



ELSEVIER

Contents lists available at ScienceDirect

Journal of Functional Analysis

www.elsevier.com/locate/jfa



On the convergence of stationary solutions in the Smoluchowski-Kramers approximation of infinite dimensional systems



Sandra Cerrai ^{a,1}, Nathan Glatt-Holtz ^{b,2}

^a *University of Maryland, College Park, United States*

^b *Tulane University, New Orleans, United States*

ARTICLE INFO

Article history:

Received 20 June 2018

Accepted 2 December 2019

Available online 13 December 2019

Communicated by K.-T. Sturm

Keywords:

Smoluchowski-Kramers approximation

Convergence of invariant measures

ABSTRACT

We prove the convergence, in the small mass limit, of statistically invariant states for a class of semi-linear damped wave equations, perturbed by an additive Gaussian noise, both with Lipschitz-continuous and with polynomial nonlinearities. In particular, we prove that the first marginals of any sequence of invariant measures for the stochastic wave equation converge in a suitable Wasserstein metric to the unique invariant measure of the limiting stochastic semi-linear parabolic equation obtained in the Smoluchowski-Kramers approximation. The Wasserstein metric is associated to a suitable distance on the space of square integrable functions, that is chosen in such a way that the dynamics of the limiting stochastic parabolic equation is contractive with respect to such a Wasserstein metric. This implies that the limiting result is a consequence of the validity of a generalized Smoluchowski-Kramers limit at fixed times. The proof of such a generalized limit requires new delicate bounds for

E-mail address: negh@tulane.edu (N. Glatt-Holtz).

¹ Partially supported by the NSF grants DMS-1407615 and DMS-1712934.

² Partially supported by the NSF grant DMS-1313272, DMS-1816551 and the Simons Foundation grant 515990.

the solutions of the stochastic wave equation, that must be uniform with respect to the size of the mass.

© 2019 Published by Elsevier Inc.

1. Introduction

We are dealing here with the following class of stochastic damped wave equations, defined on a bounded and smooth domain $D \subset \mathbb{R}^d$

$$\begin{cases} \mu \partial_t^2 u_\mu(t, \xi) = \Delta u_\mu(t, \xi) - \partial_t u_\mu(t, \xi) + b(\xi, u_\mu(t, \xi)) + \partial_t w^Q(t, \xi), \\ u_\mu(0, \xi) = x(\xi), \quad \partial_t u_\mu(0, \xi) = y(\xi), \quad \xi \in D, \quad u_\mu(t, \xi) = 0, \\ t \geq 0, \quad \xi \in \partial D, \end{cases} \quad (1.1)$$

where $0 < \mu \ll 1$. Here, $w^Q(t, \xi)$ is a cylindrical Wiener process, white in time and colored in space. The nonlinearity $b : D \times \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be either globally Lipschitz-continuous in the second variable, uniformly with respect to the first one, with its Lipschitz constant dominated by the first eigenvalue of the Laplacian, or only locally Lipschitz continuous and satisfying some polynomial growth and dissipation conditions. In both cases, the identically zero function is globally asymptotically stable in the absence of noise. In the case the nonlinearity b is Lipschitz, we do not impose any restriction on the space dimension d , while, in the case of polynomial nonlinearity, we assume that $d = 1$.

As a consequence of Newton's law, the solution $u_\mu(t, \xi)$ of equation (1.1) can be interpreted as the displacement field of the particles of a material continuum in a domain D , subject to a random external force field $\partial_t w(t, \xi)$ and a damping force proportional to the velocity field (here the proportionality constant is taken equal 1). The second order differential operator takes into account of the interaction forces between neighboring particles, in the presence of a non-linear reaction given by b . Here μ represents the constant density of the particles and we are interested in the regime when $\mu \rightarrow 0$, the so-called Smoluchowski-Kramers approximation limit (ref. [23] and [33]).

In [4] and [5], the convergence of u_μ to the solution u of the parabolic problem

$$\begin{cases} \partial_t u(t, \xi) = \Delta u(t, \xi) + b(\xi, u(t, \xi)) + \partial_t w^Q(t, \xi), \\ u(0, \xi) = x(\xi), \quad \xi \in D, \quad u(t, \xi) = 0, \quad t \geq 0, \quad \xi \in \partial D, \end{cases} \quad (1.2)$$

has been studied, under analogous conditions on the non-linearity b . Namely, it has been proven that for every $T > 0$ and $\epsilon > 0$

$$\lim_{\mu \rightarrow 0} \mathbb{P} \left(\sup_{t \in [0, T]} |u_\mu(t) - u(t)|_{L^2(D)} > \epsilon \right) = 0. \quad (1.3)$$

In fact, in [10] it has been shown that, under the condition of Lipschitz-continuity for b , the following stronger convergence holds

$$\lim_{\mu \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |u_\mu(t) - u(t)|_{L^2(D)}^p = 0, \tag{1.4}$$

for every $p \geq 1$. Note that this type of limit has been addressed in a variety of finite and infinite dimensional context (ref. [2], [6], [16], [17], [20], [21], [22], [24], [34] for the finite dimension and [4], [5], [7], [8], [9], [10], [25], [26], [27], [28], [29], [31] for the infinite dimension).

However, once one has proved the validity of the small mass limits (1.3) and (1.4) on a fixed time interval $[0, T]$, it is of interest to compare the long-time behavior of the second order system (1.1), with that of the first order system (1.2). In particular, it is desirable to identify conditions for the convergence of statistically steady states.

In [4], the long time behavior of equations (1.1) and (1.2) has been compared, under the assumption that the two systems are of gradient type. Actually, in the case of white noise in space and time (that is $Q = I$) and dimension $d = 1$, an explicit expression for the Boltzman distribution of the process $z_\mu(t) := (u_\mu(t), \partial u_\mu / \partial t(t))$ in the phase space $\mathcal{H} := L^2(0, L) \times H^{-1}(0, L)$ has been given. Of course, since in the functional space \mathcal{H} there is no analogous of the Lebesgue measure, an auxiliary Gaussian measure has been introduced, with respect to which the density of the Boltzman distribution has been written. This auxiliary Gaussian measure is the stationary measure of the linear wave equation related to problem (1.1). In particular, it has been shown that the first marginal of the invariant measure associated with the process $z_\mu(t)$ does not depend on $\mu > 0$ and coincides with the invariant measure for the heat equation (1.2).

In the case of non-gradient systems, that is when we are not dealing with space-time white noise, there is no explicit expression for the invariant measure ν_μ associated with system (1.1) and there is no reason to expect that the first marginal of ν_μ does not depend on μ or coincides with the invariant measure ν of system (1.2).

Nevertheless, in the present paper we are going to show that in the limit, as the mass μ goes to zero, the first marginal of any invariant measure ν_μ of the second order system (1.1) approaches in a suitable manner the invariant measure ν of the first order system (1.2). Namely, we are going to prove that

$$\lim_{\mu \rightarrow 0} \rho_\alpha ((\Pi_1 \nu_\mu)', \nu) = 0, \tag{1.5}$$

where $(\Pi_1 \nu_\mu)'$ is the extension to $L^2(D)$ of the first marginal of the invariant measure ν_μ in $H^1(D) \times L^2(D)$ and ρ_α is the Wasserstein metric on $P(L^2(D))$ associated with a suitable distance α on $L^2(D)$, to be determined on the basis of the type of non-linearity b we are dealing with.

Actually, the fundamental fact that we need to be true in order to be able to prove (1.5) is that there exist a distance α in $L^2(D)$ and a constant $\delta > 0$ such that

$$\rho_\alpha (P_t^* \nu_1, P_t^* \nu_2) \leq c e^{-\delta t} \rho_\alpha (\nu_1, \nu_2), \quad t \geq 0, \tag{1.6}$$

for every $\nu_1, \nu_2 \in \mathcal{Pr}(H)$. Here P_t^* is the adjoint of the transition semigroup associated with equation (1.2). As shown in [19], a possible proof of (1.6) is based on a suitable irreducibility condition for system (1.2), together with suitable large-time smoothing estimates for the Markovian dynamics related closely to the so-called *asymptotic strong Feller property* (see [18]), along with some estimates for the solution of the first variation equation of (1.2). In case of a polynomial non-linearity b , the irreducibility condition is uniform with respect to the initial condition, so that we can take α equal to the usual distance in $L^2(D)$. In case of a Lipschitz b , the irreducibility condition is uniform only with respect to initial conditions on bounded sets of $L^2(D)$, so that a Lyapunov structure has to be taken into consideration. This means in particular that we can prove (1.6) by using the method developed in [19], once we choose

$$\alpha(x, y) = \inf_{\gamma} \int_0^1 \exp(\eta |\gamma(t)|_H^2) |\gamma'(t)|_H dt,$$

where the infimum is taken over all paths γ such that $\gamma(0) = x$ and $\gamma(1) = y$ and where η is some positive constant to be determined.

Next, as a consequence of (1.6), due to the invariance of ν_μ and ν , we have

$$\rho_\alpha ((\Pi_1 \nu_\mu)', \nu) \leq \rho_\alpha ((\Pi_1 (P_t^\mu)^* \nu_\mu)', P_t^* (\Pi_1 \nu_\mu)') + c e^{-\delta t} \rho_\alpha ((\Pi_1 \nu_\mu)', \nu).$$

Therefore, if we pick $t_\star > 0$ such that $c e^{-\delta t_\star} < 1/2$, we have

$$\rho_\alpha ((\Pi_1 \nu_\mu)', \nu) \leq 2 \rho_\alpha ((\Pi_1 (P_t^\mu)^* \nu_\mu)', P_t^* (\Pi_1 \nu_\mu)'), \quad t \geq t_\star.$$

Since the Wasserstein distance ρ_α between two measures ν_1 and ν_2 is given as

$$\rho_\alpha (\nu_1, \nu_2) = \inf \mathbb{E} \alpha(X_1, X_2),$$

where $\mathcal{L}(X_i) = \nu_i$, for $i = 1, 2$, we have

$$\rho_\alpha ((\Pi_1 (P_t^\mu)^* \nu_\mu)', P_t^* (\Pi_1 \nu_\mu)') \leq \mathbb{E} \alpha(u_\mu^{\gamma_\mu}(t), u_\mu^{\theta_\mu}(t)),$$

where $\gamma_\mu = (\theta_\mu, \eta_\mu)$ is a $H^1(D) \times L^2(D)$ -valued random variable, distributed as ν_μ . In particular, this implies that the proof of (1.5) reduces to the proof of the following limit

$$\lim_{\mu \rightarrow 0} \mathbb{E} \alpha(u_\mu^{\gamma_\mu}(t), u_\mu^{\theta_\mu}(t)) = 0, \tag{1.7}$$

for a fixed time t sufficiently large.

The limit (1.7) seems analogous to the Smoluchowski-Kramers approximation (1.4) proven in [10], at least when α is the usual distance in $L^2(D)$. In fact, limit (1.7) is

considerably stronger, and more insidious to be proven, in comparison to (1.4), as (1.4) is only valid for fixed deterministic initial condition, while (1.7) has to be verified for random initial conditions, depending on μ through the invariant measure ν_μ . As a matter of fact, the proof of (1.7) is the major challenge of the paper and requires a series of delicate and completely new uniform bounds for the solution of equation (1.1), including exponential moments.

To this purpose, we would like to mention that the convergence of stationary solutions, as μ goes to zero, has been also studied in [28], in the case there is $\sqrt{\mu}$ in front of the noise. This case is of course completely different from the case we are considering here and it is also considerably easier from a mathematical point of view. Actually, it does not require the sophisticated bounds that we have proved in the present paper and does not involve invariant measures.

To conclude, we would like to remark that our result about the limiting behavior of stationary solutions of system (1.1), in the Smoluchowski-Kramers approximation, is consistent to what proven in [8] and [9] for the limiting behavior of the quasi-potential $V_\mu(u, v)$, that describes the asymptotics of the exit times and the large deviation principle for the invariant measures in equation (1.1). Actually, in [8] the quasi-potential $V_\mu(u, v)$ associated with (1.1) has been explicitly computed under the assumption that system (1.1) is of gradient type and it has been shown that for every fixed $\mu > 0$

$$V_\mu(u) := \inf_{v \in H^{-1}(D)} V_\mu(u, v) = V(u), \quad (1.8)$$

where $V(u)$ is the quasi-potential associated with (1.2). Moreover, in [9] the non-gradient case has been considered and it has been proven that in spite of the fact that there is no explicit expression for $V_\mu(u, v)$, nevertheless

$$\lim_{\mu \rightarrow 0} V_\mu(u) = V(u),$$

for every sufficiently regular u .

Organization of the paper: In Section 2, we introduce all of our notations and assumptions on the non-linearity b and on the noise w^Q and we discuss some consequences of these choices. In Section 3 we deal with the semigroup $S_\mu(t)$ associated with the linear damped wave equation and we give some generalization of results proved in previous papers. Such generalizations are needed to prove the a-priori bounds required in the proof of (1.7). In Section 4 we collect some results about equations (1.1) and (1.2) and their transition semigroups. In Section 5 we give the main result of the paper and we introduce the method of the proof. In Section 6 we prove limit (1.7) and in the following Sections 7, 8 and 9 we give a proof of all the lemmas needed in the proof of limit (1.7). Finally, in Appendix A we recall for the reader's convenience the proof of the asymptotic strong Feller property for equation (1.2).

2. Assumptions and notations

We assume that D is a bounded open domain in \mathbb{R}^d , having a smooth boundary. We denote by H the Hilbert space $L^2(D)$, endowed with the usual scalar product $\langle \cdot, \cdot \rangle_H$ and the corresponding norm $|\cdot|_H$. The norm in $L^p(D)$ is denoted by $|\cdot|_{L^p}$.

We denote by A the realization of the Laplace operator in H , endowed with Dirichlet boundary conditions. It is known that there exists a complete orthonormal system $\{e_k\}_{k \in \mathbb{N}} \subset H$ and a sequence of positive real numbers $\{\alpha_k\}_{k \in \mathbb{N}}$, diverging to ∞ such that

$$Ae_k = -\alpha_k e_k, \quad k \in \mathbb{N}. \tag{2.1}$$

Next, for any $\beta \in \mathbb{R}$, we denote by H^β the completion of $C_0(D)$ with respect to the norm

$$|x|_{H^\beta}^2 = \sum_{k=1}^{\infty} \alpha_k^\beta |\langle x, e_k \rangle_H|^2.$$

Moreover, we denote $\mathcal{H}_\beta =: H^\beta \times H^{\beta-1}$, and for every $(x, y) \in \mathcal{H}_\beta$, we set

$$\Pi_1(x, y) = x, \quad \Pi_2(x, y) = y.$$

Throughout the paper, for every $\mu > 0$ and $\beta \in \mathbb{R}$, we define

$$A_\mu(x, y) = \frac{1}{\mu}(\mu y, Ax - y), \quad (x, y) \in D(A_\mu) = \mathcal{H}_{\beta+1}. \tag{2.2}$$

It can be proven that A_μ is the generator of a strongly continuous group of bounded linear operators $\{S_\mu(t)\}_{t \geq 0}$ on each \mathcal{H}_β (for a proof see [30, Section 7.4]).

In this paper, we assume that the cylindrical Wiener process $w^Q(t, \xi)$ is white in time and colored in space, with spatial covariance Q^2 , for some $Q \in \mathcal{L}^+(H)$, where $\mathcal{L}^+(H)$ is the space of non-negative and symmetric bounded linear operators on H . This means that $w^Q(t, \xi)$ can be formally represented as the sum

$$w^Q(t, \xi) = \sum_{k=1}^{\infty} Qe_k(\xi)\beta_k(t),$$

where $\{e_k\}_{k \in \mathbb{N}}$ is the complete orthonormal system of H that diagonalizes A and $\{\beta_k(t)\}_{k \in \mathbb{N}}$ is a sequence of mutually independent Brownian motions, all defined on the same stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$.

We impose the following two conditions on Q in what follows. The first condition ensures sufficient spatial smoothness to define solutions in the right functional space and prove a-priori bounds for them. The second condition guarantees that the noise

acts directly through a large enough subspace of the phase space, in order to ensure smoothing in the Markovian dynamics associated to (1.2) (see also Remark A.1)

Hypothesis 1.

1. There exists a non-negative sequence $\{\lambda_k\}_{k \in \mathbb{N}}$ such that

$$Qe_k = \lambda_k e_k, \quad k \in \mathbb{N}. \tag{2.3}$$

Moreover, there exists $\delta > 0$ such that

$$\text{Tr} [Q^2(-A)^\delta] = \sum_{k=1}^{\infty} \lambda_k^2 \alpha_k^\delta < \infty. \tag{2.4}$$

2. There exist $\bar{n} \in \mathbb{N}$ and $c > 0$ such that

$$|Qx|_H \geq c |P_{\bar{n}} x|_H, \quad x \in H, \tag{2.5}$$

where P_n is the projection of H onto $H_n := \text{span}\{e_1, \dots, e_n\}$, for every $n \in \mathbb{N}$ (concerning \bar{n} , see (2.10)).

In what follows, for any $\mu > 0$ we define the bounded linear operator

$$Q_\mu : H \rightarrow \mathcal{H}_1, \quad x \in H \mapsto \frac{1}{\mu}(0, Qx) \in \mathcal{H}_1. \tag{2.6}$$

Concerning the non-linearity b , we assume the following conditions.

Hypothesis 2.

1. The mapping $b : D \times \mathbb{R} \rightarrow \mathbb{R}$ is measurable and $b(\xi, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is of class C^1 , for almost all $\xi \in D$. Moreover,

$$\sup_{(\xi, \sigma) \in D \times \mathbb{R}} \partial_\sigma b(\xi, \sigma) =: L_b < \infty. \tag{2.7}$$

2. There exists $\lambda \in [1, 3]$ such that for every $\sigma \in \mathbb{R}$

$$\sup_{\xi \in D} |b(\xi, \sigma)| \leq c (1 + |\sigma|^\lambda), \quad \sup_{\xi \in D} |\partial_\sigma b(\xi, \sigma)| \leq c (1 + |\sigma|^{\lambda-1}). \tag{2.8}$$

3. If $\lambda = 1$, we have

$$L_b < \alpha_1. \tag{2.9}$$

4. If $\lambda \in (1, 3]$, we have

$$L_b < \alpha_{\bar{n}}, \tag{2.10}$$

where \bar{n} is the integer introduced in (2.5). Moreover, there exist $c_1 > 0$ and $c_2 \in \mathbb{R}$ such that for every $(\xi, \sigma) \in D \times \mathbb{R}$

$$\sup_{\xi \in D} b(\xi, \sigma)\sigma \leq -c_1 |\sigma|^{\lambda+1} + c_2. \tag{2.11}$$

5. If $\lambda \in (1, 3]$, then $d = 1$ and $D = [0, L]$, for some $L > 0$.

Remark 2.1.

1. Let $a : D \rightarrow \mathbb{R}$ be a measurable function such that

$$0 < a_0 \leq a(\xi) \leq a_1, \quad \xi \in D,$$

and let $f : D \times \mathbb{R} \rightarrow \mathbb{R}$ be a measurable mapping such that $f(\xi, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz continuous, uniformly with respect to $\xi \in D$. Then, if $\lambda \in (1, 3]$, the function b defined by

$$b(\xi, \sigma) = -a(\xi)|\sigma|^{\lambda-1}\sigma + f(\xi, \sigma), \quad (\xi, \sigma) \in D \times \mathbb{R},$$

satisfies all conditions in Hypothesis 2. In case $f = 0$, b is the so-called Klein-Gordon nonlinearity.

2. If $\lambda = 1$, for every $\epsilon > 0$ we can find $c_\epsilon > 0$ such that

$$\sup_{\xi \in D} b(\xi, \sigma)\sigma \leq (L_b + \epsilon) |\sigma|^2 + c_\epsilon, \quad \sigma \in \mathbb{R}, \tag{2.12}$$

where L_b is the constant introduced in (2.7).

3. If $\lambda \in (1, 3]$, from (2.8) and (2.11) it follows that there exist $\kappa_1 > 0$ and $\kappa_2 \in \mathbb{R}$ such that for every $\sigma, \rho \in \mathbb{R}$

$$\sup_{\xi \in D} b(\xi, \sigma + \rho)\sigma \leq -\kappa_1 |\sigma|^{\lambda+1} + \kappa_2 (1 + |\rho|^{\lambda+1}). \tag{2.13}$$

4. If $\lambda = 1$, as a consequence of (2.9), it is possible to prove that for every $\xi \in D$ there exists an antiderivative $\mathfrak{b}(\xi, \cdot)$ of $b(\xi, \cdot)$ such that

$$\sup_{\xi \in D} \mathfrak{b}(\xi, \sigma) \leq L_b |\sigma|^2, \quad \sigma \in \mathbb{R}, \tag{2.14}$$

where L_b is the constant introduced in (2.7).

5. In the same way, when $\lambda \in (1, 3]$, due to (2.10) and (2.11) it is possible to prove that for every $\xi \in D$ there exists an antiderivative $\mathfrak{b}(\xi, \cdot)$ of $b(\xi, \cdot)$ such that

$$\sup_{\xi \in D} \mathfrak{b}(\xi, \sigma) \leq -\kappa |\sigma|^{\lambda+1}, \quad \sigma \in \mathbb{R}, \tag{2.15}$$

for some constant $\kappa > 0$. \square

Now, for every $x \in H$, we define

$$B(x)(\xi) =: b(\xi, x(\xi)), \quad \xi \in D.$$

Notice that if $\lambda = 1$, then $b(\xi, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz-continuous, uniformly with respect to $\xi \in D$, so that $B : H \rightarrow H$ is Lipschitz-continuous. Moreover, as shown in [32, Theorem 4], for every $u \in H^\beta$, with $0 \leq \beta \leq 1$, we have

$$|B(u)|_{H^\beta} \leq c(1 + |u|_{H^\beta}). \tag{2.16}$$

Notice moreover that in the general case, due to (2.8), for every $u \in H^1 \cap L^\infty(D)$

$$|B(u)|_{H^1} \leq |\partial_\sigma b(\cdot, u)|_{L^\infty} |u|_{H^1} \leq c(1 + |u|_{L^\infty}^{\lambda-1}) |u|_{H^1}. \tag{2.17}$$

Next, for every $\mu > 0$ and $\beta \in [0, 1]$, we set

$$B_\mu(x, y)(\xi) =: \frac{1}{\mu} (0, B(x))(\xi) = \frac{1}{\mu} (0, b(\xi, x(\xi))), \quad \xi \in D, \quad (x, y) \in \mathcal{H}_\beta. \tag{2.18}$$

If $\lambda = 1$, then the Lipschitz continuity of $B : H \rightarrow H$ implies that, for every $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ in \mathcal{H}_β ,

$$\begin{aligned} |B_\mu(z_1) - B_\mu(z_2)|_{\mathcal{H}_\beta} &= \frac{1}{\mu} |B(x_1) - B(x_2)|_{H^{\beta-1}} \leq \frac{c}{\mu} |B(x_1) - B(x_2)|_H \\ &\leq \frac{cM}{\mu} |x_1 - x_2|_H \leq \frac{cM}{\mu} |z_1 - z_2|_{\mathcal{H}_\beta}. \end{aligned}$$

This means that the mapping $B_\mu : \mathcal{H}_\beta \rightarrow \mathcal{H}_\beta$ is Lipschitz continuous, for every $\beta \in [0, 1]$ and $\mu > 0$.

If $\lambda \in (1, 3]$, the mapping $b(\xi, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is only locally Lipschitz-continuous and does not have sub-linear growth in general. Since

$$\sup_{(\xi, \sigma) \in D \times \mathbb{R}} \frac{|\mathfrak{b}(\xi, \sigma)|}{(|\sigma|^{\lambda+1} + 1)} < \infty,$$

due to (2.15), we have

$$-c(|\sigma|^{\lambda+1} + 1) \leq \mathfrak{b}(\xi, \sigma) \leq -\kappa |\sigma|^{\lambda+1}, \quad (\xi, \sigma) \in D \times \mathbb{R}. \tag{2.19}$$

On the other hand, if $\lambda = 1$, due to (2.14) we have

$$-c (1 + |\sigma|^2) \leq \mathfrak{b}(\xi, \sigma) \leq L_b |\sigma|^2, \quad (\xi, \sigma) \in D \times \mathbb{R}, \tag{2.20}$$

where L_b is the constant introduced in (2.9).

Recall that when $\lambda \in (1, 3]$ we assume $D = [0, L]$, so that $H^1 \hookrightarrow L^\infty(D) \hookrightarrow L^p(D)$, for every $p \geq 1$. This implies, together with (2.8), that for every $z = (x, y) \in \mathcal{H}_1$

$$|B_\mu(z)|_{\mathcal{H}_1} = \frac{1}{\mu} |b(\cdot, x)|_H \leq \frac{c}{\mu} \left(1 + |x|_{L^{2\lambda}(D)}^\lambda \right) \leq \frac{c}{\mu} (1 + |x|_{H^1}^\lambda) \leq \frac{c}{\mu} (1 + |z|_{\mathcal{H}_1}^\lambda),$$

for every $\mu > 0$, for some c independent of μ . Moreover, by the same arguments, for every $z_1, z_2 \in \mathcal{H}_1$

$$|B_\mu(z_1) - B_\mu(z_2)|_{\mathcal{H}_1} \leq \frac{c}{\mu} |x_1 - x_2|_{H^1} (1 + |x_1|_{H^1}^{\lambda-1} + |x_2|_{H^1}^{\lambda-1}). \tag{2.21}$$

In particular, $B_\mu : \mathcal{H}_1 \rightarrow \mathcal{H}_1$ is locally Lipschitz-continuous.

3. The approximating semigroup $S_\mu(t)$

In the present section, we present some modifications of the results proven in [10] about the approximating semigroup $S_\mu(t)$, in the slightly different setting where the motion of the particle with small mass is subject to a magnetic field and a friction. These modifications are necessary here because, in order to prove (5.7) we need some uniformity with respect to the initial conditions on bounded sets.

For every $\mu > 0$, $\beta \in \mathbb{R}$ and $(x, y) \in \mathcal{H}_\beta$, we shall denote

$$u_\mu(t) := \Pi_1 S_\mu(t)(x, y), \quad v_\mu(t) := \Pi_2 S_\mu(t)(x, y), \quad t \geq 0,$$

where $S_\mu(t)$ is the semigroup generated in \mathcal{H}_β by the operator A_μ and introduced in (2.2).

In [10, Lemma 3.1], it has been shown that for every $\mu > 0$, $\beta \in \mathbb{R}$ and $(x, y) \in \mathcal{H}_\beta$

$$\mu |v_\mu(t)|_{H^{\beta-1}}^2 + |u_\mu(t)|_{H^\beta}^2 + 2 \int_0^t |v_\mu(s)|_{H^{\beta-1}}^2 ds = \mu |y|_{H^{\beta-1}}^2 + |x|_{H^\beta}^2, \tag{3.1}$$

and

$$\mu |u_\mu(t)|_{H^\beta}^2 + |\mu v_\mu(t) + u_\mu(t)|_{H^{\beta-1}}^2 + 2 \int_0^t |u_\mu(s)|_{H^\beta}^2 ds = \mu |x|_{H^\beta}^2 + |\mu y + x|_{H^{\beta-1}}^2. \tag{3.2}$$

In particular, the two equalities above imply that for any $\mu > 0$ there exists $c_\mu > 0$ such that for any $(x, y) \in \mathcal{H}_\beta$

$$\int_0^\infty |S_\mu(t)(x, y)|_{\mathcal{H}_\beta}^2 dt \leq \frac{c_\mu}{2} |(x, y)|_{\mathcal{H}_\beta}^2,$$

and, as a consequence of the Datko theorem, we can conclude that there exist $M_{\mu,\beta}$, and $\omega_{\mu,\beta} > 0$ such that

$$\|S_\mu(t)\|_{\mathcal{L}(\mathcal{H}_\beta)} \leq M_{\mu,\beta} e^{-\omega_{\mu,\beta}t}, \quad t \geq 0.$$

Moreover, in [10, Lemma 3.2] it has been proven that for any $\mu > 0$, and for any $\beta \in \mathbb{R}$ it holds

$$|\Pi_1 S_\mu(t)(0, y)|_{H^\beta} \leq 2\mu |y|_{H^\beta}, \quad t \geq 0, \quad y \in H^\beta. \tag{3.3}$$

The following result has been proven in [10, Lemma 3.3]. Notice that in [10], estimate (3.4) is given only for $\beta = 0$, but the case $\beta > 0$ is an easy generalization and we omit its proof.

Lemma 3.1. *For any $0 < \rho < 1$ there exists a constant $c_\rho > 0$ such that for any $k \in \mathbb{N}$ and $\beta \geq 0$*

$$\sup_{\mu > 0} \int_0^\infty s^{-\rho} |\Pi_1 S_\mu(s) Q_\mu e_k|_{H^\beta}^2 ds \leq c_\rho \frac{\lambda_k^2}{\alpha_k^{1-(\rho+\beta)}}, \tag{3.4}$$

where Q_μ is as in (2.6) and λ_k and α_k denote the eigenvalues of Q and A , respectively. In particular, (3.4) implies that for every $p \geq 1$, $T > 0$ and $\delta > 0$

$$\sum_{k=1}^\infty \frac{\lambda_k^2}{\alpha_k^{1-\delta}} < \infty \implies \sup_{\mu > 0} \mathbb{E} \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^\beta}^p < +\infty, \quad \beta < \delta, \tag{3.5}$$

where $\Gamma_\mu(t)$ is the stochastic convolution as in (4.3).

Next, we denote by $S(t)$, $t \geq 0$, the semigroup generated by the operator A in H^β , for every $\beta \in \mathbb{R}$. The following two approximation results are a modification of what proven in [10, Theorem 3.5 and Corollaries 3.6 and 3.7]. Such a modification is critical for the proof of Theorem 6.1, our main finite-time convergence result below.

Lemma 3.2. *Let $T > 0$ and $\beta > 0$ be fixed. Then, for every $R > 0$ it holds*

$$\lim_{\mu \rightarrow 0} \sup_{|x|_{H^\beta} + \sqrt{\mu} |y|_H \leq R} \sup_{t \leq T} |\Pi_1 S_\mu(t)(x, y) - S(t)x|_H = 0. \tag{3.6}$$

Proof. If we denote by P_n the projection of H onto $\text{span}\{e_1, \dots, e_n\}$, we have

$$\begin{aligned} |\Pi_1 S_\mu(t)(x, y) - S(t)x|_H &\leq |\Pi_1 S_\mu(t)(0, y)|_H + |\Pi_1 S_\mu(t)(P_n x, 0) - S(t)P_n x|_H \\ &+ |\Pi_1 S_\mu(t)(x - P_n x, 0)|_H + |S(t)(x - P_n x)|_H =: I_1^\mu(t) + I_{2,n}^\mu(t) + I_{3,n}^\mu(t) + I_{4,n}(t). \end{aligned}$$

As a consequence of (3.3), for every $t \geq 0$ and $\mu > 0$ we have

$$I_1^\mu(t) \leq 2\mu |y|_H. \tag{3.7}$$

Moreover, as shown in [10, proof of Theorem 3.5], for every $n \in \mathbb{N}$ there exists some $c_T(n)$ such that for every $t \in [0, T]$ and $\mu > 0$

$$I_{2,n}^\mu(t) \leq \mu c_T(n) |x|_H. \tag{3.8}$$

Moreover, as a consequence of (3.1), for every $n \in \mathbb{N}$, $t \in [0, T]$ and $\mu > 0$ we have

$$I_{3,n}^\mu(t) + I_{4,n}(t) \leq c_T |P_n x - x|_H \leq c_T \|P_n - I\|_{\mathcal{L}(H^\beta, H)} |x|_{H^\beta}.$$

In particular, if we fix $R > 0$ and $\epsilon > 0$, we can find $\bar{n} = n(\epsilon, R, T)$ such that

$$|x|_{H^\beta} \leq R \implies I_{3,\bar{n}}^\mu(t) + I_{4,\bar{n}}(t) \leq \epsilon, \quad \mu > 0.$$

Thanks to (3.7) and (3.8), this implies that

$$|x|_{H^\beta} + \sqrt{\mu} |y|_H \leq R \implies \sup_{t \in [0, T]} |\Pi_1 S_\mu(t)(x, y) - S(t)x|_H \leq 2\sqrt{\mu} R + \mu c_T(\bar{n})R + \epsilon,$$

and, due to the arbitrariness of ϵ , (3.6) follows. \square

Lemma 3.3. *Let $T > 0$ and $\beta > 0$ be fixed. Then, for any $R > 0$*

$$\begin{aligned} \lim_{\mu \rightarrow 0} \sup_{|\psi|_{L^1(0, T; H^\beta)} \leq R} \sup_{t \in [0, T]} \left| \frac{1}{\mu} \int_0^t \Pi_1 S_\mu(t-s)(0, \psi(s)) ds - \int_0^t S(t-s)\psi(s) ds \right|_H \\ = 0. \end{aligned} \tag{3.9}$$

Proof. For every $\mu > 0$ and $t \geq 0$, we define

$$\Phi_\mu(t)y = \frac{1}{\mu} \Pi_1 S_\mu(t)(0, y) - S(t)y, \quad y \in H. \tag{3.10}$$

Then, for every $n \in \mathbb{N}$ we have

$$\begin{aligned} & \frac{1}{\mu} \int_0^t \Pi_1 S_\mu(t-s)(0, \psi(s)) ds - \int_0^t S(t-s)\psi(s) ds \\ &= \int_0^t \Phi_\mu(t-s) [\psi(s) - \psi_n(s)] ds + \int_0^t \Phi_\mu(t-s)\psi_n(s) ds, \end{aligned}$$

where

$$\psi_n(s) := I_{\{|\psi(s)|_H \leq n\}} P_n \psi(s), \quad s \geq 0.$$

As shown in [10, proof of Theorem 3.5], for every $n \in \mathbb{N}$, there exists $c_T(n)$ such that for every $t \in [0, T]$

$$|\Phi_\mu(t) P_n y|_H \leq c_T(n) \left(e^{-\frac{t}{\mu}} + \mu \right) |y|_H. \tag{3.11}$$

Moreover, due to (3.3), for every $\beta \in [0, 1]$, we have

$$\|\Phi_\mu(t)\|_{\mathcal{L}(H^\beta)} \leq M_{\beta,T}, \quad t \in [0, T]. \tag{3.12}$$

Hence

$$\begin{aligned} & \left| \frac{1}{\mu} \int_0^t \Pi_1 S_\mu(t-s)(0, \psi(s)) ds - \int_0^t S(t-s)\psi(s) ds \right|_H \\ & \leq c M_{0,T} \int_0^t |\psi(s) - \psi_n(s)|_H ds + c_T(n) \int_0^t \left(e^{-\frac{t-s}{\mu}} + \mu \right) I_{\{|\psi(s)|_H \leq n\}} |\psi(s)|_H ds \\ & \leq c M_{0,T} \|P_n - I\|_{\mathcal{L}(H^\beta, H)} \|\psi\|_{L^1(0,T;H^\beta)} + n c_T(n) (1+T) \mu. \end{aligned}$$

Thus, if we fix $\epsilon > 0$ and pick $\bar{n} = n(T, R, \epsilon) \in \mathbb{N}$ such that

$$c M_{0,T} \|P_{\bar{n}} - I\|_{\mathcal{L}(H^\beta, H)} R \leq \epsilon,$$

we get

$$\begin{aligned} & \limsup_{\mu \rightarrow 0} \sup_{|\psi|_{L^1(0,T;H^\beta)} \leq R} \left| \frac{1}{\mu} \int_0^t \Pi_1 S_\mu(t-s)(0, \psi(s)) ds - \int_0^t S(t-s)\psi(s) ds \right|_H \\ & \leq \epsilon + \lim_{\mu \rightarrow 0} \bar{n} c_T(\bar{n}) (1+T) \mu = \epsilon. \end{aligned}$$

As ϵ is arbitrary, this yields (3.9). \square

4. Preliminary results on equations (1.1) and (1.2)

If we denote $z(t) = (u(t), \partial_t u(t))$ and

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

with the notation that we have just introduced in Section 2, equation (1.1) can be rewritten as the following abstract stochastic evolution equation in \mathcal{H}_β

$$dz_\mu(t) = [A_\mu z_\mu(t) + B_\mu(z_\mu(t))] dt + Q_\mu dw(t), \quad z_\mu(0) = (x, y) \in \mathcal{H}_\beta. \tag{4.1}$$

In case $\lambda = 1$, if

$$\sum_{k=1}^{\infty} \frac{\lambda_k^2}{\alpha_k^{1-\delta}} < \infty,$$

for some $\delta \in [0, 1]$, then for every $\beta < \delta$ and any random initial condition $\gamma = (\theta, \eta) \in L^p(\Omega; \mathcal{H}_\beta)$ equation (4.1) admits a unique mild solution $z_\mu^\gamma = (u_\mu^\gamma, \partial_t u_\mu^\gamma) \in L^p(\Omega; C([0, T]; \mathcal{H}_\beta))$, for every $p \geq 1, T > 0$ and $\mu > 0$. This means that

$$z_\mu^\gamma(t) = S_\mu(t)\gamma + \int_0^t S_\mu(t-s)B_\mu(z_\mu^\gamma(s)) ds + \Gamma_\mu(t), \tag{4.2}$$

where

$$\Gamma_\mu(t) := \int_0^t S_\mu(t-s)Q_\mu dw(s), \tag{4.3}$$

(for a proof see e.g. [12, Theorem 7.2]).

In [1] and [11] it is proven that in case $\lambda \in (1, 3]$, for every $\mu > 0$ and $\gamma = (\theta, \eta) \in L^p(\Omega; \mathcal{H}_1)$ there exists a unique mild solution $z_\mu^\gamma = (u_\mu^\gamma, \partial_t u_\mu^\gamma) \in L^p(\Omega; C([0, T]; \mathcal{H}_1))$, for every $p \geq 1$ and $T > 0$.

In what follows, we shall denote by P_t^μ the transition semigroup associated with equation (4.1) in the space \mathcal{H}_1 . Namely,

$$P_t^\mu \varphi(z) = \mathbb{E} \varphi(z_\mu^z(t)), \quad t \geq 0, \quad z \in \mathcal{H}_1,$$

for every Borel bounded $\varphi : \mathcal{H}_1 \rightarrow \mathbb{R}$. In [1] and [11], the long time behavior of equation (4.1) has been studied and it has been proven that, under Hypotheses 1 and 2, for every fixed $\mu > 0$ the semigroup P_t^μ admits an invariant measure ν_μ . In both papers conditions for uniqueness of the invariant measure are also given.

Next, we consider equation (1.2). If Hypotheses 1 and 2 hold, then for any $T > 0$, $p \geq 1$ and $\theta \in L^p(\Omega; H)$, equation (1.2) admits a unique mild solution $u^\theta \in L^p(\Omega; C([0, T]; H))$. This means that if $S(t)$ denotes the semigroup generated by the operator A in H , then

$$u^\theta(t) = S(t)\theta + \int_0^t S(t-s)B(u^\theta(s)) ds + \Gamma(t),$$

where

$$\Gamma(t) := \int_0^t S(t-s)dw^Q(s). \tag{4.4}$$

Moreover, it can be shown that

$$\mathbb{E} \sup_{t \in [0, T]} |u^\theta(t)|_H^2 \leq c_T (1 + \mathbb{E} |\theta|_H^2), \tag{4.5}$$

(for a proof, see e.g. [12]).

If $\lambda = 1$, then it is possible to show that in fact for every $\beta \in [0, 1]$ and $\theta \in L^p(\Omega; H^\beta)$ equation (1.2) is well posed in $L^p(\Omega; C([0, T]; H^\beta))$ and

$$\mathbb{E} \sup_{t \in [0, T]} |u^\theta(t)|_{H^\beta}^p \leq c_{\beta, T} (1 + \mathbb{E} |\theta|_{H^\beta}^p). \tag{4.6}$$

In case $\lambda \in (1, 3]$, if $\theta \in L^p(\Omega; H^1)$, then

$$\mathbb{E} \sup_{t \in [0, T]} |u^\theta(t)|_{H^1}^p \leq c_T (1 + \mathbb{E} |\theta|_{H^1}^p) \tag{4.7}$$

(for a proof, see e.g. [12]).

Now, we denote by P_t the transition semigroup associated with equation (1.2) in H ,

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

for every Borel bounded $\varphi : H \rightarrow \mathbb{R}$. It is possible to prove that under Hypotheses 1 and 2, the semigroup P_t admits an invariant measure ν (for a proof in the case $\lambda \in (1, 3]$, see [3, Proposition 8.2.2]).

Lemma 4.1. *Assume Hypotheses 1 and 2. Then, for every $M, \epsilon > 0$ there exists $t_\star = t_\star(M, \epsilon) > 0$ such that*

$$\inf_{|x|_H \leq M} \mathbb{P} (|u^x(t)|_H < \epsilon) > 0, \quad t \geq t_\star. \tag{4.8}$$

Moreover, in the case $\lambda \in (1, 3]$ we have

$$\inf_{x \in H} \mathbb{P} (|u^x(t)|_H < \epsilon) > 0, \quad t \geq t_*, \tag{4.9}$$

for some $t_* = t_*(\epsilon) > 0$.

Proof. We give here a proof for (4.9), that holds when $\lambda \in (1, 3]$. The proof of (4.8) is left to the reader. We define $v^x(t) = u^x(t) - \Gamma(t)$, where $\Gamma(t)$ is the stochastic convolution defined in (4.4). We have that v^x solves the random equation

$$\frac{d}{dt}v^x(t) = Av(t) + B(v^x(t) + \Gamma(t)), \quad v^x(0) = x.$$

Due to (2.13), we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |v^x(t)|_H^2 &= \langle Av^x(t), v^x(t) \rangle_H + \langle B(v^x(t) + \Gamma(t)), v^x(t) \rangle_H \\ &\leq -\kappa_1 |v^x(t)|_{L^{\lambda+1}}^{\lambda+1} + \kappa_2 (1 + |\Gamma(t)|_{L^{\lambda+1}}^{\lambda+1}) \\ &\leq -\kappa_1 |D|^{\frac{\lambda-1}{2}} |v^x(t)|_H^{\lambda+1} + \kappa_2 (1 + |\Gamma(t)|_{L^{\lambda+1}}^{\lambda+1}). \end{aligned}$$

Since $\lambda + 1 > 2$, by a comparison argument (see [3, Lemma 1.2.6]) the inequality above yields

$$\sup_{x \in H} |v^x(t)|_H \leq c \max \left(\sup_{s \in [0,t]} |\Gamma(s)|_{L^{\lambda+1}}^{\frac{1}{\lambda+1}}, t^{-\frac{\lambda-1}{4}} \right), \quad t > 0,$$

for some constant $c > 0$ depending on λ . This implies that there exists $t_* = t_*(\epsilon) > 0$ such that

$$\inf_{x \in H} \mathbb{P} (|v^x(t)|_H < \epsilon/2) \geq \mathbb{P} \left(\sup_{s \in [0,t]} |\Gamma(s)|_{L^{\lambda+1}} < \left(\frac{\epsilon}{2c} \right)^{\lambda+1} \right), \quad t \geq t_*.$$

Since $|u^x(t)|_H \leq |v^x(t)|_H + |\Gamma(t)|_H \leq |v^x(t)|_H + |D|^{\frac{\lambda-1}{2(\lambda+1)}} |\Gamma(t)|_{L^{\lambda+1}}$ and we can assume $\epsilon \leq 1$, from the inequality above we obtain

$$\inf_{x \in H} \mathbb{P} (|u^x(t)|_H < \epsilon) \geq \mathbb{P} \left(\sup_{s \in [0,t]} |\Gamma(s)|_{L^{\lambda+1}} < c\epsilon^{\lambda+1} \right), \quad t \geq t_*,$$

for some constant $c > 0$. Since the law of $\Gamma(\cdot)$ is full in $C([0, t]; L^{\lambda+1}(D))$, this implies (4.9). \square

In [3] the *strong Feller property* of the semigroup P_t is studied and it proven that, under suitable non-degeneracy conditions on the covariance Q^2 , for every $\varphi \in B_b(H)$ and $t > 0$ we have $P_t\varphi \in C_b^1(H)$, and

$$\sup_{x \in H} |D(P_t\varphi)(x)|_H \leq c(t \wedge 1)^{-\rho} \sup_{x \in H} |\varphi(x)|,$$

for some constant $\rho > 0$ (see [3, Proposition 4.4.3 and Theorem 7.3.1]). If, as in the present paper, we do not want to assume those restrictive assumptions on the non-degeneracy of the noise, the strong Feller property does not hold anymore and it has to be replaced by the so-called *asymptotic strong Feller property*. As a matter of fact, thanks to condition (2.5), it is known that for every $\varphi \in C_b^1(H)$ and $t \geq 0$

$$|D(P_t\varphi)(x)|_H \leq c\sqrt{P_t|\varphi|^2(x)} + \alpha(t)\sqrt{P_t|D\varphi|_H^2(x)}, \quad x \in H, \tag{4.10}$$

for some function $\alpha(t)$ such that $\alpha(t) \rightarrow 0$, as $t \rightarrow \infty$. In Appendix A we give a proof of this bound for the reader’s convenience.

Moreover, in [3] it is proven that the solution $u^x(t)$ of equation (1.2) is differentiable with respect to the initial condition $x \in H$ and if $D_x u^x(t)h$ denotes its derivative along the direction $h \in H$, it holds

$$\sup_{x \in H} |D_x u^x(t)h|_H \leq c(t) |h|_H, \quad \mathbb{P} - \text{a.s.} \tag{4.11}$$

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

5. The main result

In the previous section, we have seen that there exists a sequence of probability measures $\{\nu_\mu\}_{\mu>0} \subset \mathcal{P}(\mathcal{H}_1)$ which are invariant for the semigroup P_t^μ . The family $\{\Pi_1\nu_\mu\}_{\mu>0} \subset \mathcal{P}(H^1)$ can be extended to the family $\{(\Pi_1\nu_\mu)'\}_{\mu>0} \subset \mathcal{P}(H)$, by setting every Borel set $A \subset H$

$$(\Pi_1\nu_\mu)'(A) = \Pi_1\nu_\mu(A \cap H^1), \quad \mu > 0.$$

By using the results proven in [19], it is possible to show that in the case we are considering conditions (4.8) and (4.9) and the gradient estimate (4.11), together with (4.10), imply that there exist a distance α on H and some constant $\delta > 0$ such that

$$\rho_\alpha(P_t^*\nu_1, P_t^*\nu_2) \leq ce^{-\delta t} \rho_\alpha(\nu_1, \nu_2), \quad t \geq 0, \tag{5.1}$$

for every $\nu_1, \nu_2 \in \mathcal{Pr}(H)$. Here $P_t^*\nu$ is the measure defined by

$$\int_H \varphi(x) dP_t^*\nu(x) = \int_H P_t\varphi(x) d\nu(x),$$

for every $\varphi \in C_b(H)$, and ρ_α is the Wasserstein distance in $\mathcal{P}(H)$ associated with α , defined by

$$\rho_\alpha(\nu_1, \nu_2) = \inf_{[\varphi]_{\text{Lip}_{H,\alpha}} \leq 1} \left| \int_{H^1} \varphi(x) d\nu_1(x) - \int_{H^1} \varphi(x) d\nu_2(x) \right|,$$

where

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

Notice that it can be proven that

$$\rho_\alpha(\nu_1, \nu_2) = \inf \mathbb{E} \alpha(X, Y), \tag{5.2}$$

over all random variables X and Y distributed as ν_1 and ν_2 , respectively (for a proof see [35]).

The choice of a distance α on H which allows to obtain the contraction property (5.1) strongly depends on the type of non-linearity b we are dealing with.

In the case the constant λ in Hypothesis 2 belongs to $(1, 3]$, the uniform condition (4.9) holds. By proceeding as in [14, Lemma 4.1] it is possible to show that (4.9) implies that for every $\epsilon > 0$ there exists $t_\star = t_\star(\epsilon) > 0$ such that

$$\inf_{x, y \in H} \sup_{\Gamma \in \mathcal{C}(P_t^\star \delta_x, P_t^\star \delta_y)} \Gamma((x', y') \in H \times H : |x' - y'|_H < \epsilon) > 0, \quad t \geq t_\star,$$

where $\mathcal{C}(P_t^\star \delta_x, P_t^\star \delta_y)$ denotes the collection of all couplings of the measures $P_t^\star \delta_x$ and $P_t^\star \delta_y$. This means that [19, Assumption 3] is satisfied. Moreover, since we have (4.10), also [19, Assumption 2] is satisfied. Therefore, thanks to [19, Theorem 2.5] we have that (5.1) holds for

$$\alpha(x, y) = |x - y|_H. \tag{5.3}$$

In the case $\lambda = 1$, bound (4.8) is uniform only on bounded sets of H . Hence a Lyapunov structure has to be taken into the picture and, as proven in [19, Theorem 3.4], condition (5.1) is true with the following distance α

$$\alpha(x, y) = \inf_{\gamma} \int_0^1 \exp(\eta |\gamma(t)|_H^2) |\gamma'(t)|_H dt, \tag{5.4}$$

where the infimum is taken over all paths γ such that $\gamma(0) = x$ and $\gamma(1) = y$ and where η is some positive constant, which is less than a value η_0 determined by $\alpha_1 - L_b$. Notice that for every $x, y \in H$

$$|x - y|_H \leq \alpha(x, y) \leq \exp(\eta(|x|_H^2 + |y|_H^2)) |x - y|_H \tag{5.5}$$

By proceeding again as in [14, Lemma 4.1], it is possible to show that (4.8) implies that for every $M, \epsilon > 0$ there exists $t_\star = t_\star(M, \epsilon) > 0$ such that

$$\inf_{x, y \in B_H(M)} \sup_{\Gamma \in \mathcal{C}(P_t^\star \delta_x, P_t^\star \delta_y)} \Gamma((x', y') \in H \times H : |x' - y'|_H < \epsilon) > 0, \quad t \geq t_\star.$$

This means that [19, Assumption 6] is verified. Moreover, if we define $U(t) = \eta |u(t)|_H^2$, from Itô's formula we have

$$dU(t) = \eta [\text{Tr}Q^2 - |u(t)|_{H^1}^2 + \langle B(u(t)), u(t) \rangle_H] dt + 2\eta |Qu(t)|_H d\beta(t),$$

for some Brownian motion $\beta(t)$. It is then immediate to check that we can apply [19, Lemma 5.1], so that [19, Assumption 4] holds for $V(x) = \exp(\eta |x|_H^2)$. Finally, due to (4.10), [19, Assumption 5] holds. This means that we can apply [19, Theorem 3.4] and conclude that (5.1) holds for α defined as in (5.4).

In what follows, we shall prove that, if ν is the unique invariant measure of equation (1.2), then the following result holds.

Theorem 5.1. *Under Hypotheses 1 and 2, we have*

$$\lim_{\mu \rightarrow 0} \rho_\alpha((\Pi_1 \nu_\mu)', \nu) = 0, \tag{5.6}$$

where α is either (5.3) or (5.4), in the case $\lambda \in (1, 3]$ and $\lambda = 1$, respectively.

Proof. The method used here is analogous to the one used for example in [14,15]. Due to the invariance of ν_μ and ν , for every $t \geq 0$ we have

$$\begin{aligned} \rho_\alpha((\Pi_1 \nu_\mu)', \nu) &= \rho_\alpha((\Pi_1 (P_t^\mu)^\star \nu_\mu)', P_t^\star \nu) \\ &\leq \rho_\alpha((\Pi_1 (P_t^\mu)^\star \nu_\mu)', P_t^\star (\Pi_1 \nu_\mu)') + \rho_\alpha(P_t^\star (\Pi_1 \nu_\mu)', P_t^\star \nu). \end{aligned}$$

Hence, thanks to (5.1), we obtain

$$\rho_\alpha((\Pi_1 \nu_\mu)', \nu) \leq \rho_\alpha((\Pi_1 (P_t^\mu)^\star \nu_\mu)', P_t^\star (\Pi_1 \nu_\mu)') + c e^{-\delta t} \rho_\alpha((\Pi_1 \nu_\mu)', \nu).$$

This means that, if we pick $t_\star > 0$ such that $c e^{-\delta t_\star} \leq 1/2$, we have

$$\rho_\alpha((\Pi_1 \nu_\mu)', \nu) \leq 2 \rho_\alpha((\Pi_1 (P_t^\mu)^\star \nu_\mu)', P_t^\star (\Pi_1 \nu_\mu)'), \quad t \geq t_\star.$$

In particular, (5.6) follows once we are able to show that

$$\lim_{\mu \rightarrow 0} \rho_\alpha((\Pi_1 (P_t^\mu)^\star \nu_\mu)', P_t^\star (\Pi_1 \nu_\mu)') = 0.$$

Due to (5.2), we have that

$$\rho_\alpha((\Pi_1(P_t^\mu)^* \nu_\mu)', P_t^*(\Pi_1 \nu_\mu)') \leq \mathbb{E} \alpha(u^{\gamma_\mu}(t), u^{\theta_\mu}(t)),$$

where $\gamma_\mu = (\theta_\mu, \eta_\mu)$ is any \mathcal{H}_1 -valued random variable such that $\mathcal{L}(\gamma_\mu) = \nu_\mu$. Therefore, (5.6) follows once we show that

$$\lim_{\mu \rightarrow 0} \mathbb{E} \alpha(u^{\gamma_\mu}(t), u^{\theta_\mu}(t)) = 0. \quad \square \tag{5.7}$$

The proof of limit (5.7), which concludes the proof of Theorem 5.1, is given in the next section.

6. Proof of limit (5.7)

In Section 4, we have seen that for every $\mu > 0$ equation (4.1) admits an invariant measure. Let us denote by $\{\nu_\mu\}_{\mu > 0}$ any family of invariant measures for equation (1.1).

Theorem 6.1. *Assume Hypotheses 1 and 2. For every $\mu > 0$, let $\gamma_\mu = (\theta_\mu, \eta_\mu)$ be a \mathcal{F}_0 -measurable \mathcal{H}_1 -valued random variable, distributed as the invariant measure ν_μ . Then, for every $T > 0$ we have*

$$\lim_{\mu \rightarrow 0} \sup_{t \in [0, T]} \mathbb{E} \alpha(u_\mu^{\gamma_\mu}(t), u^{\theta_\mu}(t)) = 0. \tag{6.1}$$

We will prove the theorem above by distinguishing the case $b(\xi, \cdot)$ is Lipschitz-continuous (that is $\lambda = 1$) and the case $b(\xi, \cdot)$ has polynomial growth (that is $\lambda \in (1, 3]$). In both cases the following fundamental lemma, whose proof is postponed to Section 7, holds.

Lemma 6.2. *Assume Hypotheses 1 and 2 holds and fix any $T > 0$.*

1. *If $\lambda = 1$, then for every $0 < \beta \leq 1$ and $\gamma \in L^2(\Omega; \mathcal{H}_\beta)$ it holds*

$$\sup_{\mu \in (0, 1]} \mathbb{E} \sup_{t \in [0, T]} |u_\mu^\gamma(t)|_{H^\beta}^2 \leq c_T (1 + \mathbb{E} |\gamma|_{H^\beta}^2). \tag{6.2}$$

2. *If $\lambda \in (1, 3]$, then for any $\gamma = (\theta, \eta) \in L^2(\Omega; \mathcal{H}_1)$, such that $\theta \in L^{\lambda+1}(\Omega; L^{\lambda+1}(D))$, and for any $R > 0$, it holds*

$$\sup_{\mu \in (0, 1]} \mathbb{P} \left(\sup_{t \in [0, T]} |u_\mu^\gamma(t)|_{H^1} \geq R \right) \leq c_T(R) (1 + \mathbb{E} |\gamma|_{\mathcal{H}_1}^2 + \mathbb{E} |\theta|_{L^{\lambda+1}}^{\lambda+1}), \tag{6.3}$$

for some function $c_T : [0, +\infty) \rightarrow [0, +\infty)$ such that

$$\lim_{R \rightarrow \infty} c_T(R) = 0.$$

6.1. Proof of Theorem 6.1 in the Lipschitz case

We recall that, due to (5.5), we have

$$\alpha(x, y) \leq \exp(\eta [|x|_H^2 + |y|_H^2]) |x - y|_H, \quad x, y \in H.$$

For every $\epsilon > 0$ we have

$$\begin{aligned} \exp(\eta [|x|_H^2 + |y|_H^2]) |x - y|_H &\leq c \exp(\eta [|x|_H^2 + |y|_H^2]) (|x|_H^2 + |y|_H^2 + 1) |x - y|_H^{1/2} \\ &\leq c_\epsilon \exp((\eta + \epsilon) [|x|_H^2 + |y|_H^2]) |x - y|_H^{1/2}, \end{aligned}$$

for some $c_\epsilon > 0$. This implies that

$$\begin{aligned} &\mathbb{E} \alpha(u_\mu^{\gamma_\mu}(t), u^{\theta_\mu}(t)) \\ &\leq c_\epsilon (\mathbb{E} \exp(2(\eta + \epsilon) [|u_\mu^{\gamma_\mu}(t)|_H^2 + |u^{\theta_\mu}(t)|_H^2]))^{1/2} (\mathbb{E} |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H)^{1/2}. \end{aligned}$$

Therefore, (6.1) holds if we can prove that

$$\lim_{\mu \rightarrow 0} \mathbb{E} (|u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H) = 0, \tag{6.4}$$

and, for some $\mu_0 > 0$ and $\bar{\eta} > 2\eta > 0$,

$$\sup_{\mu \leq \mu_0} \mathbb{E} \exp(\bar{\eta} [|u_\mu^{\gamma_\mu}(t)|_H^2 + |u^{\theta_\mu}(t)|_H^2]) < \infty. \tag{6.5}$$

Concerning (6.5), due to Lemma 9.2 and Lemma 9.3 (cf. (9.5) and (9.6)), there exists $\eta > 0$ such that for every $T > 0$ and $t \leq T$

$$\begin{aligned} &\mathbb{E} \exp(\eta [|u_\mu^{\gamma_\mu}(t)|_H^2 + |u^{\theta_\mu}(t)|_H^2]) \\ &\leq c_T \mathbb{E} \exp(\eta [\mu |\theta_\mu|_{H^1}^2 + |\theta_\mu|_H^2 + \mu^2 |\eta_\mu|_H^2 + \mu |\theta_\mu|_{L^{\lambda+1}}^{\lambda+1}]). \end{aligned}$$

Hence, as $\mathcal{L}(\gamma_\mu) = \nu_\mu$, due to (9.1) we get

$$\begin{aligned} &\sup_{\mu \leq 1/2} \mathbb{E} \exp(\eta [|u_\mu^{\gamma_\mu}(t)|_H^2 + |u^{\theta_\mu}(t)|_H^2]) \\ &\leq c_T \sup_{\mu \leq 1/2} \int_{\mathcal{H}_1} \exp(\eta [\mu |u|_{H^1}^2 + |u|_H^2 + \mu^2 |v_\mu|_H^2 + \mu |u|_{L^{\lambda+1}}^{\lambda+1}]) d\nu_\mu(u, v) < \infty, \end{aligned}$$

and (6.5) follows.

Now, let us prove (6.4). Referring back to (2.16) and (4.2) and using (3.3) and the Lipschitz continuity of B in H , we have

$$\begin{aligned}
 |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H &\leq |\Pi_1 S_\mu(t)\gamma_\mu - S(t)\theta_\mu|_H + c \int_0^t |u_\mu^{\gamma_\mu}(s) - u^{\theta_\mu}(s)|_H ds \\
 &+ \int_0^t \left| \frac{1}{\mu} \Pi_1 S_\mu(t-s)(0, B(u^{\theta_\mu}(s)) - S(t-s)B(u^{\theta_\mu}(s))) \right|_H ds + |\Pi_1 \Gamma_\mu(t) - \Gamma(t)|_H.
 \end{aligned}$$

Therefore, as a consequence of the Gronwall lemma, we have

$$\begin{aligned}
 \sup_{t \in [0, T]} |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H &\leq c_T \sup_{t \in [0, T]} |\Pi_1 S_\mu(t)\gamma_\mu - S(t)\theta_\mu|_H \\
 &+ c_T \sup_{t \in [0, T]} \int_0^t \left| \frac{1}{\mu} \Pi_1 S_\mu(t-s)(0, B(u^{\theta_\mu}(s)) - S(t-s)B(u^{\theta_\mu}(s))) \right|_H ds \\
 &+ c_T \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t) - \Gamma(t)|_H =: c_T \sum_{k=1}^3 I_{\mu, k}.
 \end{aligned}$$

In [10, proof of Theorem 4.1], it is shown that

$$\lim_{\mu \rightarrow 0} \mathbb{E} I_{\mu, 3} = 0.$$

Therefore, in order to get (6.4), we have to prove

$$\lim_{\mu \rightarrow 0} \mathbb{E} I_{\mu, k} = 0, \quad k = 1, 2.$$

We start with $I_{\mu, 1}$. For every fixed $R > 0$ and $\beta \in (0, 1]$, we have

$$\begin{aligned}
 \mathbb{E} I_{\mu, 1} &\leq \mathbb{E} (I_{\mu, 1}; |\theta_\mu|_{H^\beta} + \sqrt{\mu} |\eta_\mu|_H \leq R) \\
 &+ \mathbb{E} (I_{\mu, 1}; \{|\theta_\mu|_{H^\beta} \geq R/2\} \cup \{\sqrt{\mu} |\eta_\mu|_H \geq R/2\}) =: J_R^{\mu, 1} + J_R^{\mu, 2}.
 \end{aligned}$$

Due to (3.1), we have

$$\begin{aligned}
 J_R^{\mu, 2} &\leq \frac{c}{R} (\mathbb{E} |I_{\mu, 1}|^2)^{\frac{1}{2}} (\mathbb{E} |\theta_\mu|_{H^\beta}^2 + \mu \mathbb{E} |\eta_\mu|_H^2)^{\frac{1}{2}} \\
 &\leq \frac{c_T}{R} (\mathbb{E} |\theta_\mu|_H^2 + \mu \mathbb{E} |\eta_\mu|_{H^{-1}}^2)^{\frac{1}{2}} (\mathbb{E} |\theta_\mu|_{H^\beta}^2 + \mu \mathbb{E} |\eta_\mu|_H^2)^{\frac{1}{2}} \leq \frac{c_T}{R} (\mathbb{E} |\theta_\mu|_{H^1}^2 + \mu \mathbb{E} |\eta_\mu|_H^2) \\
 &= \frac{c_T}{R} \int_{\mathcal{H}_1} (|u|_{H^1}^2 + \mu |v|_H^2) d\nu_\mu(u, v).
 \end{aligned}$$

Therefore, thanks to (8.6), we have

$$\sup_{\mu \in (0, 1/2]} J_R^{\mu, 2} \leq \frac{c_T}{R}.$$

This means that if we fix any $\epsilon > 0$ we can find $R_\epsilon > 0$ such that $J_{R_\epsilon}^{\mu, 2} \leq \epsilon$, for every $\mu \in (0, 1/2]$. Once fixed R_ϵ , due to (3.6), we have

$$\limsup_{\mu \rightarrow 0} \mathbb{E} I_{\mu,1} \leq \epsilon + \lim_{\mu \rightarrow 0} J_{R_\epsilon}^{\mu,1} = \epsilon,$$

so that, due to the arbitrariness of ϵ , we conclude that

$$\lim_{\mu \rightarrow 0} \sup_{t \in [0,T]} |\Pi_1 S_\mu(t) \gamma_\mu - S(t) \theta_\mu|_H = \lim_{\mu \rightarrow 0} \mathbb{E} I_{\mu,1} = 0. \tag{6.6}$$

Concerning $I_{\mu,2}$, we have, for any $R > 0$,

$$\begin{aligned} \mathbb{E} I_{\mu,2} &= \mathbb{E} (I_{\mu,2}; |B(u^{\theta_\mu}(\cdot))|_{L^1(0,T;H^\beta)} \leq R) + \mathbb{E} (I_{\mu,2}; |B(u^{\theta_\mu}(\cdot))|_{L^1(0,T;H^\beta)} > R) \\ &=: J_R^{\mu,1} + J_R^{\mu,2}. \end{aligned}$$

As a consequence of (3.12), (2.16) and (4.6)

$$\begin{aligned} \mathbb{E} |I_{\mu,2}|^2 &\leq c_{\beta,T} \mathbb{E} |B(u^{\theta_\mu})|_{L^2(0,T;H^\beta)}^2 \\ &\leq c_{\beta,T} \left(1 + \mathbb{E} |u^{\theta_\mu}|_{L^2(0,T;H^\beta)}^2 \right) \leq c_{\beta,T} \left(1 + \mathbb{E} |\theta_\mu|_{H^\beta}^2 \right). \end{aligned}$$

Thanks to (8.6), this implies that

$$\sup_{\mu \in (0,1/2]} \mathbb{E} |I_{\mu,2}|^2 \leq c_{\beta,T} \sup_{\mu \in (0,1/2]} \mathbb{E} |B(u^{\theta_\mu})|_{L^2(0,T;H^\beta)}^2 < +\infty.$$

Therefore, for every $\epsilon > 0$ we can find $R_\epsilon > 0$ such that

$$\sup_{\mu \in (0,1/2]} J_{R_\epsilon}^{\mu,2} \leq \frac{c}{R_\epsilon} \leq \epsilon.$$

Once fixed R_ϵ , due to (3.9) we have

$$\lim_{\mu \rightarrow 0} J_{R_\epsilon}^{\mu,1} = 0,$$

so that we conclude that

$$\lim_{\mu \rightarrow 0} \mathbb{E} I_{\mu,2} = 0.$$

This, together with (6.6), allows to obtain (6.4).

6.2. Proof of Theorem 6.1 in the non Lipschitz case

In this case $\alpha(x, y) = |x - y|_H$, so that (6.1) follows if we prove that

$$\lim_{\mu \rightarrow 0} \mathbb{E} |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H = 0.$$

For every $R > 0$, we have

$$\begin{aligned} \mathbb{E} |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H &\leq \mathbb{E} (|u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H; \tau_R^\mu \leq t) + \mathbb{E} (|u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H; \tau_R^\mu > t) \\ &=: J_R^{\mu,1}(t) + J_R^{\mu,2}(t), \end{aligned} \tag{6.7}$$

where

$$\tau_R^\mu(t) := \inf \{s \geq 0 : \max (|u_\mu^{\gamma_\mu}(s)|_{L^\infty}, |u^{\theta_\mu}(s)|_{L^\infty}) \geq R\}.$$

Concerning the first term $J_R^{\mu,1}(t)$, we have

$$\begin{aligned} |J_R^{\mu,1}(t)|^2 &\leq \mathbb{E} |u_\mu^{\gamma_\mu}(t) - u^{\theta_\mu}(t)|_H^2 \mathbb{P} (\tau_R^\mu \leq t) \\ &\leq 2 [\mathbb{E} |u_\mu^{\gamma_\mu}(t)|_H^2 + \mathbb{E} |u^{\theta_\mu}(t)|_H^2] \left[\mathbb{P} \left(\sup_{s \leq t} |u_\mu^{\gamma_\mu}(s)|_{L^\infty} > R \right) + \mathbb{P} \left(\sup_{s \leq t} |u^{\theta_\mu}(s)|_{L^\infty} > R \right) \right] \end{aligned}$$

Thanks to (8.11), for every $\mu \leq 1$ we get

$$\sup_{t \in [0, T]} \mathbb{E} |u_\mu^{\gamma_\mu}(t)|_H^2 \leq c_T (1 + \mu \mathbb{E} |\theta_\mu|_{H^1}^2 + \mathbb{E} |\theta_\mu|_H^2 + \mu^2 \mathbb{E} |\eta_\mu|_H^2 + \mu \mathbb{E} |\theta_\mu|_{L^{\lambda+1}}^{\lambda+1}).$$

Moreover, due to (4.5) we have

$$\mathbb{E} \sup_{t \in [0, T]} |u^{\theta_\mu}(t)|_H^2 \leq c_T (1 + \mathbb{E} |\theta_\mu|_H^2).$$

Therefore, in view of (8.6), we can conclude that for every $R > 0$ and $\mu \in (0, 1/2]$

$$\sup_{t \in [0, T]} |J_R^{\mu,1}(t)|^2 \leq c_T \sup_{t \in [0, T]} \left[\mathbb{P} \left(\sup_{s \leq t} |u_\mu^{\gamma_\mu}(s)|_{L^\infty} > R \right) + \mathbb{P} \left(\sup_{s \leq t} |u(s)|_{L^\infty} > R \right) \right].$$

As shown in Lemma 6.2, for every $R > 0$ and $\mu \in (0, 1]$

$$\mathbb{P} \left(\sup_{s \leq t} |u_\mu^{\gamma_\mu}(s)|_{L^\infty} > R \right) \leq c_T(R) (1 + \mathbb{E} |\gamma_\mu|_{\mathcal{H}_1}^2 + \mathbb{E} |\theta_\mu|_{L^{\lambda+1}}^{\lambda+1}).$$

Analogously, as a consequence of (4.7), for every $\mu > 0$ we have

$$\mathbb{P} \left(\sup_{s \leq t} |u^{\theta_\mu}(s)|_{L^\infty} > R \right) \leq \frac{c_T}{R^2} (1 + \mathbb{E} |\theta_\mu|_{H^1}^2).$$

Therefore, since $c_T(R) \rightarrow 0$, as $R \rightarrow \infty$, thanks to (8.6), for any $\epsilon > 0$ there exists $R_\epsilon > 0$ such that

$$\sup_{\mu \in (0, 1/2]} \sup_{t \in [0, T]} J_{R_\epsilon}^{\mu,1}(t) \leq \epsilon.$$

Thanks to (6.7), this yields

$$\sup_{t \in [0, T]} \mathbb{E} |u_{\mu}^{\gamma}(t) - u^{\theta}(t)|_H \leq \epsilon + J_{R_{\epsilon}}^{\mu, 2}(t), \quad \mu \in (0, 1/2]. \tag{6.8}$$

Now, let us estimate $J_{R_{\epsilon}}^{\mu, 2}(t)$. For every $\mu > 0$, we have

$$\begin{aligned} u_{\mu}^{\gamma}(t) - u^{\theta}(t) &= [\Pi_1 S_{\mu}(t)(\theta_{\mu}, \eta_{\mu}) - S(t)\theta_{\mu}] \\ &+ \int_0^t [\Pi_1 S_{\mu}(t-s)B_{\mu}(z_{\mu}^{\gamma}(s)) - S(t-s)B(u^{\theta}(s))] ds + [\Pi_1 \Gamma_{\mu}(t) - \Gamma(t)] =: \sum_{i=1}^3 I_i^{\mu}(t), \end{aligned}$$

where $\Gamma_{\mu}(t)$ and $\Gamma(t)$ are the two stochastic convolutions, defined respectively in (4.3) and (4.4). This implies that

$$J_{R_{\epsilon}}^{\mu, 2}(t) \leq \mathbb{E} |I_1^{\mu}(t)|_H + \mathbb{E} |I_3^{\mu}(t)|_H + \mathbb{E} (|I_2^{\mu}(t)|_H ; \tau_{R_{\epsilon}}^{\mu} > t). \tag{6.9}$$

We have

$$\begin{aligned} I_2^{\mu}(t) &= \int_0^t \Pi_1 S_{\mu}(t-s) [B_{\mu}(z_{\mu}^{\gamma}(s)) - B_{\mu}(u^{\theta}(s), 0)] ds + \int_0^t \Phi_{\mu}(t-s)B(u^{\theta}(s)) ds \\ &=: I_2^{\mu, 1}(t) + I_2^{\mu, 2}(t), \end{aligned}$$

where $\Phi_{\mu}(t)$ is the operator introduced in (3.10).

According to (3.3), we have, in view of

$$|I_2^{\mu, 1}(t)|_H \leq 2 \int_0^t |B(u_{\mu}^{\gamma}(s)) - B(u^{\theta}(s))|_H ds.$$

For every $u, v \in L^{\infty}(D)$, we have, in view of (2.8), that

$$\begin{aligned} |B(u) - B(v)|_H &\leq c ||u - v| (1 + |u|^{\lambda-1} + |v|^{\lambda-1})|_H \\ &\leq c |u - v|_H (1 + |u|_{L^{\infty}}^{\lambda-1} + |v|_{L^{\infty}}^{\lambda-1}). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E} (|I_2^{\mu, 1}(t)|_H ; \tau_{R_{\epsilon}}^{\mu} > t) &\leq c (1 + 2 R_{\epsilon}^{\lambda-1}) \int_0^t \mathbb{E} (|u_{\mu}^{\gamma}(s) - u(s)|_H ; \tau_{R_{\epsilon}}^{\mu} > t) ds \\ &\leq c (1 + 2 R_{\epsilon}^{\lambda-1}) \int_0^t J_{R_{\epsilon}}^{\mu, 2}(s) ds. \end{aligned} \tag{6.10}$$

Combining this bounds with to (6.9), this yields

$$J_{R_\epsilon}^{\mu,2}(t) \leq \mathbb{E} |I_1^\mu(t)|_H + \mathbb{E} |I_3^\mu(t)|_H + \mathbb{E} \left(|I_2^{\mu,2}(t)|_H ; \tau_{R_\epsilon}^\mu > t \right) + c (1 + 2R_\epsilon^{\lambda-1}) \int_0^t J_{R_\epsilon}^{\mu,2}(s) ds,$$

and, due to the Gronwall lemma, for every $t \in [0, T]$

$$J_{R_\epsilon}^{\mu,2}(t) \leq c_{T,\epsilon} \sup_{t \in [0,T]} \left(\mathbb{E} |I_1^\mu(t)|_H + \mathbb{E} |I_3^\mu(t)|_H + \mathbb{E} \left(|I_2^{\mu,2}(t)|_H ; \tau_{R_\epsilon}^\mu > t \right) \right). \tag{6.11}$$

Next, if $\tau_{R_\epsilon}^\mu > t$, in view of (2.17) we have

$$|B(u^{\theta_\mu}(s))|_{H^1} \leq c (1 + |u^{\theta_\mu}(s)|_{L^\infty}^{\lambda-1}) |u^{\theta_\mu}(s)|_{H^1} \leq c (1 + R_\epsilon^{\lambda-1}) |u^{\theta_\mu}(s)|_{H^1}, \quad s \leq t,$$

so that, thanks to Lemma 8.2,

$$\left| B(u^{\theta_\mu}(\cdot)) \mathbb{I}_{\{\tau_{R_\epsilon}^\mu > t\}} \right|_{L^2([0,t];H^1)} \leq c, \quad \mathbb{P} - \text{a.s.}$$

Thanks to (3.9), this allows us to conclude that

$$\lim_{\mu \rightarrow 0} \mathbb{E} \sup_{t \in [0,T]} \left(|I_2^{\mu,2}(t)|_H ; \tau_{R_\epsilon}^\mu > t \right) = 0. \tag{6.12}$$

Moreover, in [10, Theorem 4.1] it is proven that

$$\lim_{\mu \rightarrow 0} \sup_{t \in [0,T]} \mathbb{E} |I_3^\mu(t)|_H = 0. \tag{6.13}$$

Therefore, collecting together (6.11), (6.12) and (6.13), in view of (6.6), we conclude

$$\lim_{\mu \rightarrow 0} J_{R_\epsilon}^{\mu,2}(t) = 0.$$

According to (6.8), due to the arbitrariness of $\epsilon > 0$, (6.1) follows.

7. Proof of Lemma 6.2

In our proof of Lemma 6.2, we distinguish the case $\lambda = 1$ and the case $\lambda \in (1, 3]$.

Case $\lambda = 1$. We have

$$u_\mu^\gamma(t) = \Pi_1 S_\mu(t) \gamma + \int_0^t \Pi_1 S_\mu(t-s) B_\mu(z_\mu^\gamma(s)) ds + \Pi_1 \Gamma_\mu(t).$$

Then, thanks to (3.3) and (3.1), for every $t \in [0, T]$ and $\mu \in (0, 1]$ we have

$$|u_\mu^\gamma(t)|_{H^\beta}^2 \leq c|\gamma|_{\mathcal{H}_\beta}^2 + c_T \int_0^t |B(u_\mu^\gamma(s))|_{H^\beta}^2 ds + c |\Pi_1 \Gamma_\mu(t)|_{H^\beta}^2.$$

Due to (2.16), this implies

$$|u_\mu^\gamma(t)|_{H^\beta}^2 \leq c_T \left(|\gamma|_{\mathcal{H}_\beta}^2 + \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^\beta}^2 \right) + c_T \int_0^t (|u_\mu^\gamma(s)|_{H^\beta}^2 + 1) ds,$$

so that, thanks to the Gronwall lemma and (3.5) we can conclude that (6.2) holds.

Case $\lambda \in (1, 3]$. If we denote

$$v_\mu^\gamma(t) := u_\mu^\gamma(t) - \Pi_1 \Gamma_\mu(t),$$

we have that v_μ^γ solves the equation

$$\begin{cases} \mu \partial_t^2 v_\mu^\gamma(t, \xi) = \Delta v_\mu^\gamma(t, \xi) - \partial_t v_\mu^\gamma(t, \xi) + b(\xi, u_\mu(t, \xi)), \\ v_\mu^\gamma(0, \xi) = \theta(\xi), \quad \partial_t v_\mu^\gamma(0, \xi) = \eta(\xi), \quad \xi \in D, \quad v_\mu^\gamma(t, \xi) = 0, \\ t \geq 0, \quad \xi \in \partial D. \end{cases} \tag{7.1}$$

This implies

$$\begin{aligned} & \mu \frac{d}{dt} |\partial_t v_\mu^\gamma(t)|_H^2 + \frac{d}{dt} |v_\mu^\gamma(t)|_{H^1}^2 + 2 |\partial_t v_\mu^\gamma(t)|_H^2 \\ &= 2 \langle b(\cdot, v_\mu^\gamma(t)), \partial_t v_\mu^\gamma(t) \rangle_H + 2 \langle b(\cdot, v_\mu^\gamma(t) + \Pi_1 \Gamma_\mu(t)) - b(\cdot, v_\mu^\gamma(t)), \partial_t v_\mu^\gamma(t) \rangle_H \\ &\leq 2 \frac{d}{dt} \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) d\xi + |b(\cdot, v_\mu^\gamma(t) + \Pi_1 \Gamma_\mu(t)) - b(\cdot, v_\mu^\gamma(t))|_H^2 + |\partial_t v_\mu^\gamma(t)|_H^2, \end{aligned}$$

where \mathfrak{b} is the antiderivative of b that satisfies (2.19). Thus

$$\begin{aligned} & \mu \frac{d}{dt} |\partial_t v_\mu^\gamma(t)|_H^2 + \frac{d}{dt} |v_\mu^\gamma(t)|_{H^1}^2 + |\partial_t v_\mu^\gamma(t)|_H^2 - 2 \frac{d}{dt} \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) d\xi \\ &\leq |b(\cdot, v_\mu^\gamma(t) + \Pi_1 \Gamma_\mu(t)) - b(\cdot, v_\mu^\gamma(t))|_H^2. \end{aligned} \tag{7.2}$$

Now, for every $x \in L^{2(\lambda-1)}(D)$ and $h \in L^\infty(D)$, we have

$$\begin{aligned} |b(x+h) - b(x)|_H &\leq c \left(|h| (1 + |x|^{\lambda-1} + |h|^{\lambda-1}) \right)_H \\ &\leq c |h|_{L^\infty} \left(1 + |x|_{L^{2(\lambda-1)}}^{\lambda-1} + |h|_{L^{2(\lambda-1)}}^{\lambda-1} \right). \end{aligned}$$

Hence, as $H^1 \hookrightarrow L^\infty(D)$ (we are assuming $d = 1$ here), we have

$$\begin{aligned} &|b(\cdot, v_\mu^\gamma(t) + \Pi_1 \Gamma_\mu(t)) - b(\cdot, v_\mu^\gamma(t))|_H^2 \\ &\leq c \left(1 + |\Pi_1 \Gamma_\mu(t)|_{H^1}^{2\lambda}\right) + c |\Pi_1 \Gamma_\mu(t)|_{H^1}^2 |v_\mu^\gamma(t)|_{L^{2(\lambda-1)}}^{2(\lambda-1)}. \end{aligned}$$

Since $\lambda \leq 3$, we have that $2(\lambda - 1) \leq \lambda + 1$, so that, thanks to (2.19) we get

$$\begin{aligned} &|b(\cdot, v_\mu^\gamma(t) + \Pi_1 \Gamma_\mu(t)) - b(\cdot, v_\mu^\gamma(t))|_H^2 \\ &\leq c \left(1 + |\Pi_1 \Gamma_\mu(t)|_{H^1}^{2\lambda}\right) - c |\Pi_1 \Gamma_\mu(t)|_{H^1}^2 \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) \, d\xi. \end{aligned}$$

Therefore, if we replace the inequality above in (7.2), we obtain

$$\begin{aligned} &\mu \frac{d}{dt} |\partial_t v_\mu^\gamma(t)|_H^2 + \frac{d}{dt} |v_\mu^\gamma(t)|_{H^1}^2 + |\partial_t v_\mu^\gamma(t)|_H^2 - 2 \frac{d}{dt} \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) \, d\xi \\ &\leq c \left(1 + |\Pi_1 \Gamma_\mu(t)|_{H^1}^{2\lambda}\right) - c |\Pi_1 \Gamma_\mu(t)|_{H^1}^2 \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) \, d\xi. \end{aligned}$$

Now, if we integrate with respect to $t \in [0, T]$, we obtain

$$\begin{aligned} &\mu |\partial_t v_\mu^\gamma(t)|_H^2 + |v_\mu^\gamma(t)|_{H^1}^2 + \int_0^t |\partial_t v_\mu^\gamma(s)|_H^2 \, ds - 2 \int_D \mathfrak{b}(\xi, v_\mu^\gamma(t, \xi)) \, d\xi \\ &\leq \mu |\eta|_H^2 + |\theta|_{H^1}^2 - 2 \int_D \mathfrak{b}(\xi, \theta(\xi)) \, d\xi + c \left(1 + \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1}^{2\lambda}\right) \\ &\quad - c \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1}^2 \int_0^t \int_D \mathfrak{b}(\xi, v_\mu^\gamma(s, \xi)) \, d\xi \, ds. \end{aligned}$$

Then, as a consequence of the Gronwall lemma and of (2.19) we get

$$\begin{aligned} &\sup_{t \in [0, T]} \left(\mu |\partial_t v_\mu^\gamma(t)|_H^2 + |v_\mu^\gamma(t)|_{H^1}^2 + |v_\mu^\gamma(t)|_{L^{\lambda+1}}^{\lambda+1} \right) + \int_0^T |\partial_t v_\mu^\gamma(s)|_H^2 \, ds \\ &\leq c_T \left(\mu |\eta|_H^2 + |\theta|_{H^1}^2 + |\theta|_{L^{\lambda+1}}^{\lambda+1} + c \left(1 + \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1}^{2\lambda}\right) \right) \\ &\quad \times \exp \left(T \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1}^2 \right). \end{aligned}$$

This implies that, if we define

$$\Lambda_\mu(T) := \sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1}^2,$$

for every $\mu \in (0, 1]$ we have

$$\begin{aligned} & \mathbb{P} \left(\sup_{t \in [0, T]} |v_\mu^\gamma(t)|_{H^1}^2 \geq \frac{R^2}{4} \right) \leq \mathbb{P} \left(\exp(T\Lambda_\mu(T)) \geq \frac{R}{2} \right) \\ & \quad + \mathbb{P} \left(\mu |\eta|_H^2 + |\theta|_{H^1}^2 + |\theta|_{L^{\lambda+1}}^{\lambda+1} \geq \frac{R}{4c_T} \right) + \mathbb{P} \left(\Lambda_\mu(T)^\lambda \geq \frac{R}{4cc_T} - 1 \right) \\ & \leq \mathbb{P} \left(\Lambda_\mu(T) \geq \frac{1}{T} \log \left(\frac{R}{2} \right) \right) + \mathbb{P} \left(\Lambda_\mu(T) \geq \left(\frac{R}{4cc_T} - 1 \right)^{\frac{1}{\lambda}} \right) \\ & \quad + \mathbb{P} \left(|\gamma|_{\mathcal{H}^1}^2 + |\theta|_{L^{\lambda+1}}^{\lambda+1} \geq \frac{R}{4c_T} \right). \end{aligned}$$

Due to (3.5), this implies that

$$\begin{aligned} & \sup_{\mu \in (0, 1]} \mathbb{P} \left(\sup_{t \in [0, T]} |v_\mu^\gamma(t)|_{H^1}^2 \geq \frac{R^2}{4} \right) \\ & \leq c_T \left(T \left(\log \left(\frac{R}{2} \right) \right)^{-1} + \left(\frac{R}{4cc_T} - 1 \right)^{-\frac{1}{\lambda}} \right) + \frac{4c_T}{R} \mathbb{E} (|\gamma|_{\mathcal{H}^1}^2 + |\theta|_{L^{\lambda+1}}^{\lambda+1}). \end{aligned}$$

Now, since

$$\begin{aligned} & \mathbb{P} \left(\sup_{t \in [0, T]} |u_\mu^\gamma(t)|_{H^\beta} \geq R \right) \\ & \leq \mathbb{P} \left(\sup_{t \in [0, T]} |\Pi_1 \Gamma_\mu(t)|_{H^1} \geq \frac{R}{2} \right) + \mathbb{P} \left(\sup_{t \in [0, T]} |v_\mu^\gamma(t)|_{H^1}^2 \geq \frac{R^2}{4} \right), \end{aligned}$$

by using again (3.5), the inequality above gives (6.3), for any $\lambda \in (1, 3]$.

8. Uniform bounds for the moments of equation (4.1)

For every $\mu > 0$, we denote by N_μ the Kolmogorov operator associated with equation (4.1). If $\varphi : \mathcal{H}_1 \rightarrow \mathbb{R}$ is a twice continuously differentiable mapping, with

$$\text{Tr} [D^2\varphi(z)Q_\mu Q_\mu^*] < \infty, \quad z \in \mathcal{H}_1,$$

then we have

$$N_\mu\varphi(z) = \frac{1}{2} \text{Tr} [D^2\varphi(z)Q_\mu Q_\mu^*] + \langle A_\mu z + B_\mu(z), D\varphi(z) \rangle_{\mathcal{H}}, \quad z \in D(A_\mu),$$

where Q_μ , A_μ and B_μ are defined respectively in (2.6), (2.2) and (2.18).

In Section 4, we have seen that for every $\mu > 0$, equation (4.1) admits an invariant measure ν_μ on the space \mathcal{H}_1 . In particular, due to invariance, we have

$$\int_{\mathcal{H}_1} N_\mu \varphi(z) d\nu_\mu(z) = 0. \tag{8.1}$$

Now, for every $\vartheta, \mu > 0$ we define

$$K_{\vartheta,\mu}(u, v) := \mu |u|_{H^1}^2 + (\vartheta \mu + 1/2) |u|_H^2 + \mu^2 |\vartheta u + v|_H^2 + \mu \langle u, v \rangle_H - 2 \mu \int_D \mathfrak{b}(\xi, u(\xi)) d\xi, \tag{8.2}$$

where \mathfrak{b} is the antiderivative of b that satisfies (2.19) and (2.20), in case $\lambda \in (1, 3]$ and $\lambda = 1$, respectively. We have

$$\begin{aligned} D_u K_{\vartheta,\mu}(u, v) \cdot h &= 2\mu \langle u, h \rangle_{H^1} + (2\vartheta^2 \mu^2 + 2\vartheta \mu + 1) \langle u, h \rangle_H \\ &\quad + (1 + \mu(\vartheta + 1)) \langle \mu v, h \rangle_H - 2\mu \langle \mathfrak{b}(u), h \rangle_H, \\ D_v K_{\vartheta,\mu}(u, v) \cdot h &= \mu(1 + 2\vartheta \mu) \langle u, h \rangle_H + 2\mu \langle \mu v, h \rangle_H, \\ D_v^2 K_{\vartheta,\mu}(u, v) &= 2\mu^2 I. \end{aligned} \tag{8.3}$$

Lemma 8.1. For every $\bar{\mu}, \bar{\vartheta} > 0$ there exist some $c_1, c_2 > 0$ such that for every $(u, v) \in \mathcal{H}_1$

$$c_1 K_\mu(u, v) \leq K_{\vartheta,\mu}(u, v) \leq c_2 (K_\mu(u, v) + 1), \quad \vartheta \leq \bar{\vartheta}, \quad \mu \leq \bar{\mu} \tag{8.4}$$

where

$$K_\mu(u, v) = \mu |u|_{H^1}^2 + \mu |u|_{L^{\lambda+1}}^{\lambda+1} + |u|_H^2 + \mu^2 |v|_H^2. \tag{8.5}$$

Proof. We have

$$K_{\vartheta,\mu}(u, v) = \mu |u|_{H^1}^2 + (1/2 + \vartheta \mu + \vartheta^2 \mu^2) |u|_H^2 + \mu^2 |v|_H^2 + (1 + 2\mu\vartheta) \langle u, \mu v \rangle_H - 2 \mu \int_D \mathfrak{b}(\xi, u(\xi)) d\xi.$$

Due to (2.19) and (2.20), since both μ and ϑ remain bounded, it is immediate to see that there exists some $c_2 > 0$ such that

$$K_{\vartheta_1,\mu}(u, v) \leq c_2 (K_\mu(u, v) + 1), \quad \vartheta \leq \bar{\vartheta}, \quad \mu \leq \bar{\mu}.$$

On the other hand, since

$$|\langle u, \mu v \rangle_H| \leq \frac{3}{8} |u|_H^2 + \frac{2}{3} \mu^2 |v|_H^2,$$

by using again (2.19) and (2.20), we get

$$K_{\vartheta,\mu}(u, v) \geq \mu |u|_{H^1}^2 + (1/8 + \vartheta \mu/4 + \vartheta^2 \mu^2) |u|_H^2 + \mu^2/3 |v|_H^2 + c \mu |u|_{L^{\lambda+1}}^{\lambda+1}.$$

This clearly implies that there exists some $c_1 > 0$ such that the lower bound in (8.4) is satisfied. \square

We first prove a uniform bound for some moments of the invariant measure ν_μ .

Lemma 8.2. *Under Hypotheses 1 and 2, we have*

$$\sup_{\mu \in (0,1/2]} \int_{\mathcal{H}_1} [|u|_{H^1}^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1}] d\nu_\mu(u, v) < \infty. \tag{8.6}$$

Proof. If $K_{\vartheta,\mu}$ is the function introduced in (8.2), due to (8.3) we have

$$\begin{aligned} N_\mu K_{\vartheta,\mu}(u, v) &= 2\mu \langle u, v \rangle_{H^1} + (2\vartheta^2 \mu^2 + 2\vartheta\mu + 1) \langle u, v \rangle_H + (1 + \mu(\vartheta + 1)) \langle \mu v, v \rangle_H \\ &\quad - 2\mu \langle b(u), v \rangle_H + (1 + 2\vartheta\mu) \langle u, Au - v + b(u) \rangle_H + 2 \langle \mu v, Au - v + b(u) \rangle_H + \text{Tr } Q^2 \\ &= -(1 + 2\vartheta\mu) |u|_{H^1}^2 - (1 - \mu(\vartheta + 1)) \mu |v|_H^2 + 2\vartheta^2 \mu \langle \mu v, u \rangle_H \\ &\quad + (1 + 2\vartheta\mu) \langle b(u), u \rangle_H + \text{Tr } Q^2. \end{aligned}$$

This implies that

$$N_\mu K_{\vartheta,\mu}(u, v) = -\vartheta [|u|_{H^1}^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1}] + R_{\vartheta,\mu}(u, v), \tag{8.7}$$

where

$$\begin{aligned} R_{\vartheta,\mu}(u, v) &= -(1 + 2\vartheta\mu - \vartheta) |u|_{H^1}^2 - (1 - \mu(\vartheta + 1) - \vartheta) \mu |v|_H^2 \\ &\quad + (1 + 2\vartheta\mu) \langle b(u), u \rangle_H + \vartheta |u|_{L^{\lambda+1}}^{\lambda+1} + 2\vartheta^2 \mu \langle \mu v, u \rangle_H + \text{Tr } Q^2. \end{aligned}$$

Now, in the estimate of $R_{\vartheta,\mu}(u, v)$, we distinguish the case $\lambda = 1$ and $\lambda \in (1, 3]$.

If $\lambda = 1$, due to (2.12) for every $\epsilon > 0$ we can fix $c_\epsilon > 0$ such that

$$\langle b(u), u \rangle_H \leq (L_b + \epsilon) |u|_H^2 + c_\epsilon \leq (L_b + \epsilon) \alpha_1^{-1} |u|_{H^1}^2 + c_\epsilon.$$

Therefore, if we pick $\bar{\epsilon} > 0$ such that $(L_b + \bar{\epsilon}) \alpha_1^{-1} < 1$, we get

$$\begin{aligned} R_{\vartheta,\mu}(u, v) &\leq -((1 - (L_b + \bar{\epsilon}) \alpha_1^{-1})(1 + 2\vartheta\mu) - \vartheta(\alpha_1^{-1} + 1 + \vartheta \mu^2 \alpha_1^{-1})) |u|_{H^1}^2 \\ &\quad - (1 - \mu(\vartheta + 1) - \vartheta(1 + \vartheta\mu)) \mu |v|_H^2 + \text{Tr } Q^2 + c(1 + 2\vartheta\mu). \end{aligned}$$

In particular, there exists $\bar{\vartheta} > 0$ such that

$$\sup_{\mu \in (0,1/2]} R_{\vartheta,\mu}(u, v) \leq c, \quad \vartheta \leq \bar{\vartheta}, \quad (u, v) \in \mathcal{H}_1. \tag{8.8}$$

On the other hand, if $\lambda \in (1, 3]$, due to (2.11), it is possible to prove that

$$R_{\vartheta, \mu}(u, v) \leq -(1 + 2\vartheta\mu - \vartheta(1 + \vartheta\mu^2\alpha_1^{-1}))|u|_{H^1}^2 - (1 - \mu(\vartheta + 1) - \vartheta(1 + \vartheta\mu))\mu|v|_H^2 - (c(1 + 2\vartheta\mu) - \vartheta)|u|_{L^{\lambda+1}}^{\lambda+1} + \text{Tr } Q^2 + c(1 + 2\vartheta\mu).$$

Hence, also in this case, we can find $\bar{\vartheta} > 0$ such that (8.8) holds.

Thanks to (8.7), (8.8) yields

$$N_\mu K_{\vartheta, \mu}(u, v) \leq -\vartheta[|u|_{H^1}^2 + \mu|v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1}] + c, \quad \vartheta \leq \bar{\vartheta}, \quad \mu \leq 1/2, \tag{8.9}$$

so that, from (8.1) applied to the function $K_{\bar{\vartheta}, \mu}$, we obtain

$$\int_{\mathcal{H}_1} [|u|_{H^1}^2 + \mu|v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1}] d\nu_\mu(u, v) \leq \frac{c}{\bar{\vartheta}}, \tag{8.10}$$

for every $\mu \in (0, 1/2]$. \square

Lemma 8.3. *Under Hypotheses 1 and 2, there exists $c > 0$ such that for every $\mu \in (0, 1/2)$ and $\gamma = (\theta, \eta) \in L^2(\Omega; \mathcal{H}_1)$, such that $\theta \in L^{\lambda+1}(\Omega; L^{\lambda+1}(D))$, we have*

$$\begin{aligned} &\mu \mathbb{E} |u_\mu^\gamma(t)|_{H^1}^2 + \mathbb{E} |u_\mu^\gamma(t)|_H^2 + \mu^2 \mathbb{E} |\partial_t u_\mu^\gamma(t)|_H^2 + \mu \mathbb{E} |u_\mu^\gamma(t)|_{L^{\lambda+1}}^{\lambda+1} \\ &\leq c (\mu \mathbb{E} |\theta|_{H^1}^2 + \mathbb{E} |\theta|_H^2 + \mu^2 \mathbb{E} |\eta|_H^2 + \mu \mathbb{E} |\theta|_{L^{\lambda+1}}^{\lambda+1} + t + 1). \end{aligned} \tag{8.11}$$

Proof. If we apply Itô’s formula to $(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t))$ and to the function $K_{\vartheta, \mu}$ introduced in (8.2), due to (8.3) we have

$$\begin{aligned} &dK_{\vartheta, \mu}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \\ &= N_\mu K_{\vartheta, \mu}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) dt + \langle (1 + 2\mu\theta) u_\mu^\gamma(t) + 2\mu \partial_t u_\mu^\gamma(t), Qdw(t) \rangle. \end{aligned}$$

Due to (8.9), this implies

$$\begin{aligned} &dK_{\vartheta, \mu}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \leq c dt + \langle (1 + 2\mu\theta) u_\mu^\gamma(t) + 2\mu \partial_t u_\mu^\gamma(t), Qdw(t) \rangle, \\ &\vartheta \leq \bar{\vartheta}, \quad \mu \leq 1/2. \end{aligned}$$

Integrating with respect to t and then taking expectation, we get

$$\mathbb{E} K_{\vartheta, \mu}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \leq \mathbb{E} K_{\vartheta, \mu}(\theta, \eta) + ct, \quad \vartheta \leq \bar{\vartheta}, \quad \mu \leq 1/2.$$

Therefore, (8.11) follows from estimate (8.4). \square

9. Uniform bounds for the exponential bounds of equation (4.1)

We prove here some exponential estimates for the moments of the invariant measure ν_μ , for the solutions $(u_\mu^\gamma, \partial_t u_\mu^\gamma)$ of equation (4.1), with random initial condition γ , and for the solution u^θ of equation (1.2), with random initial condition γ .

Lemma 9.1. *Under Hypotheses 1 and 2, there exists $\eta > 0$ such that*

$$\sup_{\mu \in (0, 1/2]} \int_{\mathcal{H}_1} \exp(\eta [\mu |u|_{H^1}^2 + |u|_H^2 + \mu^2 |v|_H^2 + \mu |u|_{L^{\lambda+1}}^{\lambda+1}]) \, d\nu_\mu(u, v) < \infty. \tag{9.1}$$

Proof. For every $\vartheta, \mu, \delta > 0$, we define

$$\Phi_{\vartheta, \mu, \delta}(u, v) = \exp(\delta K_{\vartheta, \mu}(u, v)), \tag{9.2}$$

where $K_{\vartheta, \mu}$ is the function introduced in (8.2). Clearly, we have

$$N_\mu \Phi_{\vartheta, \mu, \delta}(u, v) = \delta \Phi_{\vartheta, \mu, \delta}(u, v) N_\mu K_{\vartheta, \mu}(u, v) + \frac{\delta^2}{2\mu^2} \Phi_{\vartheta, \mu, \delta}(u, v) |Q(D_v K_{\vartheta, \mu}(u, v))|_H^2.$$

Due to (8.3), we have

$$|Q(D_v K_{\vartheta, \mu}(u, v))|_H^2 = \mu^2(1 + 2\vartheta\mu)^2 |Qu|_H^2 + 4\mu^4 |Qv|_H + 4\mu^3(1 + 2\vartheta\mu) \langle Qu, Qv \rangle_H.$$

Therefore, thanks to (8.9), for every $\mu \leq 1/2$ we have

$$\begin{aligned} N_\mu \Phi_{\bar{\vartheta}, \mu, \delta}(u, v) &\leq -\bar{\vartheta} \delta \Phi_{\bar{\vartheta}, \mu, \delta}(u, v) \left[|u|_{H^1}^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1} - \frac{c}{\bar{\vartheta}} \right] \\ &+ 2\delta^2 \Phi_{\bar{\vartheta}, \mu, \delta}(u, v) \left[(1/2 + \bar{\vartheta}\mu)^2 |Qu|_H^2 + \mu^2 |Qv|_H + \mu(1 + 2\bar{\vartheta}\mu) \langle Qu, Qv \rangle_H \right] \\ &\leq -\bar{\vartheta} \delta \Phi_{\bar{\vartheta}, \mu, \delta}(u, v) \left[\frac{1}{2} |u|_{H^1}^2 + \left(\frac{\alpha_1}{2} - \frac{4\delta \|Q\|^2}{\bar{\vartheta}} (1/2 + \bar{\vartheta}\mu)^2 \right) |u|_H^2 \right. \\ &\left. + \left(1 - \frac{3\delta\mu \|Q\|^2}{\bar{\vartheta}} \right) \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1} - \frac{c}{\bar{\vartheta}} \right]. \end{aligned}$$

In particular, it is immediate to check that there exist $\bar{\delta} > 0$ and $\lambda_1, \lambda_2 > 0$ such that

$$N_\mu \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \leq -\lambda_1 \bar{\vartheta} \bar{\delta} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[|u|_{H^1}^2 + |u|_H^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1} - \frac{\lambda_2}{\bar{\vartheta}} \right]. \tag{9.3}$$

Now, if we integrate both sides above with respect to the invariant measure ν_μ , according to (8.1) we obtain

$$\int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[|u|_{H^1}^2 + |u|_H^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1} - \frac{\lambda_2}{\bar{\vartheta}} \right] \, d\nu_\mu \leq 0, \quad \mu \leq 1/2. \tag{9.4}$$

Thanks to (8.4), this yields

$$\begin{aligned} & \frac{1}{c_2} \int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[K_{\bar{\vartheta}, \mu}(u, v) - \frac{c_2(\bar{\vartheta} + \lambda_2)}{\bar{\vartheta}} \right] d\nu_\mu \leq \int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[K_\mu(u, v) - \frac{\lambda_2}{\bar{\vartheta}} \right] d\nu_\mu \\ & \leq \int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[|u|_{H^1}^2 + |u|_H^2 + \mu |v|_H^2 + |u|_{L^{\lambda+1}}^{\lambda+1} - \frac{\lambda_2}{\bar{\vartheta}} \right] d\nu_\mu \leq 0, \quad \mu \leq 1/2. \end{aligned}$$

It is easy to prove that there exists some $\lambda_3 > 0$ such that

$$e^{\bar{\delta}r} \left(r - \frac{c_2(\bar{\vartheta} + \lambda_2)}{\bar{\vartheta}} \right) \geq e^{\bar{\delta}r} - \lambda_3, \quad r \geq 0,$$

so that for every $\mu \leq 1/2$

$$\int_{\mathcal{H}_1} [\Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) - \lambda_3] d\nu_\mu \leq \int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) \left[K_{\bar{\vartheta}, \mu}(u, v) - \frac{c_2(\bar{\vartheta} + \lambda_2)}{\bar{\vartheta}} \right] d\nu_\mu \leq 0.$$

By using again (8.4), this implies

$$\int_{\mathcal{H}_1} \exp(\bar{\delta}c_1 K_\mu(u, v)) d\nu_\mu \leq \int_{\mathcal{H}_1} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u, v) d\nu_\mu \leq \lambda_3, \quad \mu \leq 1/2,$$

so that (9.1) follows, with $\eta = \bar{\delta}c_1$. \square

Lemma 9.2. *Under Hypotheses 1 and 2, there exist $\eta > 0$ and $c > 0$ such that for every $\mu \in (0, 1/2)$, $T > 0$ and $\gamma = (\theta, \eta) \in L^2(\Omega; \mathcal{H}_1)$, such that $\theta \in L^{\lambda+1}(\Omega; L^{\lambda+1}(D))$, we have*

$$\begin{aligned} & \mathbb{E} \exp \left(\eta \left[\mu |u_\mu^\gamma(t)|_{H^1}^2 + |u_\mu^\gamma(t)|_H^2 + \mu^2 |\partial_t u_\mu^\gamma(t)|_H^2 + \mu |u_\mu^\gamma(t)|_{L^{\lambda+1}}^{\lambda+1} \right] \right) \\ & \leq c_T \mathbb{E} \exp \left(\eta \left[\mu |\theta|_{H^1}^2 + |\theta|_H^2 + \mu^2 |\eta|_H^2 + \mu |\theta|_{L^{\lambda+1}}^{\lambda+1} \right] \right), \quad t \leq T. \end{aligned} \tag{9.5}$$

Proof. If we apply Itô’s formula to $(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t))$ and to the function $\Phi_{\vartheta, \mu, \delta}$ introduced in (9.2), due to (8.3) we have

$$\begin{aligned} & d\Phi_{\vartheta, \mu, \delta}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) = N_\mu \Phi_{\vartheta, \mu, \delta}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) dt \\ & + \delta \Phi_{\vartheta, \mu, \delta}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \langle (1 + 2\mu\vartheta) u_\mu^\gamma(t) + 2\mu \partial_t u_\mu^\gamma(t), Qdw(t) \rangle. \end{aligned}$$

Due to (9.3), this implies that there exists $\bar{\vartheta}, \bar{\delta} > 0$ and $\lambda_1, \lambda_2 > 0$ such that

$$\begin{aligned}
 & d\Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \\
 & \leq -\lambda_1 \bar{\vartheta} \bar{\delta} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \\
 & \times \left[|u_\mu^\gamma(t)|_{H^1}^2 + |u_\mu^\gamma(t)|_H^2 + \mu |\partial_t u_\mu^\gamma(t)|_H^2 + |u_\mu^\gamma(t)|_{L^{\lambda+1}}^{\lambda+1} - \frac{\lambda_2}{\bar{\vartheta}} \right] \\
 & + \delta \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \langle (1 + 2\mu\vartheta) u_\mu^\gamma(t) + 2\mu \partial_t u_\mu^\gamma(t), Qdw(t) \rangle, \quad \mu \leq 1/2.
 \end{aligned}$$

Therefore, by integrating first with respect to time and then by taking expectation in both sides, we get

$$\mathbb{E} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \leq \mathbb{E} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(\vartheta, \eta) + \frac{\lambda_1 \lambda_2 \bar{\vartheta} \bar{\delta}}{\bar{\vartheta}} \int_0^t \mathbb{E} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(s), \partial_t u_\mu^\gamma(s)) ds,$$

and this implies

$$\mathbb{E} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(u_\mu^\gamma(t), \partial_t u_\mu^\gamma(t)) \leq c_T \mathbb{E} \Phi_{\bar{\vartheta}, \mu, \bar{\delta}}(\vartheta, \eta), \quad t \leq T.$$

Thanks to (8.5), this implies (9.5) for some $\eta > 0$. \square

By using arguments analogous but considerably simpler than those used in the proof of Lemma 9.2, we can prove that the following result holds.

Lemma 9.3. *Under Hypotheses 1 and 2, there exist $\eta > 0$ and $c > 0$ such that for every $T > 0$ and $\theta \in L^2(\Omega; \mathcal{H}_1)$, we have*

$$\mathbb{E} \exp(\eta |u^\theta(t)|_H^2) \leq c_T \mathbb{E} \exp(\eta |\theta|_H^2), \quad t \leq T. \tag{9.6}$$

Acknowledgments

This work was initiated while the two authors were visiting scholars at the Mathematical Science Research Institute (MSRI) in Berkeley, in the Fall semester 2015. They both want to thank David Herzog and Jonathan Mattingly for some interesting discussions.

Appendix A. Asymptotic strong Feller property

We want to give a proof of (4.10), namely for every $x \in H$ and every $\varphi \in C_b^1(H)$

$$|DP_t \varphi(x)|_H \leq C \sqrt{P_t |\varphi|^2(x)} + \alpha(t) \sqrt{P_t |D\varphi|_H^2(x)}, \tag{A.1}$$

for some function $\alpha(t)$ such that $\alpha(t) \rightarrow 0$, as $t \rightarrow \infty$. This bound is a time-asymptotic smoothing estimate for P_t which implies the so-called asymptotic strong Feller condition introduced in [18].

We now demonstrate how condition (2.5) on Q combined with condition (2.10) on b imply (A.1). We denote by $D_x u^x(t)h$ the derivative of $u^x(t)$ with respect to the initial condition x , along the direction $h \in H$. Moreover, we denote by $A^x(t)v$ the Malliavin derivative of $u^x(t)$ along the admissible perturbation v of the Wiener path. For any $\varphi \in C_b^1(H)$ we have

$$\begin{aligned} \langle DP_t \varphi(x), h \rangle_H &= \mathbb{E} \langle D\varphi(u^x(t)), D_x u^x(t)h \rangle_H \\ &= \mathbb{E} \langle D\varphi(u^x(t)), D_x u^x(t)h - A^x(t)v \rangle_H + \mathbb{E} \langle D\varphi(u^x(t)), A^x(t)v \rangle_H \\ &= \mathbb{E} \langle D\varphi(u^x(t)), D_x u^x(t)h - A^x(t)v \rangle_H + \mathbb{E} \varphi(u^x(t)) \int_0^t \langle v(s), dw^Q(s) \rangle_H. \end{aligned} \tag{A.2}$$

The last line follows from the Malliavin integration by parts formula and the stochastic integral is interpreted in the Skorohod sense if v is not adapted.

Now, we define $\rho(t) = \rho^{h,v}(t) := D_x u^x(t)h - A^x(t)v$ and we choose

$$v(t) = \alpha_{\bar{n}} Q^{-1} P_{\bar{n}} \rho(t), \tag{A.3}$$

where \bar{n} is the integer introduced in (2.5). Note that Q is invertible on $H_{\bar{n}} = \text{span}\{e_1, \dots, e_{\bar{n}}\}$ according to the assumption (2.5) and that this choice of v is adapted. It is immediate to check that $\rho(t)$ satisfies

$$\partial_t \rho(t) = \Delta \rho(t) + DB(u^x(t))\rho(t) - \alpha_{\bar{n}} P_{\bar{n}} \rho(t), \quad \rho(0) = h.$$

Hence, $\rho(t)$ satisfies the following estimate

$$\frac{1}{2} \frac{d}{dt} |\rho(t)|_H^2 + |\nabla \rho(t)|_H^2 + \alpha_{\bar{n}} |P_{\bar{n}} \rho(t)|_H^2 = \langle DB(u^x(t)), \rho(t) \rangle_H \leq L_b |\rho(t)|_H^2.$$

Observe that

$$\begin{aligned} |\nabla \rho(t)|_H^2 + \alpha_{\bar{n}} |P_{\bar{n}} \rho(t)|_H^2 &\geq |\nabla(I - P_{\bar{n}})\rho(t)|_H^2 + \alpha_{\bar{n}} |P_{\bar{n}} \rho(t)|_H^2 \\ &= \sum_{k=\bar{n}+1}^{\infty} \alpha_k |\langle \rho(t), e_k \rangle_H|^2 + \sum_{k=1}^{\bar{n}} \alpha_{\bar{n}} |\langle \rho(t), e_k \rangle_H|^2 \geq \alpha_{\bar{n}} |\rho(t)|_H^2, \end{aligned}$$

so that, thanks to (2.10) we obtain

$$\frac{d}{dt} |\rho(t)|_H^2 + (\alpha_{\bar{n}} - L_b) |\rho(t)|_H^2 \leq 0.$$

This implies

$$|\rho(t)|_H^2 \leq |h|_H^2 e^{-(\alpha_{\bar{n}} - L_b)t}, \tag{A.4}$$

and hence for every $t \geq 0$

$$\begin{aligned} \mathbb{E} \left| \int_0^t \langle v(s), dw^Q(s) \rangle_H \right|^2 &= \int_0^t \mathbb{E} |Qv(s)|_H^2 ds \leq \alpha_{\bar{n}}^2 |h|_H^2 \int_0^t e^{(\alpha_{\bar{n}} - L_b)s} ds \\ &\leq |h|_H^2 \frac{\alpha_{\bar{n}}^2}{\alpha_{\bar{n}} - L_b}. \end{aligned} \tag{A.5}$$

Therefore, returning now to (A.2), thanks to bounds (A.4) and (A.5) we obtain

$$\begin{aligned} |DP_t \varphi(x)|_H &\leq \sup_{|h|_H \leq 1} \mathbb{E} |D\varphi(u^x(t))|_H |\rho(t)|_H + \sup_{|h|_H \leq 1} \mathbb{E} \left| \varphi(u^x(t)) \int_0^t \langle v(s), dw^Q(s) \rangle_H \right| \\ &\leq \sqrt{P_t |D\varphi|_H^2(x)} \sup_{|h|_H \leq 1} (\mathbb{E} |\rho(t)|_H^2)^{1/2} + \sqrt{P_t |\varphi|^2(x)} \sup_{|h|_H \leq 1} \left(\mathbb{E} \left| \int_0^t \langle v(s), dw^Q(s) \rangle_H \right|^2 \right)^{1/2} \\ &\leq \frac{\alpha_{\bar{n}}}{\sqrt{\alpha_{\bar{n}} - L_b}} \sqrt{P_t |\varphi|^2(x)} + e^{-(\alpha_{\bar{n}} - L_b)t} \sqrt{P_t |D\varphi|_H^2(x)}, \end{aligned}$$

which is (A.1).

Remark A.1.

1. Observe that the asymptotic strong Feller condition holds without any conditions on Q when $L_b < \alpha_1$. In particular we obtain the contract estimate desired in our paper even when there is no noise. This is natural, in this case the $t = \infty$ dynamics contracts exponentially to zero.
2. It should be noted that the bound (A.1) would be expected to hold in a much more degenerate situation where \bar{n} does not depend on how large L_b is. Instead the condition (2.10) must be replaced with a Hormander condition, a delicate algebraic property of the interaction between the noise w^Q and the nonlinear term b . For brevity of presentation we may wish to omit such details and instead refer the reader to [18,19,13,15]. \square

References

- [1] V. Barbu, G. Da Prato, The stochastic nonlinear damped wave equation, *Appl. Math. Optim.* 46 (2002) 125–141.
- [2] J. Birrell, S. Hottovy, G. Volpe, J. Wehr, Small mass limit of a Langevin equation on a manifold, *Ann. Henri Poincaré* 18 (2017) 707–755.
- [3] S. Cerrai, Second Order PDEs in Finite and Infinite Dimensions, *Lecture Notes in Mathematics*, vol. 1762, Springer Verlag, 2001.
- [4] S. Cerrai, M. Freidlin, On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom, *Probab. Theory Relat. Fields* 135 (2006) 363–394.
- [5] S. Cerrai, M. Freidlin, Smoluchowski-Kramers approximation for a general class of SPDE’s, *J. Evol. Equ.* 6 (2006) 657–689.
- [6] S. Cerrai, M. Freidlin, Small mass asymptotics for a charged particle in a magnetic field and longtime influence of small perturbations, *J. Stat. Phys.* 144 (2011) 101–123.

- [7] S. Cerrai, M. Freidlin, M. Salins, On the Smoluchowski-Kramers approximation for SPDEs and its interplay with large deviations and long time behavior, *Discrete Contin. Dyn. Syst., Ser. A* 37 (2017) 33–76.
- [8] S. Cerrai, M. Salins, Smoluchowski-Kramers approximation and large deviations for infinite dimensional gradient systems, *Asymptot. Anal.* 88 (2013) 201–215.
- [9] S. Cerrai, M. Salins, Smoluchowski-Kramers approximation and large deviations for infinite dimensional non-gradient systems with applications to the exit problem, *Ann. Probab.* 44 (2016) 2591–2642.
- [10] S. Cerrai, M. Salins, On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom subject to a magnetic field, *Stoch. Process. Appl.* 127 (2017) 273–303.
- [11] P.L. Chow, Stochastic wave equations with polynomial nonlinearity, *Ann. Appl. Probab.* 12 (2002) 361–381.
- [12] G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.
- [13] J. Földes, N. Glatt-Holtz, G. Richards, E. Thomann, Ergodic and mixing properties of the Boussinesq equations with a degenerate random forcing, *J. Funct. Anal.* 269 (2015) 2427–2504.
- [14] J. Földes, N. Glatt-Holtz, G. Richards, Large Prandtl number asymptotics in randomly forced turbulent convection, *Nonlinear Differ. Equ. Appl. NoDEA* 26 (2019) 43, <https://link.springer.com/article/10.1007/s00030-019-0589-z>.
- [15] J. Földes, S. Friedlander, N. Glatt-Holtz, G. Richards, Asymptotic analysis for randomly forced MHD, *SIAM J. Math. Anal.* 49 (6) (2017) 4440–4469, <https://epubs.siam.org/doi/abs/10.1137/16M1071857>.
- [16] M. Freidlin, Some remarks on the Smoluchowski-Kramers approximation, *J. Stat. Phys.* 117 (2004) 617–634.
- [17] M. Freidlin, W. Hu, Smoluchowski-Kramers approximation in the case of variable friction, *J. Math. Sci.* 179 (2011) 184–207.
- [18] M. Hairer, J.C. Mattingly, Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, *Ann. Math.* 164 (2006) 993–1032.
- [19] M. Hairer, J.C. Mattingly, Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations, *Ann. Probab.* 36 (2008) 2050–2091.
- [20] D. Herzog, S. Hottovy, G. Volpe, The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction, *J. Stat. Phys.* 163 (2016) 659–673.
- [21] S. Hottovy, A. McDaniel, G. Volpe, J. Wehr, The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction, *Commun. Math. Phys.* 336 (2015) 1259–1283.
- [22] W. Hu, K. Spiliopoulos, Hypocoelliptic multiscale Langevin diffusions: large deviations, invariant measures and small mass asymptotics, *Electron. J. Probab.* 22 (2017).
- [23] H. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica* 7 (1940) 284–304.
- [24] J.J. Lee, Small mass asymptotics of a charged particle in a variable magnetic field, *Asymptot. Anal.* 86 (2014) 99–121.
- [25] Y. Lv, A. Roberts, Averaging approximation to singularly perturbed nonlinear stochastic wave equations, *J. Math. Phys.* 53 (2012) 1–11.
- [26] Y. Lv, A. Roberts, Large deviation principle for singularly perturbed stochastic damped wave equations, *Stoch. Anal. Appl.* 32 (2014) 50–60.
- [27] Y. Lv, A. Roberts, W. Wang, Approximation of the random inertial manifold of singularly perturbed stochastic wave equations, *Stoch. Dyn.* 32 (2014).
- [28] Y. Lv, W. Wang, Limiting dynamics for stochastic wave equations, *J. Differ. Equ.* 244 (2008) 1–23.
- [29] H. Nguyen, The small-mass limit and white-noise limit of an infinite dimensional generalized Langevin equation, [arXiv:1804.09682](https://arxiv.org/abs/1804.09682), 2018.
- [30] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
- [31] M. Salins, Smoluchowski-Kramers approximation for the damped stochastic wave equation with multiplicative noise in any spatial dimension, [arXiv:1801.10538](https://arxiv.org/abs/1801.10538), 2018.
- [32] W. Sickel, Superposition of functions in Sobolev spaces of fractional order. A survey, *Part. Differ. Eq. Banach Center Publ.* 27 (1992) 481–497.
- [33] M. Smoluchowski, Drei Vortage über Diffusion Brownsche Bewegung und Koagulation von Kolloidteilchen, *Phys. Z.* 17 (1916) 557–585.
- [34] K. Spiliopoulos, A note on the Smoluchowski-Kramers approximation for the Langevin equation with reflection, *Stoch. Dyn.* 7 (2007) 141–152.
- [35] C. Villani, *Optimal Transport. Old and New*, Springer-Verlag, Berlin, 2009.