambda/2$ , we obtain

$$\begin{aligned} \mathbb{E}|\rho(t)|_H^2 &\leq (1+\varepsilon) \left(1 + \frac{2c_{1,\eta}}{\lambda - 2\eta}\right) |x_1 - x_2|^2 + c \left(\frac{1+\varepsilon}{\varepsilon}\right) \left(\frac{1+\eta}{\eta}\right) |x_1 - x_2|_H^2 \\ &\quad + \frac{(1+\varepsilon)L_{b_2}^2}{\lambda} \int_0^t e^{-(\lambda-2\eta)(t-s)} \mathbb{E}|\rho(s)|_H^2 ds \\ &\quad + \frac{(1+\varepsilon)}{\varepsilon} (1+\eta)L_{g_2}^2 K_2 \\ &\quad \times \int_0^t (t-s)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/\rho_2(t-s)} \mathbb{E}|\rho(s)|_H^2 ds \end{aligned}$$

and hence, if we integrate with respect to  $t$  both sides, from the Young inequality we get

$$\begin{aligned} & \int_0^t \mathbb{E}|\rho(s)|_H^2 ds \\ & \leq \left[ \left( 1 + \frac{2c_{1,\eta}}{\lambda - 2\eta} \right) + \frac{c}{\varepsilon} \left( \frac{1 + \eta}{\eta} \right) \right] (1 + \varepsilon)t|x_1 - x_2|_H^2 \\ & \quad + (1 + \varepsilon) \left[ \frac{L_{b_2}^2}{\lambda} \int_0^t e^{-(\lambda - 2\eta)s} ds \right. \\ & \quad \quad \left. + \frac{1 + \eta}{\varepsilon} L_{g_2}^2 K_2 \int_0^t s^{-\beta_2(\rho_2 - 2)/\rho_2} e^{-\lambda(\rho_2 + 2)/\rho_2 s} ds \right] \\ & \quad \times \int_0^t \mathbb{E}|\rho(s)|_H^2 ds \\ & \leq c_{\eta,\varepsilon} t|x_1 - x_1|_H^2 + M_{\eta,\varepsilon} \int_0^t \mathbb{E}|\rho(s)|_H^2 ds, \end{aligned}$$

where

$$(3.6) \quad \begin{aligned} M_{\eta,\varepsilon} := & (1 + \varepsilon) \left[ \frac{L_{b_2}^2}{\lambda(\lambda - 2\eta)} + \frac{1 + \eta}{\varepsilon} L_{g_2}^2 \left( \frac{\beta_2}{e} \right)^{\beta_2(\rho_2 - 2)/\rho_2} \right. \\ & \left. \times \zeta_2^{(\rho_2 - 2)/\rho_2} \kappa_2^{2/\rho_2} \left( \frac{\rho_2}{\lambda(\rho_2 + 2)} \right)^{1 - \beta_2(\rho_2 - 2)/\rho_2} \right]. \end{aligned}$$

Now, by taking the minimum over  $\varepsilon > 0$ , we get

$$\int_0^t \mathbb{E}|\rho(s)|_H^2 ds \leq c_{\eta,\bar{\varepsilon}} t|x_1 - x_1|_H^2 + M_{\eta}^2 \int_0^t \mathbb{E}|\rho(s)|_H^2 ds,$$

where

$$\begin{aligned} M_{\eta} := & \frac{L_{b_2}}{\sqrt{\lambda(\lambda - 2\eta)}} + \sqrt{1 + \eta} L_{g_2} \left( \frac{\beta_2}{e} \right)^{\beta_2(\rho_2 - 2)/(2\rho_2)} \\ & \times \zeta_2^{(\rho_2 - 2)/(2\rho_2)} \kappa_2^{2/(2\rho_2)} \left( \frac{\rho_2}{\lambda(\rho_2 + 2)} \right)^{1/2 - \beta_2(\rho_2 - 2)/2\rho_2}. \end{aligned}$$

Then as in (3.1) we have assumed that  $M_0 < 1$ , we can fix  $\bar{\eta} \in (0, \lambda/2)$  such that  $M_{\bar{\eta}} < 1$ , and hence

$$\int_0^t \mathbb{E}|\rho(s)|_H^2 ds \leq \frac{c_{\bar{\eta},\bar{\varepsilon}}}{1 - M_{\bar{\eta}}^2} t|x_1 - x_1|_H^2.$$

This implies

$$\frac{1}{t} \int_0^t \mathbb{E}|\rho(s)|_H ds \leq \left( \frac{c_{3,\bar{\eta}}}{1 - M_{\bar{\eta}}^2} \right)^{1/2} |x_1 - x_1|_H$$

and the proof of the lemma is complete.  $\square$

3.2. *The diffusion coefficient  $\bar{G}$ .* If we assume that the function  $g_1 : D \times \mathbb{R}^2 \rightarrow \mathbb{R}$  is uniformly bounded, the mapping  $G_1$  is well defined from  $H$  into  $\mathcal{L}(H)$ . Moreover, for any fixed  $x, h, k \in H$  the mapping

$$z \in H \mapsto \langle G_1(x, z)h, G_1(x, z)k \rangle_H \in \mathbb{R},$$

is Lipschitz continuous. Thus, under the assumptions described above, if we take

$$(3.7) \quad \langle S(x)h, k \rangle_H = \int_H \langle G_1(x, z)h, G_1(x, z)k \rangle_H \mu^x(dz),$$

we have that  $S : H \rightarrow \mathcal{L}(H)$  and, due to (2.13), for any  $T > 0, t \geq 0$  and  $x, y, h, k \in H$

$$(3.8) \quad \left| \frac{1}{T} \int_t^{t+T} \mathbb{E} \langle G_1(x, v^{x,y}(s))h, G_1(x, v^{x,y}(s))k \rangle_H ds - \langle S(x)h, k \rangle_H \right| \leq \alpha(T)(1 + |x|_H^2 + |y|_H^2)|h|_H|k|_H$$

for some function  $\alpha(T)$  going to zero as  $T \uparrow \infty$ .

It is immediate to check that  $S(x) = S(x)^*$  and  $S(x) \geq 0$ , for any  $x \in H$ . Then as is well known, there exists an operator  $\bar{G}(x) \in \mathcal{L}(H)$  such that  $\bar{G}(x)^2 = S(x)$ . If we assume that there exists  $\delta > 0$  such that

$$\inf_{\substack{\xi \in D \\ \sigma \in \mathbb{R}^2}} g_1(\xi, \sigma) \geq \delta,$$

we have that  $S(x) \geq \delta^2$ , and hence  $\bar{G}(x) \geq \delta$ . In particular,  $\bar{G}(x)$  is invertible and

$$\|\bar{G}(x)^{-1}\|_{\mathcal{L}(H)} \leq \frac{1}{\delta}.$$

Next, we notice that for any  $x_1, x_2 \in H$

$$(3.9) \quad S(x_1)S(x_2) = S(x_2)S(x_1).$$

Actually, according to (3.7) for any  $h, k \in H$ ,

$$\begin{aligned} \langle S(x_1)S(x_2)h, k \rangle_H &= \int_H \langle G_1(x_1, z)S(x_2)h, G_1(x_1, z)k \rangle_H \mu^{x_1}(dz) \\ &= \int_H \int_H \langle G_1^2(x_2, w)h, G_1^2(x_1, z)k \rangle_H \mu^{x_2}(dw) \mu^{x_1}(dz) \\ &= \int_H \int_H \langle G_1^2(x_1, z)h, G_1^2(x_2, w)k \rangle_H \mu^{x_1}(dz) \mu^{x_2}(dw) \\ &= \langle S(x_2)S(x_1)h, k \rangle_H. \end{aligned}$$

In particular, from (3.9) for any  $x_1, x_2 \in H$ , we have

$$(3.10) \quad \bar{G}(x_1)\bar{G}(x_2) = \bar{G}(x_2)\bar{G}(x_1).$$

Now, as  $g_1(\xi, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is bounded and Lipschitz-continuous, uniformly with respect to  $\xi \in D$ , we have that  $g_1^2(\xi, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is Lipschitz-continuous as well, uniformly with respect to  $\xi \in D$ . This implies that for any  $x_1, x_2, y \in H$ ,  $h \in L^\infty(D)$  and  $k \in H$

$$\begin{aligned} & | \langle [G_1^2(x_1, v^{x_1, y}(s)) - G_1^2(x_2, v^{x_2, y}(s))]h, k \rangle_H | \\ & \leq c(|x_1 - x_2|_H + |v^{x_1, y}(s) - v^{x_2, y}(s)|_H) |h|_\infty |k|_H, \end{aligned}$$

so that according to (3.8),

$$\begin{aligned} & | \langle (S(x_1) - S(x_2))h, k \rangle_H | \\ & \leq c|h|_\infty |k|_H \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T (|x_1 - x_2|_H + \mathbb{E}|v^{x_1, y}(s) - v^{x_2, y}(s)|_H) ds \\ & = c|h|_\infty |k|_H \left( |x_1 - x_2|_H + \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbb{E}|v^{x_1, y}(s) - v^{x_2, y}(s)|_H ds \right). \end{aligned}$$

Then thanks to Lemma 3.1, we can conclude that  $S : H \rightarrow \mathcal{L}(L^\infty(D), H)$  is Lipschitz-continuous.

This implies that  $\bar{G} : H \rightarrow \mathcal{L}(L^\infty(D), H)$  is Lipschitz-continuous as well. Actually, thanks to (3.10) and to the fact that  $\bar{G}(x_1) + \bar{G}(x_2)$  is invertible, for any  $h \in L^\infty(D)$  and  $k \in H$ ,

$$(3.11) \quad \langle [\bar{G}(x_1) - \bar{G}(x_2)]h, k \rangle_H = \langle [S(x_1) - S(x_2)]h, [\bar{G}(x_1) + \bar{G}(x_2)]^{-1}k \rangle_H.$$

Then as

$$(3.12) \quad \| [\bar{G}(x_1) + \bar{G}(x_2)]^{-1} \|_{\mathcal{L}(H)} \leq \frac{1}{2\delta},$$

we obtain

$$| \langle [\bar{G}(x_1) - \bar{G}(x_2)]h, k \rangle_H | \leq c|x_1 - x_2|_H |h|_\infty \frac{1}{2\delta} |k|_H$$

and this implies the Lipschitz-continuity of  $\bar{G} : H \rightarrow \mathcal{L}(L^\infty(D), H)$ .

We conclude by showing that the operator  $\bar{G}$  introduced in Hypothesis 4 satisfies a suitable Hilbert–Schmidt property which assures the well-posedness of the stochastic convolution

$$\int_0^t e^{(t-s)A_1} \bar{G}(u(s)) dw^{\mathcal{Q}_1}(s), \quad t \geq 0$$

in  $L^p(\Omega; C([0, T]; H))$ , for any  $p \geq 1$  and  $T > 0$  and for any process  $u \in C([0, T]; L^p(\Omega; H))$ .

LEMMA 3.2. *Assume Hypotheses 1, 2 and 4. Then for any  $t > 0$  and  $x_1, x_2 \in H$ , we have*

$$\| e^{tA_1} [\bar{G}(x_1) - \bar{G}(x_2)] \mathcal{Q}_1 \|_2 \leq c(t) |x_1 - x_2|_H t^{-\beta_1(\rho_1 - 2)/(2\rho_1)}$$

for some continuous increasing function  $c(t)$ .

PROOF. According to Hypothesis 1, we have

$$\begin{aligned} & \|e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]Q_1\|_2^2 \\ &= \sum_{k=1}^{\infty} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]Q_1 e_{1,k}|_H^2 \\ &\leq \left( \sum_{k=1}^{\infty} \lambda_{1,k}^{\rho_1} |e_{1,k}|_{\infty}^2 \right)^{2/\rho_1} \\ &\quad \times \left( \sum_{k=1}^{\infty} |e_{1,k}|_{\infty}^{-4/(\rho_1-2)} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^{2\rho_1/(\rho_1-2)} \right)^{(\rho_1-2)/\rho_1} \\ &\leq c \sup_{k \in \mathbb{N}} |e_{1,k}|_{\infty}^{-4/\rho_1} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^{4/\rho_1} \\ &\quad \times \left( \sum_{k=1}^{\infty} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^2 \right)^{(\rho_1-2)/\rho_1} \end{aligned}$$

and then as  $\bar{G} : H \rightarrow \mathcal{L}(L^\infty(D), H)$  is Lipschitz-continuous, we conclude

$$\begin{aligned} (3.13) \quad & \|e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]Q_1\|_2^2 \\ & \leq c(t) |x_1 - x_2|_H^{4/\rho_1} \left( \sum_{k=1}^{\infty} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^2 \right)^{(\rho_1-2)/\rho_1} \end{aligned}$$

for some continuous increasing function  $c(t)$ .

Now, by using again the Lipschitz-continuity of  $\bar{G} : H \rightarrow \mathcal{L}(L^\infty(D), H)$ , we have

$$\begin{aligned} \sum_{k=1}^{\infty} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^2 &= \sum_{k=1}^{\infty} \sum_{h=1}^{\infty} |\langle e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}, e_{1,h} \rangle_H|^2 \\ &= \sum_{h=1}^{\infty} |[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,h}|_H^2 e^{-2t\alpha_{1,h}} \\ &\leq c |x_1 - x_2|_H^2 \sum_{h=1}^{\infty} e^{-2t\alpha_{1,h}} |e_{1,h}|_H^2 \end{aligned}$$

and then, if we take  $\beta_1$  as in Hypothesis 1, we obtain

$$\begin{aligned} & \sum_{k=1}^{\infty} |e^{tA_1}[\bar{G}(x_1) - \bar{G}(x_2)]e_{1,k}|_H^2 \\ & \leq c |x_1 - x_2|_H^2 t^{-\beta_1} \sum_{h=1}^{\infty} \alpha_{1,h}^{-\beta_1} |e_{1,h}|_H^2 \leq c |x_1 - x_2|_H^2 t^{-\beta_1}. \end{aligned}$$

Thanks to (3.13), this implies our thesis.  $\square$

3.3. *Martingale problem and mild solution of the averaged equation.* Since the mappings  $\bar{B}: H \rightarrow H$  and  $\bar{G}: H \rightarrow \mathcal{L}(L^\infty(D); H)$  are both Lipschitz-continuous and Lemma 3.2 holds, for any initial datum  $x \in H$  the averaged equation

$$(3.14) \quad du(t) = [A_1u(t) + \bar{B}(u(t))]dt + \bar{G}(u(t))dw^{\mathcal{Q}_1}(t), \quad u(0) = x,$$

admits a unique mild solution  $\bar{u}$  in  $L^p(\Omega, C([0, T]; H))$ , for any  $p \geq 1$  and  $T > 0$  (for a proof, see, e.g., [6], Section 3). This means that there exists a unique adapted process  $\bar{u} \in L^p(\Omega, C([0, T]; H))$  such that for any  $t \leq T$

$$\begin{aligned} \bar{u}(t) &= e^{tA_1}x + \int_0^t e^{(t-s)A_1} \bar{B}(\bar{u}(s)) ds \\ &\quad + \int_0^t e^{(t-s)A_1} \bar{G}(\bar{u}(s)) dw^{\mathcal{Q}_1}(s) \end{aligned}$$

or equivalently,

$$\begin{aligned} \langle \bar{u}(t), h \rangle_H &= \langle x, h \rangle_H + \int_0^t [\langle \bar{u}(s), A_1h \rangle_H + \langle \bar{B}(\bar{u}(s)), h \rangle_H] ds \\ &\quad + \int_0^t \langle \bar{G}(\bar{u}(s)) dw^{\mathcal{Q}_1}(s), h \rangle_H \end{aligned}$$

for any  $h \in D(A_1)$ .

Now, we recall the notion of martingale problem with parameters  $(x, A_1, \bar{B}, \bar{G}, \mathcal{Q}_1)$ . For any fixed  $x \in H$ , we denote by  $C_x([0, T]; H)$  the space of continuous functions  $\omega: [0, T] \rightarrow H$  such that  $\omega(0) = x$  and we denote by  $\eta(t)$  the canonical process on  $C_x([0, T]; H)$ , which is defined by

$$\eta(t)(\omega) = \omega(t), \quad t \in [0, T].$$

Moreover, we denote by  $\mathcal{E}_t$  the canonical filtration  $\sigma(\eta(s), s \leq t)$ , for  $t \in [0, T]$ , and by  $\mathcal{E}$  the canonical  $\sigma$ -algebra  $\sigma(\eta(s), s \leq T)$ .

DEFINITION 3.3. A function  $\varphi: H \rightarrow \mathbb{R}$  is a *regular cylindrical function* associated with the operator  $A_1$  if there exist  $k \in \mathbb{N}$ ,  $f \in C_c^\infty(\mathbb{R}^k)$ ,  $a_1, \dots, a_k \in H$  and  $N \in \mathbb{N}$  such that

$$\varphi(x) = f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H), \quad x \in H,$$

where  $P_N$  is the projection of  $H$  onto  $\text{span}\{e_{1,1}, \dots, e_{1,N}\}$  and  $\{e_{1,n}\}_{n \in \mathbb{N}}$  is the orthonormal basis diagonalizing  $A_1$  and introduced in Hypothesis 1.

In what follows, we shall denote the set of all regular cylindrical functions by  $\mathcal{R}(H)$ . For any  $\varphi \in \mathcal{R}(H)$  and  $x \in H$ , we define

$$\begin{aligned}
 \mathcal{L}_{av}\varphi(x) &:= \frac{1}{2} \operatorname{Tr}[(\bar{G}(x)Q_1)^* D^2\varphi(x)\bar{G}(x)Q_1] \\
 &\quad + \langle A_1 D\varphi(x), x \rangle_H + \langle D\varphi(x), \bar{B}(x) \rangle_H \\
 &= \frac{1}{2} \sum_{i,j=1}^k D_{ij}^2 f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\
 (3.15) \quad &\quad \times \langle \bar{G}(x)Q_1 P_N a_i, \bar{G}(x)Q_1 P_N a_j \rangle_H \\
 &\quad + \sum_{i=1}^k D_i f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\
 &\quad \times (\langle x, A_1 P_N a_i \rangle_H + \langle \bar{B}(x), P_N a_i \rangle_H).
 \end{aligned}$$

$\mathcal{L}_{av}$  is the Kolmogorov operator associated with the averaged equation (3.14). Notice that the expression above is meaningful, as for any  $i = 1, \dots, k$

$$Q_1 P_N a_i = \sum_{k=1}^N \lambda_{1,k} \langle a_i, e_{1,k} \rangle_H e_{1,k} \in L^\infty(D)$$

and

$$A_1 P_N a_i = - \sum_{k=1}^N \alpha_{1,k} \langle a_i, e_{1,k} \rangle_H e_{1,k} \in H.$$

DEFINITION 3.4. A probability measure  $\mathbb{Q}$  on  $(C_x([0, T]; H), \mathcal{E})$  is a solution of the martingale problem with parameters  $(x, A_1, \bar{B}, \bar{G}, Q_1)$  if the process

$$\varphi(\eta(t)) - \int_0^t \mathcal{L}_{av}\varphi(\eta(s)) ds, \quad t \in [0, T]$$

is an  $\mathcal{E}_t$ -martingale on  $(C_x([0, T]; H), \mathcal{E}, \mathbb{Q})$ , for any  $\varphi \in \mathcal{R}(H)$ .

As the coefficients  $\bar{B}$  and  $\bar{G}$  are Lipschitz-continuous, the solution  $\mathbb{Q}$  to the martingale problem with parameters  $(x, A_1, \bar{B}, \bar{G}, Q_1)$  exists, is unique and coincides with  $\mathcal{L}(\bar{u})$  (to this purpose see [9], Chapter 8, and also [26], Theorems 5.9 and 5.10).

**4. A priori bounds for the solution of system (1.6).** In the present section, we prove uniform estimates, with respect to  $\varepsilon \in (0, 1]$ , for the solution  $u_\varepsilon$  of the slow motion equation and for the solution  $v_\varepsilon$  of the fast motion equation in system (1.6). As a consequence, we will obtain the tightness of the family  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0,1]}$  in  $C([0, T]; H)$ , for any  $T > 0$ .

In what follows, for the sake of simplicity, we denote by  $|\cdot|_\theta$  the norm  $|\cdot|_{D((-A_1)^\theta)}$ . Moreover, for any  $\varepsilon > 0$ , we denote

$$(4.1) \quad \Gamma_{1,\varepsilon}(t) := \int_0^t e^{(t-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s), \quad t \geq 0.$$

LEMMA 4.1. *Under Hypotheses 1 and 2, there exists  $\bar{\theta} > 0$  and  $\bar{p} \geq 1$  such that for any  $\varepsilon > 0, T > 0, p > \bar{p}$  and  $\theta \in [0, \bar{\theta}]$*

$$(4.2) \quad \mathbb{E} \sup_{t \leq T} |\Gamma_{1,\varepsilon}(t)|_\theta^p \leq c_{T,p,\theta} \int_0^T (1 + \mathbb{E}|u_\varepsilon(r)|_H^p + \mathbb{E}|v_\varepsilon(r)|_H^p) ds$$

for some positive constant  $c_{T,p,\theta}$  which is independent of  $\varepsilon > 0$ .

PROOF. By using a factorization argument, for any  $\alpha \in (0, 1/2)$ , we have

$$\Gamma_{1,\varepsilon}(t) = c_\alpha \int_0^t (t-s)^{\alpha-1} e^{(t-s)A_1} Y_{\varepsilon,\alpha}(s) ds,$$

where

$$Y_{\varepsilon,\alpha}(s) := \int_0^s (s-r)^{-\alpha} e^{(s-r)A_1} G_{1,\varepsilon}(s) dw^{\mathcal{Q}_1}(r)$$

and

$$G_{1,\varepsilon}(s) := G_1(u_\varepsilon(s), v_\varepsilon(s)).$$

For any  $p > 1/\alpha$  and  $\theta > 0$ , we have

$$(4.3) \quad \begin{aligned} \sup_{s \leq t} |\Gamma_{1,\varepsilon}(s)|_\theta^p &\leq c_{\alpha,p} \left( \int_0^t s^{(\alpha-1)p/(p-1)} ds \right)^{p-1} \int_0^t |Y_{\varepsilon,\alpha}(s)|_\theta^p ds \\ &= c_{\alpha,p} t^{\alpha p-1} \int_0^t |Y_{\varepsilon,\alpha}(s)|_\theta^p ds. \end{aligned}$$

According to the Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned} \mathbb{E} \int_0^t |Y_{\varepsilon,\alpha}(s)|_\theta^p ds &\leq c_p \int_0^t \mathbb{E} \left( \int_0^s (s-r)^{-2\alpha} \|(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) \mathcal{Q}_1\|_2^2 dr \right)^{p/2} ds. \end{aligned}$$

By the same arguments as those used in the proof of Lemma 3.1, we have

$$\begin{aligned} &\|(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) \mathcal{Q}_1\|_2^2 \\ &\leq \sup_{k \in \mathbb{N}} |(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) e_{1,k}|_H^{4/\rho_1} |e_{1,k}|_H^{-4/\rho_1} \end{aligned}$$

$$\begin{aligned} & \times \kappa_1^{2/\rho_1} \left( \sum_{k=1}^{\infty} |(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) e_{1,k}|_H^2 \right)^{(\rho_1-2)/\rho_1} \\ & \leq c_\theta \kappa_1^{2/\rho_1} (s-r)^{-4\theta/\rho_1} \|G_{1,\varepsilon}(s)\|_{\mathcal{L}(L^\infty(D),H)}^{4/\rho_1} \\ & \quad \times \left( \sum_{k=1}^{\infty} |(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) e_{1,k}|_H^2 \right)^{(\rho_1-2)/\rho_1}. \end{aligned}$$

By proceeding again as in the proof of Lemma 3.1 we have

$$\begin{aligned} & \sum_{k=1}^{\infty} |(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) e_{1,k}|_H^2 \\ & \leq \|G_{1,\varepsilon}(s)\|_{\mathcal{L}(L^\infty(D),H)}^2 \sum_{k=1}^{\infty} |e_{1,k}|_0^2 \alpha_{1,k}^\theta e^{-\alpha_{1,k}(s-r)} \end{aligned}$$

and then thanks to (3.4), we get

$$\begin{aligned} & \sum_{k=1}^{\infty} |(-A_1)^\theta e^{(s-r)A_1} G_{1,\varepsilon}(s) e_{1,k}|_H^2 \\ & \leq \left( \frac{\beta_1 + \theta}{e} \right)^{\beta_1 + \theta} \zeta_1 (s-r)^{-(\beta_1 + \theta)} \|G_{1,\varepsilon}(s)\|_{\mathcal{L}(L^\infty(D),H)}^2. \end{aligned}$$

Therefore, if we set

$$K_{1,\theta} := c_\theta \left( \frac{\beta_1 + \theta}{e} \right)^{(\beta_1 + \theta)(\rho_1 - 2)/\rho_1} \zeta_1^{(\rho_1 - 2)/\rho_1} \kappa_1^{2/\rho_1}$$

and fix  $\bar{\theta} > 0$  such that

$$\frac{\beta_1(\rho_1 - 2) + \bar{\theta}(\rho_1 + 2)}{\rho_1} < 1$$

for any  $\theta \in [0, \bar{\theta}]$ , we have

$$\begin{aligned} & \mathbb{E} \int_0^t |Y_{\varepsilon,\alpha}(s)|_H^p ds \\ & \leq c_p K_{1,\theta}^{p/2} \int_0^t \mathbb{E} \left( \int_0^s (s-r)^{-(2\alpha + (\beta_1(\rho_1 - 2) + \theta(\rho_1 + 2))/\rho_1)} \right. \\ & \quad \left. \times \|G_{1,\varepsilon}(r)\|_{\mathcal{L}(L^\infty(D),H)}^2 dr \right)^{p/2} ds. \end{aligned}$$

Hence, if we choose  $\bar{\alpha} > 0$ , such that

$$2\bar{\alpha} + \frac{\beta_1(\rho_1 - 2) + \bar{\theta}(\rho_1 + 2)}{\rho_1} < 1$$

and  $p > \bar{p} := 1/\bar{\alpha}$ , by the Young inequality this yields for  $t \in [0, T]$

$$\begin{aligned} & \mathbb{E} \int_0^t |Y_{\varepsilon, \bar{\alpha}}(s)|_{\theta}^p ds \\ & \leq c_p K_{1, \theta}^{p/2} \left( \int_0^t s^{-(2\bar{\alpha} + (\beta_1(\rho_1 - 2) + \theta(\rho_1 + 2))/\rho_1)} ds \right)^{p/2} \\ & \quad \times \mathbb{E} \int_0^t \|G_{1, \varepsilon}(s)\|_{\mathcal{L}(L^\infty(D), H)}^p ds \\ & \leq c_{T, p} \int_0^t (1 + \mathbb{E}|u_\varepsilon(s)|_H^p + \mathbb{E}|v_\varepsilon(s)|_H^p) ds. \end{aligned}$$

Thanks to (4.3), this implies (4.2).  $\square$

Now, we can prove the first a priori bounds for the solution  $u_\varepsilon$  of the slow motion equation and for the solution  $v_\varepsilon$  of the fast motion equation in system (1.6).

PROPOSITION 4.2. *Under Hypotheses 1 and 2, for any  $T > 0$  and  $p \geq 1$ , there exists a positive constant  $c(p, T)$  such that for any  $x, y \in H$  and  $\varepsilon \in (0, 1]$*

$$(4.4) \quad \mathbb{E} \sup_{t \in [0, T]} |u_\varepsilon(t)|_H^p \leq c(p, T)(1 + |x|_H^p + |y|_H^p)$$

and

$$(4.5) \quad \int_0^T \mathbb{E}|v_\varepsilon(t)|_H^p dt \leq c(p, T)(1 + |x|_H^p + |y|_H^p).$$

Moreover, there exists some  $c_T > 0$  such that

$$(4.6) \quad \sup_{t \in [0, T]} \mathbb{E}|v_\varepsilon(t)|_H^2 \leq c_T(1 + |x|_H^2 + |y|_H^2).$$

PROOF. Let  $\varepsilon > 0$  and  $x, y \in H$  be fixed once for all and let  $\Gamma_{1, \varepsilon}(t)$  be the process defined in (4.1). If we set  $\Lambda_{1, \varepsilon}(t) := u_\varepsilon(t) - \Gamma_{1, \varepsilon}(t)$ , we have

$$\frac{d}{dt} \Lambda_{1, \varepsilon}(t) = A_1 \Lambda_{1, \varepsilon}(t) + B_1(\Lambda_{1, \varepsilon}(t) + \Gamma_{1, \varepsilon}(t), v_\varepsilon(t)), \quad \Lambda_{1, \varepsilon}(0) = x$$

and then for any  $p \geq 2$  we have

$$\begin{aligned} \frac{1}{p} \frac{d}{dt} |\Lambda_{1, \varepsilon}(t)|_H^p &= \langle A_1 \Lambda_{1, \varepsilon}(t), \Lambda_{1, \varepsilon}(t) \rangle_H |\Lambda_{1, \varepsilon}(t)|_H^{p-2} \\ & \quad + \langle B_1(\Lambda_{1, \varepsilon}(t) + \Gamma_{1, \varepsilon}(t), v_\varepsilon(t)) \\ & \quad \quad - B_1(\Gamma_{1, \varepsilon}(t), v_\varepsilon(t)), \Lambda_{1, \varepsilon}(t) \rangle_H |\Lambda_{1, \varepsilon}(t)|_H^{p-2} \\ & \quad + \langle B_1(\Gamma_{1, \varepsilon}(t), v_\varepsilon(t)), \Lambda_{1, \varepsilon}(t) \rangle_H |\Lambda_{1, \varepsilon}(t)|_H^{p-2} \\ & \leq c_p |\Lambda_{1, \varepsilon}(t)|_H^p + c_p |B_1(\Gamma_{1, \varepsilon}(t), v_\varepsilon(t))|_H^p \\ & \leq c_p |\Lambda_{1, \varepsilon}(t)|_H^p + c_p (1 + |\Gamma_{1, \varepsilon}(t)|_H^p + |v_\varepsilon(t)|_H^p). \end{aligned}$$

This implies that

$$|\Lambda_{1,\varepsilon}(t)|_H^p \leq e^{c_p t} |x|_H^p + c_p \int_0^t e^{c_p(t-s)} (1 + |\Gamma_{1,\varepsilon}(s)|_H^p + |v_\varepsilon(s)|_H^p) ds,$$

so that, for any  $t \in [0, T]$ ,

$$\begin{aligned} |u_\varepsilon(t)|_H^p &\leq c_p |\Gamma_{1,\varepsilon}(t)|_H^p + c_p e^{c_p t} |x|_H^p \\ &\quad + c_p \int_0^t e^{c_p(t-s)} (1 + |\Gamma_{1,\varepsilon}(s)|_H^p + |v_\varepsilon(s)|_H^p) ds \\ &\leq c_{T,p} \left( 1 + |x|_H^p + \sup_{s \leq t} |\Gamma_{1,\varepsilon}(s)|_H^p + \int_0^t |v_\varepsilon(s)|_H^p ds \right). \end{aligned}$$

According to (4.2) (with  $\theta = 0$ ), we obtain

$$\begin{aligned} \mathbb{E} \sup_{s \leq t} |u_\varepsilon(s)|_H^p &\leq c_{T,p} (1 + |x|_H^p) + c_{T,p} \int_0^t \mathbb{E} |v_\varepsilon(s)|_H^p ds \\ &\quad + c_{T,p} \int_0^t \left( 1 + \mathbb{E} \sup_{r \leq s} |u_\varepsilon(r)|_H^p \right) ds \end{aligned}$$

and hence by comparison,

$$(4.7) \quad \mathbb{E} \sup_{s \leq t} |u_\varepsilon(s)|_H^p \leq c_{T,p} \left( 1 + |x|_H^p + \int_0^t \mathbb{E} |v_\varepsilon(s)|_H^p ds \right).$$

Now, we have to estimate

$$\int_0^t \mathbb{E} |v_\varepsilon(s)|_H^p ds.$$

If we define

$$\Gamma_{2,\varepsilon}(t) := \frac{1}{\sqrt{\varepsilon}} \int_0^t e^{(t-s)/\varepsilon A_2} G_2(u_\varepsilon(s), v_\varepsilon(s)) dw^{Q_2}(s)$$

and set  $\Lambda_{2,\varepsilon}(t) := v_\varepsilon(t) - \Gamma_{2,\varepsilon}(t)$ , we have

$$\frac{d}{dt} \Lambda_{2,\varepsilon}(t) = \frac{1}{\varepsilon} [A_2 \Lambda_{2,\varepsilon}(t) + B_2(u_\varepsilon(t), \Lambda_{2,\varepsilon}(t) + \Gamma_{2,\varepsilon}(t))], \quad \Lambda_{2,\varepsilon}(0) = y.$$

Hence, as before, for any  $p \geq 1$ , we have

$$\begin{aligned} \frac{1}{p} \frac{d}{dt} |\Lambda_{2,\varepsilon}(t)|_H^p &= \frac{1}{\varepsilon} \langle A_2 \Lambda_{2,\varepsilon}(t), \Lambda_{2,\varepsilon}(t) \rangle_H |\Lambda_{2,\varepsilon}(t)|_H^{p-2} \\ &\quad + \frac{1}{\varepsilon} \langle B_2(u_\varepsilon(t), \Lambda_{2,\varepsilon}(t) + \Gamma_{2,\varepsilon}(t)) \\ &\quad \quad - B_1(u_\varepsilon(t), \Gamma_{2,\varepsilon}(t)), \Lambda_{2,\varepsilon}(t) \rangle_H |\Lambda_{2,\varepsilon}(t)|_H^{p-2} \\ &\quad + \frac{1}{\varepsilon} \langle B_2(u_\varepsilon(t), \Gamma_{2,\varepsilon}(t)), \Lambda_{2,\varepsilon}(t) \rangle_H |\Lambda_{2,\varepsilon}(t)|_H^{p-2} \\ &\leq -\frac{\lambda - L_{b_2}}{2\varepsilon} |\Lambda_{2,\varepsilon}(t)|_H^p + \frac{c_p}{\varepsilon} (1 + |u_\varepsilon(t)|_H^p + |\Gamma_{2,\varepsilon}(t)|_H^p). \end{aligned}$$

By comparison this yields

$$\begin{aligned}
 |v_\varepsilon(t)|_H^p &\leq c_p |\Lambda_{2,\varepsilon}(t)|_H^p + c_p |\Gamma_{2,\varepsilon}(t)|_H^p \\
 &\leq c_p e^{-p(\lambda-L_{b_2})/(2\varepsilon)t} |y|_H^p \\
 (4.8) \quad &+ \frac{c_p}{\varepsilon} \int_0^t e^{-p(\lambda-L_{b_2})/(2\varepsilon)(t-s)} (1 + |u_\varepsilon(s)|_H^p + |\Gamma_{2,\varepsilon}(s)|_H^p) ds \\
 &+ c_p |\Gamma_{2,\varepsilon}(t)|_H^p.
 \end{aligned}$$

Therefore, by integrating with respect to  $t$ , we easily obtain

$$(4.9) \quad \int_0^t |v_\varepsilon(s)|_H^p ds \leq c_p \left( \varepsilon |y|_H^p + \int_0^t |\Gamma_{2,\varepsilon}(s)|_H^p ds + \int_0^t |u_\varepsilon(s)|_H^p ds + 1 \right).$$

According to the Burkholder–Davis–Gundy inequality and to (3.5), we have

$$\begin{aligned}
 (4.10) \quad &\mathbb{E}|\Gamma_{2,\varepsilon}(s)|_H^p \\
 &\leq c_p \varepsilon^{-p/2} \mathbb{E} \left( \int_0^s \|e^{(s-r)/\varepsilon A_2} G_2(u_\varepsilon(r), v_\varepsilon(r)) Q_2\|_2^2 dr \right)^{p/2} \\
 &\leq c_p K_2^{p/2} \varepsilon^{-p/2} \mathbb{E} \left( \int_0^s \left( \frac{s-r}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(s-r)} \right. \\
 &\quad \left. \times \|G_2(u_\varepsilon(r), v_\varepsilon(r))\|_{\mathcal{L}(L^\infty(D), H)}^2 dr \right)^{p/2} \\
 &\leq c_p K_2^{p/2} \varepsilon^{-p/2} \mathbb{E} \left( \int_0^s \left( \frac{s-r}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(s-r)} \right. \\
 &\quad \left. \times (1 + |u_\varepsilon(r)|_H^2 + |v_\varepsilon(r)|_H^{2\gamma}) dr \right)^{p/2},
 \end{aligned}$$

so that

$$(4.11) \quad \int_0^t \mathbb{E}|\Gamma_{2,\varepsilon}(s)|_H^p ds \leq c_p \int_0^t (1 + \mathbb{E}|u_\varepsilon(s)|_H^p + \mathbb{E}|v_\varepsilon(s)|_H^{p\gamma}) ds.$$

Due to (4.9), this allows to conclude

$$\int_0^t \mathbb{E}|v_\varepsilon(s)|_H^p ds \leq c_p \left( \varepsilon |y|_H^p + \int_0^t (1 + \mathbb{E}|u_\varepsilon(s)|_H^p) ds + \int_0^t \mathbb{E}|v_\varepsilon(s)|_H^{p\gamma} ds + 1 \right)$$

and then as  $\gamma$  is assumed to be strictly less than 1, if  $\varepsilon \in (0, 1]$  and  $t \in [0, T]$ , we obtain

$$\int_0^t \mathbb{E}|v_\varepsilon(s)|_H^p ds \leq \frac{1}{2} \int_0^t \mathbb{E}|v_\varepsilon(s)|_H^p ds + c_p |y|_H^p + c_p \int_0^t \mathbb{E}|u_\varepsilon(s)|_H^p ds + c_{p,T}.$$

This yields

$$(4.12) \quad \int_0^t \mathbb{E}|v_\varepsilon(s)|_H^p ds \leq c_p |y|_H^p + c_p \int_0^t \mathbb{E} \sup_{r \leq s} |u_\varepsilon(r)|_H^p ds + c_{p,T}.$$

Hence, if we plug (4.12) into (4.7), we get

$$\mathbb{E} \sup_{s \leq t} |u_\varepsilon(s)|_H^p \leq c_{T,p}(1 + |x|_H^p + |y|_H^p) + c_{T,p} \int_0^t \mathbb{E} \sup_{r \leq s} |u_\varepsilon(r)|_H^p ds$$

and from the Gronwall lemma (4.4) follows. Now, in view of estimates (4.4) and (4.11), from (4.9), we obtain (4.5).

Finally, let us prove (4.6). From (4.10), with  $p = 2$ , we get

$$\sup_{t \leq T} \mathbb{E} |\Gamma_{2,\varepsilon}(t)|_H^2 \leq c_2 |y|_H^2 + c_2 \left( 1 + \sup_{t \leq T} \mathbb{E} |u_\varepsilon(t)|_H^2 + \sup_{t \leq T} \mathbb{E} |v_\varepsilon(t)|_H^{2\gamma} \right)$$

and then if we substitute in (4.8), we obtain

$$\mathbb{E} |v_\varepsilon(t)|_H^2 \leq c_2 \left( 1 + |y|_H^2 + \sup_{t \leq T} \mathbb{E} |u_\varepsilon(t)|_H^2 \right) + c_2 \sup_{t \leq T} \mathbb{E} |v_\varepsilon(t)|_H^{2\gamma}.$$

As  $\gamma < 1$ , for any  $\eta > 0$ , we can fix  $c_\eta > 0$  such that

$$c_2 \sup_{t \leq T} \mathbb{E} |v_\varepsilon(t)|_H^{2\gamma} \leq \eta \sup_{t \leq T} \mathbb{E} |v_\varepsilon(t)|_H^2 + c_\eta.$$

Therefore, if we take  $\eta \leq 1/2$ , we obtain

$$\frac{1}{2} \sup_{t \leq T} \mathbb{E} |v_\varepsilon(t)|_H^2 \leq c_2 \left( 1 + |y|_H^2 + \sup_{t \leq T} \mathbb{E} |u_\varepsilon(t)|_H^2 \right)$$

and (4.6) follows from (4.4).  $\square$

Next, we prove uniform bounds for  $u_\varepsilon$  in  $L^\infty(0, T; D((-A_1)^\alpha))$ , for some  $\alpha > 0$ .

**PROPOSITION 4.3.** *Under Hypotheses 1 and 2, there exists  $\bar{\alpha} > 0$  such that for any  $T > 0$ ,  $p \geq 1$ ,  $x \in D((-A_1)^\alpha)$ , with  $\alpha \in [0, \bar{\alpha})$ , and  $y \in H$*

$$(4.13) \quad \sup_{\varepsilon \in (0, 1]} \mathbb{E} \sup_{t \leq T} |u_\varepsilon(t)|_\alpha^p \leq c_{T,\alpha,p}(1 + |x|_\alpha^p + |y|_H^p)$$

for some positive constant  $c_{T,\alpha,p}$ .

**PROOF.** Assume that  $x \in D((-A_1)^\alpha)$ , for some  $\alpha \geq 0$ . We have

$$\begin{aligned} u_\varepsilon(t) &= e^{tA_1} x + \int_0^t e^{(t-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \\ &\quad + \int_0^t e^{(t-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{Q_1}(s). \end{aligned}$$

If  $\alpha < 1/2, t \leq T$  and  $p \geq 2$

$$\begin{aligned} & \left| \int_0^t e^{(t-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \right|_\alpha^p \\ & \leq c_{p,\alpha} \left( \int_0^t (t-s)^{-\alpha} |B_1(u_\varepsilon(s), v_\varepsilon(s))|_H ds \right)^p \\ & \leq c_{p,\alpha} \left( \int_0^t (t-s)^{-\alpha} (1 + |u_\varepsilon(s)|_H + |v_\varepsilon(s)|_H) ds \right)^p \\ & \leq c_{p,\alpha} \left( 1 + \sup_{s \leq T} |u_\varepsilon(s)|_H^p \right) T^{(1-\alpha)p} \\ & \quad + c_{p,\alpha} \left( \int_0^T s^{-2\alpha} ds \right)^{p/2} \left( \int_0^T |v_\varepsilon(s)|_H^p ds \right) T^{(p-2)/2}, \end{aligned}$$

so that, thanks to (4.4) and (4.5),

$$(4.14) \quad \mathbb{E} \sup_{t \leq T} \left| \int_0^t e^{(t-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \right|_\alpha^p \leq c_{T,\alpha,p} (1 + |x|_H^p + |y|_H^p).$$

Concerning the stochastic term  $\Gamma_{1,\varepsilon}(t)$ , due to Lemma 4.1 and to (4.4), there exists  $\bar{\theta} > 0$  such that for any  $\alpha \leq \bar{\theta}$  and  $p \geq 1$

$$(4.15) \quad \mathbb{E} \sup_{t \leq T} |\Gamma_{1,\varepsilon}(t)|_\alpha^p \leq c_{T,\alpha,p} (1 + |x|_H^p + |y|_H^p).$$

Hence, if we choose  $\bar{\alpha} := \bar{\theta} \wedge 1/2$ , thanks to (4.14) and (4.15), for any  $p \geq 2$  and  $\alpha < \bar{\alpha}$  we have

$$\begin{aligned} \mathbb{E} \sup_{t \leq T} |u_\varepsilon(t)|_\alpha^p & \leq \sup_{t \leq T} |e^{tA_1} x|_\alpha^p + \mathbb{E} \sup_{t \leq T} \left| \int_0^t e^{(t-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \right|_\alpha^p \\ & \quad + \mathbb{E} \sup_{t \leq T} |\Gamma_{1,\varepsilon}(t)|_\alpha^p \\ & \leq c_{T,\alpha,p} (1 + |x|_\alpha^p + |y|_H^p). \end{aligned} \quad \square$$

Next, we prove uniform bounds for the increments of the mapping  $t \in [0, T] \mapsto u_\varepsilon(t) \in H$ .

**PROPOSITION 4.4.** *Under Hypotheses 1 and 2, for any  $\alpha > 0$ , there exists  $\beta(\alpha) > 0$  such that for any  $T > 0, p \geq 2, x \in D((-A_1)^\alpha)$  and  $y \in H$  it holds*

$$(4.16) \quad \sup_{\varepsilon \in (0,1]} \mathbb{E} |u_\varepsilon(t) - u_\varepsilon(s)|_H^p \leq c_{T,\alpha,p} |t - s|^{\beta(\alpha)p} (|x|_\alpha^p + |y|_H^p + 1),$$

$s, t \in (0, T].$

PROOF. For any  $t, h \geq 0$ , with  $t, t + h \in [0, T]$ , we have

$$\begin{aligned} u_\varepsilon(t + h) - u_\varepsilon(t) &= (e^{hA_1} - I)u_\varepsilon(t) \\ &\quad + \int_t^{t+h} e^{(t+h-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \\ &\quad + \int_t^{t+h} e^{(t+h-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s). \end{aligned}$$

In view of (4.13), if we fix  $\alpha \in [0, \bar{\alpha}]$  and  $p \geq 1$ , we have

$$(4.17) \quad \mathbb{E}|(e^{hA_1} - I)u_\varepsilon(t)|_H^p \leq c_p h^{\alpha p} \mathbb{E}|u_\varepsilon(t)|_\alpha^p \leq c_{T,\alpha,p} h^{\alpha p} (1 + |x|_\alpha^p + |y|_H^p).$$

In view of (4.4) and (4.5),

$$\begin{aligned} (4.18) \quad &\mathbb{E} \left| \int_t^{t+h} e^{(t+h-s)A_1} B_1(u_\varepsilon(s), v_\varepsilon(s)) ds \right|_H^p \\ &\leq ch^{p-1} \int_t^{t+h} (1 + \mathbb{E}|u_\varepsilon(s)|_H^p + \mathbb{E}|v_\varepsilon(s)|_H^p) ds \\ &\leq c_T h^p \left( 1 + \sup_{s \leq T} \mathbb{E}|u_\varepsilon(s)|_H^p \right) + ch^{p-1} \int_0^T \mathbb{E}|v_\varepsilon(s)|_H^p ds \\ &\leq c_{T,p} (1 + |x|_H^p + |y|_H^p) h^{p-1}. \end{aligned}$$

Finally, for the stochastic term, by using (3.5), for any  $t \leq T$  and  $p \geq 1$ , we have

$$\begin{aligned} &\mathbb{E} \left| \int_t^{t+h} e^{(t+h-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s) \right|_H^p \\ &\leq c_p \mathbb{E} \left( \int_t^{t+h} \|e^{(t+h-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) \mathcal{Q}_1\|_2^2 ds \right)^{p/2} \\ &\leq c_p K_1^{p/2} \mathbb{E} \left( \int_t^{t+h} (t+h-s)^{-\beta_2(\rho_2-2)/\rho_2} \right. \\ &\quad \left. \times \|G_1(u_\varepsilon(s), v_\varepsilon(s))\|_{\mathcal{L}(L^\infty(D), H)}^2 ds \right)^{p/2}. \end{aligned}$$

Then, if we take  $\bar{p} \geq 1$  such that

$$\frac{\beta_2(\rho_2 - 2)}{\rho_2} \frac{\bar{p}}{\bar{p} - 2} < 1$$

for any  $p \geq \bar{p}$ , we have

$$\begin{aligned} &\mathbb{E} \left| \int_t^{t+h} e^{(t+h-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s) \right|_H^p \\ &\leq c_{T,p} h^{(p-2)/2 - \beta_2(\rho_2-2)/\rho_2 p/2} \int_0^T (1 + \mathbb{E}|u_\varepsilon(s)|_H^p + \mathbb{E}|v_\varepsilon(s)|_H^p) ds \end{aligned}$$

and, thanks to (4.4) and (4.5), we conclude

$$(4.19) \quad \mathbb{E} \left| \int_t^{t+h} e^{(t+h-s)A_1} G_1(u_\varepsilon(s), v_\varepsilon(s)) dw^{\mathcal{Q}_1}(s) \right|_H^p \leq c_{T,p} (1 + |x|_H^p + |y|_H^p) h^{(1-2/\bar{p}-\beta_2(\rho_2-2)/\rho_2)p/2}.$$

Therefore, collecting together (4.17), (4.18) and (4.19), we obtain

$$\begin{aligned} & \mathbb{E}|u_\varepsilon(t+h) - u_\varepsilon(t)|_H^p \\ & \leq c_{T,\alpha,p} h^{\alpha p} (1 + |x|_\alpha^p + |y|_H^p) \\ & \quad + c_{T,p} (h^{(1-2/\bar{p}-\beta_2(\rho_2-2)/\rho_2)p/2} + h^{p-1}) (1 + |x|_H^p + |y|_H^p) \end{aligned}$$

and, as we are assuming  $|h| \leq 1$ , (4.16) follows for any  $p \geq \bar{p}$  by taking

$$\beta(\alpha) := \min \left\{ \alpha, \frac{1}{2} \left( 1 - \frac{2}{\bar{p}} - \frac{\beta_2(\rho_2 - 2)}{\rho_2} \right) \right\}.$$

Estimate (4.16) for  $p < \bar{p}$  follows from the Hölder inequality.  $\square$

As a consequence of Propositions 4.3 and 4.4, we have the following fact.

**COROLLARY 4.5.** *Under Hypotheses 1 and 2, for any  $T > 0$ ,  $x \in D((-A_1)^\alpha)$ , with  $\alpha > 0$ , and  $y \in H$  the family  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0,1]}$  is tight in  $C([0, T]; H)$ .*

**PROOF.** Let  $\alpha > 0$  be fixed and let  $x \in D((-A_1)^\alpha)$  and  $y \in H$ . According to (4.16), in view of the Garcia–Rademich–Rumsey theorem, there exists  $\bar{\beta} > 0$  such that for any  $p \geq 1$

$$\sup_{\varepsilon \in (0,1]} \mathbb{E}|u_\varepsilon|_{C^{\bar{\beta}}([0,T];H)}^p \leq c_{T,p} (1 + |x|_\alpha^p + |y|_H^p).$$

Due to Proposition 4.3, this implies that for any  $\eta > 0$  we can find  $R_\eta > 0$  such that

$$\mathbb{P}(u_\varepsilon \in K_{R_\eta}) \geq 1 - \eta, \quad \varepsilon \in (0, 1],$$

where, by the Ascoli–Arzelà theorem,  $K_{R_\eta}$  is the compact subset of  $C([0, T]; H)$  defined by

$$K_{R_\eta} := \left\{ u \in C([0, T]; H) : |u|_{C^{\bar{\beta}}([0,T];H)} + \sup_{t \in [0,T]} |u(t)|_\alpha \leq R_\eta \right\}.$$

This implies that the family of probability measures  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0,1]}$  is tight in  $C([0, T]; H)$ .  $\square$

We conclude this section by noticing that with arguments analogous to those used in the proof of Propositions 4.2, 4.3 and 4.4, we can obtain a priori bounds also for the conditional second momenta of the  $H$ -norms of  $u_\varepsilon$  and  $v_\varepsilon$ .

PROPOSITION 4.6. *Assume Hypotheses 1 and 2. Then for any  $0 \leq s < t \leq T$  and any  $\varepsilon \in (0, 1]$  the following facts holds.*

1. *There exists  $\bar{\alpha} > 0$  such that for any  $x \in D((-A_1)^\alpha)$ , with  $\alpha \in [0, \bar{\alpha}]$ , and  $y \in H$*

$$\mathbb{E}(|u_\varepsilon(t)|_\alpha^2 | \mathcal{F}_s) \leq c_{T,\alpha}(1 + |u_\varepsilon(s)|_\alpha^2 + |v_\varepsilon(s)|_H^2), \quad \mathbb{P}\text{-a.s.}$$

*for some constant  $c_{T,\alpha}$  independent of  $\varepsilon$ .*

2. *For any  $x, y \in H$*

$$(4.20) \quad \mathbb{E}(|v_\varepsilon(t)|_H^2 | \mathcal{F}_s) \leq c_T(1 + |u_\varepsilon(s)|_H^2 + |v_\varepsilon(s)|_H^2), \quad \mathbb{P}\text{-a.s.}$$

*for some constant  $c_T$  independent of  $\varepsilon$ .*

3. *For any  $\alpha > 0$ , there exists  $\beta(\alpha) > 0$  such that for any  $x \in D((-A_1)^\alpha)$  and  $y \in H$*

$$\mathbb{E}(|u_\varepsilon(t) - u_\varepsilon(s)|_H^2 | \mathcal{F}_s) \leq c_{T,\alpha}(t - s)^{2\beta(\alpha)}(|u_\varepsilon(s)|_\alpha^2 + |v_\varepsilon(s)|_H^2 + 1), \quad \mathbb{P}\text{-a.s.}$$

*for some constant  $c_{T,\alpha}$  independent of  $\varepsilon$ .*

**5. The key lemma.** We introduce the Kolmogorov operator associated with the slow motion equation, with frozen fast component, by setting for any  $\varphi \in \mathcal{R}(H)$  and  $x, y \in H$

$$\begin{aligned} \mathcal{L}_{sl}\varphi(x, y) &= \frac{1}{2} \text{Tr}[Q_1 G_1(x, y) D^2 \varphi(x) G_1(x, y) Q_1] \\ &\quad + \langle A_1 D \varphi(x), x \rangle_H + \langle D \varphi(x), B_1(x, y) \rangle_H \\ (5.1) \quad &= \frac{1}{2} \sum_{i,j=1}^k D_{ij}^2 f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\ &\quad \times \langle G_1(x, y) Q_{1,Na_i}, G_1(x, y) Q_{1,Na_j} \rangle_H \\ &\quad + \sum_{i=1}^k D_i f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\ &\quad \times (\langle x, A_{1,Na_i} \rangle_H + \langle B_1(x, y), P_N a_i \rangle_H). \end{aligned}$$

LEMMA 5.1. *Assume Hypotheses 1–4 and fix  $x \in D((-A_1)^\alpha)$ , with  $\alpha > 0$ , and  $y \in H$ . Then for any  $\varphi \in \mathcal{R}(H)$  and  $0 \leq t_1 < t_2 \leq T$ ,*

$$(5.2) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_{t_1}^{t_2} \mathbb{E}(\mathcal{L}_{sl}\varphi(u_\varepsilon(r), v_\varepsilon(r)) - \mathcal{L}_{av}\varphi(u_\varepsilon(r)) | \mathcal{F}_{t_1}) dr \right| = 0.$$

PROOF. By using the Khasminskii idea introduced in [16], we realize a partition of  $[0, T]$  into intervals of size  $\delta_\varepsilon > 0$ , to be chosen later on, and for each  $\varepsilon > 0$

we denote by  $\hat{v}_\varepsilon(t)$  the solution of the problem

$$\begin{aligned}
 \hat{v}_\varepsilon(t) &= e^{(t-k\delta_\varepsilon)A_2/\varepsilon} v_\varepsilon(k\delta_\varepsilon) + \frac{1}{\varepsilon} \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} B_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) ds \\
 &+ \frac{1}{\sqrt{\varepsilon}} \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} G_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) dw^{Q_2}(s),
 \end{aligned}
 \tag{5.3}$$

$t \in [k\delta_\varepsilon, (k+1)\delta_\varepsilon)$

for  $k = 0, \dots, [T/\delta_\varepsilon]$ . In what follows, we shall set  $\zeta_\varepsilon := \delta_\varepsilon/\varepsilon$ .

*Step 1.* Now, we prove that there exist  $\kappa_1, \kappa_2 > 0$  such that if we set

$$\zeta_\varepsilon = \left( \log \frac{1}{\varepsilon^{\kappa_2}} \right)^{\kappa_1},$$

then

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T]} \mathbb{E} |\hat{v}_\varepsilon(t) - v_\varepsilon(t)|_H^2 = 0.
 \tag{5.4}$$

If we fix  $k = 0, \dots, [T/\delta_\varepsilon]$  and take  $t \in [k\delta_\varepsilon, (k+1)\delta_\varepsilon)$ , we have

$$\begin{aligned}
 v_\varepsilon(t) &= e^{(t-k\delta_\varepsilon)A_2/\varepsilon} v_\varepsilon(k\delta_\varepsilon) + \frac{1}{\varepsilon} \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} B_2(u_\varepsilon(s), v_\varepsilon(s)) ds \\
 &+ \frac{1}{\sqrt{\varepsilon}} \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} G_2(u_\varepsilon(s), v_\varepsilon(s)) dw^{Q_2}(s),
 \end{aligned}$$

so that

$$\begin{aligned}
 &\mathbb{E} |\hat{v}_\varepsilon(t) - v_\varepsilon(t)|_H^2 \\
 &\leq \frac{2\delta_\varepsilon}{\varepsilon^2} \int_{k\delta_\varepsilon}^t \mathbb{E} |B_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - B_2(u_\varepsilon(s), v_\varepsilon(s))|_H^2 ds \\
 &+ \frac{2}{\varepsilon} \mathbb{E} \left| \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} [G_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - G_2(u_\varepsilon(s), v_\varepsilon(s))] dw^{Q_2}(s) \right|_H^2.
 \end{aligned}$$

For the first term, we have

$$\begin{aligned}
 &\frac{\delta_\varepsilon}{\varepsilon^2} \int_{k\delta_\varepsilon}^t \mathbb{E} |B_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - B_2(u_\varepsilon(s), v_\varepsilon(s))|_H^2 ds \\
 &\leq \frac{c}{\varepsilon} \int_{k\delta_\varepsilon}^t \zeta_\varepsilon (\mathbb{E} |u_\varepsilon(k\delta_\varepsilon) - u_\varepsilon(s)|_H^2 + \mathbb{E} |\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2) ds.
 \end{aligned}
 \tag{5.5}$$

For the second term, by proceeding as in the proof of Proposition 4.2, we obtain

$$\begin{aligned}
 &\mathbb{E} \left| \int_{k\delta_\varepsilon}^t e^{(t-s)A_2/\varepsilon} [G_2(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - G_2(u_\varepsilon(s), v_\varepsilon(s))] dw^{Q_2}(s) \right|_H^2 \\
 &\leq c \int_{k\delta_\varepsilon}^t \left( \frac{t-s}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(t-s)} \\
 &\quad \times (\mathbb{E} |u_\varepsilon(k\delta_\varepsilon) - u_\varepsilon(s)|_H^2 + \mathbb{E} |\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2) ds.
 \end{aligned}
 \tag{5.6}$$

In view of (4.16), we have

$$\begin{aligned} & \frac{1}{\varepsilon} \int_{k\delta_\varepsilon}^t \left[ \left( \frac{t-s}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(t-s)} + \zeta_\varepsilon \right] \mathbb{E} |u_\varepsilon(k\delta_\varepsilon) - u_\varepsilon(s)|_H^2 ds \\ & \leq \frac{c_T}{\varepsilon} \int_{k\delta_\varepsilon}^t \left[ \left( \frac{t-s}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(t-s)} + \zeta_\varepsilon \right] \\ & \quad \times (s - k\delta_\varepsilon)^{2\beta(\alpha)} ds (1 + |x|_\alpha^2 + |y|_H^2) \\ & \leq c_T \delta_\varepsilon^{2\beta(\alpha)} (1 + \zeta_\varepsilon^2) (1 + |x|_\alpha^2 + |y|_H^2). \end{aligned}$$

Moreover,

$$\begin{aligned} & \frac{1}{\varepsilon} \int_{k\delta_\varepsilon}^t \left[ \left( \frac{t-s}{\varepsilon} \right)^{-\beta_2(\rho_2-2)/\rho_2} e^{-\lambda(\rho_2+2)/(\varepsilon\rho_2)(t-s)} + \zeta_\varepsilon \right] \mathbb{E} |\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2 ds \\ & \leq \frac{C}{\varepsilon} (\varepsilon^{\beta_2(\rho_2-2)/\rho_2} + \zeta_\varepsilon \delta_\varepsilon^{\beta_2(\rho_2-2)/\rho_2}) \\ & \quad \times \int_{k\delta_\varepsilon}^t (t-s)^{-\beta_2(\rho_2-2)/\rho_2} \mathbb{E} |\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2 ds. \end{aligned}$$

Then thanks to (5.5) and (5.6), we obtain

$$\begin{aligned} \mathbb{E} |\hat{v}_\varepsilon(t) - v_\varepsilon(t)|_H^2 & \leq c_T \delta_\varepsilon^{2\beta(\alpha)} (1 + \zeta_\varepsilon^2) (1 + |x|_\alpha^2 + |y|_H^2) \\ (5.7) \quad & + c \varepsilon^{\beta_2(\rho_2-2)/\rho_2-1} (1 + \zeta_\varepsilon^{1+\beta_2(\rho_2-2)/\rho_2}) \\ & \times \int_{k\delta_\varepsilon}^t (t-s)^{-\beta_2(\rho_2-2)/\rho_2} \mathbb{E} |\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2 ds. \end{aligned}$$

Now, we recall the following simple fact (for a proof, see, e.g., [11]).

LEMMA 5.2. *If  $M, L, \theta$  are positive constants and  $g$  is a nonnegative function such that*

$$g(t) \leq M + L \int_{t_0}^t (t-s)^{\theta-1} g(s) ds, \quad t \geq t_0,$$

then

$$g(t) \leq M + \frac{ML}{\theta} (t-t_0)^\theta + L^2 \int_0^1 r^{\theta-1} (1-r)^{\theta-1} dr \int_{t_0}^t (t-s)^{2\theta-1} g(s) ds, \quad t \geq t_0.$$

Notice that if we iterate the lemma above  $n$ -times, we find

$$g(t) \leq c_{1,n,\theta} M (1 + L^{2^n-1} (t-t_0)^{2^n-1}) + c_{2,n,\theta} L^{2^n} \int_{t_0}^t (t-s)^{2^n\theta-1} g(s) ds$$

for some positive constants  $c_{1,n,\theta}$  and  $c_{2,n,\theta}$ .

If we apply  $\bar{n}$ -times Lemma 5.2 to (5.7), with  $\bar{n} \in \mathbb{N}$  such that

$$2^{\bar{n}}\theta := 2^{\bar{n}}\left(1 - \frac{\beta_2(\rho_2 - 2)}{\rho_2}\right) > 1,$$

we get

$$\begin{aligned} & \mathbb{E}|\hat{v}_\varepsilon(t) - v_\varepsilon(t)|_H^2 \\ & \leq c_T \delta_\varepsilon^{2\beta(\alpha)} (1 + \zeta_\varepsilon^2) (1 + |x|_\alpha^2 + |y|_H^2) \\ & \quad \times (1 + \varepsilon^{-(2^{\bar{n}}-1)\theta} (1 + \zeta_\varepsilon^{(2^{\bar{n}}-1)(2-\theta)}) \delta_\varepsilon^{(2^{\bar{n}}-1)\theta}) \\ & \quad + c\varepsilon^{-2^{\bar{n}}\theta} (1 + \zeta_\varepsilon^{2^{\bar{n}}(2-\theta)}) \int_{k\delta_\varepsilon}^t (t-s)^{2^{\bar{n}}\theta-1} \mathbb{E}|\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2 ds \\ & \leq c_T \delta_\varepsilon^{2\beta(\alpha)} (1 + |x|_\alpha^2 + |y|_H^2) (1 + \zeta_\varepsilon^{2^{\bar{n}+1}}) \\ & \quad + \frac{c}{\delta_\varepsilon} (1 + \zeta_\varepsilon^{2^{\bar{n}+1}}) \int_{k\delta_\varepsilon}^t \mathbb{E}|\hat{v}_\varepsilon(s) - v_\varepsilon(s)|_H^2 ds. \end{aligned}$$

From the Gronwall lemma this yields

$$\mathbb{E}|\hat{v}_\varepsilon(t) - v_\varepsilon(t)|_H^2 \leq c_T \delta_\varepsilon^{2\beta(\alpha)} (1 + |x|_\alpha^2 + |y|_H^2) (1 + \zeta_\varepsilon^{2^{\bar{n}+1}}) \exp(c\zeta_\varepsilon^{2^{\bar{n}+1}}).$$

Now, since we have

$$\exp(c\zeta_\varepsilon^{2^{\bar{n}+1}}) = \exp\left(c\left(\log \frac{1}{\varepsilon^{\kappa_2}}\right)^{\kappa_1 2^{\bar{n}+1}}\right),$$

if we take  $\kappa_1 := 2^{-(\bar{n}+1)}$  and  $\kappa_2 < 2\beta(\alpha)c^{-1}$ , we conclude that (5.4) holds.

Moreover, as for  $t \in [k\delta_\varepsilon, (k+1)\delta_\varepsilon]$  the process  $\hat{v}_\varepsilon(t)$  is the mild solution of the problem

$$\begin{aligned} dv(t) &= \frac{1}{\varepsilon} [A_2 v(t) + B_2(u_\varepsilon(k\delta_\varepsilon), v(t))] dt + \frac{1}{\sqrt{\varepsilon}} G_2(u_\varepsilon(k\delta_\varepsilon), v(t)) dw^{Q_2}(t), \\ & v(k\delta_\varepsilon) = v_\varepsilon(k\delta_\varepsilon) \end{aligned}$$

with the same arguments as those used to prove (4.6) and (4.20), we obtain

$$(5.8) \quad \sup_{t \in [0, T]} \mathbb{E}|\hat{v}_\varepsilon(t)|_H^2 \leq c_T (1 + |x|_H^2 + |y|_H^2)$$

and, for any  $t \in [k\delta_\varepsilon, (k+1)\delta_\varepsilon]$ ,

$$\mathbb{E}(|\hat{v}_\varepsilon(t)|_H^2 | \mathcal{F}_{k\delta_\varepsilon}) \leq c(1 + |u_\varepsilon(k\delta_\varepsilon)|_H^2 + |v_\varepsilon(k\delta_\varepsilon)|_H^2), \quad \mathbb{P}\text{-a.s.}$$

Step 2. Now, we fix  $\varphi \in \mathcal{R}(H)$ . We can assume that

$$\varphi(x) = f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H)$$

for some  $f \in C_c^\infty(\mathbb{R}^k)$  and  $k, N \in \mathbb{N}$ . According to (3.15) and (5.1), we have

$$\mathcal{L}_{sl}\varphi(u_\varepsilon(r), v_\varepsilon(r)) - \mathcal{L}_{av}\varphi(u_\varepsilon(r)) = \frac{1}{2} \sum_{i,j=1}^k I_{ij}^\varepsilon(r) + \sum_{i=1}^k J_i^\varepsilon(r),$$

where

$$\begin{aligned} I_{ij}^\varepsilon &:= D_{ij}^2 f(\langle u_\varepsilon, P_N a_1 \rangle_H, \dots, \langle u_\varepsilon, P_N a_k \rangle_H) \\ &\quad \times (\langle G_1(u_\varepsilon, v_\varepsilon) Q_{1,Na_i}, G_1(u_\varepsilon, v_\varepsilon) Q_{1,Na_j} \rangle_H \\ &\quad - \langle \bar{G}(u_\varepsilon) Q_{1,Na_i}, \bar{G}(u_\varepsilon) Q_{1,Na_j} \rangle_H) \end{aligned}$$

and

$$J_i^\varepsilon = D_i f(\langle u_\varepsilon, P_N a_1 \rangle_H, \dots, \langle u_\varepsilon, P_N a_k \rangle_H) \langle B_1(u_\varepsilon, v_\varepsilon) - \bar{B}(u_\varepsilon), P_N a_i \rangle_H.$$

Hence, if we prove that for any  $i, j = 1, \dots, k$ ,

$$(5.9) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_{t_1}^{t_2} \mathbb{E}(I_{ij}^\varepsilon(r) | \mathcal{F}_{t_1}) dr \right| = 0$$

and

$$(5.10) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_{t_1}^{t_2} \mathbb{E}(J_i^\varepsilon(r) | \mathcal{F}_{t_1}) dr \right| = 0,$$

we immediately get (5.2).

We have

$$\int_{t_1}^{t_2} \mathbb{E}(I_{ij}^\varepsilon(r) | \mathcal{F}_{t_1}) dr = \sum_{l=1}^3 \int_{t_1}^{t_2} \mathbb{E}(I_{l,ij}^\varepsilon(r) | \mathcal{F}_{t_1}) dr,$$

where

$$\begin{aligned} I_{1,ij}^\varepsilon(r) &:= D_{ij}^2 f(\langle u_\varepsilon(r), P_N a_1 \rangle_H, \dots, \langle u_\varepsilon(r), P_N a_k \rangle_H) \\ &\quad \times \langle G_1(u_\varepsilon(r), v_\varepsilon(r)) Q_{1,Na_i}, G_1(u_\varepsilon(r), v_\varepsilon(r)) Q_{1,Na_j} \rangle_H \\ &\quad - D_{ij}^2 f(\langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_1 \rangle_H, \dots, \langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_k \rangle_H) \\ &\quad \times \langle G_1(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(r)) Q_{1,Na_i}, \\ &\quad \quad G_1(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(r)) Q_{1,Na_j} \rangle_H \end{aligned}$$

and

$$\begin{aligned} I_{2,ij}^\varepsilon(r) &:= D_{ij}^2 f(\langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_1 \rangle_H, \dots, \langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_k \rangle_H) \\ &\quad \times (\langle G_1(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(r)) Q_{1,Na_i}, G_1(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(r)) Q_{1,Na_j} \rangle_H \\ &\quad - \langle \bar{G}(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)) Q_{1,Na_i}, \bar{G}(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)) Q_{1,Na_j} \rangle_H) \end{aligned}$$

and

$$\begin{aligned}
 I_{3,ij}^\varepsilon(r) &:= D_{ij}^2 f(\langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_1 \rangle_H, \dots, \langle u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon), P_N a_k \rangle_H) \\
 &\quad \times \langle \bar{G}(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)) Q_{1,Na_i}, \bar{G}(u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)) Q_{1,Na_j} \rangle_H \\
 &\quad - D_{ij}^2 f(\langle u_\varepsilon(r), P_N a_1 \rangle_H, \dots, \langle u_\varepsilon(r), P_N a_k \rangle_H) \\
 &\quad \times \langle \bar{G}(u_\varepsilon(r)) Q_{1,Na_i}, \bar{G}(u_\varepsilon(r)) Q_{1,Na_j} \rangle_H.
 \end{aligned}$$

It is immediate to check that

$$\begin{aligned}
 &|I_{1,ij}^\varepsilon(r)| + |I_{3,ij}^\varepsilon(r)| \\
 &\quad \leq c(|u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon) - u_\varepsilon(r)|_H + |v_\varepsilon(r) - \hat{v}_\varepsilon(r)|_H) \\
 &\quad \quad \times (1 + |u_\varepsilon(r)|_H^2 + |v_\varepsilon(r)|_H^2 + |u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)|_H + |\hat{v}_\varepsilon(r)|_H),
 \end{aligned}$$

so that

$$\begin{aligned}
 &\left( \mathbb{E} \int_{t_1}^{t_2} [|I_{1,ij}^\varepsilon(r)| + |I_{3,ij}^\varepsilon(r)|] dr \right)^2 \\
 &\quad \leq c \int_{t_1}^{t_2} [\mathbb{E}|u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon) - u_\varepsilon(r)|_H^2 + \mathbb{E}|v_\varepsilon(r) - \hat{v}_\varepsilon(r)|_H^2] dr \\
 &\quad \quad \times \int_{t_1}^{t_2} [1 + \mathbb{E}|u_\varepsilon(r)|_H^4 + \mathbb{E}|v_\varepsilon(r)|_H^4 + \mathbb{E}|u_\varepsilon([r/\delta_\varepsilon]\delta_\varepsilon)|_H^2 + \mathbb{E}|\hat{v}_\varepsilon(r)|_H^2] dr.
 \end{aligned}$$

According to (4.16), (4.4), (4.5) and (5.8), we conclude

$$\begin{aligned}
 &\left( \mathbb{E} \int_{t_1}^{t_2} [|I_{1,ij}^\varepsilon(r)| + |I_{3,ij}^\varepsilon(r)|] dr \right)^2 \\
 &\quad \leq c_T \left( \delta_\varepsilon^{2\beta(\alpha)} + \sup_{t \in [0, T]} \mathbb{E}|v_\varepsilon(t) - \hat{v}_\varepsilon(t)|_H^2 \right) (1 + |x|_H^4 + |y|_H^4 + |x|_\alpha^2),
 \end{aligned}$$

so that due to (5.4),

$$(5.11) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \int_{t_1}^{t_2} [|I_{1,ij}^\varepsilon(r)| + |I_{3,ij}^\varepsilon(r)|] dr = 0.$$

Next, let us estimate  $I_{2,ij}$ . We have

$$\begin{aligned}
 &\int_{t_1}^{t_2} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{t_1}) dr \\
 &\quad = \sum_{k=[t_1/\delta_\varepsilon]+1}^{[t_2/\delta_\varepsilon]-1} \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E}(\mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{k\delta_\varepsilon}) | \mathcal{F}_{t_1}) dr \\
 &\quad \quad + \int_{t_1}^{([t_1/\delta_\varepsilon]+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{t_1}) dr \\
 &\quad \quad + \int_{[t_2/\delta_\varepsilon]\delta_\varepsilon}^{t_2} \mathbb{E}(\mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{[t_2/\delta_\varepsilon]\delta_\varepsilon}) | \mathcal{F}_{t_1}) dr.
 \end{aligned}$$

The distribution of the process

$$v_{1,\varepsilon}(r) := \hat{v}_\varepsilon(k\delta_\varepsilon + r), \quad r \in [0, \delta_\varepsilon],$$

coincides with the distribution of the process

$$v_{2,\varepsilon}(r) := \tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}(r/\varepsilon), \quad r \in [0, \delta_\varepsilon],$$

where  $\tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}$  is the solution of problem (2.8) with random frozen slow component  $u_\varepsilon(k\delta_\varepsilon)$ , random initial datum  $v_\varepsilon(k\delta_\varepsilon)$  and noise  $\tilde{w}^{Q_2}$  independent of  $u_\varepsilon(k\delta_\varepsilon)$  and  $v_\varepsilon(k\delta_\varepsilon)$ . Then if we set

$$h(x) := D_{ij}^2 f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H), \quad x \in H$$

for any  $k = [t_1/\delta_\varepsilon] + 1, \dots, [t_2/\delta_\varepsilon] - 1$ , we have

$$\begin{aligned} & \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{k\delta_\varepsilon}) dr \\ &= \int_0^{\delta_\varepsilon} \mathbb{E}(h(u_\varepsilon(k\delta_\varepsilon)) [ \langle G_1(u_\varepsilon(k\delta_\varepsilon), v_{1,\varepsilon}(r)) Q_{1,Na_i}, \\ & \quad G_1(u_\varepsilon(k\delta_\varepsilon), v_{1,\varepsilon}(r)) Q_{1,Na_j} \rangle_H \\ & \quad - \langle \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_i}, \\ & \quad \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_j} \rangle_H ] | \mathcal{F}_{k\delta_\varepsilon}) dr \\ &= \int_0^{\delta_\varepsilon} \mathbb{E}(h(u_\varepsilon(k\delta_\varepsilon)) [ \langle G_1(u_\varepsilon(k\delta_\varepsilon), v_{2,\varepsilon}(r)) Q_{1,Na_i}, \\ & \quad G_1(u_\varepsilon(k\delta_\varepsilon), v_{2,\varepsilon}(r)) Q_{1,Na_j} \rangle_H \\ & \quad - \langle \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_i}, \\ & \quad \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_j} \rangle_H ] | \mathcal{F}_{k\delta_\varepsilon}) dr \end{aligned}$$

and, with a change of variables,

$$\begin{aligned} & \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{k\delta_\varepsilon}) dr \\ &= \varepsilon \int_0^{\delta_\varepsilon/\varepsilon} \mathbb{E}(h(u_\varepsilon(k\delta_\varepsilon)) \\ & \quad \times [ \langle G_1(u_\varepsilon(k\delta_\varepsilon), \tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}(r)) Q_{1,Na_i}, \\ & \quad G_1(u_\varepsilon(k\delta_\varepsilon), \tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}(r)) Q_{1,Na_j} \rangle_H \\ & \quad - \langle \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_i}, \bar{G}(u_\varepsilon(k\delta_\varepsilon)) Q_{1,Na_j} \rangle_H ] | \mathcal{F}_{k\delta_\varepsilon}) dr. \end{aligned}$$

Therefore, due to the Markov property, we obtain

$$\begin{aligned} & \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{k\delta_\varepsilon}) dr \\ &= \varepsilon \int_0^{\delta_\varepsilon/\varepsilon} (\mathbb{E} \langle G_1(x, v^{x,y}(r)) Q_{1,Na_i}, \\ & \quad G_1(x, v^{x,y}(r)) Q_{1,Na_j} \rangle_{Hh}(x) \\ & \quad - \langle \bar{G}(x) Q_{1,Na_i}, \bar{G}(x) Q_{1,Na_j} \rangle_{Hh}(x)) |_{x=u_\varepsilon(k\delta_\varepsilon), y=v_\varepsilon(k\delta_\varepsilon)} dr \end{aligned}$$

and hence according to (2.17),

$$\left| \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{k\delta_\varepsilon}) dr \right| \leq c_{ij} \delta_\varepsilon \alpha(\delta_\varepsilon/\varepsilon) (1 + |u_\varepsilon(k\delta_\varepsilon)|_H^2 + |v_\varepsilon(k\delta_\varepsilon)|_H^2),$$

ℙ-a.s.

Analogously,

$$\begin{aligned} & \left| \int_{t_1}^{([t_1/\delta_\varepsilon]+1)\delta_\varepsilon} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{t_1}) dr \right| \\ & \leq c_{ij} \delta_\varepsilon (1 - \{t_1/\delta_\varepsilon\}) \alpha((1 - \{t_1/\delta_\varepsilon\})\delta_\varepsilon/\varepsilon) (1 + |u_\varepsilon(t_1)|_H^2 + |v_\varepsilon(t_1)|_H^2), \end{aligned}$$

ℙ-a.s.

and

$$\begin{aligned} & \left| \int_{[t_2/\delta_\varepsilon]\delta_\varepsilon}^{t_2} \mathbb{E}(I_{2,ij}(r) | \mathcal{F}_{[t_2/\delta_\varepsilon]\delta_\varepsilon}) dr \right| \\ & \leq c_{ij} \delta_\varepsilon \{t_2/\delta_\varepsilon\} \alpha(\{t_2/\delta_\varepsilon\}\delta_\varepsilon/\varepsilon) (1 + |u_\varepsilon([t_2/\delta_\varepsilon]\delta_\varepsilon)|_H^2 + |v_\varepsilon([t_2/\delta_\varepsilon]\delta_\varepsilon)|_H^2), \end{aligned}$$

ℙ-a.s.

Thanks to (4.4) and (4.6), the three inequality above imply

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_{t_1}^{t_2} \mathbb{E}(I_{2,ij}^\varepsilon(r) | \mathcal{F}_{t_1}) dr \right| = 0,$$

so that from (5.11), we conclude that (5.9) holds.

In an analogous way [just by replacing assumption (2.17) with assumption (2.16)], we can prove that (5.10) holds and then combining together (5.9) with (5.10), we obtain (5.2). □

**6. The averaging limit.** Before concluding with the proof of the averaging limit, we introduce an approximating slow motion equation.

For any  $n \in \mathbb{N}$ , we define

$$A_{1,n} := A_1 P_n, \quad Q_{1,n} := Q_1 P_n,$$

where  $P_n$  is the projection of  $H$  into  $\text{span}\langle e_{1,1}, \dots, e_{1,n} \rangle$ , and we denote by  $u_{\varepsilon,n}$  the solution of the problem

$$(6.1) \quad \begin{aligned} du(t) &= [A_{1,n}u(t) + B_1(u(t), v_\varepsilon(t))]dt + G_1(u(t), v_\varepsilon(t))dw^{\mathcal{Q}_{1,n}}(t), \\ u(0) &= x. \end{aligned}$$

Notice that as  $A_{1,n} \in \mathcal{L}(H)$  and  $\mathcal{Q}_{1,n}$  has finite rank,  $u_{\varepsilon,n}$  is a strong solution to (6.1), that is,

$$(6.2) \quad \begin{aligned} u_{\varepsilon,n}(t) &= x + \int_0^t [A_{1,n}u_{\varepsilon,n}(s) + B_1(u_{\varepsilon,n}(s), v_\varepsilon(s))]ds \\ &\quad + \int_0^t G_1(u_{\varepsilon,n}(s), v_\varepsilon(s))dw^{\mathcal{Q}_{1,n}}(s). \end{aligned}$$

By standard arguments, it is possible to show that for any  $p \geq 1$  and  $\varepsilon > 0$ ,

$$(6.3) \quad \lim_{n \rightarrow \infty} \mathbb{E} \sup_{t \in [0, T]} |u_\varepsilon(t) - u_{\varepsilon,n}(t)|_H^2 = 0.$$

Moreover, for any  $p \geq 1$  and  $\varepsilon > 0$ , it holds

$$(6.4) \quad \sup_{n \in \mathbb{N}} \mathbb{E} \sup_{t \in [0, T]} |u_{\varepsilon,n}(t)|_H^p < \infty.$$

In analogy to (5.1), we introduce the Kolmogorov operator associated with the approximating slow motion equation (6.1), with frozen fast component  $y \in H$ , by setting

$$\begin{aligned} \mathcal{L}_{sl}^n \varphi(x, y) &= \frac{1}{2} \text{Tr}[\mathcal{Q}_{1,n}G_1(x, y)D^2\varphi(x)G_1(x, y)\mathcal{Q}_{1,n}] \\ &\quad + \langle A_{1,n}D\varphi(x), x \rangle_H + \langle D\varphi(x), B_1(x, y) \rangle_H \\ &= \frac{1}{2} \sum_{i,j=1}^k D_{ij}^2 f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\ &\quad \times \langle G_1(x, y)\mathcal{Q}_{1,N \wedge n} a_i, G_1(x, y)\mathcal{Q}_{1,N \wedge n} a_j \rangle_H \\ &\quad + \sum_{i=1}^k D_i f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H) \\ &\quad \times (\langle x, A_{1,N \wedge n} a_i \rangle_H + \langle B_1(x, y), P_N a_i \rangle_H). \end{aligned}$$

In the next lemma, we show that the Kolmogorov operator  $\mathcal{L}_{sl}^n$  approximates in a proper way the Kolmogorov operator  $\mathcal{L}_{sl}$ .

LEMMA 6.1. *Assume Hypotheses 1 and 2. Then for any  $\varphi \in \mathcal{R}(H)$  and  $\varepsilon > 0$ ,*

$$(6.5) \quad \lim_{n \rightarrow \infty} \mathbb{E} \sup_{t \in [0, T]} |\mathcal{L}_{sl}^n \varphi(u_{\varepsilon,n}(t), v_\varepsilon(t)) - \mathcal{L}_{sl} \varphi(u_\varepsilon(t), v_\varepsilon(t))| = 0.$$

PROOF. Let

$$\varphi(x) = f(\langle x, P_N a_1 \rangle_H, \dots, \langle x, P_N a_k \rangle_H), \quad x \in H$$

for some  $k, N \in \mathbb{N}$ ,  $a_1, \dots, a_k \in H$  and  $f \in C_c^\infty(\mathbb{R}^k)$ . If  $n \geq N$ , then

$$\mathcal{L}_{sl}^n \varphi(x, y) - \mathcal{L}_{sl} \varphi(x, y) = 0, \quad x, y \in H,$$

so that for any  $\varepsilon > 0$  and  $n \geq N$ ,

$$\begin{aligned} &\mathcal{L}_{sl}^n \varphi(u_{\varepsilon,n}(t), v_\varepsilon(t)) - \mathcal{L}_{sl} \varphi(u_\varepsilon(t), v_\varepsilon(t)) \\ &= \mathcal{L}_{sl} \varphi(u_{\varepsilon,n}(t), v_\varepsilon(t)) - \mathcal{L}_{sl} \varphi(u_\varepsilon(t), v_\varepsilon(t)). \end{aligned}$$

Now, due to the assumptions on the coefficients  $B_1$  and  $G_1$  and on the function  $f$ , it is immediate to check that for  $x_1, x_2, y \in H$ ,

$$|\mathcal{L}_{sl} \varphi(x_1, y) - \mathcal{L}_{sl} \varphi(x_2, y)| \leq c|x_1 - x_2|_H(1 + |x_1|_H^2 + |x_2|_H^2 + |y|_H^2).$$

Then

$$\begin{aligned} &\sup_{t \in [0, T]} |\mathcal{L}_{sl} \varphi(u_{\varepsilon,n}(t), v_\varepsilon(t)) - \mathcal{L}_{sl} \varphi(u_\varepsilon(t), v_\varepsilon(t))| \\ &\leq c \sup_{t \in [0, T]} |u_{\varepsilon,n}(t) - u_\varepsilon(t)|_H \\ &\quad \times \left( 1 + \sup_{t \in [0, T]} |u_{\varepsilon,n}(t)|_H^2 + \sup_{t \in [0, T]} |u_\varepsilon(t)|_H^2 + \sup_{t \in [0, T]} |v_\varepsilon(t)|_H^2 \right). \end{aligned}$$

According to (6.3) and (6.4), this implies (6.5).  $\square$

Finally, we conclude with the proof of the averaging limit.

**THEOREM 6.2.** *Assume Hypotheses 1–4 and fix any  $x \in D((-A_1)^\alpha)$ , with  $\alpha > 0$ , and any  $y \in H$ . Then if  $\bar{u}$  is the solution of the averaged equation (3.14), for any  $T > 0$ ,*

$$w - \lim_{\varepsilon \rightarrow 0} \mathcal{L}(u_\varepsilon) = \mathcal{L}(\bar{u}), \quad \text{in } C([0, T]; H).$$

PROOF. Due to the tightness of the sequence  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0, 1]}$  in  $\mathcal{P}(C_x([0, T]; H), \mathcal{E})$  (see Proposition 4.5), there exists a sequence  $\{\varepsilon_k\}_{k \in \mathbb{N}} \downarrow 0$  such that the sequence  $\{\mathcal{L}(u_{\varepsilon_k})\}_{k \in \mathbb{N}}$  converges weakly to some  $\mathbb{Q}$ . If we are able to identify  $\mathbb{Q}$  with  $\mathcal{L}(\bar{u})$ , where  $\bar{u}$  is the unique mild solution of the averaged equation (3.14), then we conclude that the whole sequence  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0, 1]}$  weakly converges to  $\mathcal{L}(\bar{u})$  in  $C([0, T]; H)$ .

As  $u_{\varepsilon,n}$  verifies (6.2), for any  $\varphi \in \mathcal{R}(H)$  we can apply Itô’s formula to  $\varphi(u_{\varepsilon,n})$ , and we obtain that the process

$$t \in [0, T] \mapsto \varphi(u_{\varepsilon,n}(t)) - \varphi(x) - \int_0^t \mathcal{L}_{sl}^n \varphi(u_{\varepsilon,n}(s), v_\varepsilon(s)) ds$$

is a martingale with respect to  $\{\mathcal{F}_t\}_{t \in [0, T]}$ . Then by taking the limit as  $n$  goes to infinity, due to (6.3) and to (6.5), we conclude that for any  $\varepsilon > 0$  the process

$$t \in [0, T] \mapsto \varphi(u_\varepsilon(t)) - \varphi(x) - \int_0^t \mathcal{L}_{sl}\varphi(u_\varepsilon(s), v_\varepsilon(s)) ds$$

is an  $\mathcal{F}_t$ -martingale. In particular, for any  $0 \leq s \leq t \leq T$  and any bounded  $\mathcal{F}_s$ -measurable random variable  $\Psi$

$$(6.6) \quad \mathbb{E}\left(\Psi\left[\varphi(u_\varepsilon(t)) - \varphi(u_\varepsilon(s)) - \int_s^t \mathcal{L}_{sl}\varphi(u_\varepsilon(r), v_\varepsilon(r)) dr\right]\right) = 0.$$

We denote by  $\mathbb{E}^\mathbb{Q}$  and  $\mathbb{E}^{\mathbb{Q}_k}$  the expectations in  $(C_x([0, T]; H), \mathcal{E})$  with respect to the probability measures  $\mathbb{Q}$  and  $\mathbb{Q}_k$ , where  $\mathbb{Q}_k = \mathcal{L}(u_{\varepsilon_k})$ , and we denote by  $\eta(t)$  the canonical process in  $(C_x([0, T]; H), \mathcal{E})$ . Then for any bounded  $\mathcal{E}_s$ -measurable random variable

$$\Phi = F(\eta(t_1), \dots, \eta(t_N))$$

with  $F \in C_b(\mathbb{R}^N)$  and  $0 \leq t_1 < \dots < t_N$ , any function  $\varphi \in \mathcal{R}(H)$  and any  $0 \leq s \leq t \leq T$  we have

$$\begin{aligned} & \mathbb{E}^\mathbb{Q}\left(\Phi\left[\varphi(\eta(t)) - \varphi(\eta(s)) - \int_s^t \mathcal{L}_{av}\varphi(\eta(r)) dr\right]\right) \\ &= \lim_{k \rightarrow \infty} \mathbb{E}^{\mathbb{Q}_k}\left(\Phi\left[\varphi(\eta(t)) - \varphi(\eta(s)) - \int_s^t \mathcal{L}_{av}\varphi(\eta(r)) dr\right]\right) \\ &= \lim_{k \rightarrow \infty} \mathbb{E}\left(\Phi \circ u_{\varepsilon_k}\left[\varphi(u_{\varepsilon_k}(t)) - \varphi(u_{\varepsilon_k}(s)) - \int_s^t \mathcal{L}_{av}\varphi(u_{\varepsilon_k}(r)) dr\right]\right). \end{aligned}$$

In view of (6.6), this implies

$$\begin{aligned} & \mathbb{E}^\mathbb{Q}\left(\Phi\left[\varphi(\eta(t)) - \varphi(\eta(s)) - \int_s^t \mathcal{L}_{av}\varphi(\eta(r)) dr\right]\right) \\ &= \lim_{k \rightarrow \infty} \mathbb{E}\left(\Phi \circ u_{\varepsilon_k} \int_s^t [\mathcal{L}_{sl}\varphi(u_{\varepsilon_k}(r), v_{\varepsilon_k}(r)) - \mathcal{L}_{av}\varphi(u_{\varepsilon_k}(r))] dr\right). \end{aligned}$$

We have

$$\begin{aligned} & \left| \mathbb{E}\left(\Phi \circ u_{\varepsilon_k} \int_s^t [\mathcal{L}_{sl}\varphi(u_{\varepsilon_k}(r), v_{\varepsilon_k}(r)) - \mathcal{L}_{av}\varphi(u_{\varepsilon_k}(r))] dr\right) \right| \\ &= \left| \mathbb{E}\left(\Phi \circ u_{\varepsilon_k} \int_s^t \mathbb{E}(\mathcal{L}_{sl}\varphi(u_{\varepsilon_k}(r), v_{\varepsilon_k}(r)) - \mathcal{L}_{av}\varphi(u_{\varepsilon_k}(r)) | \mathcal{F}_s) dr\right) \right| \\ &\leq \|\Phi\|_\infty \mathbb{E}\left|\int_s^t \mathbb{E}(\mathcal{L}_{sl}\varphi(u_{\varepsilon_k}(r), v_{\varepsilon_k}(r)) - \mathcal{L}_{av}\varphi(u_{\varepsilon_k}(r)) | \mathcal{F}_s) dr\right|. \end{aligned}$$

Hence, according to (5.2), we can conclude that

$$\mathbb{E}^\mathbb{Q}\left(\Phi\left[\varphi(\eta(t)) - \varphi(\eta(s)) - \int_s^t \mathcal{L}_{av}\varphi(\eta(r)) dr\right]\right) = 0.$$

This means that  $\mathbb{Q}$  solves the martingale problem with parameters  $(x, A_1, \bar{B}, \bar{G}, Q_1)$ , and due to what we have seen in Section 3.3,  $\mathbb{Q} = \mathcal{L}(\bar{u})$ .  $\square$

6.1. *Averaging limit in probability.* In the case the diffusion coefficient  $g_1$  in the slow motion equation does not depend on the fast variable, it is possible to prove that the sequence  $\{u_\varepsilon\}_{\varepsilon \in (0,1]}$  converges in probability to  $\bar{u}$  and not just in weak sense.

To this purpose, we need to replace Hypothesis 3 with the following stronger condition.

**HYPOTHESIS 5.** There exists a mapping  $\bar{B}_1 : H \rightarrow H$  such that for any  $T > 0$ ,  $t \geq 0$  and  $x, y, h \in H$

$$(6.7) \quad \mathbb{E} \left| \frac{1}{T} \int_t^{t+T} \langle B_1(x, v^{x,y}(s)), h \rangle_H ds - \langle \bar{B}_1(x), h \rangle_H \right| \leq \alpha(T)(1 + |x|_H + |y|_H)|h|_H$$

for some  $\alpha(T)$  such that

$$\lim_{T \rightarrow \infty} \alpha(T) = 0.$$

In Section 2.1, by referring to our previous paper [7], we have seen that if the dissipativity constant of the operator  $A_1$  is large enough and/or the Lipschitz constants  $L_{b_2}$  and  $L_{g_2}$  and the constants  $\zeta_2$  and  $\kappa_2$  introduced in Hypothesis 1 are small enough [in this spirit see condition (3.1)], then the fast transition semigroup admits a unique invariant measure  $\mu^x$ , which is strongly mixing and such that (2.13) holds.

In Lemma 2.3, we have seen that this implies that for any  $\varphi \in \text{Lip}(H)$ ,  $T > 0$ ,  $x, y \in H$  and  $t \geq 0$ ,

$$\mathbb{E} \left| \frac{1}{T} \int_t^{t+T} \varphi(v^{x,y}(s)) ds - \int_H \varphi(z) \mu^x(dz) \right| \leq \frac{c}{\sqrt{T}} ([\varphi]_{\text{Lip}(H)}(1 + |x|_H + |y|_H) + |\varphi(0)|).$$

Then if we apply the inequality above to  $\varphi = \langle B_1(x, \cdot), h \rangle_H$  and if we set  $\bar{B}_1(x) = \langle B_1(x, \cdot), \mu^x \rangle$ , we have that Hypothesis 5 holds.

As  $u_\varepsilon$  is the mild solution of the slow motion equation in system (1.6) [see also (2.7) for its abstract version] and  $g_1$  does not depend on  $v_2$ , for any  $h \in D(A_1) \cap L^\infty(D)$ , we have

$$(6.8) \quad \begin{aligned} \langle u_\varepsilon(t), h \rangle_H &= \langle x, h \rangle_H + \int_0^t \langle u_\varepsilon(s), A_1 h \rangle_H ds + \int_0^t \langle \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \\ &\quad + \int_0^t \langle G_1(u_\varepsilon(s))h, dw^{Q_1}(s) \rangle_H + R_\varepsilon(t), \end{aligned}$$

where

$$R_\varepsilon(t) := \int_0^t \langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds.$$

In order to prove the averaging limit, we need the following key lemma, which is the counterpart of Lemma 5.1.

LEMMA 6.3. *Assume Hypotheses 1, 2 and 5 and fix  $T > 0$ . Then for any  $x \in D((-A_1)^\alpha)$ , with  $\alpha > 0$ , and  $y, h \in H$ , we have*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |R_\varepsilon(t)| = 0.$$

PROOF. *Step 1.* We prove that

$$(6.9) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \right| = 0,$$

where  $\hat{v}_\varepsilon(t)$  is the solution of problem (5.3). Let  $k = 0, \dots, [T/\delta_\varepsilon]$  be fixed. If we take a noise  $\tilde{w}^{Q_2}(t)$  independent of  $u_\varepsilon(k\delta_\varepsilon)$  and  $v_\varepsilon(k\delta_\varepsilon)$  in the fast motion equation (2.8), it is immediate to check that the process

$$z_{1,\varepsilon}(s) := \tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}(s/\varepsilon), \quad s \in [0, \delta_\varepsilon],$$

coincides in distribution with the process

$$z_{2,\varepsilon}(s) := \hat{v}_\varepsilon(k\delta_\varepsilon + s), \quad s \in [0, \delta_\varepsilon].$$

This means that

$$\begin{aligned} & \mathbb{E} \left| \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \langle B_1(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right| \\ &= \mathbb{E} \left| \int_0^{\delta_\varepsilon} \langle B_1(u_\varepsilon(k\delta_\varepsilon), z_{2,\varepsilon}(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right| \\ &= \mathbb{E} \left| \int_0^{\delta_\varepsilon} \langle B_1(u_\varepsilon(k\delta_\varepsilon), z_{1,\varepsilon}(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right| \\ &= \delta_\varepsilon \mathbb{E} \left| \frac{1}{\zeta_\varepsilon} \int_0^{\zeta_\varepsilon} \langle B_1(u_\varepsilon(k\delta_\varepsilon), \tilde{v}^{u_\varepsilon(k\delta_\varepsilon), v_\varepsilon(k\delta_\varepsilon)}(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right|. \end{aligned}$$

Hence, according to Hypothesis 5, due to (4.4) and (4.6), we have

$$\begin{aligned} & \mathbb{E} \left| \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \langle B_1(u_\varepsilon(k\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right| \\ & \leq \delta_\varepsilon \alpha(\zeta_\varepsilon) (1 + \mathbb{E}|u_\varepsilon(k\delta_\varepsilon)|_H + \mathbb{E}|v_\varepsilon(k\delta_\varepsilon)|_H) |h|_H \\ & \leq c_T (1 + |x|_H + |y|_H) |h|_H \delta_\varepsilon \alpha(\zeta_\varepsilon). \end{aligned}$$

This allows to obtain (6.9). Actually, we have

$$\begin{aligned} & \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \right| \\ & \leq \sum_{k=0}^{[T/\delta_\varepsilon]} \mathbb{E} \left| \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \langle B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)), h \rangle_H ds \right| \\ & \quad + \sum_{k=0}^{[T/\delta_\varepsilon]} \int_{k\delta_\varepsilon}^{(k+1)\delta_\varepsilon} \mathbb{E} |\langle \bar{B}_1(u_\varepsilon(k\delta_\varepsilon)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H| ds \end{aligned}$$

and, as  $\bar{B}_1$  is Lipschitz continuous, thanks to (4.16) we get

$$\begin{aligned} & \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \right| \\ & \leq c_T [T/\delta_\varepsilon] (1 + |x|_H + |y|_H) \delta_\varepsilon \alpha(\zeta_\varepsilon) \\ & \quad + c_T [T/\delta_\varepsilon] (1 + |x|_\alpha + |y|_H) |h|_H \delta_\varepsilon^{1+\beta(\alpha)} \end{aligned}$$

and (6.9) follows.

*Step 2.* It holds

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T]} \mathbb{E} |R_\varepsilon(t)| = 0.$$

Thanks to (6.9), we have

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \right| \\ & \leq \limsup_{\varepsilon \rightarrow 0} \mathbb{E} \int_0^T |\langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)), h \rangle_H| ds \\ & \quad + \lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)) - \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds \right| \\ & = \limsup_{\varepsilon \rightarrow 0} \mathbb{E} \int_0^T |\langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)), h \rangle_H| ds. \end{aligned}$$

By using (4.16), we have

$$\begin{aligned} & \mathbb{E} \int_0^T |\langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)), h \rangle_H| ds \\ & \leq c \int_0^T (\mathbb{E} |u_\varepsilon(s) - u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon)|_H + \mathbb{E} |v_\varepsilon(s) - \hat{v}_\varepsilon(s)|_H) ds |h|_H \\ & \leq c_T (1 + |x|_\alpha + |y|_H) |h|_H \delta_\varepsilon^{\beta(\alpha)} + T \sup_{t \in [0, T]} \mathbb{E} |v_\varepsilon(t) - \hat{v}_\varepsilon(t)|_H |h|_H \end{aligned}$$

and then due to (5.4), we have

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \int_0^T | \langle B_1(u_\varepsilon(s), v_\varepsilon(s)) - B_1(u_\varepsilon([s/\delta_\varepsilon]\delta_\varepsilon), \hat{v}_\varepsilon(s)), h \rangle_H | ds = 0.$$

This allows to conclude.  $\square$

**THEOREM 6.4.** *Assume that the diffusion coefficient  $g_1$  in the slow motion equation does not depend on the fast variable  $v_\varepsilon$  and fix  $x \in D((-A_1)^\alpha)$ , for some  $\alpha > 0$ , and  $y \in H$ . Then under Hypotheses 1, 2 and 5 for any  $T > 0$  and  $\eta > 0$ , we have*

$$(6.10) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{P}(|u_\varepsilon - \bar{u}|_{C([0,T];H)} > \eta) = 0,$$

where  $\bar{u}$  is the solution of the averaged equation (3.14).

**PROOF.** We have seen that the family of probability measures  $\{\mathcal{L}(u_\varepsilon)\}_{\varepsilon \in (0,1]}$  is tight in  $\mathcal{P}(C([0, T]; H))$ , for any fixed  $T > 0$ . Then if we fix any two sequences  $\{\varepsilon_n\}_{n \in \mathbb{N}}$  and  $\{\varepsilon_m\}_{m \in \mathbb{N}}$  which converge to zero, due to the Skorokhod theorem, we can find two subsequences  $\{\varepsilon_{n(k)}\}_{k \in \mathbb{N}}$  and  $\{\varepsilon_{m(k)}\}_{k \in \mathbb{N}}$  and a sequence

$$\{X_k\}_{k \in \mathbb{N}} := \{(u_1^k, u_2^k, \hat{w}_k^{Q_1})\}_{k \in \mathbb{N}} \subset \mathcal{C} := [C([0, T]; H)]^2 \times C([0, T]; \mathcal{D}'(D)),$$

defined on some probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ , such that

$$(6.11) \quad \mathcal{L}(X_k) = \mathcal{L}((u_{\varepsilon_{n(k)}} , u_{\varepsilon_{m(k)}} , w^{Q_1})), \quad k \in \mathbb{N}$$

and  $X_k$  converges  $\hat{\mathbb{P}}$  almost surely to some  $X := (u_1, u_2, \hat{w}^{Q_1}) \in \mathcal{C}$  and  $\hat{w}^{Q_1}$  is a cylindrical Wiener process. In particular,  $(\hat{w}_k^{Q_1}, \hat{\mathbb{P}})$  is a cylindrical Wiener process, for any  $k \in \mathbb{N}$ , and

$$\lim_{k \rightarrow \infty} \sup_{t \in [0, T]} | \langle \hat{w}_k^{Q_1}(t) - \hat{w}^{Q_1}(t), h \rangle_H | = 0$$

for any  $h \in H$ .

Now, for  $k \in \mathbb{N}$  and  $i = 1, 2$ , we define

$$(6.12) \quad \begin{aligned} R_i^k(t) &:= \langle u_i^k(t), h \rangle_H - \langle x, h \rangle_H - \int_0^t \langle u_i^k(s), A_1 h \rangle_H ds \\ &\quad - \int_0^t \langle \bar{B}_1(u_i^k(s)), h \rangle_H ds - \int_0^t \langle G_1(u_i^k(s))h, d\hat{w}_k^{Q_1}(s) \rangle_H. \end{aligned}$$

In (6.8), we have defined

$$\begin{aligned} R_\varepsilon(t) &:= \langle u_\varepsilon(t), h \rangle_H - \langle x, h \rangle_H - \int_0^t \langle u_\varepsilon(s), A_1 h \rangle_H ds \\ &\quad - \int_0^t \langle \bar{B}_1(u_\varepsilon(s)), h \rangle_H ds - \int_0^t \langle G_1(u_\varepsilon(s))h, dw^{Q_1}(s) \rangle_H \end{aligned}$$

and in Lemma 6.3 we have proved that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |R_\varepsilon(t)| = 0.$$

In view of (6.11), this implies

$$\lim_{k \rightarrow \infty} \hat{\mathbb{E}} \sup_{t \in [0, T]} |R_i^k(t)| = 0, \quad i = 1, 2.$$

On the other hand, since the sequence  $\{X_k\}_{k \in \mathbb{N}} \subset \mathcal{C}$  converges  $\hat{\mathbb{P}}$ -a.s. to  $X = (u_1, u_2, \hat{w}^{\mathcal{Q}_1})$ , it is possible to show that for  $i = 1, 2$  there exists the  $\hat{\mathbb{P}}$ -a.s. limit of the right-hand side in (6.12) and it coincides with

$$\begin{aligned} &\langle u_i(t), h \rangle_H - \langle x, h \rangle_H - \int_0^t \langle u_i(s), A_1 h \rangle_H ds \\ &- \int_0^t \langle \bar{B}_1(u_i(s)), h \rangle_H ds - \int_0^t \langle G_1(u_i(s))h, d\hat{w}^{\mathcal{Q}_1}(s) \rangle_H \end{aligned}$$

(for a detailed proof of this fact, see, e.g., [22], Lemmas 4.3 and 4.4).

Then if we pass possibly to a subsequence, we can take the  $\hat{\mathbb{P}}$ -almost sure limit in both sides of (6.12), and we get that both  $u_1$  and  $u_2$  solve the problem

$$\begin{aligned} \langle u(t), h \rangle_H &= \langle x, h \rangle_H + \int_0^t \langle u(s), A_1 h \rangle_H ds \\ &+ \int_0^t \langle \bar{B}_1(u(s)), h \rangle_H ds + \int_0^t \langle G_1(u(s))h, d\hat{w}^{\mathcal{Q}_1}(s) \rangle_H \end{aligned}$$

for any  $h \in D(A_1) \cap L^\infty(D)$ . This means that  $u_1 = u_2$ , as they coincide with the unique mild solution of equation

$$du(t) = [A_1 u(t) + \bar{B}_1(u(t))] dt + G_1(u(t)) d\hat{w}^{\mathcal{Q}_1}(t), \quad u(0) = x$$

and then the sequences  $\{\mathcal{L}(u_{\varepsilon_{n(k)}})\}$  and  $\{\mathcal{L}(u_{\varepsilon_{m(k)}})\}$  weakly converge to the same limit.

This allows to conclude that (6.10) is true. As a matter of fact, in Gyöngy and Krylov [15], Lemma 1.1, it is proved that if  $\{Z_n\}$  is a sequence of random element in a Polish space  $X$ , then  $\{Z_n\}$  converges in probability to a  $X$ -valued random element if and only if for every pair of subsequences  $\{Z_l\}$  and  $\{Z_m\}$  there exists a subsequence  $v_k = (Z_{l_k}, Z_{m_k})$  converging weakly to a random element  $v$  supported on the diagonal of  $X \times X$ .  $\square$

### REFERENCES

- [1] ARNOLD, V. I., KOZLOV, V. V. and NEISHTADT, A. I. (2006). *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd ed. *Encyclopaedia of Mathematical Sciences* **3**. Springer, Berlin. MR2269239
- [2] BAKHTIN, V. and KIFER, Y. (2004). Diffusion approximation for slow motion in fully coupled averaging. *Probab. Theory Related Fields* **129** 157–181. MR2063374

- [3] BOGOLIUBOV, N. N. and MITROPOLSKY, Y. A. (1961). *Asymptotic Methods in the Theory of Nonlinear Oscillations*. Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York. [MR0141845](#)
- [4] BRIN, M. and FREIDLIN, M. (2000). On stochastic behavior of perturbed Hamiltonian systems. *Ergodic Theory Dynam. Systems* **20** 55–76. [MR1747030](#)
- [5] CERRAI, S. (2001). *Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach. Lecture Notes in Math.* **1762**. Springer, Berlin. [MR1840644](#)
- [6] CERRAI, S. (2003). Stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term. *Probab. Theory Related Fields* **125** 271–304. [MR1961346](#)
- [7] CERRAI, S. (2006). Asymptotic behavior of systems of stochastic partial differential equations with multiplicative noise. In *Stochastic Partial Differential Equations and Applications—VII. Lecture Notes in Pure and Applied Mathematics* **245** 61–75. Chapman and Hall/CRC, Boca Raton, FL. [MR2227220](#)
- [8] CERRAI, S. and FREIDLIN, M. (2008). Averaging principle for a class of SPDE's. In *Probability Theory and Related Fields*. To appear.
- [9] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and its Applications* **44**. Cambridge Univ. Press, Cambridge. [MR1207136](#)
- [10] DA PRATO, G. and ZABCZYK, J. (1996). *Ergodicity for Infinite-dimensional Systems. London Mathematical Society Lecture Note Series* **229**. Cambridge Univ. Press, Cambridge. [MR1417491](#)
- [11] FONTBONA, J. (2003). Nonlinear martingale problems involving singular integrals. *J. Funct. Anal.* **200** 198–236. [MR1974095](#)
- [12] FREIDLIN, M. I. (2001). On stable oscillations and equilibriums induced by small noise. *J. Statist. Phys.* **103** 283–300. [MR1828731](#)
- [13] FREIDLIN, M. I. and WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer, New York. [MR1652127](#)
- [14] FREIDLIN, M. I. and WENTZELL, A. D. (2006). Long-time behavior of weakly coupled oscillators. *J. Statist. Phys.* **123** 1311–1337. [MR2253881](#)
- [15] GYÖNGY, I. and KRYLOV, N. (1996). Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields* **103** 143–158. [MR1392450](#)
- [16] KHASHMINSKII, R. Z. (1968). On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)* **4** 260–279. [MR0260052](#)
- [17] KIFER, Y. (2004). Some recent advances in averaging. In *Modern Dynamical Systems and Applications* 385–403. Cambridge Univ. Press, Cambridge. [MR2093312](#)
- [18] KIFER, Y. (2001). Averaging and climate models. In *Stochastic Climate Models (Chorin, 1999)*. *Progr. Probab.* **49** 171–188. Birkhäuser, Basel. [MR1948296](#)
- [19] KIFER, Y. (2001). Stochastic versions of Anosov's and Neistadt's theorems on averaging. *Stoch. Dyn.* **1** 1–21. [MR1837154](#)
- [20] KUKSIN, S. B. and PIATNITSKI, A. L. (2008). Khasminskii–Whitham averaging for randomly perturbed KdV equation. *J. Math. Pures Appl.* (9) **89** 400–428. [MR2401144](#)
- [21] MASLOWSKI, B., SEIDLER, J. and VRKOČ, I. (1991). An averaging principle for stochastic evolution equations. II. *Math. Bohem.* **116** 191–224. [MR1112004](#)
- [22] MILLET, A. and MORIEN, P.-L. (2001). On a nonlinear stochastic wave equation in the plane: existence and uniqueness of the solution. *Ann. Appl. Probab.* **11** 922–951. [MR1865028](#)
- [23] NEISHTADT, A. I. (1976). Averaging in multifrequency systems. *Soviet Physics Doklady* **21** 80–82.
- [24] PAPANICOLAOU, G. C., STROOCK, D. and VARADHAN, S. R. S. (1977). Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference (Duke Univ.,*

- Durham, NC*, 1976), Paper No. 6 *Duke Mathematical Journal* **3** 120. Duke Univ. Press, Durham, NC. [MR0461684](#)
- [25] SEIDLER, J. and VRKOČ, I. (1990). An averaging principle for stochastic evolution equations. *Časopis Pěst. Mat.* **115** 240–263. [MR1071056](#)
- [26] TESSITORE, G. and ZABCZYK, J. (2006). Wong–Zakai approximations of stochastic evolution equations. *J. Evol. Equ.* **6** 621–655. [MR2267702](#)
- [27] VERETENNIKOV, A. Y. (1991). On an averaging principle for systems of stochastic differential equations. *Mat. Sb.* **69** 271–284. [MR1046602](#)
- [28] VOLOSOV, V. M. (1962). Averaging in systems of ordinary differential equations. *Russian Math. Surveys* **17** 1–126. [MR0146454](#)

DIPARTIMENTO DI MATEMATICA PER LE DECISIONI  
UNIVERSITÀ DI FIRENZE  
VIA C. LOMBROSO 6/17  
I-50134 FIRENZE  
ITALY  
E-MAIL: [sandra.cerrai@dmd.unifi.it](mailto:sandra.cerrai@dmd.unifi.it)