IOP Publishing Nonlinearity
Nonlinearity **20** (2007) 1721–1744 doi:10.1088/0951-7715/20/7/009

Multiscale analysis for stochastic partial differential equations with quadratic nonlinearities

D Blömker[1], **M Hairer**[2] **and G A Pavliotis**[3]

[1] Institut für Mathematik, Universität Augsburg, Germany
[2] Mathematics Institute, The University of Warwick, UK
[3] Department of Mathematics, Imperial College London, UK

E-mail: dirk.bloemker@math.uni-augsburg.de, m.hairer@warwick.ac.uk and g.pavliotis@imperial.ac.uk

Received 17 November 2006, in final form 17 May 2007
Published 6 June 2007
Online at stacks.iop.org/Non/20/1721

Recommended by E Vanden Eijnden

Abstract

In this paper we derive rigorously amplitude equations for stochastic partial differential equations with quadratic nonlinearities, under the assumption that the noise acts only on the stable modes and for an appropriate scaling between the distance from bifurcation and the strength of the noise. We show that, due to the presence of two distinct timescales in our system, the noise (which acts only on the fast modes) gets transmitted to the slow modes and, as a result, the amplitude equation contains both additive and multiplicative noise.

As an application we study the case of the one-dimensional Burgers equation forced by additive noise in the orthogonal subspace to its dominant modes. The theory developed in the present paper thus allows us to explain theoretically some recent numerical observations on stabilization with additive noise.

Mathematics Subject Classification: 60H10, 60H15

1. Introduction

Stochastic partial differential equations (SPDEs) with quadratic nonlinearities arise in various applications in physics. As examples we mention the use of the stochastic Burgers equation in the study of closure models for hydrodynamic turbulence [CY95] and the use of the stochastic Kuramoto–Sivashinsky equation or similar models [CB95, LCM96, BGR02, RML+00] for the modelling of surface phenomena. Very often SPDEs have two widely separated characteristic timescales and it is desirable to obtain a simplified equation which governs the evolution of the dominant modes of the SPDE and captures the dynamics of the infinite-dimensional stochastic system at the slow timescale. The purpose of this paper is to derive rigorously such an *amplitude equation* for a quite general class of SPDEs (cf (2.1)) with quadratic nonlinearities and, furthermore, to obtain sharp error estimates, in the case where there is only one dominant mode (i.e. the amplitude equation is a one-dimensional SDE).

Consider, as a working example of the class of SPDEs that we will consider in this paper, the following variation on the Burgers equation

$$\partial_t u = \partial_x^2 u + u\partial_x u + (1+\gamma)u + \sigma\phi \tag{1.1}$$

subject to external forcing $\sigma\phi$ and to Dirichlet boundary conditions on $[0, \pi]$. Since we are working very far from the inviscid regime, the solutions to this equation in the absence of forcing would decay quickly to 0 were it not for the extra linear instability $(1+\gamma)u$. The constant 1 appearing in this term is taken to be equal to the Poincaré constant for the Dirichlet Laplacian on $[0, \pi]$ and is designed to render the first mode $\sin(x)$ linearly neutral. The constant γ therefore describes the linearized behaviour of that mode. The aim of this paper is to study the behaviour of solutions to (1.1) for small γ over (large) timescales of order γ^{-1}.

It is well known [Blö05, CH93] (see also [Sch94, Sch01] for general results on unbounded domains) that in the absence of forcing, the solution to (1.1) is of the type

$$u(t, x) = \sqrt{\gamma}a(\gamma t)\sin(x) + \mathcal{O}(\gamma), \tag{1.2}$$

where the amplitude a solves the deterministic Landau equation

$$\partial_t a = a - \tfrac{1}{12}a^3.$$

If the forcing ϕ is taken to be white in time (actually, any stochastic process with sufficiently good mixing properties would also do), then, provided that $\sigma = \mathcal{O}(\gamma)$, the solution to (1.1) is still of the type (1.2), but a now solves a stochastic version of the Landau equation:

$$\partial_t a = a - \tfrac{1}{12}a^3 + \tilde{\sigma}\xi(t)\ ,$$

where ξ is white noise in time and the constant $\tilde{\sigma}$ is proportional one the one hand to the ratio σ/γ and on the other hand to the size of the projection of ϕ onto the 'slow' subspace spanned by the mode $\sin(x)$ [Blö05]. In particular, one gets $\tilde{\sigma} = 0$ if the projection of ϕ onto that subspace vanishes.

This naturally raises the question of the behaviour of solutions to (1.1) when the external forcing acts only on the orthogonal complement of the 'slow' subspace. Roberts [Rob03] considered for example noise acting only on the second mode $\sin(2x)$. Using formal expansions relying on centre manifold type arguments, he derived a reduced model describing the amplitude of the dominant mode. Moreover he demonstrated numerically that additive noise is capable of stabilizing the dominant mode, i.e. the noise eliminates a small linear instability. In turns out that, in order to have a non-trivial effect on the limiting amplitude equation, the strength of the noise should be chosen to scale like $\sqrt{\gamma}$, i.e. $\sigma = \mathcal{O}(\sqrt{\gamma})$. We show that in this case, one has (after integrating against smooth test functions)

$$u(t, x) = \sqrt{\gamma}a(\gamma t)\sin(x) + \mathcal{O}(\gamma^{5/8}). \tag{1.3}$$

To be more precise, we have additional noise terms of order $\sqrt{\gamma}$ on higher modes that average out when integrated against test functions, i.e. they are small in some appropriate weak (averaged) sense.

The amplitude a solves a stochastic differential equation of Stratonovich type

$$\mathrm{d}a = (1+\delta_1)a\,\mathrm{d}t - \tfrac{1}{12}a^3\,\mathrm{d}t + \sqrt{\delta_2 + \delta_3 a^2} \circ \mathrm{d}B(t). \tag{1.4}$$

Here, the constants δ_i, $i = 1, 2, 3$ are proportional to σ^2/γ, σ^4/γ^2, and σ^2/γ, respectively, with proportionality constants depending on the exact nature of the noise. The Wiener process B can be constructed explicitly from the external forcing ϕ, but unless $\delta_2 = 0$, it is *not* given by a simple rescaling.

In the particular case, where $\phi(x, t) = \sin(2x)\xi(t)$ with ξ a white noise, one has $\delta_1 = -\sigma^2/88\gamma$, $\delta_2 = 0$ and $\delta_3 = \sigma^2/36\gamma$. Note that δ_1 is negative, so that if $\sigma^2 > 88\gamma$, the solution to (1.4) converges to 0 almost surely. This explains the stabilization effect observed in [Rob03].

In this paper, we justify rigorously expressions of the form (1.3) for PDEs of the form (1.1) and we obtain formulae for the coefficients in the amplitude equation (1.4). Unlike [Blö05] we are interested in the situation where the noise *does not* act on the slow degrees of freedom directly but gets transmitted to them through the nonlinear interaction with the fast degrees of freedom. From a technical point of view, one of the main novelties of this paper is that it provides explicit error bounds on the difference between the solution of the original SPDE and the solution of the approximating amplitude equation; this is a key requirement in tackling the infinite-dimensional problem. Thus, our result is stronger in that respect than weak convergence type results in the spirit of, e.g. [EK86, Kur73]. Furthermore, we provide an explicit coupling between the two solutions, which is not trivial in the sense that, unlike in the case where the noise acts on the slow variables directly, the white noise driving the resulting amplitude equation is not a simple rescaling of the noise driving the original equation.

In the case where there are more than one dominant modes and the amplitude equation is a vector valued SDE such an explicit identification of the noise which drives the limiting dynamics does not seem to be possible. As a result, we are not able to prove a strong convergence result, neither are we able to obtain error estimates. We can still, however, prove a weak convergence result. It is important to emphasise that, even in the multidimensional case, the amplitude equation is (although higher dimensional) of the form (1.4), where the cubic and quadratic term are just replaced by terms defined by trilinear or bilinear maps on some $\mathbb{R}^N$. Furthermore explicit formulae for all the coefficients that appear in the amplitude equation are obtained using our approach.

The approximation of the original SPDE by a finite-dimensional stochastic system has various interesting applications. First, we can show that additive noise can have a stabilizing effect to solutions of SPDEs. As mentioned earlier, this is proved in this paper for the case of the Burgers equation with only the second mode being forced by noise. In connection to this, detailed properties of the solution to the SPDE can be inferred through the study of the amplitude equation. Furthermore, the amplitude equation can be used to show that, for ε sufficiently small, with very high probability the solution to the SPDE exists over relatively long time intervals.

Finite-dimensional SDEs with quadratic nonlinearities and two characteristic, widely separated, timescales were analysed systematically by Majda, Timofeyev and Vanden Eijnden in a series of papers [MTVE01, MTVE99]. The SDEs that were studied by these authors can be thought of as finite-dimensional approximations of stochastic PDEs with quadratic nonlinearities of the form (1.1) (in fact, the authors consider finite-dimensional approximations of deterministic PDEs and they introduce stochastic effects by replacing the quadratic self-interaction terms of the unresolved variables by an appropriate stochastic process). In these papers, techniques from the theory of singular perturbation theory for Markov processes were used to derive stochastic amplitude equations with additive and/or multiplicative noise, which can be either stable or unstable. The results obtained by formal multiscale asymptotics can be in principle justified rigorously using the theorem of Kurtz [Kur73], see also [EK86, theorem 3.1, chapter 12]. However, since these results lack explicit error estimates, it is not clear *a priori* whether they can be applied to the infinite-dimensional situation that we study in this paper.

The rest of the paper is organized as follows. In section 2 we state the assumptions that we make, and present our main result. Sections 3 to 6 are devoted to the proof of our main theorem. In particular, in section 5 we prove the weak convergence result for the case where there is a finite number of dominant modes. On the other hand, in section 6 we prove the strong approximation theorem, together with the error bounds, for the case where the amplitude equation is one dimensional. Finally, in section 7 we apply our theory to the stochastic Burgers equation.

2. Notations, assumptions and main result

The main object of study of the present paper is the following SPDE written in the form (cf [DPZ92]):

$$\mathrm{d}u = \left(-Lu + B(u,u) + \nu\varepsilon^2 u\right)\mathrm{d}t + \varepsilon Q\,\mathrm{d}W(t). \tag{2.1}$$

Remark 2.1. We could easily allow for a deterministic forcing term εf acting on the fast modes. We omit this for simplicity of presentation.

Throughout this paper, we make the following assumptions.

Assumption 2.2. *L is a nonnegative definite self-adjoint operator with compact resolvent in some real Hilbert space $\mathcal{H}$.*

Let $\|\cdot\|$ be the norm and $\langle 1\,1\,1\rangle$ be the inner product in $\mathcal{H}$. We denote by $\{e_k\}_{k=1}^{\infty}$ and $\{\lambda_k\}_{k=1}^{\infty}$ an orthonormal basis of eigenfunctions and the corresponding (ordered) eigenvalues.

We will use the notation $\mathcal{N} := \ker L \approx \mathbb{R}^n$, and P_c for the orthogonal projection $P_c : \mathcal{H} \to \mathcal{N}$. Furthermore, $P_s := I - P_c$. Before stating our assumptions on the nonlinearity B, we introduce the following interpolation spaces. For $\alpha > 0$, we will denote by $\mathcal{H}^\alpha$ the domain of $L^{\alpha/2}$ endowed with the scalar product $\langle u, v\rangle_\alpha = \langle u, (1+L)^\alpha v\rangle$ and the corresponding norm $\|\cdot\|_\alpha$. Furthermore, we identify $\mathcal{H}^{-\alpha}$ with the dual of $\mathcal{H}^\alpha$ with respect to the inner product in $\mathcal{H}$. With this notation at hand we can state our next assumption.

Assumption 2.3. *There exists $\alpha \in (0,2)$ and $\beta \in (\alpha-2,\alpha]$ such that $B(u,v)\colon \mathcal{H}^\alpha \times \mathcal{H}^\alpha \to \mathcal{H}^\beta$ is a bounded symmetric bilinear map with*

$$P_c B(e_k, e_k) = 0\,, \tag{2.2}$$

for every $k > n$. (Recall that n is the dimension of $\ker L$.*)*

Finally, we assume that the Wiener process driving equation (2.1) satisfies:

Assumption 2.4. *W is a cylindrical Wiener process on $\mathcal{H}$. The covariance operator Q is symmetric, bounded, commutes with L, and satisfies.*

$$\langle e_k, Qe_k\rangle = 0\,, \tag{2.3}$$

for every $k \leqslant n$. Furthermore, $Q^2L^{\alpha-1}$ is trace class in $\mathcal{H}$, where the value of α is the same as in assumption 2.3.

For some of the results that we prove in this paper it will furthermore be necessary to make the following assumption.

Assumption 2.5. *The kernel of L is one-dimensional, i.e. $\lambda_1 = 0$ and $\lambda_2 > 0$.*

Remark 2.6. This assumption will not be necessary to obtain the convergence of the 'slow' dynamics to the solution of a limiting amplitude equation, see theorem 5.1. It is only necessary in order to get the strong control over the error bounds and the explicit expression for the driving noise of the limiting equation given in theorem 6.4.

We are interested in studying the behaviour of small solutions to (2.1) on timescales of order ε^{-2}. To this end, we define v through $\varepsilon v(\varepsilon^2 t) = u(t)$, so that v is the solution to

$$\mathrm{d}v = \left(-\varepsilon^{-2}Lv + \nu v + \varepsilon^{-1}B(v,v)\right)\mathrm{d}t + \varepsilon^{-1}Q\,\mathrm{d}W(t). \tag{2.4}$$

Remark 2.7. The scaling in ε in equations (2.1) and (2.4) below is dictated by the symmetry assumptions (2.2) and (2.3). If either of these assumptions were to fail, the scaling considered in this paper would not yield a meaningful limit.

Note also that we made an abuse of notation in that the Wiener process W appearing in (2.4) is actually a rescaled version of the one appearing in (2.1), but it has the same distribution. Now we are ready to state the main result of this paper.

Theorem 2.8. *Let L, B, and Q satisfy assumptions 2.2–2.5. Fix a terminal time $T > 0$, a number R, as well as constants $p > 0$ and $\kappa > 0$. Then there exists $C > 0$ such that, for every $0 < \varepsilon < 1$ and every solution v of (2.4) with initial condition $v_0 \in \mathcal{H}^\alpha$ and $\|v_0\|_\alpha \leqslant R$, there exists a stopping time τ and a Wiener process B such that*

$$\mathbb{E} \sup_{t\in[0,\tau]} \|P_c v(t) - a(t)e_1\|_\alpha^p \leqslant C\varepsilon^{p/4-\kappa} , \qquad \mathbb{P}(\tau < T) \leqslant C\varepsilon^p.$$

Here, $a(t)$ is the solution to the stochastic amplitude equation

$$\mathrm{d}a(t) = \big(\tilde{\nu}a(t) - \tilde{\eta}a(t)^3\big)\,\mathrm{d}t + \sqrt{\sigma_b + \sigma_a a(t)^2}\,\mathrm{d}B, \qquad a(0) = \langle v_0, e_1\rangle, \tag{2.5}$$

where the coefficients $\tilde{\nu}$, $\tilde{\eta}$, σ_a and σ_b are given by equations (4.7), (4.8), (6.2), respectively.

Furthermore, the fast Ornstein–Uhlenbeck process $z(t)$ solving

$$\mathrm{d}z = -\varepsilon^{-2}Lz\,\mathrm{d}t + \varepsilon^{-1}Q\,\mathrm{d}W(t), \qquad z(0) = P_s v_0 ,$$

satisfies $\mathbb{E}\sup_{t\in[0,\tau]} \|P_s v(t) - z(t)\|_\alpha^p \leqslant C\varepsilon^{p-\kappa}$.

Remark 2.9. In a weak norm in time (for example H^{-1}) one can show that $z(t)$ is well approximated by white in time and coloured in space noise of order ε. Formally, we can write

$$\begin{aligned} z(t) &= \mathrm{e}^{-tL\varepsilon^{-2}} P_s v_0 + \varepsilon^{-1}\int_0^t \mathrm{e}^{-(t-s)L\varepsilon^{-2}} Q\,\mathrm{d}W(s) \\ &\approx \varepsilon L^{-1} Q\partial_t W . \end{aligned}$$

Of course, for small transient timescales of order $\mathcal{O}(\varepsilon^2)$ the initial value $P_s v_0$ of $z(t)$ has a contribution of order $\mathcal{O}(1)$. Thus estimates of the error uniformly in time are out of reach.

Remark 2.10. An immediate corollary of our result is that, under the assumptions of theorem 2.8, we can write

$$\mathbb{E} \sup_{t\in[0,\tau\varepsilon^{-2}]} \|u(t) - \varepsilon a(\varepsilon^2 t)e_1 - \varepsilon R(t)\|_\alpha^p \leqslant C\varepsilon^{\frac{5p}{4}-\kappa},$$

where $u(t)$ is the solution to (2.1) with $u(0) = \mathcal{O}(\varepsilon)$, $a(t)$ is the solution to the amplitude equation (2.5) with $\varepsilon a(0) = \langle u(0), e_1\rangle$ and $R(t) = z(\varepsilon^2 t)$ is the solution to

$$\mathrm{d}R = -LR + Q\,\mathrm{d}W, \qquad \varepsilon R(0) = P_s u(0).$$

The noise that appears in the equation for R is a rescaled version of the noise that appears in the equation for z.

Let us discuss briefly the main steps in the proof of this result. We first decompose the solution of (2.1) into a slow and a fast part:

$$v(t) = P_c v(t) + P_s v(t) =: x(t) + y(t) , \tag{2.6}$$

to obtain a system of SDEs for (x, y), equation (3.1). Our next step is to apply Itô's formula to suitably chosen functions of x and y in order to eliminate the $\mathcal{O}(1/\varepsilon)$ terms from (3.1). We furthermore show that we can replace the fast process y by an appropriate Ornstein–Uhlenbeck process z. In this way, we obtain an SDE for x that involves only x and the (infinite-dimensional) Ornstein–Uhlenbeck process z. This is done in section 3, see proposition 3.9.

A general averaging result (with error estimates) for deterministic integrals that involve monomials of the infinite-dimensional OU process z, see corollary 4.5, enables us to eliminate or simplify various terms in the equation for (x, z) and to reduce the evolution of x to the integral equation

$$x(t) = x(0) + \tilde{\nu} \int_0^t x(s)\, \mathrm{d}s - \tilde{\eta} \int_0^t \big(x(s)\big)^3\, \mathrm{d}s + M(t) + R(t), \tag{2.7}$$

where $R(t) = \mathcal{O}(\varepsilon^{1/2-\kappa})$ (for arbitrary $\kappa > 0$) and $M(t)$ is a martingale whose quadratic variation has an explicit expression in terms of (x, z). (We made an abuse of notation here and wrote $\tilde{\eta}x^3$ for what should in general really be a trilinear form acting on x.) This is done in section 4.

The final step in the reduction procedure is to show that the martingale $M(t)$ can be approximated (pathwise) by the stochastic integral

$$\tilde{M}(t) = \int_0^t \sqrt{\sigma_b + \sigma_a a^2(s)}\, \mathrm{d}B(s)\ ,$$

where $B(t)$ is a suitable one-dimensional Brownian motion and a is the solution to the amplitude equation (2.5). This is done in sections 6.1 and 6.2. We remark that, whereas the derivation of equation (2.7) is independent of the dimensionality of $x(t)$ (provided that it is finite), the third part of the proof is valid only in the case where the kernel of L is one dimensional. This is the price we have to pay in order to obtain rigorous explicit error estimates on the validity of the amplitude equation.

However, it is relatively straightforward, using standard techniques, to show from (2.7) that the process x converges in law as $\varepsilon \to 0$ to the solution of a finite-dimensional SDE whose coefficients can be expressed explicitly. This will be done in section 5. It does not seem possible, however, to obtain pathwise convergence in this situation by using our approach, since the time change employed in the proof of lemma 6.1 works only in one-dimension. Neither does it seem straightforward to modify the present proof in such a way that one can obtain explicit error estimates without using lemma 6.1. In the case where the amplitude $a(t)$ is a Brownian motion on $\mathbb{R}^k$, one could consider one-dimensional projections, as was done in [HP04]. It is not clear however how to adapt the argument used in that paper to the case where the amplitude $a(t)$ is the solution of a general SDE. Furthermore, the error estimate obtained in [HP04] scales like ε^{c/n^2} for some appropriate small constant c, which is clearly far from being optimal.

3. The reduction to finite dimensions

Let us fix a terminal time T and constants $\kappa > 0$ and $p > 0$. Note that these constants are not necessarily the same as the ones appearing in the statement of theorem 6.2, but can get 'worse' in the course of the proof.

Note first that one has

Lemma 3.1. *Under assumptions 2.2–2.4, equation (2.4) has a unique local (mild) solution v in $\mathcal{H}^\alpha$ for every initial condition $v_0 \in \mathcal{H}^\alpha$. Furthermore, v has continuous paths in $\mathcal{H}^\alpha$.*

Proof. This follows from an application of Picard's iteration scheme for the mild solution, see for example [DPZ92]. One can check that the assumption $\beta < \alpha - 2$, together with the continuity assumption on $B(\cdot\,,\,\cdot)$, imply that the solution map has the required contraction properties for sufficiently small time. The fact that the stochastic convolution takes values in $\mathcal{H}^\alpha$ is a consequence of assumption 2.4. $\square$

Remark 3.2. Note that we do not rely on a dissipativity assumption of the underlying SPDE (2.4). Thus we can only establish the existence of local solutions. The existence of solutions on a sufficiently long timescale will be shown later to follow from the dissipativity of the approximating equations.

Substituting the decomposition (2.6) into (2.4), we obtain the following system of equations:

$$dx = \nu x\, dt + 2\varepsilon^{-1} P_c B(x, y)\, dt + \varepsilon^{-1} P_c B(y, y)\, dt \tag{3.1a}$$

$$dy = \big(\nu - \varepsilon^{-2} L\big) y\, dt + \varepsilon^{-1} P_s B(x + y, x + y)\, dt + \varepsilon^{-1} Q\, dW(t). \tag{3.1b}$$

Since lemma 3.1 does not rule out the possibility of a finite time blow up in $\mathcal{H}^\alpha$ for the quite general system (3.1), we introduce the stopping time

$$\tau^* = T \wedge \inf\{t > 0 \mid \|v(t)\|_\alpha \geqslant \varepsilon^{-\kappa}\}\ . \tag{3.2}$$

Note that lemma 3.1 ensures that, for a fixed initial condition v_0 and for ε sufficiently small, one has $\tau^* > 0$ almost surely.

Let us fix now some notation.

Definition 3.3. *For a real-valued family of processes $\{X_\varepsilon(t)\}_{t\geqslant 0}$ we say $X_\varepsilon = \mathcal{O}(f_\varepsilon)$, if for every $p \geqslant 1$ there exists a constant C_p such that*

$$\mathbb{E}\Big(\sup_{t\in[0,\tau^*)} |X_\varepsilon(t)|^p\Big) \leqslant C_p f_\varepsilon^p\ . \tag{3.3}$$

We say that $X_\varepsilon = \mathcal{O}(f_\varepsilon)$ (uniformly) on $[0, T]$, if we can replace the stopping time τ^* by the constant time T in (3.3). If X_ε is a random variable independent of time, we use the same notation without supremum in time, i.e. $X_\varepsilon = \mathcal{O}(f_\varepsilon)$ if $\mathbb{E}|X_\varepsilon|^p \leqslant C_p f_\varepsilon^p$.

We use the notation $X_\varepsilon = \mathcal{O}(f_\varepsilon^-)$ if $X_\varepsilon = \mathcal{O}(f_\varepsilon \varepsilon^{-\kappa})$ for every $\kappa > 0$.

3.1. Approximation of the stable part by an Ornstein–Uhlenbeck process

In this subsection we show that the 'fast' process $y(t)$ is actually close to an Ornstein–Uhlenbeck process, at least up to time τ^*. We have the following result.

Lemma 3.4. *Let $z(t)$ be the $\mathcal{N}^\perp$-valued process solving the SDE*

$$dz(t) = -\varepsilon^{-2} L z\, dt + \varepsilon^{-1} Q\, dW(t), \qquad z(0) = y(0). \tag{3.4}$$

Then one has $\|y(\cdot) - z(\cdot)\|_\alpha = \mathcal{O}(\varepsilon^{1-})$.

Proof. It follows from the mild formulation of (2.4) that

$$y(t) = z(t) + \frac{1}{\varepsilon} \int_0^t e^{-(t-s)L\varepsilon^{-2}} N(x(s), y(s))\, ds, \tag{3.5}$$

where we have used the notation $N(x, y) = \varepsilon\nu y + P_s B(x + y, x + y)$. From the properties of L we deduce that there exist positive constants $C,\ c$ such that

$$\|e^{-Lt} P_s\|_{\mathcal{H}^\beta\to\mathcal{H}^\alpha} \leqslant \begin{cases} C t^{(\beta-\alpha)/2} & \text{for } t \leqslant 1, \\ C e^{-ct} & \text{for } t \geqslant 1. \end{cases} \tag{3.6}$$

Since on the other hand assumption 2.3 implies that

$$\|N(x, y)\|_\beta \leqslant C(1 + \|x + y\|_\alpha)^2\ ,$$

the claim follows from the definition of τ^* and the fact that the right-hand side of (3.6) is integrable for $\beta > \alpha - 2$. □

The above approximation result enables us to obtain estimates on the statistics of the stopping time τ^*. For this we will need an estimate on the Ornstein–Uhlenbeck process (3.4).

Lemma 3.5. *Suppose that assumption 2.4 holds. Then there is a version of z which is almost surely $\mathcal{H}^\alpha$-valued with continuous sample paths. Furthermore, for every $\kappa_0 > 0$ and every $p > 0$, there exists a constant C such that*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} \|z(t)\|_\alpha^p\Big) \leqslant C\varepsilon^{-\kappa_0}. \tag{3.7}$$

Proof. It follows for example from the proof of [DPZ92, theorem 5.9]. □

An immediate corollary of lemmas 3.4 and 3.5 is that the process $y(t)$ is 'almost bounded' in ε. More precisely:

Corollary 3.6. *Assume that the conditions of lemma 3.4 and 3.5 hold. Then, for every $\kappa_0 > 0$ and every $p > 0$, there exists a constant C such that*

$$\mathbb{E}\Big(\sup_{t\in[0,\tau^*]} \|y(t)\|_\alpha^p\Big) \leqslant C\varepsilon^{-\kappa_0}. \tag{3.8}$$

Proof. Follows from lemma 3.4, equation (3.7), and the triangle inequality. □

Note that the value of κ_0 appearing in the statement above can be chosen independently of the value κ appearing in the definition of τ^*. Thus, with high probability, the event $\tau^* < T$ is caused by $x(t)$ getting too large. To be more precise:

Corollary 3.7. *Under assumptions 2.2 and 2.3, for every $p > 0$ and for every $K > 1$, there exists a constant C such that*

$$\mathbb{P}(\tau^* < T) \leqslant \mathbb{P}\left(K|x(\tau^*)| \geqslant \varepsilon^{-\kappa}\right) + C\varepsilon^p \qquad \text{for } \varepsilon \in (0, 1).$$

Proof. Follows from corollary 3.6 and the Chebyshev inequality. □

3.2. Elimination of the $\mathcal{O}(\frac{1}{\varepsilon})$ terms [4]

Let us first introduce some notation. Given a Hilbert space $\mathcal{H}$ we denote by $\mathcal{H} \otimes_s \mathcal{H}$ its symmetric tensor product. Similarly, we use the notation $v_1 \otimes_s v_2 = \frac{1}{2}\big(v_1 \otimes v_2 + v_2 \otimes v_1\big)$ for the symmetric tensor product of two elements and $(A \otimes_s B)(x \otimes y) = \frac{1}{2}\big(Ax \otimes By + By \otimes Ax\big)$ for the symmetric tensor product of two linear operators.

Let us recall that the scalar product in the tensor product space $\mathcal{H}_\alpha \otimes \mathcal{H}_\beta$ is given by $\langle u_1 \otimes v_1, u_2 \otimes v_2\rangle_{\alpha,\beta} := \langle u_1, u_2\rangle_\alpha \langle v_1, v_2\rangle_\beta$. With a slight abuse of notation, we write $\langle\cdot,\cdot\rangle_\alpha := \langle\cdot,\cdot\rangle_{\alpha,\alpha}$. Furthermore, we extend the bilinear form B to the tensor product space by $B(u \otimes v) = B(u, v)$.

With this notation, one can check that[5]:

[4] The analysis presented in this section is quite standard, however it requires many tedious calculations and estimates. For the sake of brevity we will omit most of the details.

[5] Since L has a zero eigenvalue, one should interpret $(I \otimes_s L)^{-1}$ and L^{-1} as pseudo-inverses, where, for instance, $L^{-1} = 0$ on its kernel $\mathcal{N}$. We will only apply these two operators to elements in the orthogonal complement to their respective kernels so that this is of no consequence.

Lemma 3.8. *The operator* $(I \otimes_s L)^{-1}$ *is bounded from* $\mathcal{H}_\gamma \otimes_s \mathcal{H}_\gamma$ *to* $\mathcal{H}_{\gamma+1} \otimes_s \mathcal{H}_{\gamma+1}$ *for any* $\gamma \in \mathbb{R}$.

Proof. It suffices to note that $I \otimes_s L$ is diagonal with eigenvalues $(\lambda_j + \lambda_k)/2$ in the basis $e_j \otimes_s e_k$. Note that $\ker(I \otimes_s L) = \mathcal{N} \otimes_s \mathcal{N}$. □

Now we are ready to present the main result of this subsection.

Proposition 3.9. *Let* x *and* z *be as above and let assumptions 2.2–2.4 hold. Then, there exists a process* $R = \mathcal{O}(\varepsilon^{1-})$ *such that, for every stopping time* t *with* $t \leqslant \tau^*$ *almost surely, one has*

$$\begin{aligned}
x(t) = x(0) &+ \nu \int_0^t x \,\mathrm{d}s + 4 \int_0^t P_c B\big(P_c B(x,z), L^{-1} z\big)\,\mathrm{d}s \\
&+ 2 \int_0^t P_c B\big(x, L^{-1} P_s B(x+z, x+z)\big)\,\mathrm{d}s + 2 \int_0^t P_c B\big(P_c B(z,z), L^{-1} z\big)\,\mathrm{d}s \\
&+ \int_0^t P_c B (I \otimes_s L)^{-1} \big(z \otimes_s Q \,\mathrm{d}W(t)\big) + 2 \int_0^t P_c B\big(x, L^{-1} Q \,\mathrm{d}W(s)\big) \\
&+ \int_0^t P_c B (I \otimes_s L)^{-1} \big(z \otimes_s P_s B(x+z, x+z)\big)\,\mathrm{d}s + R(t).
\end{aligned} \tag{3.9}$$

An immediate corollary is

Corollary 3.10. *Under the assumptions of proposition 3.9, define the process* $x_R(t)$ *by* $x_R(t) = x(t) - R(t)$. *Then for every* $p > 0$ *and every* $\tilde{\alpha} < 1/2$, *one has*

$$\sup_{0 \leqslant s < t \leqslant \tau^*} \frac{|x_R(t) - x_R(s)|}{|t-s|^{\tilde{\alpha}}} = \mathcal{O}(\varepsilon^{0-}) \,.$$

Proof. It follows immediately from (3.9), using the definition of τ^*. The condition $\tilde{\alpha} < 1/2$ arises from the two stochastic integrals in the right-hand side of (3.9). □

Proof of proposition 3.9. Applying Itô's formula to $B(x, L^{-1} y)$, we get the following identity in $\mathcal{H}^\beta$:

$$\begin{aligned}
\mathrm{d}B(x, L^{-1} y) = {} & 2\nu B(x, L^{-1} y)\,\mathrm{d}t + 2\varepsilon^{-1} B\big(P_c B(x,y), L^{-1} y\big)\,\mathrm{d}t \\
&+ \varepsilon^{-1} B\big(P_c B(y,y), L^{-1} y\big)\,\mathrm{d}t - \varepsilon^{-2} B(x,y)\,\mathrm{d}t \\
&+ \varepsilon^{-1} B\big(x, L^{-1} P_s B(x+y, x+y)\big)\,\mathrm{d}t + \varepsilon^{-1} B\big(x, L^{-1} Q \,\mathrm{d}W(t)\big).
\end{aligned}$$

Combining this with lemma 3.4 and the continuity properties of B stated in assumption 2.3, it follows that, for every stopping time t with $t \leqslant \tau^*$ almost surely, one has

$$\begin{aligned}
2 \int_0^t B(x,y)\,\mathrm{d}s = 4\varepsilon & \int_0^t B\big(P_c B(x,z), L^{-1} z\big)\,\mathrm{d}s + 2\varepsilon \int_0^t B\big(x, L^{-1} Q \,\mathrm{d}W(s)\big) \\
&+ 2\varepsilon \int_0^t B\big(x, L^{-1} P_s B(x+z, x+z)\big)\,\mathrm{d}s \\
&+ 2\varepsilon \int_0^t B\big(P_c B(z,z), L^{-1} z\big)\,\mathrm{d}s + R_1(t),
\end{aligned} \tag{3.10}$$

where $R_1(t) = \mathcal{O}(\varepsilon^{2-})$. Applying Itô's formula to $\frac{1}{2}(y \otimes y)$, we get the following identity in $\mathcal{H}^{\alpha-2} \otimes \mathcal{H}^{\alpha-2}$:

$$\begin{aligned}
\tfrac{1}{2}\mathrm{d}(y \otimes y) = {} & \nu\big(y \otimes y\big)\,\mathrm{d}t - \varepsilon^{-2}\big(y \otimes_s L y\big)\,\mathrm{d}t \\
&+ \varepsilon^{-1} y \otimes_s P_s B(x+y, x+y)\,\mathrm{d}t + \varepsilon^{-1} y \otimes_s Q \,\mathrm{d}W(t) \\
&+ \varepsilon^{-2} \sum_{i=1}^{\infty} Q e_i \otimes Q e_i \,\mathrm{d}t.
\end{aligned} \tag{3.11}$$

Note however that all terms but the second one actually belong to $\mathcal{H}^{\alpha-1} \otimes \mathcal{H}^{\alpha-1}$. Since B is bounded from $\mathcal{H}^{\alpha} \otimes \mathcal{H}^{\alpha}$ into $\mathcal{H}^{\beta}$ and since $(I \otimes_s L)^{-1}$ is bounded from $\mathcal{H}^{\alpha-1} \otimes \mathcal{H}^{\alpha-1}$ to $\mathcal{H}^{\alpha} \otimes \mathcal{H}^{\alpha}$, we can apply $P_c B(I \otimes_s L)^{-1}$ to both sides of (3.11). Noting that $P_c B(e_i \otimes e_i) = 0$ by assumption 2.3, we get

$$\int_0^t P_c B(y, y)\,\mathrm{d}s = \varepsilon \int_0^t P_c B(I \otimes_s L)^{-1}\big(z \otimes_s P_s B(x+z, x+z)\big)\,\mathrm{d}s$$
$$+ \varepsilon \int_0^t P_c B(I \otimes_s L)^{-1}\big(z \otimes_s Q\,\mathrm{d}W(t)\big) + R_2(t),$$

where $R_2(t) = \mathcal{O}(\varepsilon^{2-})$. Collecting both terms and inserting them into (3.1a) concludes the proof.

4. Averaging over the fast Ornstein–Uhlenbeck process

In this section, we simplify the equation for x further by showing that one can eliminate all terms in (3.9) that contain odd powers of z. Furthermore, concerning the terms that are quadratic in z, there exists a constant $\widehat{Q} \in \mathcal{H}^{\alpha} \otimes_s \mathcal{H}^{\alpha}$ so that one can make the formal substitution $z \otimes z \mapsto \widehat{Q}$.

We start with a number of bounds on the integrals of products of Ornstein–Uhlenbeck processes.

4.1. An averaging result with explicit error bounds

Recall that QW can (at least on a formal level) be written as $\sum_{k=2}^{\infty} q_k e_k w_k(t)$ for some independent standard Wiener processes w_k. For $\varepsilon > 0$ and $k > 1$, we define $\hat{z}_k(t)$ to be the stationary solution of

$$\mathrm{d}\hat{z}_k = -\lambda_k \varepsilon^{-2} \hat{z}_k\,\mathrm{d}t + q_k \varepsilon^{-1}\,\mathrm{d}w_k(t)\ .$$

This is a Gaussian process with covariance

$$\mathbb{E}\left(\hat{z}_k(s)\hat{z}_k(t)\right) = \frac{q_k^2}{2\lambda_k} \mathrm{e}^{-\frac{\lambda_k |t-s|}{\varepsilon^2}}. \tag{4.1}$$

Since $\hat{z}_k(t)$ fluctuates very rapidly, one would expect from the law of large numbers that as $\varepsilon \to 0$, one has $\hat{z}_k \to 0$ weakly, but $(\hat{z}_k)^2 \to \frac{q_k^2}{2\lambda_k}$ weakly. This is made precise by the following bounds:

Lemma 4.1. *For every $p > 0$ there exists a constant C_p such that, for every $t > s > 0$ and every $k, \ell, m > 1$, the bounds*

$$\mathbb{E}\Big(\int_s^t \hat{z}_k(r)\,\mathrm{d}r\Big)^{2p} \leqslant C_p \Big(\frac{q_k^2}{\lambda_k}\Big)^p (t-s)^p \varepsilon^{2p},$$
$$\mathbb{E}\Big(\int_s^t \Big(\hat{z}_k(r)\hat{z}_\ell(r) - \frac{q_k^2}{2\lambda_k}\delta_{kl}\Big)\,\mathrm{d}r\Big)^{2p} \leqslant C_p \Big(\frac{q_k^2 q_\ell^2}{\lambda_k \lambda_\ell}\Big)^p (t-s)^p \varepsilon^{2p}\ ,$$
$$\mathbb{E}\Big(\int_s^t \hat{z}_k(r)\hat{z}_\ell(r)\hat{z}_m(r)\,\mathrm{d}r\Big)^{2p} \leqslant C_p \Big(\frac{q_k^2 q_\ell^2 q_m^2}{\lambda_k \lambda_\ell \lambda_m}\Big)^p (t-s)^p \varepsilon^{2p}\ ,$$

hold. Here we denoted by δ_{kl} the Kronecker symbol.

Proof. The first bound can be checked explicitly in the case $p = 1$ by using (4.1). The case $p > 1$ follows immediately from the fact that $\int_s^t \hat{z}_k(r)\,\mathrm{d}r$ is Gaussian.

In order to obtain the other bounds, we recall first the fact that for an $\mathbb{R}^{2p}$-valued Gaussian random variable $X = (X_1, \dots, X_{2p})$, we have

$$\mathbb{E}X_1 \cdot \ldots \cdot X_n = \sum_{\sigma \in \Sigma(2p)} \prod_{i=1}^{p} \mathbb{E}X_{\sigma_{2i-1}} X_{\sigma_{2i}}$$

where $\Sigma(2p)$ is the set of all permutations of $\{1, \dots, 2p\}$.

Turning to the second claim, consider first the case where $k \neq \ell$, so that $\hat{z}_k$ and $\hat{z}_\ell$ are independent. Thus

$$\begin{aligned}\mathbb{E}\prod_{i=1}^{2p} \hat{z}_k(t_i)\hat{z}_\ell(t_i) &= \mathbb{E}\prod_{i=1}^{2p} \hat{z}_k(t_i)\mathbb{E}\prod_{i=1}^{2p} \hat{z}_\ell(t_i) \\ &\leqslant C\Big(\frac{q_k^2 q_\ell^2}{\lambda_k \lambda_\ell}\Big)^p \sum_{\sigma \in \Sigma(2p)} \prod_{i=1}^{p} \exp\Big(-\frac{c}{\varepsilon^2}|t_{\sigma_{2i}} - t_{\sigma_{2i-1}}|\Big) ,\end{aligned}$$

where $c = \lambda_1$ is the smallest non-zero eigenvalue of L. The bound then follows by integrating over $t_1, \dots, t_{2p}$ and using the fact that $\int_s^t \int_s^t \exp(-c|r-u|/\varepsilon^2)\, dr\, du \leqslant C\varepsilon^2(t-s)$.

In the case where $k = \ell$, we have

$$\begin{aligned}\mathbb{E}\prod_{i=1}^{2p}\Big((\hat{z}_k)^2(t_i) - \frac{q_k^2}{2\lambda_k}\Big) &= \sum_{A \subset \{1,\dots,2p\}} \Big(-\frac{q_k^2}{2\lambda_k}\Big)^{2p-|A|} \mathbb{E}\prod_{i \in A} \hat{z}_k(t_i)^2 \\ &= \sum_{A \subset \{1,\dots,2p\}} \Big(-\frac{q_k^2}{2\lambda_k}\Big)^{2p-|A|} \sum_{\tau \in \Sigma^2(A)} \prod_{i=1}^{|A|} \mathbb{E}\hat{z}_k(t_{\tau_{2i}})\hat{z}_k(t_{\tau_{2i-1}}) ,\end{aligned}$$

where the sum runs over $\Sigma^2(A)$, the space of all permutations of numbers in A, where each number is allowed to appear twice. Now it is possible to check that all terms in the double sum where $|A| < 2p$ are cancelled by a term with a larger $\tilde{A}$, where $t_{\tau_{2i}} = t_{\tau_{2i-1}}$ for some i. All remaining terms have $|A| = 2p$. It follows from (4.1) that there exists a constant C such that

$$\mathbb{E}\prod_{i=1}^{2p}\Big((\hat{z}_k)^2(t_i) - \frac{q_k^2}{2\lambda_k}\Big) \leqslant C\Big(\frac{q_k^2}{\lambda_k}\Big)^{2p} \sum_{\substack{\sigma \in \Sigma(2p) \\ \sigma_i \neq i}} \exp\Big(-c\frac{|t_{\sigma_1} - t_1| + \ldots + |t_{\sigma_{2p}} - t_{2p}|}{\varepsilon^2}\Big) .$$

The bound then follows immediately by integrating over $t_1, \dots, t_p$. The last term can be bounded in a similar way. $\square$

Let now $\widehat{Q} \in \mathcal{H}^\alpha \otimes_s \mathcal{H}^\alpha$ be given by

$$\widehat{Q} = \sum_{k=2}^{\infty} \frac{q_k^2}{2\lambda_k}\big(e_k \otimes e_k\big) . \tag{4.2}$$

The fact that it does indeed belong to $\mathcal{H}^\alpha \otimes_s \mathcal{H}^\alpha$ is a consequence of assumption 2.4. Writing $\hat{z}(t) = \sum_{k=1}^\infty \hat{z}_k(t)e_k$, we have the following corollary of lemma 4.1:

Corollary 4.2. *For every $p > 0$ there exists a constant C_p such that the bounds*

$$\begin{aligned}\mathbb{E}\Big\|\int_s^t \hat{z}(r)\, dr\Big\|_\alpha^{2p} &\leqslant C_p(t-s)^p\varepsilon^{2p} , \\ \mathbb{E}\Big\|\int_s^t \big(\hat{z}(r) \otimes \hat{z}(r) - \widehat{Q}\big)\, dr\Big\|_\alpha^{2p} &\leqslant C_p(t-s)^p\varepsilon^{2p} , \\ \mathbb{E}\Big\|\int_s^t \big(\hat{z}(r) \otimes \hat{z}(r) \otimes \hat{z}(r)\big)\, dr\Big\|_\alpha^{2p} &\leqslant C_p(t-s)^p\varepsilon^{2p} ,\end{aligned}$$

hold for every $t > s > 0$.

Proof. It is a straightforward consequence of the following fact. Let $\{v_k\}$ be a sequence of real-valued random variables such that there exists a sequence $\{\gamma_k\}$ and, for every $p \geqslant 1$, a constant C_p such that $\mathbb{E}|v_k|^p \leqslant C_p \gamma_k^p$. Then, for every $p \geqslant 1$ there exists a constant C'_p such that

$$\mathbb{E}\Big(\sum_{k=1}^{\infty} \lambda_k^\alpha v_k^2\Big)^p \leqslant C'_p \Big(\sum_{k=1}^{\infty} \lambda^\alpha \gamma_k^2\Big)^p .$$

The result now follows immediately from lemma 4.1 and from assumption 2.4 which states that the sequence $q_k^2 \lambda_k^{\alpha-1}$ is summable. □

Lemma 4.3. *Let G_ε be a family of processes in some Hilbert space $\mathcal{K}$ such that, for every $p \geqslant 1$ and every $\kappa > 0$ there exists a constant C such that*

$$\mathbb{E}\Big\| \int_s^t G_\varepsilon(r)\,\mathrm{d}r \Big\|^{2p} \leqslant C(t-s)^p \varepsilon^{2p} , \tag{4.3}$$

holds for every $0 \leqslant s < t \leqslant 1$. Then, for every $p > 0$ and every $\kappa > 0$, there exists a constant C such that

$$\mathbb{E}\Big(\sup_{n<N} \Big\| \int_{n\delta}^{(n+1)\delta} G_\varepsilon(s)\,\mathrm{d}s \Big\|\Big)^{2p} \leqslant C N^\kappa \delta^p \varepsilon^{2p} ,$$

holds for every $N > 0$, every $\delta \in (0, (N+1)^{-1})$, and every $\varepsilon > 0$.

Proof. It follows from (4.3) that, for every $q > 0$, there exists a constant C such that

$$\mathbb{P}\Big(\sup_{n<N} \Big\| \int_{n\delta}^{(n+1)\delta} G_\varepsilon(s)\,\mathrm{d}s \Big\| > K\Big) \leqslant C N \frac{\delta^{q/2}}{K^q} \varepsilon^q , \tag{4.4}$$

holds for every $K > 0$. Note now that if a positive random variable X satisfies $\mathbb{P}(X > x) \leqslant \bar{C}/x^q$ for every $x > 0$, then, for $p < q$, one has

$$\begin{aligned}\mathbb{E}X^p &= p \int_0^\infty x^{p-1} \mathbb{P}(X > x)\,\mathrm{d}x \leqslant p \int_0^{\bar{C}^{1/q}} x^{p-1}\,\mathrm{d}x + \bar{C} p \int_{\bar{C}^{1/q}}^\infty x^{p-q-1}\,\mathrm{d}x \\ &= \frac{q}{q-p} \bar{C}^{p/q} .\end{aligned}$$

Combining this with (4.4) and choosing q sufficiently large yields the required bound. □

Proposition 4.4. *Let $\mathcal{K}$ be a Hilbert space, let f be a $\mathcal{K}$-valued random process with almost surely $\tilde{\alpha}$-Hölder continuous trajectories, let G_ε be a family of $\mathcal{K}$-valued processes satisfying (4.3), and let*

$$F_\varepsilon(t) := \int_0^t \langle G_\varepsilon(s), f(s)\rangle\,\mathrm{d}s.$$

Assume furthermore that, for every $\kappa > 0$ and every $p > 0$, there exists a constant C such that

$$\mathbb{E} \sup_{t\in[0,1]} \|G_\varepsilon(t)\|^p \leqslant C\varepsilon^{-\kappa} . \tag{4.5}$$

Then, for every $\gamma < 2\tilde{\alpha}/(1+2\tilde{\alpha})$, there exists a constant C depending only on p and γ such that

$$\mathbb{E}\|F_\varepsilon\|_{C^{1-\tilde{\alpha}}}^p \leqslant C\big(\mathbb{E}\|f\|_{C^{\tilde{\alpha}}}^{2p}\big)^{1/2} \varepsilon^{\gamma p} ,$$

where $\|\cdot\|_{C^{\tilde{\alpha}}}$ denotes the $\tilde{\alpha}$-Hölder norm for $\mathcal{K}$-valued functions on $[0, 1]$.

Note that, if we can choose $\tilde{\alpha} < 1/2$, but arbitrarily close, then we can choose $\gamma < 1/2$, but arbitrarily close, too.

Proof. We focus only on the Hölder part of the norm. The L^∞ part follows easily, as $F_\varepsilon(0) = 0$, e.g. by taking $s = 0$ in the following.

Choose $\delta > 0$ to be fixed later. Moreover, for every pair s and t in $[0, 1]$, set $\bar{s} = \delta[\delta^{-1}s]$ and $\bar{t} = \delta[\delta^{-1}t]$. We furthermore define $f_\delta(t) = f(\bar{t})$. One then has

$$\begin{aligned}\left|\int_s^t \langle G_\varepsilon(r), f(r)\rangle\,\mathrm{d}r\right| &\leqslant \left|\int_s^t \langle G_\varepsilon(r), \big(f(r) - f_\delta(r)\big)\rangle\,\mathrm{d}r\right| + \left|\int_s^t \langle G_\varepsilon(r), f_\delta(r)\rangle\,\mathrm{d}r\right| \\ &\leqslant \|G_\varepsilon\|_{L^\infty(\mathcal{K})}\delta^{\tilde{\alpha}}|t-s|\|f\|_{C^{\tilde{\alpha}}} + \left|\int_s^t \langle G_\varepsilon(r), f_\delta(r)\rangle\,\mathrm{d}r\right|.\end{aligned}$$

The second term in the right-hand side can be bounded by

$$\mathbf{1}_{|t-s|\geqslant\delta}\left|\int_{\bar{s}}^{\bar{t}} \langle G_\varepsilon(r), f_\delta(r)\rangle\,\mathrm{d}r\right| + \min\{|t-s|, 2\delta\}\|f\|_{C^{\tilde{\alpha}}}\|G_\varepsilon\|_{L^\infty(\mathcal{K})}.$$

The first term of this expression is in turn bounded by

$$\frac{2|t-s|}{\delta}\Big(\sup_{n<\delta^{-1}}\Big\|\int_{n\delta}^{(n+1)\delta} G_\varepsilon(r)\,\mathrm{d}r\Big\|\Big)\|f\|_{C^{\tilde{\alpha}}}\ .$$

Collecting all these expressions yields

$$\|F_\varepsilon\|_{C^{1-\tilde{\alpha}}} \leqslant C\|G_\varepsilon\|_{L^\infty(\mathcal{K})}\|f\|_{C^{\tilde{\alpha}}}\delta^{\tilde{\alpha}} + \delta^{-1}\Big(\sup_{n<\delta^{-1}}\Big\|\int_{n\delta}^{(n+1)\delta} G_\varepsilon(r)\,\mathrm{d}r\Big\|\Big)\|f\|_{C^{\tilde{\alpha}}}.$$

Choosing $\delta = \varepsilon^{2/(1+2\tilde{\alpha})}$, applying lemma 4.3, and using (4.5) easily concludes the proof. □

We are actually going to use the following corollary of proposition 4.4.

Corollary 4.5. *Let $\hat{z}$ be as above and let α be as in assumptions 2.3 and 2.4. Fix $T > 0$ and let f_i with $i \in \{1, 2, 3\}$ be $\tilde{\alpha}$-Hölder continuous functions on $[0, T]$ with values in $\big((\mathcal{H}^\alpha)^{\otimes i}\big)^*$, respectively. Let F_ε be given by*

$$F_\varepsilon(t) := \int_0^t \Big(\big(f_1(s)\big)(\hat{z}) + \big(f_2(s)\big)(\hat{z}\otimes\hat{z} - \widehat{Q}) + \big(f_3(s)\big)(\hat{z}\otimes\hat{z}\otimes\hat{z})\Big)\,\mathrm{d}s.$$

Then, for every $\gamma < 2\tilde{\alpha}/(1+2\tilde{\alpha})$, there exists a constant C depending only on p and γ such that

$$\mathbb{E}\sup_{t\in[0,T]}|F_\varepsilon(t)|^p \leqslant C\varepsilon^{\gamma p}\big(\mathbb{E}\big(\|f_1\|_{C^{\tilde{\alpha}}} + \|f_2\|_{C^{\tilde{\alpha}}} + \|f_3\|_{C^{\tilde{\alpha}}}\big)^{2p}\big)^{1/2}\ ,$$

where $\|\cdot\|_{C^{\tilde{\alpha}}}$ denotes the $\tilde{\alpha}$-Hölder norm for $\big((\mathcal{H}^\alpha)^{\otimes i}\big)^*$-valued functions on $[0, \tau^*]$.

Proof. Note that $\hat{z}$ satisfies (4.5) with $\mathcal{K} = \mathcal{H}^\alpha$. This follows for example from the proof of [DPZ92, theorem 5.9]. The statement is then a consequence of corollary 4.2 and of proposition 4.4. □

4.2. The reduction of the slow modes

We now use the results of the previous subsection in order to show that most of the terms that appear on the right-hand side of equation (3.9) are of order $\mathcal{O}(\varepsilon^{1/2-})$. Note first that we can replace all occurrences of z in (3.9) by the stationary process $\hat{z}$ without changing the order of magnitude of the remaining term R. We are now ready to state the main result of this section.

Proposition 4.6. *Under assumptions 2.2–2.4 and with x and $\hat{z}$ defined as above, we obtain*

$$
\begin{aligned}
x(t) = x(0) &+ \int_0^t P_c B(I \otimes_s L)^{-1}\big(\hat{z}(s) \otimes_s Q\,\mathrm{d}W(s)\big) \\
&+ 2\int_0^t P_c B\big(x(s), L^{-1} Q\,\mathrm{d}W(s)\big) \\
&+ \int_0^t \tilde{\nu} x(s)\,\mathrm{d}s - \int_0^t \tilde{\eta}\big(x(s), x(s), x(s)\big)\,\mathrm{d}s + R(t)\,,
\end{aligned} \tag{4.6}
$$

where $\|R(\cdot)\| = \mathcal{O}(\varepsilon^{1/2-})$. The linear map $\tilde{\nu}: \mathcal{N} \to \mathcal{N}$ and the trilinear map $\tilde{\eta}: \mathcal{N}^3 \to \mathcal{N}$ are given by

$$
\begin{aligned}
\tilde{\nu} x = \nu x + 2B_c\Big((I \otimes_s L)^{-1}(B_s \otimes_s I) &+ (I \otimes L^{-1} B_s) \\
&+ 2(B_c \otimes L^{-1})\Big)(x \otimes \widehat{Q})\,,
\end{aligned} \tag{4.7}
$$

$$
\tilde{\eta}(x, x, x) = -2B_c\big(x, L^{-1} B_s(x, x)\big)\,. \tag{4.8}
$$

Here, we used the notation $B_s := P_s B$ and $B_c := P_c B$.

Proof. First we replace all instances of z by $\hat{z}$ on the right-hand side of equation (3.9), which results in an error of order $\mathcal{O}(\varepsilon^{1-})$ which is absorbed into R. This is a straightforward but somewhat lengthy calculation which we do not reproduce here. We rely on

$$
z(t) - \hat{z}(t) = \mathrm{e}^{-tL\varepsilon^{-2}}(P_s v_0 - z(0)) = \mathrm{e}^{-tL\varepsilon^{-2}}\mathcal{O}(\varepsilon^{0-})\,.
$$

Note that obviously $\|z - \hat{z}\| \neq \mathcal{O}(\varepsilon^{1-})$, as bounds uniformly in time are not available due to transient effects on timescales smaller than $\mathcal{O}(\varepsilon^2)$. Nevertheless, we only bound the error in integrated form, which is sufficient for our application. Actually, it is possible to check that the error terms which result from this substitution are of $\mathcal{O}(\varepsilon^{2-})$. The only exception to this is the stochastic integral, where we apply the Burkholder–Davies–Gundy inequality in order to get a remainder term of $\mathcal{O}(\varepsilon^{1-})$.

Once this substitution has been performed, the proposition follows from an application of corollary 4.5 to the modified equation (3.9). The fact that, for every $\tilde{\alpha} > 1/2$, the various integrands are indeed $\tilde{\alpha}$-Hölder continuous functions with values in $\big((\mathcal{H}^\alpha)^{\otimes i}\big)^*$ and Hölder norm of order $\mathcal{O}(\varepsilon^{0-})$ follows from corollary 3.10. □

Remark 4.7. An alternative way of formulating propositions 3.9 and 4.6 would be to take t deterministic, but to change the definition of x so that x remains constant once the stopping time τ^* is reached. Since we only claim bounds on the error $R(t)$ for $t \leqslant \tau^*$, equation (4.6) would then be satisfied for that modified process x up to the deterministic time t.

5. Convergence to a limiting equation

We are now going to show that the process x converges in law, as $\varepsilon \to 0$, to the solution of a finite-dimensional SDE, without requiring the one dimensionality of the kernel of L. The price that we have to pay is that we do not obtain any convergence rate. This is because the argument that we use is a tightness argument to show the existence of convergent subsequences, combined with the characterization of the solution of the limiting SDE as the (unique in law) solution to the appropriate martingale problem. Since this technique is fairly standard (see for example [PSV77], [BLP78, chapter 3]) we do not enter into exhaustive details. The main deviation from the standard procedure concerns the fact that we do not assume global existence of solutions to the original equation, which leads to some technical difficulties.

In order to formulate the main theorem of this section, we introduce first the function $F:\mathcal{N}\to\mathcal{N}$ given by

$$F(x) = \tilde{\nu}x - \tilde{\eta}(x,x,x)\;,$$

where $\tilde{\nu}$ and $\tilde{\eta}$ are defined in (4.7) and (4.8). We furthermore define the symmetric, $L(\mathcal{N},\mathcal{N})$-valued map Σ by

$$\langle y, \Sigma(x)y\rangle = 4\sum_{k>n} q_k^2 \langle y, B_c(e_k,x)\rangle^2 + \sum_{k,\ell>n} \frac{q_k^2 q_\ell^2}{2\big(\lambda_k+\lambda_\ell\big)^2\lambda_\ell}\langle y, B_c(e_k,e_\ell)\rangle^2\;.$$

(Recall that the kernel of L is spanned by the first n eigenvectors $e_1,\dots,e_n$.) With these notations in place, we have

Theorem 5.1. *Let v_ε^0 be a bounded sequence of elements in $\mathcal{H}^\alpha$ such that $P_c v_\varepsilon^0 = a_0$ is independent of ε. Fix a time $T>0$ and denote by $\mathbb{P}_\varepsilon$ the law of the solution of (3.1), stopped at the stopping time τ^* defined by (3.2). Denote furthermore by $\pi:\mathcal{C}([0,T],\mathcal{H}_\alpha)\to\mathcal{C}([0,T],\mathcal{N})$ the projection onto the kernel of L and assume that $\tilde{\eta}$ is such that $\langle x,\tilde{\eta}(x,x,x)\rangle>0$ for every $x\in\mathcal{N}$.*

Then, the sequence of measures $\pi^*\mathbb{P}_\varepsilon$ converges weakly to the measure $\mathbb{P}$, the law of the solution to the SDE

$$\mathrm{d}a = F(a)\,\mathrm{d}t + \Sigma^{1/2}(a)\,\mathrm{d}B(t),\qquad a(0)=a_0\;. \tag{5.1}$$

Here, B denotes an $\mathcal{N}$-valued Wiener process.

Proof. Fix a sequence v_ε^0 of initial conditions as described in the assumption and denote by x_ε the projection of the solution v_ε onto the kernel of L. Denote furthermore as before by y_ε its projection onto the orthogonal complement of $\mathcal{N}$.

Denote by $\mathbb{P}_\varepsilon$ the measure on $\mathcal{C}([0,T],\mathcal{H}_\alpha)$ given by the law of the process v_ε stopped at τ^* (with τ^* defined as in (3.2) above) and by $\mathbb{P}$ the measure on $\mathcal{C}([0,T],\mathcal{N})$ given by the law of the solution to the amplitude equation (5.1).

The claim of the theorem (that $\pi^*\mathbb{P}_\varepsilon$ converges weakly to $\mathbb{P}$ as $\varepsilon\to 0$) is then an immediate consequence of the following two claims whose proofs will be given below:

1. The sequence of probability measures $\pi^*\mathbb{P}_\varepsilon$ is tight.
2. Every accumulation point of $(\pi^*\mathbb{P}_\varepsilon)_{\varepsilon>0}$ is a solution to the martingale problem associated with (5.1).

Let us start with the proof of tightness of $\pi^*\mathbb{P}_\varepsilon$. Consider (4.6) and define $x_R(t) = x(t^*)-R(t^*)$, where R is as in (4.6) and where we used the notation $t^* = t\wedge\tau^*$. Since $\mathbb{E}\sup_{t\in[0,T]}\|R(t^*)\| = \mathcal{O}(\varepsilon^{1/2-})$, it follows from the version of Kolmogorov's continuity test given in [RY99] that tightness of $\pi^*\mathbb{P}_\varepsilon$ follows if we can show that

$$\mathbb{E}\|x_R(t)-x_R(s)\|^4 \leqslant C|t-s|^2\;, \tag{5.2}$$

with a constant C independent of ε. We first show the boundedness of $\|x_R\|^{2p}$ in expectation for every integer $p\geqslant 1$. Applying Itô's formula to (4.6), we get

$$\begin{aligned}\mathbb{E}\|x_R(t)\|^{2p} - \|x_R(0)\|^{2p} &\leqslant 2p\,\mathbb{E}\int_0^{t^*} \|x_R(s)\|^{2p-2}\langle x_R(s), F(x(s))\rangle\,\mathrm{d}s\\ &\quad + C\mathbb{E}\int_0^{t^*}\|x_R(s)\|^{2p-2}\big(\|x_R(s)\|^2+\|R(s)\|^2+\|\hat{z}(s)\|_\alpha^2\big)\,\mathrm{d}s\;,\end{aligned}$$

It then follows from the bounds on R, combined with the fact that moments of $\hat{z}$ are uniformly bounded and that, by assumption, $\langle x,\tilde{\eta}(x,x,x)\rangle\geqslant c\|x\|^4$ for some $c>0$, that

$$\mathbb{E}\|x_R(t)\|^{2p} - \|x_R(0)\|^{2p} \leqslant Ct - \mathbb{E}\int_0^{t^*}\|x_R(s)\|^{2p}\,\mathrm{d}s\;.$$

It follows from Gronwall's inequality that $\mathbb{E}\|x_R(t)\|^{2p} < C$ for every $t \in [0, T]$, uniformly in $\varepsilon > 0$. The bound (5.2) now follows easily from (4.6), combined with the Burkholder–Davis–Gundy inequality.

Let $\phi \in \mathcal{C}_0^\infty(\mathcal{N})$ now be a smooth and compactly supported function. Applying Itô's formula to (4.6) and defining $x_R(t) = x(t^*) - R(t^*)$ as before, we find that, under $\mathbb{P}_\varepsilon$, one has the almost sure identity

$$\begin{aligned}\phi(x_R(t)) - \phi(x(0)) = M^\phi(t) + \int_0^{t^*} \langle D\phi(x_R(s)), F(x(s))\rangle \,\mathrm{d}s \\ + \sum_{k>n} \frac{q_k^2}{2} \int_0^{t^*} \big(D^2\phi(x_R(s))\big)\big(\mathcal{T}_k(x,\hat z), \mathcal{T}_k(x,\hat z)\big)\,\mathrm{d}s\ ,\end{aligned} \tag{5.3}$$

for some martingale M^ϕ with $\mathcal{T}_k(x,\hat z) = B_c(I \otimes_s L)^{-1}(\hat z \otimes_s e_k) + 2B_c(x(s), L^{-1}e_k)$.

Combining the bound obtained above on x with corollary 3.7, we obtain

$$\lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon\big(\tau^* \neq T\big) = 0.$$

It follows furthermore form corollary 4.5 that the second term in (5.3) converges to

$$\frac{1}{2}\int_0^{t^*} \mathrm{tr}\big(D^2\phi(x(s))\big)\Sigma(x(s))\,\mathrm{d}s,$$

where we identify $D^2\phi$ with a linear map from $\mathcal{N}$ to $\mathcal{N}$. This, together with the existing bounds on R and the smoothness of the test function ϕ, shows that there exists a process $\hat R_\varepsilon$ and a martingale $\hat M_\varepsilon^\phi$ (under $\mathbb{P}_\varepsilon$ and with respect to the filtration generated by x) such that

$$\begin{aligned}\phi(x(t)) - \phi(x(0)) = \hat M_\varepsilon^\phi(t) + \int_0^t \langle D\phi(x(s)), F(x(s))\rangle\,\mathrm{d}s \\ + \frac{1}{2}\int_0^t \mathrm{tr}\big(D^2\phi(x(s))\big)\Sigma(x(s))\,\mathrm{d}s + \hat R_\varepsilon(t)\ .\end{aligned} \tag{5.4}$$

If we stop the martingale $\hat M_\varepsilon^\phi$ at the time τ^* (which is also the time at which the process x is stopped), we can verify that the stopped error term $\hat R_\varepsilon$ satisfies in that case

$$\lim_{\varepsilon\to 0} \mathbb{E}_{\varepsilon,\delta}\Big(\sup_{t\leqslant T} |\hat R_\varepsilon(t)|\Big) = 0. \tag{5.5}$$

Define now the (continuous) function $M^\phi : \mathcal{C}([0,T],\mathcal{N}) \to \mathcal{C}([0,T],\mathbb{R})$ by

$$\big(M^\phi(x)\big)(t) = \phi(x(t)) - \phi(x(0)) - \int_0^t \big(\mathcal{L}\phi\big)(x(s))\,\mathrm{d}s\ ,$$

were $\mathcal{L}$ is the generator of (5.1). Let $\hat{\mathbb{P}}$ be an arbitrary accumulation point of $\mathbb{P}_\varepsilon$ and choose a sequence $\varepsilon_n \to 0$ such that $\mathbb{P}_{\varepsilon_n} \to \hat{\mathbb{P}}$ weakly. If we can show that M^ϕ is a $\hat{\mathbb{P}}$-martingale, the proof of the theorem is complete. Fix $t > s > 0$ and choose an arbitrary continuous function $f : \mathcal{C}([0,T],\mathcal{N}) \to \mathbb{R}$ which is also $\mathcal{F}_s$-measurable. Then

$$\begin{aligned}\hat{\mathbb{E}}\big(M^\phi(x)\big)(t)f(x) &= \lim_{n\to\infty} \mathbb{E}_{\varepsilon_n}\big(M^\phi(x)\big)(t)f(x) \\ &= \lim_{n\to\infty} \mathbb{E}_{\varepsilon_n}\big(M^\phi(x) - \hat R_{\varepsilon_n}\big)(t)f(x) \\ &= \lim_{n\to\infty} \mathbb{E}_{\varepsilon_n}\big(M^\phi(x) - \hat R_{\varepsilon_n}\big)(s)f(x) \\ &= \lim_{n\to\infty} \mathbb{E}_{\varepsilon_n}\big(M^\phi(x)\big)(s)f(x) = \hat{\mathbb{E}}\big(M^\phi(x)\big)(s)f(x).\end{aligned} \tag{5.6}$$

The first equality is a consequence of weak convergence and the continuity of both M^ϕ and f (in x). The second equality follows from (5.5) and the third equality follows from the fact that $\hat M_{\varepsilon_n}^\phi$ is a martingale under $\mathbb{P}_{\varepsilon_n}$. Since (5.6) holds for every continuous f, this shows that

$M^\phi(x)$ is a martingale under $\hat{\mathbb{P}}$. Therefore $\hat{\mathbb{P}}$ is a solution to the martingale problem associated with (5.1). Since this solution is unique [SV06], we conclude that $\hat{\mathbb{P}} = \mathbb{P}$. $\square$

6. Approximation of the martingale term

This section deals with the final reduction step for the general system (3.1) in the case where the kernel of L is one dimensional. We start by eliminating the stochastic integral of the type $\int_0^t \hat{z} \otimes Q\,\mathrm{d}W(s)$ from (4.6). In fact, we show that we can replace the martingale part in equation (4.6) by a single stochastic integral of the type

$$\int_0^t \sqrt{\sigma_a + \sigma_b a^2(s)}\,\mathrm{d}B(s),$$

against a *one-dimensional* Wiener process B. Note that this section is superfluous in the particular case where the first stochastic integral in (4.6) vanishes. This is the case for example in the situation considered in [Rob03]. See theorem 7.1 in the next section.

We emphasize again that the argument presented in the remainder of this paper is valid *only* under the assumption that the kernel of L is one dimensional.

6.1. An abstract martingale approximation result

We start with the following lemma; we will use it to approximate the martingale part of equation (4.6) by a stochastic integral against a one-dimensional Brownian motion.

Lemma 6.1. *Let $M(t)$ be a continuous $\mathcal{F}_t$-martingale with quadratic variation f and let g be an arbitrary $\mathcal{F}_t$-adapted increasing process with $g(0) = 0$. Then, there exists a filtration $\tilde{\mathcal{F}}_t$ with $\mathcal{F}_t \subset \tilde{\mathcal{F}}_t$ and a continuous $\tilde{\mathcal{F}}_t$-martingale $\tilde{M}(t)$ with quadratic variation g such that, for every $\gamma < 1/2$ there exists a constant C with*

$$\begin{aligned}\mathbb{E} \sup_{t\in[0,T]} |M(t) - \tilde{M}(t)|^p \leqslant C\big(\mathbb{E}g(T)^{2p}\big)^{1/4}\big(\mathbb{E} \sup_{t\in[0,T]} |f(t) - g(t)|^p\big)^\gamma \\ + C\mathbb{E} \sup_{t\in[0,T]} |f(t) - g(t)|^{p/2}.\end{aligned}$$

Proof. Define the adapted increasing process h by

$$h(t) = f(t) + \inf_{s\leqslant t}\big(g(s) - f(s)\big).$$

note that one has $h \leqslant g$ almost surely. Furthermore, one has

$$0 \leqslant h(t) - h(s) \leqslant f(t) - f(s)$$

for every $t \geqslant s$, so that one has $0 \leqslant (\mathrm{d}h/\mathrm{d}f) \leqslant 1$ almost surely. Define a martingale $\hat{M}(t)$ with quadratic variation h by the Itô integral

$$\hat{M}(t) = \int_0^t \sqrt{\frac{\mathrm{d}h}{\mathrm{d}f}(s)}\,\mathrm{d}M(s).$$

Define now an increasing sequence of random times T_t by

$$T_t = \inf\big\{s \geqslant 0\,|\,h(s) \geqslant g(t)\big\} \geqslant t\ .$$

Note that since $h \leqslant g$ almost surely, the times T_t are actually stopping times with respect to $\mathcal{F}_t$, so that the time-changed process $\tilde{M}(t) = \hat{M}(T_t)$ is a martingale with quadratic variation g. Note that $\tilde{M}(t)$ is a martingale with respect to the filtration $\tilde{\mathcal{F}}_t$ induced by the stopping times T_t. Note also that $\mathcal{F}_t \subset \tilde{\mathcal{F}}_t$ as a consequence of the fact that $T_t \geqslant t$ almost surely.

It remains to show that $\tilde{M}$ satisfies the required bound. Let us start by defining the martingale Δ as the difference $\Delta = M - \hat{M}$. The quadratic variation $\langle \Delta \rangle$ of Δ is then bounded by

$$\begin{aligned}\langle \Delta \rangle(t) &= \int_0^t \left(1 - \sqrt{\frac{\mathrm{d}h}{\mathrm{d}f}(s)}\right)^2 \mathrm{d}f(s) \\ &\leqslant \int_0^t \left(1 - \frac{\mathrm{d}h}{\mathrm{d}f}(s)\right) \mathrm{d}f(s) = f(t) - h(t) \\ &= \sup_{s\leqslant t}\bigl(f(s) - g(s)\bigr) .\end{aligned}$$

Applying the Burkholder–Davies–Gundy inequalities [RY99, corollary IV (4.2)] to this bound yields the existence of a universal constant C such that

$$\mathbb{E} \sup_{t\in[0,T]} |\Delta(t)|^p \leqslant C\mathbb{E} \sup_{t\in[0,T]} |f(t) - g(t)|^{p/2} .$$

Before we turn to bounding the difference between $\hat{M}$ and $\tilde{M}$, we show that if F is an arbitrary positive random variable, B is a Brownian motion, $\gamma < 1/2$, and $q > p > 1$, then there exists a constant C depending only on p, q and γ, such that

$$\mathbb{E}\|B\|_{\gamma,F}^p \leqslant C\bigl(\mathbb{E}F^q\bigr)^{p/2q} , \tag{6.1}$$

where we defined

$$\|B\|_{\gamma,F} = \sup_{0\leqslant s<t\leqslant F} \frac{|B(t) - B(s)|}{|t-s|^\gamma} .$$

One has indeed for every $K > 0$ and every $L > 0$ the bound

$$\mathbb{P}\bigl(\|B\|_{\gamma,F} > K\bigr) \leqslant \mathbb{P}(F \geqslant L) + \mathbb{P}\bigl(\|B\|_{\gamma,L} > K\bigr) .$$

Applying Chebyshev's inequality and using the Brownian scaling together with the fact that the γ-Hölder norm of a Brownian motion on $[0, 1]$ has moments of all orders, this yields for every $q > 0$ the existence of a constant C such that

$$\mathbb{P}\bigl(\|B\|_{\gamma,F} > K\bigr) \leqslant \inf_{L>0}\Bigl(\frac{\mathbb{E}F^q}{L^q} + \frac{\mathbb{E}\|B\|_{\gamma,1}^{2q} L^q}{K^{2q}}\Bigr) \leqslant C\frac{(\mathbb{E}F^q)^{1/2}}{K^q}.$$

The bound (6.1) then follows immediately from the fact that if a positive random variable X satisfies $\mathbb{P}(X > K) \leqslant (a/K)^q$ for some a, some q and every $K > 0$ then, for every $p < q$, there exists a constant C such that $\mathbb{E}|X|^p \leqslant Ca^p$.

Note now that it follows from our construction that there exists a Brownian motion B such that $\hat{M}(t) = B(h(t))$ and $\tilde{M}(t) = B(g(t))$. Noting that $h \leqslant g$ and setting $G = g(T)$, we have

$$\begin{aligned}\mathbb{E} \sup_{t\in[0,T]} |\hat{M}(t) - \tilde{M}(t)|^p &\leqslant \mathbb{E}\|B\|_{\gamma,G}^p \sup_{t\in[0,T]} |h(t) - g(t)|^{\gamma p} \\ &\leqslant \mathbb{E}\|B\|_{\gamma,G}^p \sup_{t\in[0,T]} |f(t) - g(t)|^{\gamma p}\end{aligned}$$

and the result follows from (6.1) and Young's inequality. □

6.2. *Application to the SPDE*

Before we state the next result, we introduce some notation. Let $\gamma \in \mathcal{H}$ and $\Gamma : \mathcal{H}^\alpha \to \mathcal{H}$ be defined by

$$\langle y, \gamma\rangle = 2\langle e_1, B(e_1, L^{-1}Qy)\rangle, \quad \langle y, \Gamma z\rangle = \langle e_1, B(I \otimes_s L)^{-1}(z \otimes Qy)\rangle.$$

The facts that $\gamma \in \mathcal{H}$ and Γ is bounded follow from lemma 3.8 together with the fact that assumption 2.4 implies in particular that Q is a bounded operator from $\mathcal{H}$ to $\mathcal{H}^{\alpha-1}$. Note that Γ is actually bounded as an operator from $\mathcal{H}^{\alpha-1}$ to $\mathcal{H}$, but we will not need this fact.

Theorem 6.2. *Let assumptions 2.2–2.5 hold and let $(x(t), y(t))$ be the solution of (3.1). Let $\tilde{\nu}$, $\tilde{\eta}$ be given by (4.7) and (4.8), respectively, where we identify $\tilde{\nu}$, $\tilde{\eta} \in \mathbb{R}$ and assume that $\tilde{\eta} > 0$.*

Define

$$\sigma_a = \|\gamma\|^2 , \qquad \sigma_b = \mathrm{tr}\big(\Gamma \widehat{Q} \Gamma^*\big) , \tag{6.2}$$

where we identify $\widehat{Q}$ from (4.2) with the corresponding operator[6] from $(\mathcal{H}^\alpha)^*$ to $\mathcal{H}^\alpha$.

Define finally $X(t) = \langle x(t), e_1\rangle$. Then, there exists a Brownian motion B such that, if a is the solution to the SDE

$$\mathrm{d}a(t) = \tilde{\nu}a(t) - \tilde{\eta}a^3(t) + \sqrt{\sigma_b + \sigma_a a^2(t)}\,\mathrm{d}B(t), \qquad a(0) = X(0), \tag{6.3}$$

then, for every $p > 0$ and every $\kappa > 0$, there exists a constant C such that

$$\mathbb{E} \sup_{t\in[0,\tau^*]} |X(t) - a(t)|^p \leqslant C\varepsilon^{p/4-\kappa} ,$$

for every $\varepsilon \in (0, 1)$, where τ^* is defined in (3.2).

Proof. From proposition 4.6 we have that, with the notations introduced above,

$$X(t) = X(0) + \tilde{\nu}\int_0^t X(s)\,\mathrm{d}s - \tilde{\eta}\int_0^t X^3(s)\,\mathrm{d}s + \int_0^t \big\langle \Gamma\hat{z}(s), \mathrm{d}W(s)\big\rangle + \int_0^t X(s)\langle\gamma, \mathrm{d}w(s)\rangle + R_2(t) ,$$

where $R_2 = \mathcal{O}y(\varepsilon^{1/2-\kappa})$. Denote by $M(t)$ the martingale

$$M(t) = \int_0^t \langle\Gamma\hat{z}(s), \mathrm{d}w(s)\rangle\rangle + \int_0^t X(s)\langle\gamma, \mathrm{d}w(s)\rangle .$$

Its quadratic variation is given by

$$f(t) = \int_0^t \|\gamma X(s) + \Gamma\hat{z}(s)\|^2\,\mathrm{d}s . \tag{6.4}$$

It now follows from corollary 4.5 that

$$|f(\cdot) - g(\cdot)| = \mathcal{O}(\varepsilon^{1/2-}) , \quad \text{where} \quad g(t) = \int_0^t \big(\sigma_a X^2(s) + \sigma_b\big)\,\mathrm{d}s . \tag{6.5}$$

Denote by $\tilde{M}(t)$ the martingale with quadratic variation $g(t)$ given by lemma 6.1 and by $\tilde{x}$ the solution to

$$\mathrm{d}\tilde{x} = \tilde{\nu}\tilde{x}\,\mathrm{d}t - \tilde{\eta}\tilde{x}^3\,\mathrm{d}t + \mathrm{d}\tilde{M}(t), \qquad \tilde{x}(0) = x(0).$$

It follows from lemma 6.1 that $M(t) - \tilde{M}(t) = \mathcal{O}(\varepsilon^{1/4-})$. Therefore, using a standard estimate stated below in lemma 6.3,

$$x(t) - \tilde{x}(t) = \mathcal{O}(\varepsilon^{1/4-}) . \tag{6.6}$$

The martingale representation theorem [RY99, theorem V.3.9] ensures that one can enlarge the original probability space in such a way that there exists a filtration $\tilde{\mathcal{F}}_t$, and an $\tilde{\mathcal{F}}_t$-Brownian motion $B(t)$, such that both $x(t)$ and $\tilde{x}(t)$ are $\tilde{\mathcal{F}}_t$-adapted and such that

$$\mathrm{d}\tilde{x}(t) = \tilde{\nu}\tilde{x}(t)\,\mathrm{d}t - \tilde{\eta}\tilde{x}^3(t)\,\mathrm{d}t + \sqrt{\sigma_b + \sigma_a X^2(t)}\,\mathrm{d}B(t).$$

[6] An element $u \otimes v$ of $\mathcal{H}^\alpha \otimes \mathcal{H}^\alpha$ defines an operator from $(\mathcal{H}^\alpha)^*$ to $\mathcal{H}^\alpha$ by $(u \otimes v)(f) = u\langle f, v\rangle$.

Note that in general the σ-algebra $\tilde{\mathcal{F}}_t$ is strictly larger than the one generated by the Wiener process W up to time t. This is a consequence of the construction of lemma 6.1 where one has to 'look into the future' in order to construct $\tilde{M}$.

We finally define the process a as the solution to the SDE

$$\mathrm{d}a(t) = \tilde{\nu}a(t)\,\mathrm{d}t - \tilde{\eta}a^3(t)\,\mathrm{d}t + \sqrt{\sigma_b + \sigma_a a^2(t)}\,\mathrm{d}B(t).$$

Denote $\rho = a - \tilde{x}$ and $G(x) = \sqrt{\sigma_b + \sigma_a x^2}$. Then, one has

$$\mathrm{d}\rho^2(t) \leqslant 2\tilde{\nu}\rho^2(t)\,\mathrm{d}t + \big(G(a(t)) - G(X(t))\big)^2\,\mathrm{d}t + 2\rho\big(G(a(t)) - G(X(t))\big)\,\mathrm{d}B(t).$$

Using the fact that G is globally Lipschitz, this yields the existence of a constant C such that

$$\mathrm{d}\rho^2(t) \leqslant C\rho^2(t)\,\mathrm{d}t + C|X(t) - \tilde{x}(t)|^2\,\mathrm{d}t + 2\rho\big(G(a(t)) - G(X(t))\big)\,\mathrm{d}B(t).$$

It is now easy to verify, using Itô's formula, (6.6), and the Burkholder–Davies–Gundy inequality, that

$$\rho = \mathcal{O}(\varepsilon^{1/4-}) \quad \text{and thus} \quad X(t) - a(t) = \mathcal{O}(\varepsilon^{1/4-})\ ,$$

which is the required result. □

Let us finally state the *a priori* estimate used in the previous proof.

Lemma 6.3. *Fix $\nu \in \mathbb{R}$ and $\eta > 0$. Let $M_i(t)$ be martingales (not necessarily with respect to the same filtration), and $x_i(t)$, $i = 1, 2$ be the solution of the following SDEs:*

$$\mathrm{d}x_i(t) = \nu x_i(t)\,\mathrm{d}t - \eta x_i^3(t)\,\mathrm{d}t + \mathrm{d}M_i(t), \tag{6.7}$$

with $x_1(0) - x_2(0) = \mathcal{O}(\varepsilon^{1/4-})$ and $x_1(0) = \mathcal{O}(\varepsilon^{0-})$. Suppose furthermore $M_i(t) = \mathcal{O}(\varepsilon^{0-})$ and $M_1(t) - M_2(t) = \mathcal{O}(\varepsilon^{1/4-})$, then

$$x_1(t) - x_2(t) = \mathcal{O}(\varepsilon^{1/4-})\ .$$

Proof. This is a straightforward *a priori* estimate which relies on the stable cubic nonlinearity in (6.7). First, one easily sees from Itô's formula that $x_i(t) = \mathcal{O}(\varepsilon^{0-})$. Then using the transformation $\hat{x}_i(t) = x_i(t) - M_i(t)$ to random ODEs for $\hat{x}_i(t)$, we can write down an ODE for the difference $\hat{x}_1(t) - \hat{x}_2(t)$, which we can bound pathwise by direct *a priori* estimates. We will omit the details. □

6.3. Main result

Let us finally put the results obtained in this and the previous two sections together to obtain our final result for the system of SDEs (3.1).

Theorem 6.4. *Let assumptions 2.2–2.5 be true. Let $(x(t), y(t))$ be a solution of (3.1). Furthermore, let $z(t)$ be the OU process defined in (3.4) and τ^* the stopping time from (3.2). Let finally σ_a, σ_b, $\tilde{\nu}$, and $\tilde{\eta}$ be defined in (6.2), (4.7) and (4.8) respectively. Identify $\tilde{\nu}$, $\tilde{\eta} \in \mathbb{R}$ and assume that $\tilde{\eta} > 0$.*

Then there exists a Brownian motion $B(t)$ such that, if $a(t)$ is a solution of

$$\mathrm{d}a(t) = \tilde{\nu}a(t) - \tilde{\eta}a^3(t) + \sqrt{\sigma_b + \sigma_a a^2(t)}\,\mathrm{d}B(t)\ , \quad a(0) = \langle x(0), e_1\rangle, \tag{6.8}$$

then for all $T > 0$, $R > 0$, $p > 0$ and $\kappa > 0$ there is a constant C such that for all $\varepsilon \in (0, 1)$ and all $\|x(0)\|_\alpha < R$ and $\|y(0)\|_\alpha < R$ we have that

$$\mathbb{P}(\tau^* > T) > 1 - C\varepsilon^p\ ,$$

$$\mathbb{E}\sup_{t\in[0,\tau^*]}\|x(t) - a(t)e_1\|_\alpha^p \leqslant C\varepsilon^{p/4-\kappa}\ , \quad \text{and} \quad \mathbb{E}\sup_{t\in[0,\tau^*]}\|y(t) - z(t)\|_\alpha^p \leqslant C\varepsilon^{p-\kappa}\ .$$

Proof. The approximation of $y(t)$ by $z(t)$ is already verified in corollary 3.7 and lemma 3.4. The approximation of $x(t)$ by $a(t)$ follows from theorem 6.2. The bound on the stopping time τ^* follows then easily from the fact that $a(t) = \mathcal{O}(1)$ and $z(t) = \mathcal{O}(1)$ uniformly on $[0, T]$. □

Remark 6.5. With the notation $B_{k\ell m} = \langle B(e_k, e_\ell), e_m\rangle$, formulae (6.2), (4.7), and (4.8) for the coefficients in the amplitude equation can be written in the form:

$$\tilde{\nu} = \nu + \sum_{k=2}^{\infty} \frac{2B_{k11}^2 q_k^2}{\lambda_k^2} + \sum_{k,\ell=2}^{\infty} \frac{B_{k11} B_{\ell\ell k} q_\ell^2}{\lambda_k \lambda_\ell} + \sum_{k,\ell=2}^{\infty} \frac{2B_{k\ell 1} B_{k1\ell}}{\lambda_k + \lambda_\ell} \frac{q_k^2}{\lambda_k}, \tag{6.9a}$$

$$\tilde{\eta} = -\sum_{k=2}^{\infty} \frac{2B_{k11} B_{11k}}{\lambda_k}, \tag{6.9b}$$

$$\sigma_a = \sum_{k=2}^{\infty} \frac{4B_{k11}^2 q_k^2}{\lambda_k^2}, \qquad \sigma_b = \sum_{m,k=2}^{\infty} \frac{2B_{km1}^2 q_k^2 q_m^2}{(\lambda_k + \lambda_m)^2 \lambda_k}. \tag{6.9c}$$

If one chooses to expand the solution in a basis which is not normalised, i.e. one takes basis vectors $\tilde{e}_k = c_k e_k$, then the coefficients appearing in the right-hand side of the equation for the expansion transform according to

$$\tilde{B}_{k\ell m} = B_{k\ell m} \frac{c_k c_\ell}{c_m}, \qquad \tilde{q}_k = \frac{q_k}{c_k}.$$

It is straightforward to see that $\tilde{\nu}$ and σ_a are unchanged under this transformation, whereas $\tilde{\eta}$ is mapped to $c_1^2 \tilde{\eta}$ and σ_b is mapped to σ_b / c_1^2 as expected.

7. Application: the stochastic Burgers equation

In this section we apply our results to a modified stochastic Burgers equation:

$$\mathrm{d}u = (\partial_x^2 + 1)u\,\mathrm{d}t + u\partial_x u\,\mathrm{d}t + \varepsilon^2 \nu u\,\mathrm{d}t + \varepsilon\, Q\,\mathrm{d}w, \tag{7.1}$$

on the interval $[0, \pi]$, with Dirichlet boundary conditions. We take

$$\mathcal{H} = L^2([0,\pi]), \qquad e_k(x) = \sqrt{\frac{2}{\pi}} \sin(kx),$$

$$B(u, v) = \frac{1}{2}\partial_x(uv), \qquad L = -\partial_x^2 - 1, \qquad \lambda_k = k^2 - 1.$$

We also take $W(t)$ to be a cylindrical Wiener process on $\mathcal{H}$ and Q a bounded operator with $Qe_1 = 0$ and $Qe_k = q_k e_k$ for $k \geqslant 2$. It follows that $\mathcal{H}^\alpha = H_0^\alpha([0,\pi])$ is the standard fractional Sobolev space defined by the Dirichlet Laplacian on $[0, \pi]$.

With this choice, using the notation from remark 6.5, we get

$$B_{k\ell m} = \frac{1}{2\sqrt{2\pi}} \big(|k+\ell| \delta_{k+\ell, m} - |k-\ell| \delta_{|k-\ell|, m} \big), \tag{7.2}$$

where $\delta_{k\ell}$ is the Kronecker delta symbol.

It is possible to check that assumptions 2.2, 2.5 and 2.3 are satisfied for any $\alpha \geqslant 0$ since, for smooth functions u, v, and w, one has for example

$$\left| \int_0^\pi (uv)'(x)\, w(x)\,\mathrm{d}x \right| = \left| \int_0^\pi u(x) v(x) w'(x)\,\mathrm{d}x \right| \leqslant \|u\| \|v\| \|w'\|_{\mathrm{L}^\infty}$$
$$\leqslant C \|u\| \|v\| \|w\|_{\mathcal{H}^\gamma},$$

provided that one takes $\gamma < -3/2$. (Values of α other than 0 can be obtained in a similar way by using different Sobolev embeddings, see also [DPDT94].) Whether the trace-class assumption on $Q^2 L^{\alpha-1}$ is satisfied or not depends of course in a crucial way on the coefficients $\{q_k\}_{k=1}^{\infty}$.

The following result justifies the formal asymptotic calculations presented in [Rob03].

Theorem 7.1. *Let u be a continuous $H_0^1([0,\pi])$-valued solution of (7.1) with initial condition $u(0)$ such that $\|u(0)\|_\alpha \leqslant K\varepsilon$ for some K, and assume that the driving noise W is given by $\sigma\sin(2x)w(t)$ for a standard one-dimensional Wiener process w. Then there are Brownian motions $B(t)$ and $\beta(t)$ (not necessarily independent) such that if a is the solution of*

$$\mathrm{d}a = \left(\nu - \frac{\sigma^2}{88}\right) a\,\mathrm{d}t - \frac{1}{12}a^3\,\mathrm{d}t + \frac{\sigma}{6}|a| \circ \mathrm{d}B, \qquad \varepsilon a(0) = \frac{2}{\pi}\big(u(0), \sin(\cdot)\big)_{L^2}$$

and

$$R(t) = \frac{1}{\varepsilon}\mathrm{e}^{-Lt} P_s u(0) + \left(\int_0^t \mathrm{e}^{-3(t-s)}\mathrm{d}\beta(s)\right)\sin(2\cdot),$$

then for all $\kappa, p > 0$ there is a constant C (depending on K but otherwise not on $u(0)$) such that

$$\mathbb{P}\Big(\sup_{t\in[0,T\varepsilon^{-2}]} \|u(t) - \varepsilon a(\varepsilon^2 t)\sin(\cdot) - \varepsilon R(t)\|_{H^1} \leqslant \varepsilon^{\frac{3}{2}-\kappa}\Big) \geqslant 1 - C\varepsilon^p$$

for all $\varepsilon \in (0, 1]$.

Let us remark that in the statement of this theorem, we replaced the expectation over $\sup_{t\in[0,\tau^*]}$ by the probabilities of $\sup_{t\in[0,T\varepsilon^{-2}]}$, which is possible due to the bounds on the stopping time τ^* from theorem 6.4 and the fact that the Burgers equation cannot blow up in finite time.

Proof. Note first that assumption 2.4 is true for all α is this case, so that all the assumptions of theorem 6.4 are satisfied. Furthermore, we can use formulae (6.9) to obtain [7],

$$\tilde\eta = \frac{1}{12}, \qquad \sigma_a = \frac{\sigma^2}{36}\,, \qquad \sigma_b = 0, \qquad \tilde\nu = \nu + \frac{\sigma^2}{72} - \frac{\sigma^2}{88}.$$

Note that the second term in the expression for $\tilde\nu$ gives the Itô–Stratonovich correction.

However, the claim does not follow immediately, since we wish to get an error estimate of order $\varepsilon^{3/2}$ instead of $\varepsilon^{5/4}$. Retracing the proof of theorem 6.4, we see that the claim follows if we can show that $|f - g| = \mathcal{O}(\varepsilon^-)$, where f and g are as in (6.4) and (6.5). In our particular case, one has $\gamma = 0$, so that

$$f(t) = \int_0^t \|\Gamma\hat z(s)\|^2\,\mathrm{d}s, \qquad g(t) = \sigma_b t.$$

The result now follows from lemma 7.2 below. □

Lemma 7.2. *Let $\hat z$ be as in corollary 4.2. Then, for every final time T, every $p > 0$ and every $\kappa > 0$ there exists a constant C such that*

$$\mathbb{E}\sup_{t\in[0,T]}\left\|\int_0^t \big(\hat z(r)\otimes\hat z(r) - \widehat{Q}\big)\,\mathrm{d}r\right\|_\alpha^{2p} \leqslant C\varepsilon^{2p-\kappa}\,. \tag{7.3}$$

[7] Note that a is the amplitude of the mode $\sin(x)$ which is not normalized. This is in order to be consistent with earlier works on the stochastic Burgers equation. The modification of the formulae for the constants that appear in the amplitude equation in this situation is given by remark 6.5.

Proof. We subdivide the interval $[0, T]$ into N subintervals of length T/N, and we use the notation $t_k = kT/N$. Using exactly the same argument as in the proof of lemma 4.3, we see that, for every $p > 0$ and every $\kappa > 0$ there exists a constant C such that

$$\mathbb{E} \sup_{k\in\{0,\ldots,N\}} \left\| \int_0^{t_k} \left(\hat{z}(r) \otimes \hat{z}(r) - \widehat{Q}\right) \mathrm{d}r \right\|_\alpha^{2p} \leqslant CN^\kappa \varepsilon^{2p} .$$

On the other hand, we know from lemma 3.5 that the $\mathcal{H}^\alpha$-norm of the integrand in (7.3) is of order $\mathcal{O}(\varepsilon^{0-})$ uniformly in time. The claim then follows by taking $N \approx \varepsilon^{-1}$. □

Remark 7.3. In the case where only the second mode is forced by noise, one can actually take $\beta(t) = B(t)$, and $\beta(t)$ could be chosen to be a rescaled version of the Brownian motion that appears in equation (7.1).

The following theorem covers the case where $W(t)$ generates a noise that acts on all modes with equal strength except the first one. This corresponds to the case $q_k = \sigma$ for all $k \geqslant 2$. It is easy to check that assumption 2.4 is satisfied for all $\alpha < \frac{1}{2}$.

We again use remark 6.5 with $c_k = \sqrt{\pi/2}$ in order to compute the coefficients. We obtain

$$\tilde{\eta} = \frac{1}{12}, \qquad \sigma_a = \frac{\sigma^2}{18\pi},$$

$$\sigma_b = c_b\sigma^4, \qquad c_b = \frac{1}{2\pi^2}\sum_{k=2}^{\infty} \frac{1}{(2k^2+2k+1)(k^2-1)(k^2+2k)},$$

and finally

$$\tilde{\nu} - \nu = \frac{\sigma^2}{36\pi} - \frac{\sigma^2}{4\pi}\sum_{k=2}^{\infty}\left(\frac{1}{k-1} - \frac{1}{k(k+1)}\right)\frac{1}{2k^2+2k-1} .$$

Note again that the first term in this expression is the Stratonovich correction for the multiplicative noise term. This finally leads to

Theorem 7.4. *Assume that $\alpha \in [0, \frac{1}{2})$ and let u be a continuous $H_0^\alpha([0,\pi])$-valued solution of (7.1) with initial condition $u(0)$ such that $\|u(0)\|_\alpha \leqslant K\varepsilon$ for some K. Assume furthermore that the covariance of the noise satisfies $q_k = \sigma$ for $k \geqslant 2$. Then there is a Brownian motion $B(t)$ such that if $a(t)$ is a solution of*

$$\mathrm{d}a(t) = \tilde{\nu}a(t) - \tilde{\eta}a^3(t) + \sqrt{\sigma_b + \sigma_a a^2(t)}\,\mathrm{d}B(t), \qquad \varepsilon a(0) = \frac{2}{\pi}\big(u(0), \sin(\cdot)\big)_{L^2}$$

where the constants are defined above, and

$$R(t) = \frac{1}{\varepsilon}\mathrm{e}^{-tL}P_s u(0) + \int_0^t \mathrm{e}^{-(t-s)L} Q\,\mathrm{d}W(s),$$

then for all $\kappa, p > 0$ there is a constant C (again depending on K but otherwise not on $u(0)$) such that

$$\mathbb{P}\Big(\sup_{t\in[0,T\varepsilon^{-2}]} \|u(t) - \varepsilon a(\varepsilon^2 t)\sin(\cdot) - \varepsilon R(t)\|_{H_0^\alpha} \leqslant \varepsilon^{\frac{5}{4}-\kappa}\Big) \geqslant 1 - C\varepsilon^p$$

for all $\varepsilon \in (0, 1]$.

Acknowledgments

The authors are grateful to the suggestions of an anonymous referee which greatly improved this paper. MH was supported by an EPSRC Fellowship EP/DO71593/10.

References

[BGR02] Blömker D, Gugg C and Raible M 2002 Thin-film-growth models: roughness and correlation functions *Eur. J. Appl. Math.* **13** 385–402

[Blö05] Blömker D 2005 Approximation of the stochastic Rayleigh–Bénard problem near the onset of convection and related problems *Stoch. Dyn.* **5** 441–74

[BLP78] Bensoussan A, Lions J and Papanicolaou G 1978 *Asymptotic Analysis for Periodic Structures* (North-Holland: Amsterdam)

[CB95] Cuerno Z and Barabási A-L 1995 Dynamic scaling of ion-sputtered surfaces *Phys. Rev. Lett.* **74** 4746–9

[CH93] Cross M and Hohenberg P 1993 Pattern formation outside of equilibrium *Rev. Mod. Phys.* **65** 851–1112

[CY95] Chekhlov A and Yakhot V 1995 Kolmogorov turbulence in a random-force-driven Burgers equation: anomalous scaling and probability density functions *Phys. Rev.* E **52** 5681–4

[DPDT94] Da Prato G, Debussche A and Temam R 1994 Stochastic Burgers' equation *Nonlin. Diff. Eqns. Appl.* **1** 389–402

[DPZ92] Da Prato G and Zabczyk J 1992 Stochastic equations in infinite dimensions *Encyclopedia of Mathematics and its Applications* vol 44 (Cambridge: Cambridge University Press)

[EK86] Ethier S and Kurtz T 1986 *Markov Processes (Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics)* (New York: Wiley)

[HP04] Hairer M and Pavliotis G 2004 Periodic homogenization for hypoelliptic diffusions *J. Stat. Phys.* **117** 261–79

[Kur73] Kurtz T 1973 A limit theorem for perturbed operator semigroups with applications to random evolutions *J. Funct. Anal.* **12** 55–67

[LCM96] Lauritsen K, Cuerno R and Makse H 1996 Noisy Kuramoto–Sivashinsky equation for an erosion model *Phys. Rev.* E **54** 3577–80

[MTVE01] Majda A, Timofeyev I and Vanden Eijnden E 2001 A mathematical framework for stochastic climate models *Commun. Pure Appl. Math.* **54** 891–974

[MTVE99] Majda A J, Timofeyev I and Vanden Eijnden E 1999 Models for stochastic climate prediction *Proc. Natl Acad. Sci. USA* **96** 14687–91

[PSV77] Papanicolaou G C, Stroock D and Varadhan S R S 1977 Martingale approach to some limit theorems *Papers from the Duke Turbulence Conference (Duke University, Durham, NC, USA, 1976) Paper No 6 (Duke Univ. Math. Ser.) vol III* (Durham, NC: Duke University)

[RML+00] Raible M, Mayr S, Linz S, Moske M, Hänggi P and Samwer K 2000 Amorphous thin film growth: theory compared with experiment *Europhys. Lett.* **50** 61–7

[Rob03] Roberts A J 2003 A step towards holistic discretisation of stochastc partial differential equations *ANZIAM J.* **45** C1–15

[RY99] Revuz D and Yor M 1999 Continuous martingales and Brownian motion *Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences)* vol 293, 3rd edn (Berlin: Springer)

[Sch94] Schneider G 1994 A new estimate for the Ginzburg–Landau approximation on the real axis *J. Nonlin. Sci.* **4** 23–34

[Sch01] Schneider G 2001 Bifurcation theory for dissipative systems on unbounded cylindrical domains—an introduction to the mathematical theory of modulation equations *Z. Angew. Math. Mech.* **81** 507–22

[SV06] Stroock D and Varadhan S 2006 *Multidimensional Diffusion Processes* (*Classics in Mathematics*) (Berlin: Springer)