A Theory of Hypoellipticity and Unique Ergodicity for Semilinear Stochastic PDEs

Martin Hairer[1][*] **Jonathan C. Mattingly**[2][†]

[1]Mathematics Institute, The University of Warwick, CV4 7AL, UK
Email: `M.Hairer@Warwick.ac.uk`
[2]Mathematics Department, Center of Theoretical and Mathematical Science,
and Center of Nonlinear and Complex Systems, Duke University, Durham NC,
USA
Email: `jonm@math.duke.edu`

Abstract

We present a theory of hypoellipticity and unique ergodicity for semilinear parabolic stochastic PDEs with "polynomial" nonlinearities and additive noise, considered as abstract evolution equations in some Hilbert space. It is shown that if Hörmander's bracket condition holds at every point of this Hilbert space, then a lower bound on the Malliavin covariance operator $\mathcal{M}_t$ can be obtained. Informally, this bound can be read as "Fix any finite-dimensional projection Π on a subspace of sufficiently regular functions. Then the eigenfunctions of $\mathcal{M}_t$ with small eigenvalues have only a very small component in the image of Π."

We also show how to use a priori bounds on the solutions to the equation to obtain good control on the dependency of the bounds on the Malliavin matrix on the initial condition. These bounds are sufficient in many cases to obtain the asymptotic strong Feller property introduced in [HM06].

One of the main novel technical tools is an almost sure bound from below on the size of "Wiener polynomials," where the coefficients are possibly non-adapted stochastic processes satisfying a Lipschitz condition. By exploiting the polynomial structure of the equations, this result can be used to replace Norris' lemma, which is unavailable in the present context.

We conclude by showing that the two-dimensional stochastic Navier-Stokes equations and a large class of reaction-diffusion equations fit the framework of our theory.

Contents

1 Introduction

The overarching goal of this article is to prove the unique ergodicity of a class of nonlinear stochastic partial differential equations (SPDEs) driven by a finite number of Wiener processes. The present greatly extends the articles [MP06, HM06, BM07] allowing one to consider general polynomial nonlinearities and more general forcing. To the best of our knowledge, this is the first infinite-dimensional generalization of Hörmander's "sum of squares" hypoellipticity theorem for a general class of parabolic SPDEs. Our goal is not to present any particularly compelling examples from the applied perspective, but rather give a sufficiently general framework which can be applied in many settings. At the end, we do give some examples to serve as roadmaps for the application of the results in this article. In this section, we give an overview of the setting and the results to come later without descending into all of the technical assumptions required to make everything precise. This imprecision will be rectified starting with Section 3 where the setting and basic assumptions will be detailed.

In this article we will investigate non-linear equations of the form

$$\partial_t u(x,t) + Lu(x,t) = N(u)(x,t) + \sum_{k=1}^{d} g_k(x)\dot{W}_k(t)\,. \qquad (1.1)$$

Here L will be some positive selfadjoint operator. Typical examples arising in applications are $L = -\Delta$ or $L = \Delta^2$. N will be assumed to be a "polynomial" nonlinearity in the sense that $N(u) = \sum_{k=1}^{m} N_k(u)$, where N_k is k-multilinear. Examples of admissible nonlinearities are the Navier-Stokes nonlinearity $(u \cdot \nabla)u$ or a reaction term such as $u - u^3$. The g_k are a collection of smooth, time independent functions which dictate the "directions" in which the randomness is injected. The $\{\dot{W}_k : k = 1, \ldots, d\}$ are a collection of mutually independent one-dimensional white noises which are understood as the formal derivatives of independent Wiener processes through the Itô calculus. We assume that the possible loss of regularity due to the nonlinearity is controlled by the smoothing properties of the semigroup generated by L. See Assumption A.1 below for a precise meaning.

On one hand, our concentration on a finite number of driving Wiener processes avoids technical difficulties generated by spatially rough solutions since $W(x,t) = \sum g_k(x)W_k(t)$ has the same regularity in x as the g_k which we take to be relatively smooth. On the other hand, the fact that W contains only a finite number of Wiener processes means that our dynamic is very far from being uniformly elliptic in any sense since for fixed t, $u(\,\cdot\,, t)$ is an infinite-dimensional random variable and the noise acts only onto a finite number of degrees of freedom. To prove an ergodic theorem, we must understand how the randomness injected by W in the directions $\{g_k : k = 1, \ldots, d\}$ spreads through the infinite dimensional phase space. To do this, we prove the non-degeneracy of the Malliavin covariance matrix under an assumption that the linear span of the successive Lie brackets[1] of vector fields associated to N and the g_k is dense in the ambient (Hilbert) space at each point. This is very reminiscent of the condition in the "weak" version of Hörmander's "sum of squares" theorem. It ensures that the randomness spreads to a dense set of direction despite being injected in only a finite number of directions. This is possible since although the randomness is injected in a finite number of directions it is injected over the entire interval of time from zero to the current time. The conditions which ensure the spread of randomness is closely related to Chow's theorem and controllability, open or solid. As such Section 6 is related to recent work on controllability of projections of PDEs studied in [AS05, AS08] and results proving the existence of densities for projections in [AKSS07] which build on these ideas. However, these results do not seem to be sufficient to prove an ergodic result which is the principal aim of this work. One seems to need quantitative control of the spectrum of the Malliavin matrix (or the Gramian matrix in control theory terms).

In finite dimensions, bounds on the norm of the inverse of the Malliavin matrix are the critical ingredient in proving ergodic theorems for diffusions which are only hypoelliptic rather than uniformly elliptic. This then shows that the system has a smooth density with respect to Lebesgue measure. In infinite dimensions, there is no measure which plays the "universal" role of Lebesgue measure. One must therefore pass through a different set of ideas. Furthermore, it is not so obvious how to generalise the notion of the 'inverse' of the Malliavin matrix. In finite dimension, a linear map has dense range if and only if it admits a bounded inverse. In infinite dimensions, these two

---

[1]Recall that, when it is defined, the Lie bracket $[G, H](u) = (\mathrm{D}H)(u)G(u) - (\mathrm{D}G)(u)H(u)$ for two functions G, H from the ambient Hilbert space $\mathcal{H}$ to itself. Here D is the Fréchet derivative.

notions are very far from equivalent and, while it is possible in some cases to show that the Malliavin matrix has dense range, it is hardly ever possible in a hypoelliptic setting to show that it is invertible, or at least to characterise its range in a satisfactory manner (See [MSVE07] for a linear example in which it is possible).

The important fact which must be established is that nearby points act similarly from a "measure theoretic perspective." One classical way to make this precise is to prove that the Markov process in question has the strong Feller property. For a continuous time Markov process this is equivalent to proving that the transition probabilities are continuous in the total variation norm. While this concept is useful in finite dimensions, it is much less useful in infinite dimensions. In particular, there are many natural infinite dimensional Markov processes whose transition probabilities do *not* converge in total variation to the system's unique invariant measure. (See examples 3.14 and 3.15 from [HM06] for more discussion of this point.) In these settings, this fact also precludes the use of "minorization" conditions such as $\inf_{x \in \mathcal{C}} \mathcal{P}_t(x, \cdot) \geq c\nu(\cdot)$ for some fixed probability measure ν and "small set" $\mathcal{C}$. (see [MT93, GM06] for more and examples were this can be used.)

1.1   Ergodicity in infinite dimensions and main result

In [HM06], the authors introduced the concept of an *Asymptotic Strong Feller* diffusion. Loosely speaking, it ensures that transition probabilities are uniformly continuous in a sequence of 1-Wasserstein distances which converge to the total variation distance as time progresses. For the precise definitions, we refer the reader to [HM06]. For our present purpose, it is sufficient to recall the following proposition:

Proposition 1.1 (Proposition 3.12 from [HM06]) *Let t_n and δ_n be two positive sequences with $\{t_n\}$ non-decreasing and $\{\delta_n\}$ converging to zero. A semigroup $\mathcal{P}_t$ on a Hilbert space $\mathcal{H}$ is asymptotically strong Feller if, for all $\varphi : \mathcal{H} \to \mathbf{R}$ with $\|\varphi\|_\infty$ and $\|\mathrm{D}\varphi\|_\infty$ finite one has*

$$\|\mathrm{D}\mathcal{P}_{t_n}\varphi(u)\| \leq C(\|u\|)(\|\varphi\|_\infty + \delta_n\|\mathrm{D}\varphi\|_\infty) \tag{1.2}$$

for all n and $u \in \mathcal{H}$, where $C : \mathbf{R}_+ \to \mathbf{R}$ is a fixed non-decreasing function.

The importance of the asymptotic strong Feller property is given by the following result which states that in this case, any two distinct ergodic invariant measures must have disjoint topological supports. Recalling that u belongs to the support of a measure μ (denoted supp(μ)) if $\mu(B_\delta(u)) > 0$ for every $\delta > 0$[2], we have:

Theorem 1.2 (Theorem 3.16 from [HM06]) *Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space $\mathcal{X}$ admitting two distinct ergodic invariant measures μ and ν. If $\mathcal{P}_t$ has the asymptotic strong Feller property, then* supp(μ) $\cap$ supp(ν) *is empty.*

To better understand how the asymptotic strong Feller property can be used to connect topological properties and ergodic properties of $\mathcal{P}_t$, we introduce the following form of topological irreducibility.

Definition 1.3 We say that a Markov semigroup $\mathcal{P}_t$ is *weakly topologically irreducible* if for all $u_1, u_2 \in \mathcal{H}$ there exists a $v \in \mathcal{H}$ so that for any A open set containing v there exists $t_1, t_2 > 0$ with $\mathcal{P}_{t_i}(u_i, A) > 0$.

---

[2]Here $B_\delta(u) = \{v : \|u - v\| < \delta\}$

Also recall that $\mathcal{P}_t$ is said to be *Feller* if $\mathcal{P}_t\varphi$ is continuous whenever φ is bounded and continuous. We then have the following corollary to Theorem 1.2 whose proof is given in Section 2.

Corollary 1.4 *Any Markov semigroup $\mathcal{P}_t$ on Polish space which is Feller, weakly topologically irreducible and asymptotically strong Feller admits at most one invariant probability measure.*

The discussion of this section shows that unique ergodicity can be obtained for a Markov semigroup by showing that:

1. It satisfies the asymptotic strong Feller property.
2. There exists an "accessible point" which must belong to the topological support of every invariant probability measure.

It turns out that if one furthermore has some control on the speed at which solution return to bounded regions of phase space, one can prove the existence of spectral gaps in weighted Wasserstein-1 metrics [HM08, HMS10, HM10].

The present article will mainly concentrate on the first point. This is because, by analogy with the finite-dimensional case, one can hope to find a clean and easy way to verify condition along the lines of Hörmander's bracket condition that ensures a regularisation property like the asymptotic strong Feller property. Concerning the accessibility of points however, although one can usually use the Stroock-Varadhan support theorem to translate this into a deterministic question of approximate controllability, it can be a very hard problem even in finite dimensions. While geometric control theory can give qualitative information about the set of reachable points [Jur97, AS04], the verification of the existence of accessible points seems to rely in general on *ad hoc* considerations, even in apparently simple finite-dimensional problems. We will however verify in Section 8.4 below that for the stochastic Ginzburg-Landau equation there exist accessible points under very weak conditions on the forcing.

With this in mind, the aim of this article is to prove the following type of 'meta-theorem':

Meta-Theorem 1.5 *Consider the setting of (1.1) on some Hilbert space $\mathcal{H}$ and define a sequence of subsets of $\mathcal{H}$ recursively by $\mathrm{A}_0 = \{g_j \,:\, j = 1, \ldots, d\}$ and*

$$\mathrm{A}_{k+1} = \mathrm{A}_k \cup \{N_m(h_1, \ldots, h_m) \,:\, h_j \in \mathrm{A}_k\}\,.$$

Under additional stability and regularity assumptions, if the linear span of $\mathrm{A}_\infty \overset{\text{def}}{=} \bigcup_{n>0} \mathrm{A}_n$ is dense in $\mathcal{H}$, then the Markov semigroup $\mathcal{P}_t$ associated to (1.1) has the asymptotic strong Feller property.

The precise formulation of Meta-Theorem 1.5 will be given in Theorem 8.1 below, which in turn will be a consequence of the more general results given in Theorems 5.5 and 6.7. Note that our general results are slightly stronger than what is suggested in Meta-Theorem 1.5 since it also allows to consider arbitrary "non-constant" Lie brackets between the driving noises and the drift, see (1.4) below. As further discussed in Section 1.5 or 3.1, $N_m(h_1, \ldots, h_m)$ is proportional to $\mathrm{D}_{h_1} \cdots \mathrm{D}_{h_m} N(u)$ where D_h is the Fréchet derivative in the direction h. In turn, this is equal to the successive Lie-brackets of N with the constant vector fields in the directions h_1 to h_m.

Under the same structural assumptions as Meta-Theorem 1.5, the existence of densities for the finite dimensional projections of $\mathcal{P}_t(x, \cdot)$ was proven in [BM07]. The

smoothness of these densities was also discussed in [BM07], but unfortunately there were two errors in the proof of that article. While the arguments presented in the present article are close in sprit to those in [BM07], they diverge at the technical level. Our results on the smoothness of densities will be given in Sections 6 and 7.

The remainder of this section is devoted to a short discussion of the main techniques used in the proof of such a result and in particular on how to obtain a bound of the type 1.2 for a parabolic stochastic PDE.

1.2   A roadmap for the impatient

Readers eager to get to the heart of this article but understandably reluctant to dig into too many technicalities may want finish reading Section 1, then jump directly to Section 5 and read up to the end of Section 5.3 to get a good idea of how (1.2) is obtained from bounds on the Malliavin matrix. Then they may want to go to the beginning of Section 6 and read to the end of Section 6.4 to see how these bounds can be obtained.

1.3   How to obtain a smoothing estimate

A more technical overview of the techniques will be given in Section 5.2 below. In a nutshell, our aim is to generalise the arguments from [HM06] and the type of Malliavin calculus estimates developed in [MP06] to a large class of semilinear parabolic SPDEs with polynomial nonlinearities. Both previous works relied on the particular structure of the Navier-Stokes equations. The technique of proof can be interpreted as an "infinitesimal" version of techniques developed in [EMS01, KS00] and extended in [BKL01, MY02, Mat02, Hai02, BM05] combined with detailed lower bounds on the Malliavin covariance matrix of the solution.

In [EMS01] the idea was the following: take two distinct initial conditions u_0 and u_0' for (1.1) and a realisation W for the driving noise. Try then to find a shift v belonging to the Cameron-Martin space of the driving process and such that $\|u(t) - u'(t)\| \to 0$ as $t \to \infty$, where u' is the solution to (1.1) driven by the shifted noise $W' = W + v$. Girsanov's theorem then ensures that the two initial conditions induce equivalent measures on the infinite future. This in turn implies the unique ergodicity of the system. (See also [Mat08] for more details.)

The idea advocated in [HM06] is to consider an infinitesimal version of this construction. Fix again an initial condition u_0 and Wiener trajectory W but consider now an *infinitesimal* perturbation ξ to the initial condition instead of considering a second initial condition at distance $\mathcal{O}(1)$. This produces an infinitesimal variation in the solution u_t given by its Fréchet derivative $\mathrm{D}_\xi u_t$ with respect to u_0. Similarly to before, one can then consider the "control problem" of finding an *infinitesimal* variation of the Wiener process in a direction h from the Cameron-Martin space which, for large times t, compensates the effect of the variation ξ. Since the effect on u_t of an infinitesimal variation in the Wiener process is given by the Malliavin derivative of u_t in the direction h, denoted by $\mathcal{D}_h u_t$, the problem in this setting is to find an $h(\xi, W) \in L^2([0, \infty], \mathbf{R}^d)$ with

$$\mathbf{E}\|\mathrm{D}_\xi u_t - \mathcal{D}_h u_t\| \to 0 \text{ as } t \to \infty \,, \tag{1.3}$$

and such that the expected "cost" of h_t is finite. Here, the Malliavin derivative $\mathcal{D}_h u_t$ is given by the derivative in ε at $\varepsilon = 0$ of $u_t(W + \varepsilon v)$, with $v(t) = \int_0^t h(s)\,ds$. If h is adapted to the filtration generated by W, then the expected cost is simply $\int_0^\infty \mathbf{E}\|h_s\|^2 ds$. If it is not adapted, one must estimate directly $\limsup \mathbf{E}\| \int_0^t h_s dW_s\|$ where the integral is a Skorokhod integral. As will be explained in detail in Section 5.2,

once one establishes (1.3) with a finite expected cost h, the crucial estimate given in (1.2) (used to prove the asymptotic strong Feller property) follows by a fairly general procedure.

As this discussion makes clear, one of our main tasks will be to construct a shift h having the property (1.3). We will distinguish three cases of increasing generality (and technical difficulty). In the first case, which will be referred to as *strongly contracting* (see Section 5.1.1), the linearised dynamics contracts pathwise without modification (all Lyapunov exponents are negative). Hence h can be taken to be identically zero. The next level of complication comes when the system possesses a number of directions which are unstable on average. The simplest way to deal with this assumption is to assume that the complement of the span of the forced directions (the g_k's) is contracting on average. This was the case in [EMS01, KS00, BKL01, MY02, Mat02, Hai02, BM05]. We refer to this as the "essentially elliptic" setting since the directions essential to determine the system's long time behavior, the unstable directions, are directly forced. This is a reflection of the maxim in dynamical systems that the long time behavior is determined by the behavior in the unstable directions. Since the noise affects all of these directions, it is not surprising that the system is uniquely ergodic, see Section 4.5 of [HM06] for more details.

The last case (i.e. when the set of forced directions does not contain all of the unstable directions) is the main concern of the present paper. In this setting, we study the interaction between the drift and the forced directions to understand precisely how randomness spreads to the system. The condition ensuring that one can gain sufficient control over the unstable directions, requires that the g_k together with a collection of Lie brackets (or commutators) of the form

$$[F, g_k], \; [[F, g_k], g_j], \; [[F, g_k], F], \; [[[F, g_k], g_j], g_l], \cdots \tag{1.4}$$

span all of the unstable direction. This condition will be described more precisely in Section 6.2 below. In finite dimensions, when this collection of Lie brackets spans the entire tangent space at every point, the system is said to satisfy the "weak Hörmander" condition. When this assumption holds for the unstable directions (along with some additional technical assumptions), we can ensure that the noise spreads sufficiently to the unstable directions to find a h capable of counteracting the expansion in the unstable directions and allowing one to prove (1.3) with a cost whose expectation is finite.

We will see however that the control h used will not be adapted to the filtration generated by the increments of the driving Wiener process, thus causing a number of technical difficulties. This stems from the seemingly fundamental fact that because we need some of the "bracketed directions" (1.4) in order to control the dynamic, we need to work on a time scale longer than the instantaneous one. In the "essentially elliptic" setting, on the other hand, we were able to work instantaneously and hence obtain an adapted control h and avoid this technicality.

1.4 The role of the Malliavin matrix

Since the Malliavin calculus was developed in the 1970's and 1980's mainly to give a probabilistic proof of Hörmander's "sum of squares" theorem under the type of bracket conditions we consider, it is not surprising that the Malliavin matrix $\mathcal{M}_t = \mathcal{D}u_t \mathcal{D}u_t^*$ plays a major role in the construction of the variation h in the "weak Hörmander" setting. A rapid introduction to Malliavin calculus in our setting is given in Section 4. In finite dimensions, the key to the proof of existence and smoothness of densities is the finiteness of moments of the inverse of the Malliavin matrix. This estimate encapsulates the fact the noise effects all of the directions with a controllable cost. In infinite

dimensions while it is possible to prove that the Malliavin matrix is almost surely non-degenerate it seems very difficult to characterise its range. (With the exception of the linear case [DPZ96]. See also [DPEZ95, FM95, Cer99, EH01] for situations where the Malliavin matrix can be shown to be invertible on the range of the Jacobian.) However, in light of the preceding section, it is not surprising that we essentially only need the invertibility of the Malliavin matrix on the space spanned by the unstable directions, which is finite dimensional in all of our examples. More precisely, we need information about the likelihood of eigenvectors with sizable projections in the unstable directions to have small eigenvalues. Given a projection Π whose range includes the unstable directions we will show that the Malliavin matrix $\mathcal{M}_t$ satisfies an estimate of the form

$$\mathbf{P}\Big(\inf_{\substack{\varphi \in \mathcal{H} \\ \|\Pi\varphi\| \geq \frac{1}{2}\|\varphi\|}} \langle \mathcal{M}_t\varphi, \varphi \rangle > \varepsilon\|\varphi\|^2\Big) = o(\varepsilon^p) \qquad (1.5)$$

for all $p \geq 1$. Heuristically, this means we have control of the probabilistic cost to create motion in all of the directions in the range of Π without causing a too large effect in the complementary directions. We will pair such an estimate with the assumption that the remaining directions are stable in that the Jacobian (the linearization of the SPDE about the trajectory u_t) satisfies a contractive estimate for the directions perpendicular to the range of Π. Together, these assumptions will let us build an infinitesimal Wiener shift h which approximately compensates for the component of the infinitesimal shift caused by the variation in the initial condition in the unstable directions. Once the variation in the unstable directions have been decreased, the assumed contraction in the stable directions will ensure that the variation in the stable directions will also decrease until it is commiserate in size with the remaining variation in the unstable directions. Iterating this argument we can drive the variation to zero.

Note that one feature of the bound (1.5) is that all the norms and scalar products appearing there are the same. This is a strengthening of the result from [MP06] which fixes an error in [HM06], see Section 6 for more details.

The basic structure of the sections on Malliavin calculus follows the presentation in [BM07] which built on the ideas and techniques from [MP06, Oco88]. As in all three of these works, as well as the present article, the time reversed adjoint linearization is used to develop an alternative representation of the Malliavin Covariance matrix. In [Oco88], only the case of linear drift and linear multiplicative noise was considered. In [MP06], a nonlinear equation with a quadratic nonlinearity and additive noise was considered. In [BM07], the structure was abstracted and generalized so that it was amenable to general polynomial nonlinearities. We follow that structure and basic line of argument here while strengthening the estimates and correcting some important errors.

Most existing bounds on the inverse of the Malliavin matrix in a hypoelliptic situation make use of some version of Norris' lemma [KS84, KS85a, Nor86, MP06, BH07]. In its form taken from [Nor86], it states that if a semimartingale $Z(t)$ is small and one has some control on the roughness of both its bounded variation part $A(t)$ and its quadratic variation process $Q(t)$, then both A and Q taken separately must be small. While the versions of Norris' lemma given in [MP06, BM07, BH07] are not precisely of this form (in both cases, one cannot reduce the problem to semimartingales, either because of the infinite-dimensionality of the problem or because one considers SDEs driven by processes that are not Wiener processes), they have the same flavour in that they state that if a process is composed of a "regular" part and an "irregular" part, then these two parts cannot cancel each other. This harkens back to the more explicit

estimates based on estimates of modulus of continuity found in [KS85b, Str83]. The replacement for Norris' lemma used in the present work covers the case where one is given a finite collection of Wiener process W_j and a collection of not necessarily adapted Lipschitz continuous processes $A_\alpha(t)$ (for α a multi-index) and considers the process

$$Z(t) = A_\phi(t) + \sum_{\ell=1}^{M} \sum_{|\alpha|=\ell} A_\alpha(t) W_{\alpha_1}(t) \cdots W_{\alpha_\ell}(t) \ .$$

It then states that if Z is small, this implies that all of the A_α's with $|\alpha| \leq M$ are small. For a precise formulation, see Section 7 below. It is in order to be able to use this result that we are restricted to equations with polynomial nonlinearities. This result on Wiener polynomials is a descendant of the result proven in [MP06] for polynomials of degree one. In [BM07], a result for general Wiener polynomials was also proven. Is was show there that if $Z(t) = 0$ for $t \in [0, T]$ then then $A_\alpha(t) = 0$ for $t \in [0, T]$. This was used to prove the existence of a density for the finite dimensional projections of the transition semigroup. In the same article, the same quantitative version of this result as proven in the present article was claimed. Unfortunately, there was a error in the proof. Nonetheless the techniques used here are built on and refine those developed in [BM07].

1.5 Satisfying the Hörmander-like assumption

At first glance the condition that the collection of functions given in (1.4) are dense in our state space may seem hopelessly strong. However, we will see that it is often not difficult to ensure. Recall that the nonlinearity N is a polynomial of order m, and hence, it has a leading order part which is m-homogeneous. We can view this leading order part as a symmetric m-linear map which we will denote by N_m. Then, at least formally, the Lie bracket of N with m constant vector fields is proportional to N_m, evaluated at the constant vector fields, that is $N_m(h_1, \cdots, h_m) \propto [\cdots [[F, h_1], \cdots], h_m]$, which is again a constant vector field. While the collection of vector fields generated by brackets of this form are only a subset of the possible brackets, it is often sufficient to obtain a set of dense vector fields. For example, if $N(u) = u - u^3$ then $N_3(v_1, v_2, v_3) = v_1 v_2 v_3$ and if the forced directions $\{g_1, \cdots, g_d\}$ are $\mathcal{C}^\infty$ then $N_3(h_1, h_2, h_3) \in \mathcal{C}^\infty$ for $h_i \in \{g_1, \cdots, g_d\}$. As observed in [BM07], to obtain a simple sufficient criteria for the brackets to be dense, suppose that $\Lambda \subset \mathcal{C}^\infty$ is a finite set of functions that generates, as a multiplicative algebra, a dense subset of the phase space. Then, if the forced modes $A_0 = \{g_1, \cdots, g_d\}$ contain the set $\{h, h\bar{h} : h, \bar{h} \in \Lambda\}$, the set A_∞ constructed as in Meta-Theorem 1.5 will span a dense subset of phase space.

1.6 Probabilistic and dynamical view of smoothing

Implicit in (1.3) is the "transfer of variation" from the initial condition to the Wiener path. This is the heart of "probabilistic smoothing" and the source of ergodicity when it is fundamentally probabilistic in nature. The unique ergodicity of a dynamical system is equivalent to the fact that it "forgets its initial condition" with time. The two terms appearing on the right-hand side of (1.2) represent two different sources of this loss of memory. The first is due to the randomness entering the system. This causes nearby points to end up at the same point at a later time because they are following different noise realisations. The fact that different stochastic trajectories can arrive at the same point and hence lead to a loss of information is the hallmark of diffusions and unique

ergodicity due to randomness. From the coupling point of view, since different realizations lead to the same point yet start at different initial conditions, one can couple in finite time.

The second term in (1.2) is due to "dynamical smoothing" and is one of the sources of unique ergodicity in deterministic contractive dynamical systems. If two trajectories converge towards each other over time then the level of precision needed to determine which initial condition corresponds to which trajectory also increases with time. This is another type of information loss and equally leads to unique ergodicity. However, unlike "probabilistic smoothing", the information loss is never complete at any finite time. Another manifestation of this fact is that the systems never couples in finite time, only at infinity. In Section 5.1.1 about the strongly dissipative setting, the case of pure dynamical smoothing is considered. In this case one has (1.2) with only the second term present. When both terms exist, one has a mixture of probabilistic and dynamical smoothing leading to a loss of information about the initial condition. In Section 2.2 of [HM08] it is shown how (1.2) can be used to construct a coupling in which nearby initial conditions converge to each other at time infinity. The current article takes a "forward in time" perspective, while [EMS01, BM05] pull the initial condition back to minus infinity. The two points of view are essentially equivalent. One advantage to moving forward in time is that it makes proving a spectral gap for the dynamic more natural. We provide such an estimate in Section 8.4 for the stochastic Ginzburg-Landau equation.

1.7 Structure of the article

The structure of this article is as follows. In Section 2, we give a few abstract ergodic results both proving the results in the introduction and expanding upon them. In Section 3, we introduce the functional analytic setup in which our problem will be formulated. This setup is based on Assumption A.1 which ensures that all the operations that will be made later (differentiation with respect to initial condition, representation for the Malliavin derivative, etc) are well-behaved. Section 4 is a follow-up section where we define the Malliavin matrix and obtain some simple upper bounds on it. We then introduce some additional assumptions in Section 6.1 which ensure that we have suitable control on the size of the solutions and on the growth rate of its Jacobian.

In Section 5, we obtain the asymptotic strong Feller property under a partial invertibility assumption on the Malliavin matrix and some additional partial contractivity assumptions on the Jacobian. Section 6.3 then contains the proof that assumptions on the Malliavin matrix made in Section 5 are justified and can be verified for a large class of equations under a Hörmander-type condition. The main ingredient of this proof, a lower bound on Wiener polynomials, is proved in Section 7. Finally, we conclude in Section 8 with two examples for which our conditions can be verified. We consider the Navier-Stokes equations on the two-dimensional sphere and a general reaction-diffusion equation in three or less dimensions.

Acknowledgements

2 Abstract ergodic results

We now expand upon the abstract ergodic theorems mentioned in the introduction which build on the asymptotic strong Feller property. We begin by giving the proof

of Corollary 1.4 from the introduction and then give a slightly different result (but with the same flavour) which will be useful in the investigation of the Ginzburg-Landau equation in Section 8.4. Throughout this section, $\mathcal{P}_t$ will be a Markov semigroup on a Hilbert space $\mathcal{H}$ with norm $\| \cdot \|$.

Proof of Corollary 1.4. Since $\mathcal{P}_t$ is Feller, we know that for any $u \in \mathcal{H}$ and open set A with $\mathcal{P}_t(u, A) > 0$ there exists an open set B containing u so that

$$\inf_{u \in B} \mathcal{P}_t(u, A) > 0 \,.$$

Combining this fact with the weak topological irreducibility, we deduce that for all $u_1, u_2 \in \mathcal{H}$ there exists $v \in \mathcal{H}$ so that for any $\epsilon > 0$ there exists a $\delta, t_1, t_2 > 0$ with

$$\inf_{z \in B_\delta(u_i)} \mathcal{P}_{t_i}(z, B_\epsilon(v)) > 0 \tag{2.1}$$

for $i = 1, 2$.

Now assume by contradiction that we can find two distinct invariant probability measures μ_1 and μ_2 for $\mathcal{P}_t$. Since any invariant probability measure can be written as a convex combination of ergodic measures, we can take them to be ergodic without loss of generality. Picking $u_i \in \text{supp}(\mu_i)$, by assumption there exists a v so that for any $\epsilon > 0$ there exists t_1, t_2 and $\delta > 0$ so that (2.1) holds. Since $u_i \in \text{supp}(\mu_i)$ we know that $\mu_i(B_\delta(u_i)) > 0$ and hence

$$\mu_i(B_\epsilon(v)) = \int_{\mathcal{H}} \mathcal{P}_{t_i}(z, B_\epsilon(v))\mu_i(dz) \geq \int_{B_\delta(u_i)} \mathcal{P}_{t_i}(z, B_\epsilon(v))\mu_i(dz)$$
$$\geq \mu_i(B_\delta(u_i)) \inf_{z \in B_\delta(u_i)} \mathcal{P}_{t_i}(z, B_\epsilon(v)) > 0 \,.$$

Since ϵ was arbitrary, this shows that $v \in \text{supp}(\mu_1) \cap \text{supp}(\mu_2)$, which by Theorem 1.2 gives the required contradiction. □

We now give a more quantitative version of Theorem 1.2. It shows that if one has access to the quantitative information embodied in (1.2), as opposed to only the asymptotic strong Feller property, then not only are the supports of any two ergodic invariant measures disjoint but they are actually separated by a distance which is directly related to the function C from (1.2).

Theorem 2.1 *Let $\{\mathcal{P}_t\}$ be a Markov semigroup on a separable Hilbert space $\mathcal{H}$ such that (1.2) holds for some non-decreasing function C. Let μ_1 and μ_2 be two distinct ergodic invariant probability measures for $\mathcal{P}_t$. Then, the bound $\|u_1 - u_2\| \geq 1/C(\|u_1\| \vee \|u_2\|)$ holds for any pair of points (u_1, u_2) with $u_i \in \text{supp}\,\mu_i$.*

Proof. The proof is a variation on the proof of Theorem 3.16 in [HM06]. We begin by defining for $u, v \in \mathcal{H}$ the distance $d_n(u, v) = 1 \wedge (\sqrt{\delta_n} \|u - v\|)$ where δ_n is the sequence of positive numbers from (1.2). As shown in the proof of Theorem 3.12 in [HM06], one has

$$d_n(\mathcal{P}_t^* \delta_{u_1}, \mathcal{P}_t^* \delta_{u_2}) \leq \|u_1 - u_2\|C(\|u_1\| \vee \|u_2\|)(1 + \sqrt{\delta_n}) \tag{2.2}$$

where d_n is the 1-Wasserstein distance[3] on probability measures induced by the metric d_n. Observe that for all $u, v \in \mathcal{H}$, $d_n(u, v) \leq 1$ and $\lim d_n(u, v) = \mathbf{1}_{\{u\}}(v)$. Hence by

---

[3] $d_n(\nu_1, \nu_2) = \sup \int \varphi d\nu_1 - \int \varphi d\nu_2$ where the supremum runs over functions $\varphi : \mathcal{H} \to \mathbf{R}$ which have Lipschitz constant one with respect to the metric d_n.

in Lemma 3.4 of [HM06], for any probability measures μ and ν, $\lim_{n\to\infty} d_n(\mu, \nu) = d_{\text{TV}}(\mu, \nu)$ where $d_{\text{TV}}(\mu, \nu)$ is the total variation distance[4].

Let μ_1 and μ_2 be two ergodic invariant measures with $\mu_1 \neq \mu_2$. By Birkhoff's ergodic theorem, we know that they are mutually singular and thus $d_{\text{TV}}(\mu_1, \mu_2) = 1$. We now proceed by contradiction. We assume that there exists a pair of points (u_1, u_2) with $u_i \in \text{supp}(\mu_i)$ such that $\|u_1 - u_2\| < C(\|u_1\| \vee \|u_2\|)$. We will conclude by showing that this implies that $d_{\text{TV}}(\mu_1, \mu_2) < 1$ and hence μ_1 and μ_2 are not singular which will be a contradiction.

Our assumption on u_1 and u_2 implies that there exists a set A containing u_1 and u_2 such that $\alpha \stackrel{\text{def}}{=} \min(\mu_1(A), \mu_2(A)) > 0$ and $\beta \stackrel{\text{def}}{=} \sup\{\|u - v\| : u, v \in A\}C(\|u_1\| \vee \|u_2\|) < 1$. As shown in the proof of Theorem 3.16 in [HM06], for any n one has

$$d_n(\mu_1, \mu_2) \leq 1 - \alpha(1 - \sup_{v_i \in A} d_n(\mathcal{P}_t^* \delta_{v_1}, \mathcal{P}_t^* \delta_{v_2}))$$
$$\leq 1 - \alpha(1 - \beta(1 + \sqrt{\delta_n}))$$

where the last inequality used the bound in equation (2.2). Taking $n \to \infty$ produces $d_{\text{TV}}(\mu_1, \mu_2) \leq 1 - \alpha(1 - \beta)$. Since $\alpha \in (0, 1)$ and $\beta < 1$ we concluded that $d_{\text{TV}}(\mu_1, \mu_2) < 1$. This implies a contradiction since μ_1 and μ_2 are mutually singular measures. □

Paired with this stronger version of Theorem 1.2, we have the following version of Corollary 1.4 which uses an even weaker form of irreducibility. This is a general principle. If one has a stronger from of the asymptotic strong Feller property, one can prove unique ergodicity under a weaker form of topological irreducibility. The form of irreducibility used in Corollary 2.2 allows the point where two trajectories approach to move about, depending on the degree of closeness required. To prove unique ergodicity, the trade-off is that one needs some control of the "smoothing rate" implied by asymptotic strong Feller at different points in phase space.

Corollary 2.2 *Let $\{\mathcal{P}_t\}$ be as in Theorem 2.1. Suppose that, for every $R_0 > 0$, it is possible to find $R > 0$ and $T > 0$ such that, for every $\varepsilon > 0$, there exists a point v with $\|v\| \leq R$ such that $\mathcal{P}_T(u, \mathcal{B}_\varepsilon(v)) > 0$ for every $\|u\| \leq R_0$. Then, $\mathcal{P}_t$ can have at most one invariant probability measure.*

Proof. Assume by contradiction that there exist two ergodic invariant probability measures μ_1 and μ_2 for $\mathcal{P}_t$. Then, choosing R_0 large enough so that the open ball of radius R_0 intersects the supports of both μ_1 and μ_2, it follows form the assumption, by similar reasoning as in the proof of Corollary 1.4, that supp μ_i intersects $\mathcal{B}_\varepsilon(v)$. Since $\|v\|$ is bounded uniformly in ε, making ε sufficiently small yields a contradiction with Theorem 2.1 above. □

3 Functional analytic setup

In this section, we introduce the basic function analytic set-up for the rest of the paper. We will develop the needed existence and regularity theory to place the remainder of the paper of a firm foundation. We consider semilinear stochastic evolution equations

---

[4]Different communities normalize the total variation distance differently. Our d_{TV} is half of the total variation distance as defined typically in analysis. The definition we use is common in probability as it is normalised in such a way that $d_{\text{TV}}(\mu, \nu) = 1$ for mutually singular probability measures.

with additive noise in a Hilbert space $\mathcal{H}$ (with norm $\|\cdot\|$ and innerproduct $\langle\,\cdot\,,\,\cdot\,\rangle$) of the form

$$du = -Lu\,dt + N(u)\,dt + \sum_{k=1}^{d} g_k\,dW_k(t)\,, \quad u_0 \in \mathcal{H}\,. \tag{3.1}$$

Here, the W_k are independent real-valued standard Wiener processes over some probability space $(\Omega, \mathbf{P}, \mathcal{F})$. Our main standing assumption throughout this article is that L generates an analytic semigroup and that the nonlinearity N results in a loss of regularity of a powers of L for some $a < 1$. More precisely, we have:

Assumption A.1 *There exists $a \in [0,1)$ and $\gamma_\star, \beta_\star > -a$ (either of them possibly infinite) with $\gamma_\star + \beta_\star > -1$ such that:*

1. *The operator $L\colon \mathcal{D}(L) \to \mathcal{H}$ is selfadjoint and satisfies $\langle u, Lu \rangle \geq \|u\|^2$. We denote by $\mathcal{H}_\alpha$, $\alpha \in \mathbf{R}$ the associated interpolation spaces (i.e. $\mathcal{H}_\alpha$ with $\alpha > 0$ is the domain of L^α endowed with the graph norm and $\mathcal{H}_{-\alpha}$ is its dual with respect to the pairing in $\mathcal{H}$). Furthermore, $\mathcal{H}_\infty$ is the Fréchet space $\mathcal{H}_\infty = \bigcap_{\alpha>0} \mathcal{H}_\alpha$ and $\mathcal{H}_{-\infty}$ is its dual.*

2. *There exists $n \geq 1$ such that the nonlinearity N belongs to $\mathrm{Poly}^n(\mathcal{H}_{\gamma+a}, \mathcal{H}_\gamma)$ for every $\gamma \in [-a, \gamma_\star)$ (see the definition of Poly in Section 3.1 below). In particular, from the definition of $\mathrm{Poly}(\mathcal{H}_{\gamma+a}, \mathcal{H}_\gamma)$, it follows that it is continuous from $\mathcal{H}_\infty$ to $\mathcal{H}_\infty$.*

3. *For every $\beta \in [-a, \beta_\star)$ there exists $\gamma \in [0, \gamma_\star + 1)$ such that the adjoint (in $\mathcal{H}$) $DN^*(u)$ of the derivative DN of N at u (see again the definition in Section 3.1 below) can be extended to a continuous map from $\mathcal{H}_\gamma$ to $\mathcal{L}(\mathcal{H}_{\beta+a}, \mathcal{H}_\beta)$.*

4. *One has $g_k \in \mathcal{H}_{\gamma_\star+1}$ for every k.*

Remark 3.1 If $\gamma_\star \geq 0$, then the range $\beta \in [-a, 0]$ for Assumption A.1.3 follows directly from Assumption A.1.2, since A.1.3 simply states that for $u \in \mathcal{H}_\gamma$, $DN(u)$ is a continuous linear map from $\mathcal{H}_{-\beta}$ to $\mathcal{H}_{-\beta-a}$.

Remark 3.2 The assumption $\langle u, Lu \rangle \geq \|u\|^2$ is made only for convenience so that L^γ is well-defined as a positive selfadjoint operator for every $\gamma \in \mathbf{R}$. It can always be realized by subtracting a suitable constant to L and adding it to N.

Similarly, non-selfadjoint linear operators are allowed if the antisymmetric part is sufficiently "dominated" by the symmetric part, since one can then consider the antisymmetric part as part of the nonlinearity N.

Remark 3.3 It follows directly from the Calderón-Lions interpolation theorem [RS80, Appendix to IX.4] that if $N \in \mathrm{Poly}(\mathcal{H}_0, \mathcal{H}_{-a}) \cap \mathrm{Poly}(\mathcal{H}_{\gamma_\star+a}, \mathcal{H}_{\gamma_\star})$ for some $\gamma_\star > -a$, then $N \in \mathrm{Poly}(\mathcal{H}_{\gamma+a}, \mathcal{H}_\gamma)$ for every $\gamma \in [-a, \gamma_\star]$. This can be seen by interpreting N as a sum of *linear* maps from $\mathcal{H}_{\gamma+a}^{\otimes n}$ to $\mathcal{H}_\gamma$ for suitable values of n.

It will be convenient in the sequel to define F by

$$F(u) = -Lu + N(u)\,. \tag{3.2}$$

Note that F is in $\text{Poly}^n(\mathcal{H}_{\gamma+1}, \mathcal{H}_\gamma)$ for every $\gamma \in [-1, \gamma_\star)$. We also define a linear operator $G \colon \mathbf{R}^d \to \mathcal{H}_\infty$ by

$$Gv = \sum_{k=1}^{d} v_k g_k \,,$$

for $v = (v_1, \ldots, v_d) \in \mathbf{R}^d$. With these notations, we will sometimes rewrite (3.1) as

$$du = F(u)\,dt + G\,dW(t) \,, \quad u_0 \in \mathcal{H} \,, \tag{3.3}$$

for $W = (W_1, \ldots, W_d)$ a standard d-dimensional Wiener process.

3.1 Polynomials

We now describe in what sense we mean that N is a "polynomial" vector field. Given a Fréchet space X, we denote by $\mathcal{L}_s^n(X)$ the space of continuous symmetric n-linear maps from X to itself. We also denote by $\mathcal{L}(X, Y)$ the space of continuous linear maps from X to Y. For the sake of brevity, we will make use of the two equivalent notations $P(u)$ and $P(u^{\otimes n})$ for $P \in \mathcal{L}^n(X)$.

Given $Q \in \mathcal{L}_s^k$, its *derivative* is given by the following $n - 1$-linear map from X to $\mathcal{L}(X, X)$:

$$DQ(u)v = kQ(u^{\otimes(k-1)} \otimes v) \,.$$

We will also use the notation $DQ^* \colon X \to \mathcal{L}(X', X')$ for the dual map given by

$$\langle w, DQ^*(u)v \rangle = \langle v, DQ(u)w \rangle = k\langle v, Q(u^{\otimes(k-1)} \otimes w) \rangle \,.$$

Given $P \in \mathcal{L}_s^k$ and $Q \in \mathcal{L}_s^\ell$, we define the derivative $DQ\,P$ of Q in the direction P as a continuous map from $X \times X$ to X by

$$DQ(u)\,P(v) = \ell Q(u^{\otimes(\ell-1)} \otimes P(v)) \,.$$

Note that by polarisation, $u \mapsto DQ(u)P(u)$ uniquely defines an element on $\mathcal{L}_s^{k+\ell-1}$. This allows us to define a "Lie bracket" $[P, Q] \in \mathcal{L}_s^{k+\ell-1}$ between P and Q by

$$[P, Q](u) = DQ(u)\,P(u) - DP(u)\,Q(u) \,.$$

We also define $\text{Poly}^n(X)$ as the set of continuous maps $P \colon X \to X$ of the form

$$P(u) = \sum_{k=0}^{n} P^{(k)}(u) \,,$$

with $P^{(k)} \in \mathcal{L}_s^k(X)$ (here $\mathcal{L}_s^0(X)$ is the space of constant maps and can be identified with X). We also set $\text{Poly}(X) = \bigcup_{n \geq 0} \text{Poly}^n(X)$. The Lie bracket defined above extends to a map from $\text{Poly}(X) \times \text{Poly}(X) \to \text{Poly}(X)$ by linearity.

3.1.1 Polynomials over $\mathcal{H}$

We now specialize to polynomials over $\mathcal{H}$. We begin by choosing X equal to the Fréchet space $\mathcal{H}_\infty$, the intersection of $\mathcal{H}_a$ over all $a > 0$. Next we define the space $\text{Poly}(\mathcal{H}_a, \mathcal{H}_b) \subset \text{Poly}(\mathcal{H}_\infty)$ as the set of polynomials $P \in \text{Poly}(\mathcal{H}_\infty)$ such that there exists a continuous map $\hat{P} \colon \mathcal{H}_a \to \mathcal{H}_b$ with $\hat{P}(u) = P(u)$ for all $u \in \mathcal{H}_\infty$. Note that in general (unlike $\text{Poly}(\mathcal{H}_\infty)$), $P, Q \in \text{Poly}(\mathcal{H}_a, \mathcal{H}_b)$ does not necessarily imply $[P, Q] \in \text{Poly}(\mathcal{H}_a, \mathcal{H}_b)$. We will make an abuse of notation and use the same symbol for both P and $\hat{P}$ in the sequel.

3.1.2 Taylor expansions and Lie brackets

We now consider the Taylor expansion of a polynomial Q in a direction g belonging to $\text{span}\{g_1, \cdots, g_d\} \subset \mathcal{H}_{\gamma_\star+1}$. Fix $Q \in \text{Poly}^m(\mathcal{H}_\gamma, \mathcal{H}_\beta)$ for some $\gamma \leq \gamma_\star + 1$ and any $\beta \in \mathbf{R}$. For $v \in \mathcal{H}_\gamma$ and $w = (w_1, \ldots, w_d) \in \mathbf{R}^d$, observe that there exist polynomials Q_α such that

$$Q\Big(v + \sum_{k=1}^d g_k w_k\Big) = \sum_\alpha Q_\alpha(v) w_\alpha , \quad w_\alpha = w_{\alpha_1} \cdots w_{\alpha_\ell} , \qquad (3.4)$$

where the summation runs over all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, $\ell \geq 0$ with values in the index set $\{1, \ldots, d\}$. It can be checked that the polynomials $Q_\alpha \in \text{Poly}^{m-|\alpha|}(\mathcal{H}_\gamma, \mathcal{H}_\beta)$ are given by the formula

$$Q_\alpha(v) = \frac{1}{\alpha!}[[\ldots [Q, g_{\alpha_1}]\ldots], g_{\alpha_\ell}] = \frac{1}{\alpha!} D^{|\alpha|}Q(v)(g_{\alpha_1}, \ldots, g_{\alpha_\ell}) . \qquad (3.5)$$

Here, $\alpha!$ is defined by $\alpha! = \alpha(1)! \cdots \alpha(d)!$, where $\alpha(j)$ counts the number of occurences of the index j in α. (By convention, we set $Q_\phi = Q$ and $Q_\alpha = 0$ if $|\alpha| > m$.)

We emphasize that multi-indices are *unordered* collections of $\{1, \ldots, d\}$ where repeated elements are allowed. As such, the union of two multi-indices is a well-defined operation, as is the partial ordering given by inclusion.[5]

3.2 *A priori* bounds on the solution

This section is devoted to the proof that Assumption A.1 is sufficient to obtain not only unique solutions to (3.1) (possibly up to some explosion time), but to obtain further regularity properties for both the solution and its derivative with respect to the initial condition. We do not claim that the material presented in this section is new, but while similar frameworks can be found in [DPZ92, Fla95], the framework presented here does not seem to appear in this form in the literature. Since the proofs are rather straightforward, we choose to present them for the sake of completeness, although in a rather condensed form.

We first start with a local existence and uniqueness result for the solutions to (3.1):

Proposition 3.4 *For every initial condition $u_0 \in \mathcal{H}$, there exists a stopping time $\tau > 0$ such that (3.1) has a unique mild solution u up to time τ, that is to say u almost surely satisfies*

$$u_t = e^{-Lt}u_0 + \int_0^t e^{-L(t-s)}N(u_s)\,ds + \int_0^t e^{-L(t-s)}G\,dW(s) , \qquad (3.6)$$

for all stopping times t with $t \leq \tau$. Furthermore u is adapted to the filtration generated by W and is in $C([0, \tau), \mathcal{H})$ with probability one.

Remark 3.5 Since we assume that N is locally Lipschitz continuous from $\mathcal{H}$ to $\mathcal{H}_{-a}$ for some $a < 1$ and since the bound $\|e^{-Lt}\|_{\mathcal{H}_{-a} \to \mathcal{H}} \leq Ct^{-a}$ holds for $t \leq T$, the first integral appearing in (3.6) does converge in $\mathcal{H}$. Therefore the right hand side of (3.6) makes sense for every continuous $\mathcal{H}$-valued process u.

---

[5]To be precise, one could identify a multi-index with its counting function $\alpha\colon \{1, \ldots, d\} \to \mathbf{N}$. With this identification, the union of two multi-indices corresponds to the sums of their counting functions, while $\alpha \subset \beta$ means that $\alpha(k) \leq \beta(k)$ for every k.

For notational convenience, we denote by $W_L(s,t) = \int_s^t e^{-L(t-r)} G \, dW(r)$ the "stochastic convolution." Since we assumed that $g_k \in \mathcal{H}_{\gamma_*+1}$, it is possible to obtain bounds on all exponential moments of $\sup_{0 \le s < t \le T} \|W_L(s,t)\|_\gamma$ for every $T > 0$ and every $\gamma \le \gamma_* + 1$.

Proof. Given a function $\xi \colon \mathbf{R}_+ \to \mathcal{H}$ and a time $T > 0$, define a map $\Phi_{T,\xi} \colon \mathcal{H} \times \mathcal{C}([0,T],\mathcal{H}) \to \mathcal{C}([0,T],\mathcal{H})$ (endowed with the supremum norm) by

$$(\Phi_{T,\xi}(u_0,u))_t = e^{-Lt} u_0 + \xi(t) + \int_0^t e^{-L(t-s)} N(u_s) \, ds . \tag{3.7}$$

Since $N \in \mathrm{Poly}(\mathcal{H}, \mathcal{H}_{-a})$ by setting $\gamma = -a$ in Assumption A.1.2, and suppressing the dependence on u_0, there exists a positive constant C such that

$$\|\Phi_{T,\xi}(u) - \Phi_{T,\xi}(\tilde{u})\| \le \sup_{t \in [0,T]} C \int_0^t (t-s)^{-a} \|u_s - \tilde{u}_s\| (1 + \|u_s\| + \|\tilde{u}_s\|)^{n-1} \, ds$$

$$\le C \|u - \tilde{u}\| (1 + \|u\| + \|\tilde{u}\|)^{n-1} T^{1-a} .$$

Recall that n is the degree of the polynomial nonlinearity N. It follows that, for every ξ, there exists $T > 0$ and $R > 0$ such that $\Phi_{T,\xi}(u_0, \cdot)$ is a contraction in the ball of radius R around $e^{-Lt} u_0 + \xi(t)$. Setting $\xi(t) = W_L(0,t)$, this yields existence and uniqueness of the solution to (3.6) for almost every noise path $W_L(0,t)$ by the Banach fixed point theorem. The largest such T is a stopping time since it only depends on the norm of u_0 and on ξ up to time T. It is clear that $\Phi_{T,\xi}(u_0, u)_t$ only depends on the noise W_L up to time t, so that the solution is adapted to the filtration generated by W, thus concluding the proof of the proposition. $\qquad \square$

The remainder of this section is devoted to obtaining further regularity properties of the solutions.

Proposition 3.6 *Fix $T > 0$. For every $\gamma \in [0, \gamma_* + 1)$ there exist exponents $p_\gamma \ge 1$ and $q_\gamma \ge 0$ and a constant C such that*

$$\|u_t\|_\gamma \le C t^{-q_\gamma} \big(1 + \sup_{s \in [\frac{t}{2}, t]} \|u_s\| + \sup_{\frac{t}{2} \le s < r \le t} \|W_L(s,r)\|_\gamma \big)^{p_\gamma} \tag{3.8}$$

for all $t \in (0, T \wedge \tau]$, where $\tau = \sup\{t > 0 \,:\, \|u_t\| < \infty\}$. In particular, if $\gamma = \sum_{j=0}^k \delta_j$ for some $k \in \mathbf{N}$ and $\delta_j \in (0, 1-a)$ then $q_\gamma \le \sum_{j=1}^k \delta_j n^{j-1}$.

Proof. The proof follows a standard "bootstrapping argument" on γ in the following way. The statement is obviously true for $\gamma = 0$ with $p_\gamma = 1$ and $q_\gamma = 0$. Assume that, for some $\alpha = \alpha_0 \in [1/2, 1)$ and for some $\gamma = \gamma_0 \in [0, \gamma_* + a)$, we have the bound

$$\|u_t\|_\gamma \le C t^{-q_\gamma} \big(1 + \sup_{s \in [\alpha t, t]} \|u_s\| + \sup_{\alpha t \le s < r \le t} \|W_L(s,r)\|_\gamma \big)^{p_\gamma} , \tag{3.9}$$

for all $t \in (0, T]$.

We will then argue that, for any arbitrary $\delta \in (0, 1-a)$, the statement (3.9) also holds for $\gamma = \gamma_0 + \delta$ (and therefore also for all intermediate values of γ) and $\alpha = \alpha_0^2$. Since it is possible to go from $\gamma = 0$ to any value of $\gamma < \gamma_* + 1$ in a finite number of steps (making sure that $\gamma \le 1 + a$ in every intermediate step) and since we are allowed to choose α as close to 1 as we wish, the claim follows at once.

Using the mild formulation (3.6), we have

$$u_t = e^{-(1-\alpha)Lt} u_{\alpha t} + \int_{\alpha t}^{t} e^{-L(t-s)} N(u_s)\, ds + W_L(\alpha t, t)\,.$$

Since $\gamma \in [0, \gamma_\star + a)$, one has $N \in \mathrm{Poly}(\mathcal{H}_\gamma, \mathcal{H}_{\gamma-a})$ by Assumption A.1.2. Hence, for $t \in (0, T]$,

$$
\begin{aligned}
\|u_t\|_{\gamma+\delta} &\le Ct^{-\delta}\|u_{\alpha t}\|_\gamma + \|W_L(\alpha t, t)\|_\gamma + C\int_{\alpha t}^{t} (t-s)^{-(\delta+a)}(1 + \|u_s\|_\gamma)^n\, ds \\
&\le C(t^{-\delta} + t^{1-\delta-a}) \sup_{\alpha t \le s \le t}(1 + \|u_s\|_\gamma^n) + \|W_L(\alpha t, t)\|_\gamma \\
&\le Ct^{-\delta} \sup_{\alpha t \le s \le t}(1 + \|u_s\|_\gamma^n) + \|W_L(\alpha t, t)\|_\gamma\,.
\end{aligned}
$$

Here, the constant C depends on everything but t and u_0. Using the induction hypothesis, this yields the bound

$$\|u_t\|_{\gamma+\delta} \le Ct^{-\delta-nq_\gamma}\Big(1 + \sup_{s\in[\alpha^2 t, t]}\|u_s\| + \sup_{\alpha^2 \le s < r \le t}\|W_L(s, r)\|_\gamma\Big)^{np_\gamma} + \|W_L(\alpha t, t)\|_\gamma\,,$$

thus showing that (3.9) holds for $\gamma = \gamma_0 + \delta$ and $\alpha = \alpha_0^2$ with $p_{\gamma+\delta} = np_\gamma$ and $q_{\gamma+\delta} = \delta + nq_\gamma$. This concludes the proof of Proposition 3.6. $\qquad\square$

3.3 Linearization and its adjoint

In this section, we study how the solutions to (3.1) depend on their initial conditions. Since the map from (3.7) used to construct the solutions to (3.1) is Fréchet differentiable (it is actually infinitely differentiable) and since it is a contraction for sufficiently small values of t, we can apply the implicit functions theorem (see for example [RR04] for a Banach space version) to deduce that for every realisation of the driving noise, the map $u_s \mapsto u_t$ is Fréchet differentiable, provided that $t > s$ is sufficiently close to s.

Iterating this argument, one sees that, for any $s \le t < \tau$, the map $u_s \mapsto u_t$ given by the solutions to (3.1) is Fréchet differentiable in $\mathcal{H}$. Inspecting the expression for the derivative given by the implicit functions theorem, we conclude that the derivative $J_{s,t}\varphi$ in the direction $\varphi \in \mathcal{H}$ satisfies the following random linear equation in its mild formulation:

$$\partial_t J_{s,t}\varphi = -L J_{s,t}\varphi + DN(u_t)J_{s,t}\varphi\,, \quad J_{s,s}\varphi = \varphi\,. \tag{3.10}$$

Note that, by the properties of monomials, it follows from Assumption A.1.2 that

$$\|DN(u)v\|_\gamma \le C(1 + \|u\|_{\gamma+a})^{n-1}\|v\|_{\gamma+a}\,,$$

for every $\gamma \in [-a, \gamma_\star)$. A fixed point argument similar to the one in Proposition 3.4 shows that the solution to (3.10) is unique, but note that it does not allow us to obtain bounds on its moments. We only have that for any T smaller than the explosion time to the solutions of (3.1), there exists a (random) constant C such that

$$\sup_{0 \le s < t < T} \sup_{\|\varphi\| \le 1} \|J_{s,t}\varphi\| \le C\,. \tag{3.11}$$

The constant C depends exponentially on the size of the solution u in the interval $[0, T]$. However, if we obtain better control on $J_{s,t}$ by some means, we can then use the following bootstrapping argument:

Proposition 3.7 *For every $\gamma \in (0, \gamma_\star + 1)$, there exists an exponent $\tilde{q}_\gamma \geq 0$, and constants $C > 0$ and $\gamma_0 < \gamma$ such that we have the bound*

$$\|J_{t,t+s}\varphi\|_\gamma \leq Cs^{-\gamma} \sup_{r \in [\frac{s}{2}, s]} (1 + \|u_{t+r}\|_{\gamma_0})^{\tilde{q}_\gamma} \|J_{t,t+r}\varphi\| , \qquad (3.12)$$

for every $\varphi \in \mathcal{H}$ and every $t, s > 0$. If $\gamma < 1 - a$, then one can choose $\gamma_0 = 0$ and $\tilde{q}_\gamma = n - 1$.

Since an almost identical argument will be used in the proof of Proposition 3.9 below, we refer the reader there for details. We chose to present that proof instead of this one because the presence of an adjoint causes slight additional complications.

For $s \leq t$, let us define operators $K_{s,t}$ via the solution to the (random) PDE

$$\partial_s K_{s,t}\varphi = LK_{s,t}\varphi - DN^*(u_s)K_{s,t}\varphi , \quad K_{t,t}\varphi = \varphi , \quad \varphi \in \mathcal{H} . \qquad (3.13)$$

Note that this equation runs *backwards* in time and is random through the solution u_t of (3.1). Here, $DN^*(u)$ denotes the adjoint in $\mathcal{H}$ of the operator $DN(u)$ defined earlier. Fixing the terminal time t and setting $\varphi_s = K_{t-s,t}\varphi$, we obtain a more usual representation for φ_s:

$$\partial_s \varphi_s = -L\varphi_s + DN^*(u_{t-s})\varphi_s . \qquad (3.14)$$

The remainder of this subsection will be devoted to obtaining regularity bounds on the solutions to (3.13) and to the proof that $K_{s,t}$ is actually the adjoint of $J_{s,t}$. We start by showing that, for γ sufficiently close to (but less than) $\gamma_\star + 1$, (3.13) has a unique solution for every path $u \in \mathcal{C}(\mathbf{R}, \mathcal{H}_\gamma)$ and $\varphi \in \mathcal{H}$.

Proposition 3.8 *There exists $\gamma < \gamma_\star + 1$ such that, for every $\varphi \in \mathcal{H}$, equation (3.13) has a unique continuous $\mathcal{H}$-valued solution for every $s < t$ and every $u \in \mathcal{C}(\mathbf{R}, \mathcal{H}_\gamma)$. Furthermore, $K_{s,t}$ depends only on u_r for $r \in [s, t]$ and the map $\varphi \mapsto K_{s,t}\varphi$ is linear and bounded.*

Proof. As in Proposition 3.4, we define a map $\Phi_{T,u} \colon \mathcal{H} \times \mathcal{C}([0, T], \mathcal{H}) \to \mathcal{C}([0, T], \mathcal{H})$ by

$$\left(\Phi_{T,u}(\varphi_0, \varphi)\right)_t = e^{-Lt}\varphi_0 + \int_0^t e^{-L(t-s)}(DN^*(u_s))\varphi_s \, ds .$$

It follows from Assumption A.1.3 with $\beta = -a$ that there exists $\gamma < \gamma_\star + 1$ such that $DN^*(u) \colon \mathcal{H} \to \mathcal{H}_{-a}$ is a bounded linear operator for every $u \in \mathcal{H}_\gamma$. Proceeding as in the proof of Proposition 3.4, we see that Φ is a contraction for sufficiently small T. $\square$

Similarly to before, we can use a bootstrapping argument to show that $K_{s,t}\varphi$ actually has more regularity than stated in Proposition 3.8.

Proposition 3.9 *For every $\beta \in (0, \beta_\star + 1)$, there exists $\gamma < \gamma_\star + 1$, an exponent $\bar{q}_\beta > 0$, and a constant C such that*

$$\|K_{t-s,t}\varphi\|_\beta \leq Cs^{-\beta} \sup_{r \in [\frac{s}{2}, s]} (1 + \|u_{t-r}\|_\gamma)^{\bar{q}_\beta} \|K_{t-r,t}\varphi\| , \qquad (3.15)$$

for every $\varphi \in \mathcal{H}$, every $t, s > 0$, and every $u \in \mathcal{C}(\mathbf{R}, \mathcal{H}_\gamma)$.

Proof. Fix $\beta < \beta_\star + a$ and $\delta \in (0, 1 - a)$ and assume that the bound (3.15) holds for $\|K_{s,t}\varphi\|_\beta$. Since we run s "backwards in time" from $s = t$, we consider again t as fixed and use the notation $\varphi_s = K_{t-s,t}\varphi$. We then have, for arbitrary $\alpha \in (0, 1)$,

$$\|\varphi_s\|_{\beta+\delta} \leq C s^{-\delta} \|\varphi_{\alpha s}\|_\beta + C \int_{\alpha s}^s (s - r)^{-(\delta+a)} \|DN^*(u_{t-r})\varphi_r\|_{\beta-a} \, dr \,,$$

provided that γ is sufficiently close to $\gamma_\star + 1$ such that $DN^* \colon \mathcal{H}_\gamma \to \mathcal{L}(\mathcal{H}_\beta, \mathcal{H}_{\beta-a})$ by Assumption A.1.3. Furthermore, the operator norm of $DN^*(v)$ is bounded by $C(1 + \|v\|_\gamma)^{n-1}$, yielding

$$\|\varphi_s\|_{\beta+\delta} \leq C s^{-\delta} \|\varphi_{\alpha s}\|_\beta + C s^{-(\delta+a-1)} \sup_{r \in [\alpha s, s]} (1 + \|u_r\|_\gamma)^{n-1} \|\varphi_r\|_\beta$$

$$\leq C s^{-\delta} \sup_{r \in [\alpha s, s]} (1 + \|u_r\|_\gamma)^{n-1} \|\varphi_r\|_\beta \,.$$

Iterating these bounds as in Proposition 3.6 concludes the proof. $\qquad\square$

The following lemma appears also in [MP06, BM07]. It plays a central role in establishing the representation of the Malliavin matrix given in (4.11) on which this article as well as [MP06, BM07] rely heavily.

Proposition 3.10 *For every $0 \leq s < t$, $K_{s,t}$ is the adjoint of $J_{s,t}$ in $\mathcal{H}$, that is $K_{s,t} = J_{s,t}^*$.*

Proof. Fixing $0 \leq s < t$ and $\varphi, \psi \in \mathcal{H}_\infty$, we claim that the expression

$$\langle J_{s,r}\varphi, K_{r,t}\psi \rangle \,, \tag{3.16}$$

is independent of $r \in [s, t]$. Evaluating (3.16) at both $r = s$ and $r = t$ then concludes the proof.

We now prove that (3.16) is independent of r as claimed. It follows from (3.13) and Proposition 3.6 that, with probability one, the map $r \mapsto K_{r,t}\varphi$ is continuous with values in $\mathcal{H}_{\beta+1}$ and differentiable with values in $\mathcal{H}_\beta$, provided that $\beta < \beta_\star$. Similarly, the map $r \mapsto J_{s,r}\psi$ is continuous with values in $\mathcal{H}_{\gamma+1}$ and differentiable with values in $\mathcal{H}_\gamma$, provided that $\gamma < \gamma_\star$. Since $\gamma_\star + \beta_\star > -1$ by assumption, it thus follows that (3.16) is differentiable in r for $r \in (s, t)$ with

$$\partial_r \langle J_{s,r}\varphi, K_{r,t}\psi \rangle = \langle (L + DN(u_r))J_{s,r}\varphi, K_{r,t}\psi \rangle$$
$$- \langle J_{s,r}\varphi, (L + DN^*(u_r))K_{r,t}\psi \rangle = 0 \,.$$

Since furthermore both $r \mapsto K_{r,t}\varphi$ and $r \mapsto J_{s,r}\psi$ are continuous in r on the closed interval, the proof is complete. See for example [DL92, p. 477] for more details. $\qquad\square$

3.4 Higher order variations

We conclude this section with a formula for the higher-order variations of the solution. This will mostly be useful in Section 8 in order to obtain the smoothness of the density for finite-dimensional projections of the transition probabilities.

For integer $n \geq 2$, let $\varphi = (\varphi_1, \cdots, \varphi_n) \in \mathcal{H}^{\otimes n}$ and $s = (s_1, \cdots, s_n) \in [0, \infty)^n$ and define $\vee s = s_1 \vee \cdots \vee s_n$. We will now define the n-th variation of the equation $J_{s,t}^{(n)}\varphi$ which intuitively is the cumulative effect on u_t of varying the value of u_{s_k} in the direction φ_k.

If $I = \{n_1 < \ldots < n_{|I|}\}$ is an ordered subset of $\{1, \ldots, n\}$ (here $|I|$ means the number of elements in I), we introduce the notation $s_I = (s_{n_1}, \ldots, s_{n_{|I|}})$ and $\varphi_I = (\varphi_{n_1}, \ldots, \varphi_{n_{|I|}})$. Now the n-th variation of the equation $J_{s,t}^{(n)}\varphi$ solves

$$\partial_t J_{s,t}^{(n)}\varphi = -L J_{s,t}^{(n)}\varphi + DN(u(t))J_{s,t}^{(n)}\varphi + \mathcal{G}_{s,t}^{(n)}(u(t), \varphi), \quad t > \vee s, \qquad (3.17)$$
$$J_{s,t}^{(n)}\varphi = 0, \quad t \leq \vee s,$$

where

$$\mathcal{G}_{s,t}^{(n)}(u, \varphi) = \sum_{\nu=2}^{m \wedge n} \sum_{I_1, \ldots, I_\nu} D^{(\nu)}N(u)\Big(J_{s_{I_1}, t}^{(|I_1|)}\varphi_{I_1}, \ldots, J_{s_{I_\nu}, t}^{(|I_\nu|)}\varphi_{I_\nu}\Big), \qquad (3.18)$$

and the second sum runs over all partitions of $\{1, \ldots, n\}$ into disjoint, ordered non-empty sets $I_1, \ldots, I_\nu$.

The variations of constants formula then implies that

$$J_{s,t}^{(n)}\varphi = \int_0^t J_{r,t} \mathcal{G}_{s,r}^{(n)}(u_r, \varphi)dr , \qquad (3.19)$$

see also [BM07]. We obtain the following bound on the higher-order variations:

Proposition 3.11 *If $\beta_\star > a - 1$ then there exists $\gamma < \gamma_\star + 1$ such that, for every $n > 0$, there exist exponents N_n and M_n such that*

$$\|J_{s,t}^{(n)}\varphi\| \leq C \sup_{r \in [0,t]} (1 + \|u_r\|_\gamma)^{N_n} \sup_{0 \leq u < v \leq t} (1 + \|J_{u,v}\|)^{M_n} ,$$

uniformly over all n-uples φ with $\|\varphi_k\| \leq 1$ for every k.

Proof. We proceed by induction. As a shorthand, we set

$$\mathcal{E}(M, N) = \sup_{r \in [0,t]} (1 + \|u_r\|_\gamma)^N \sup_{0 \leq u < v \leq t} (1 + \|J_{u,v}\|)^M .$$

The result is trivially true for $n = 1$ with $M_1 = 1$ and $N_1 = 0$. For $n > 1$, we combine (3.19) and (3.18), and we use Assumption A.1, part 2., to obtain

$$\|J_{s,t}^{(n)}\varphi\| \leq C \int_0^t \|J_{r,t}\|_{-a \to 0}\Big(1 + \|u_r\|^n + \sum_I \|J_{s_I, r}^{|I|}\varphi_I\|^n\Big) dr$$

$$\leq C\mathcal{E}(nM_{n-1}, n(N_{n-1} + 1)) \int_0^t \|K_{r,t}\|_{0 \to a} dr .$$

To go from the first to the second line, we used the induction hypothesis, the fact that $K_{r,t} = J_{r,t}^*$, and the duality between $\mathcal{H}_a$ and $\mathcal{H}_{-a}$.

It remains to apply Proposition 3.9 with $\beta = a$ to obtain the required bound. $\qquad\square$

4 Malliavin calculus

In this section, we show that the solution to the SPDE (3.1) has a Malliavin derivative and we give an expression for it. Actually, since we are dealing with additive noise, we show the stronger result that the solution is Fréchet differentiable with respect to the driving noise. In this section, we will make the standing assumption that the explosion time τ from Proposition 3.4 is infinite.

4.1 Malliavin derivative

In light of Proposition 3.4, for fixed initial condition $u_0 \in \mathcal{H}$ there exists an "Itô map" $\Phi_t^{u_0} : \mathcal{C}([0, t], \mathbf{R}^d) \to \mathcal{H}$ with $u_t = \Phi_t^{u_0}(W)$. We have:

Proposition 4.1 *For every $t > 0$ and every $u \in \mathcal{H}$, the map Φ_t^u is Fréchet differentiable and its Fréchet derivative $D\Phi_t^u v$ in the direction $v \in \mathcal{C}(\mathbf{R}_+, \mathbf{R}^d)$ satisfies the equation*

$$d\, D\Phi_t^u v = -L D\Phi_t^u v\, dt + DN(u_t) D\Phi_t^u v\, dt + G dv(t) \tag{4.1}$$

in the mild sense.

Remark 4.2 Note that (4.1) has a unique $\mathcal{H}$-valued mild solution for every continuous function v because it follows from our assumptions that $Gv \in \mathcal{C}(\mathbf{R}_+, \mathcal{H}_\gamma)$ for some $\gamma > 0$ and therefore $\int_0^t e^{-L(t-s)} G\, dv(s) = Gv(t) - e^{-Lt} Gv(0) - \int_0^t L e^{-L(t-s)} Gv(s)\, ds$ is a continuous $\mathcal{H}$-valued process.

Proof of Proposition 4.1. The proof works in exactly the same way as the arguments presented in Section 3.3: it follows from Remark 4.2 that for any given $u_0 \in \mathcal{H}$ and $t > 0$, the map

$$(W, u) \mapsto e^{-Lt} u_0 + \int_0^t e^{-L(t-s)} N(u(s))\, ds + \int_0^t e^{-L(t-s)} G\, dW(s)$$

is Fréchet differentiable in $\mathcal{C}([0, t], \mathbf{R}^d) \times \mathcal{C}([0, t], \mathcal{H})$. Furthermore, for t sufficiently small (depending on u and W), it satisfies the assumptions of the implicit functions theorem, so that the claim follows in this case. The claim for arbitrary values of t follows by iterating the statement. □

As a consequence, it follows from Duhamel's formula and the fact that $J_{s,t}$ is the unique solution to (3.10) that

Corollary 4.3 *If v is absolutely continuous and of bounded variation, then*

$$D\Phi_t^u v = \int_0^t J_{s,t} G dv(s)\,, \tag{4.2}$$

where the integral is to be understood as a Riemann-Stieltjes integral and the Jacobian $J_{s,t}$ is as in (3.10).

In particular, (4.2) holds for every v in the *Cameron-Martin space*

$$\mathcal{CM} = \{v : \partial_t v \in L^2([0, \infty), \mathbf{R}^d),\quad v(0) = 0\}\,,$$

which is a Hilbert space endowed with the norm $\|v\|_{\mathcal{CM}}^2 = \int_0^\infty |\partial_t v(t)|_{\mathbf{R}^d}^2\, dt \stackrel{\text{def}}{=} \|\partial_t v\|^2$. Obviously, $\mathcal{CM}$ is isometric to $\mathcal{CM}' = L^2([0, \infty), \mathbf{R}^d)$, so we will in the sequel use the notation

$$\mathcal{D}_h \Phi_t^u \stackrel{\text{def}}{=} D\Phi_t^u v = \int_0^t J_{s,t} G dv(s) = \int_0^t J_{s,t} G h(s)\, ds\,, \quad \text{if } \partial_t v = h\,. \tag{4.3}$$

The representation (4.2) is still valid for arbitrary stochastic processes h such that $h \in \mathcal{CM}'$ almost surely.

Since $G : \mathbf{R}^d \to \mathcal{H}_{\gamma_*+1}$ is a bounded operator whose norm we denote $\|G\|$, we obtain the bound

$$\|\mathcal{D}_h \Phi_t^u\| \le \|G\| \int_0^t \|J_{s,t}\| \, |h(s)| \, ds \le C \|J_{\cdot,t}\|_{L^2(0,t,\mathcal{H})} \|h\| \,,$$

valid for every $h \in \mathcal{CM}'$. In particular, by Riesz's representation theorem, this shows that there exists a (random) element $\mathcal{D}\Phi_t^u$ of $\mathcal{CM}' \otimes \mathcal{H}$ such that

$$\mathcal{D}_h \Phi_t^u = \langle \mathcal{D}\Phi_t^u, h \rangle_{\mathcal{CM}'} = \int_0^\infty \mathcal{D}_s \Phi_t^u h(s) \, ds \,, \qquad (4.4)$$

for every $h \in \mathcal{CM}'$. This abuse of notation is partially justified by the fact that, at least formally, $\mathcal{D}_s \Phi_t^u = \mathcal{D}_h \Phi_t^u$ with $h(r) = \delta(s-r)$. In our particular case, it follows from (4.2) that one has

$$\mathcal{D}_s \Phi_t^u = J_{s,t} G \in \mathbf{R}^d \otimes \mathcal{H} \,, \quad t > s \,,$$

and $\mathcal{D}_s \Phi_t^u = 0$ for $s > t$. With this notation, the identity (4.2) can be rewritten as $\mathcal{D}_h u_t = \int_0^t \mathcal{D}_s u_t \, h(s) ds$. It follows from the theory of Malliavin calculus, see for example [Mal97, Nua95] that, for any Hilbert space $\mathcal{H}$, there exists a closed unbounded linear operator $\mathcal{D} \colon L^2(\Omega, \mathbf{R}) \otimes \mathcal{H} \to L_{\text{ad}}^2(\Omega, \mathcal{F}_t, \mathcal{CM}') \otimes \mathcal{H}$ such that $\mathcal{D}\Phi_t$ coincides with the object described above whenever Φ_t is the solution map to (3.1). Here, $\mathcal{F}_t$ is the σ-algebra generated by the increments of W up to time t and L_{ad}^2 denotes the space of L^2 functions adapted to the filtration $\{\mathcal{F}_t\}$.

The operator $\mathcal{D}$ simply acts as the identity on the factor $\mathcal{H}$, so that we really interpret it as an operator from $L^2(\Omega, \mathbf{R})$ to $L^2(\Omega, \mathcal{CM}')$. The operator $\mathcal{D}$ is called the "Malliavin derivative."

We define a family of random linear operators $\mathcal{A}_t \colon \mathcal{CM}' \to \mathcal{H}$ (depending also on the initial condition $u_0 \in \mathcal{H}$ for (3.1)) by $h \mapsto \langle \mathcal{D}\Phi_t^u, h \rangle$. It follows from (4.3) that their adjoints $\mathcal{A}_t^* \colon \mathcal{H} \to \mathcal{CM}'$ are given for $\xi \in \mathcal{H}$ by

$$(\mathcal{A}_t^* \xi)(s) = \begin{cases} G^* J_{s,t}^* \xi = G^* K_{s,t} \xi & \text{for } s \le t \,, \\ 0 & \text{for } s > t \,. \end{cases} \qquad (4.5)$$

Similarly, we define $\mathcal{A}_{s,t} \colon \mathcal{CM}' \to \mathcal{H}$ by $\mathcal{A}_{t,s} h \stackrel{\text{def}}{=} \mathcal{A}_t(h\mathbf{1}_{[t,s]}) = \langle \mathcal{D}u_t, h\mathbf{1}_{[t,s]} \rangle = \int_s^t J_{r,t} G h_r \, dr$. Observe that $\mathcal{A}_{s,t}^* \colon \mathcal{H} \to \mathcal{CM}'$ is given for $\xi \in \mathcal{H}$ by $(\mathcal{A}_{s,t}^* \xi)(r) = G^* J_{r,t}^* \xi = G^* K_{r,t} \xi$ for $r \in [s,t]$ and zero otherwise.

Recall that the Skorokhod integral $h \mapsto \int_0^t h(s) \cdot dW(s) \stackrel{\text{def}}{=} \mathcal{D}^* h$ is *defined* as the adjoint of the Malliavin derivative operator (or rather of the part acting on $L^2(\Omega, \mathcal{F}_t, \mathbf{R})$ and not on $\mathcal{H}$). In other words, one has the following identity between elements of $\mathcal{H}$:

$$\mathbf{E}\mathcal{D}_h \Phi_t^u = \mathbf{E}\langle \mathcal{D}\Phi_t^u, h \rangle = \mathbf{E}\Big(\Phi_t^u \int_0^t h(s) \cdot dW(s) \Big) \,, \qquad (4.6)$$

for every $h \in L^2(\Omega, \mathcal{CM}')$ belonging to the domain of $\mathcal{D}^*$.

It is well-established [Nua95, Ch. 1.3] that the Skorokhod integral has the following two important properties:

1. Every adapted process h with $\mathbf{E}\|h\|^2 < \infty$ belongs to the domain of $\mathcal{D}^*$ and the Skorokhod integral then coincides with the usual Itô integral.

2. For non-adapted processes h, if $h(s)$ belongs to the domain of $\mathcal{D}$ for almost every s and is such that

$$\mathbf{E} \int_0^t \int_0^t |\mathcal{D}_s h(r)|_{\mathbf{R}^d}^2 \, ds = \mathbf{E} \int_0^t \|\mathcal{D}_s h\|^2 \, ds \, < \infty \, ,$$

then one has the following modification of the Itô isometry:

$$\mathbf{E}\Big(\int_0^t h(s) \cdot dW(s)\Big)^2 = \mathbf{E} \int_0^t |h(s)|_{\mathbf{R}^d}^2 \, ds$$
$$+ \mathbf{E} \int_0^t \int_0^t \operatorname{tr} \mathcal{D}_s h(r) \mathcal{D}_r h(s) \, ds \, dr \, . \quad (4.7)$$

Note here that since $h(s) \in \mathbf{R}^d$, we interpret $\mathcal{D}_r h(s)$ as a $d \times d$ matrix.

4.2 Malliavin derivative of the Jacobian

By iterating the implicit functions theorem, we can see that the map that associates a given realisation of the Wiener process W to the Jacobian $J_{s,t}\varphi$ is also Fréchet (and therefore Malliavin) differentiable. Its Malliavin derivative $\mathcal{D}_h J_{s,t}\varphi$ in the direction $h \in \mathcal{CM}'$ is given by the unique solution to

$$\partial_t \mathcal{D}_h J_{s,t}\varphi = -L\mathcal{D}_h J_{s,t}\varphi + DN(u_t)\mathcal{D}_h J_{s,t}\varphi + D^2 N(u_t)(\mathcal{D}_h u_t, J_{s,t}\varphi) \, ,$$

endowed with the initial condition $\mathcal{D}_h J_{s,s}\varphi = 0$. Just as the Malliavin derivative of the solution was related to its derivative with respect to the initial condition, the Malliavin derivative of $J_{s,t}$ can be related to the second derivative of the flow with respect to the initial condition in the following way. Denoting by $J_{s,t}^{(2)}(\varphi, \psi)$ the second derivative of u_t with respect to u_0 in the directions φ and ψ, we see that as in (3.17), $J_{s,t}^{(2)}(\varphi, \psi)$ is the solution to

$$\partial_t J_{s,t}^{(2)}(\varphi, \psi) = -L J_{s,t}^{(2)}(\varphi, \psi) + DN(u_t) J_{s,t}^{(2)}(\varphi, \psi) + D^2 N(u_t)(J_{s,t}\psi, J_{s,t}\varphi) \, ,$$

endowed with the initial condition $J_{s,s}^{(2)}(\varphi, \psi) = 0$.

Assuming that h vanishes outside of the interval $[s, t]$ and using the identities $J_{r,t} J_{s,r} = J_{s,t}$ and $\mathcal{D}_h u_t = \int_s^t J_{r,t} Gh(r) \, dr$, we can check by differentiating both sides and identifying terms that one has the identity

$$\mathcal{D}_h J_{s,t}\varphi = \int_s^t J_{r,t}^{(2)}(Gh(r), J_{s,r}\varphi) \, dr \, , \quad (4.8)$$

which we can rewrite as

$$\mathcal{D}_r J_{s,t}\varphi = J_{r,t}^{(2)}(G, J_{s,r}\varphi) \quad (4.9)$$

This identity is going to be used in Section 5.

4.3 Malliavin covariance matrix

We now define and explore the properties of the Malliavin covariance matrix, whose non-degeneracy is central to our constructions.

Definition 4.4 Assume that the explosion time $\tau = \infty$ for every initial condition in $\mathcal{H}$. Then, for any $t > 0$, the Malliavin matrix $\mathcal{M}_t \colon \mathcal{H} \to \mathcal{H}$ is the linear operator defined by

$$\mathcal{M}_t \varphi = \sum_{k=1}^{d} \int_0^t \langle J_{s,t} g_k, \varphi \rangle J_{s,t} g_k \, ds \ . \tag{4.10}$$

Observe that this is equivalent to

$$\mathcal{M}_t = \mathcal{A}_t \mathcal{A}_t^* = \int_0^t J_{s,t} G G^* J_{s,t}^* \, ds = \int_0^t J_{s,t} G G^* K_{s,t} \, ds \ ,$$

thus motivating the definition $\mathcal{M}_{s,t} = \mathcal{A}_{s,t} \mathcal{A}_{s,t}^*$ for arbitrary time intervals $0 \le s < t$. From this it is clear that $\mathcal{M}_{s,t}$ is a symmetric positive operator with

$$\langle \mathcal{M}_t \varphi, \varphi \rangle = \sum_{k=1}^{d} \int_0^t \langle J_{s,t} g_k, \varphi \rangle^2 \, ds = \sum_{k=1}^{d} \int_0^t \langle g_k, K_{s,t} \varphi \rangle^2 \, ds \tag{4.11}$$

for all $\varphi \in \mathcal{H}$.

The meaning of the Malliavin covariance matrix defined in (4.10) is rather intuitive, especially for the diagonal elements $\langle \mathcal{M}_t \varphi, \varphi \rangle$. If $\langle \mathcal{M}_t \varphi, \varphi \rangle > 0$ then there exists some variation in the Wiener process on the time interval $[0, t]$ which creates a variation of u_t in the direction φ.

It is also useful to understand on what spaces the operator norm of $\mathcal{M}_t$ is bounded. As a simple consequence of Proposition 3.7, we have:

Proposition 4.5 *For every $T > 0$ and $\gamma \in [0, (1 - a) \wedge \frac{1}{2})$, $\mathcal{M}_T$ can be extended to a bounded (random) linear operator from $\mathcal{H}_{-\gamma}$ to $\mathcal{H}_\gamma$ with probability one. In particular, $\mathcal{M}_T$ is almost surely a positive, self-adjoint linear operator on $\mathcal{H}$ such that the bound*

$$\sup_{\substack{\varphi, \psi \in \mathcal{H}_{-\gamma} \\ \|\varphi\|_{-\gamma} = \|\psi\|_{-\gamma} = 1}} \langle \mathcal{M}_T \varphi, \psi \rangle \le T \, C \sup_{0 \le s < t \le T} \sup_k \, (1 + \|u_t\|)^{2n-2} \|J_{s,t} g_k\|^2$$

holds with some deterministic constant C.

Remark 4.6 If the linear operator L happens to have compact resolvent, which will be the case in most of the examples to which our theory applies, then the operator $\mathcal{M}_T$ is automatically compact, since the embedding $\mathcal{H}_\gamma \hookrightarrow \mathcal{H}$ is then compact for every $\gamma > 0$.

Proof. From (4.10) we have that

$$\sup_{\substack{\varphi, \psi \in \mathcal{H}_{-\gamma} \\ \|\varphi\|_{-\gamma} = \|\psi\|_{-\gamma} = 1}} \langle \mathcal{M}_t \varphi, \psi \rangle \le \sum_{k=1}^{d} \int_0^t \|J_{s,t} g_k\|_\gamma^2 \, ds \ .$$

Since the g_k belong to $\mathcal{H}$ by assumption, the required bound now follows from Proposition 3.7, noting that the singularity at $s = t$ is integrable by the assumption $\gamma < \frac{1}{2}$. $\qquad\square$

5 Smoothing in infinite dimensions

We now turn our study of (3.3) to one of the principal goals of this article. As in the preceding section, we shall assume that all solutions are global in time and that the standing assumptions from Assumption A.1 continue to hold. The aim of this section is to prove "smoothing" estimates for the corresponding Markov semigroup $\mathcal{P}_t$ whose action on bounded test functions $\varphi : \mathcal{H} \to \mathbf{R}$ is defined by

$$\mathcal{P}_t\varphi(v) = \mathbf{E}_v\varphi(u_t) \, .$$

Here, the subscript in the expectation refers to the initial condition for the solution u_t to (3.3). We begin with a brief discussion of the type of estimates we will prove and the ideas used in their proof. A long discussion on this can be found in [HM06] in which a number of the tools of this paper were developed or [Mat08] which has a longer motivating discussion.

Recall also that the Malliavin covariance matrix $\mathcal{M}_t : \mathcal{H} \to \mathcal{H}$ for the solution to (3.3) was defined in (4.10) as $\mathcal{M}_t = \mathcal{A}_t\mathcal{A}_t^*$ and that it is a random, self-adjoint operator on $\mathcal{H}$. Since $\mathcal{H}$ is assumed to be infinite-dimensional, $\mathcal{M}_t$ will in general not be invertible. However as discussed in the introduction we will only need it to be "approximately invertible" on some subspace paired with a assumption that the dynamics is counteractive off this subspace. The assumption of "approximate invertibility" on some subspace is formulated in Assumption B.1 below and the contractivity assumption is formulated in Assumption B.4. These are the two fundamental structural assumptions needed for this theory. In between the statement of these two assumption two other assumptions are given. They are more technical in nature and ensure that we can control various quantities.

Assumption B.1 (Malliavin matrix) *There exists a function $U \colon \mathcal{H} \to [1, \infty)$ and an orthogonal projection operator $\Pi \colon \mathcal{H} \to \mathcal{H}$ such that, for every $\alpha > 0$, the bound*

$$\mathbf{P}\Big(\inf_{\|\Pi\varphi\| \geq \alpha\|\varphi\|} \frac{\langle \varphi, \mathcal{M}_1\varphi \rangle}{\|\varphi\|^2} \leq \varepsilon \Big) \leq C(\alpha, p)\, U^p(u_0)\, \varepsilon^p \, , \tag{5.1}$$

holds for every $\varepsilon \leq 1$, $p \geq 1$ and $u_0 \in \mathcal{H}$. Furthermore for some $\bar{q} \geq 2$, there exist a constant C_U so that for every initial condition $u_0 \in \mathcal{H}$, the bound

$$\mathbf{E} U^{\bar{q}}(u_n) \leq C_U^{\bar{q}} U^{\bar{q}}(u_0) \, ,$$

holds uniformly in $n \geq 0$.

We are also going to assume in this section that the solutions to (3.3) have the following Lyapunov-type structure, which is stronger than Assumption C.1 used in the previous section:

Assumption B.2 (Lyapunov structure) *Equation (3.3) has global solutions for every initial condition. Furthermore, there exists a function $V \colon \mathcal{H} \to \mathbf{R}_+$ such that there exist constants $C_L > 0$ and $\eta' \in [0, 1)$ such that*

$$\mathbf{E} \exp(V(u_1)) \leq \exp(\eta' V(u_0) + C_L) \, . \tag{5.2}$$

Assumption B.3 (Jacobian) *The Jacobian $J_{s,t}$ and the second variation $J_{s,t}^{(2)}$ satisfy the bounds*

$$\mathbf{E}\|J_{s,t}\|^{\bar{p}} \leq \exp(\bar{p}\eta V(u_0) + \bar{p}C_J) \, ,$$

$$\mathbf{E}\|J_{s,t}^{(2)}\|^{\bar{p}} \le \exp(\bar{p}\eta V(u_0) + \bar{p}C_J^{(2)}) \,,$$

for all $0 \le s \le t \le 1$ and for some constants $\bar{p} \ge 10$ and $\eta > 0$ with $\bar{p}\eta < 1 - \eta'$ and $2/\bar{q} + 10/\bar{p} \le 1$, where η' is the constant from Assumption B.2 and $\bar{q}$ the constant from Assumption B.1.

Remark 5.1 When we write $\|J^{(2)}\|$ we mean the operator norm from $\mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$, namely $\sup_{\varphi,\psi\in\mathcal{H}} \|J^{(2)}(\varphi,\psi)\|/(\|\varphi\|\|\psi\|)$.

We finally assume that the Jacobian of the solution has some "smoothing properties" in the sense that if we apply it to a function that belongs to the image of the orthogonal complement $\Pi^{\perp} = 1 - \Pi$ of the projection operator Π then, at least for short times, its norm will on average be reduced:

Assumption B.4 (Smoothing) *One has the bound*

$$\mathbf{E}\|J_{0,1}\Pi^{\perp}\|^{\bar{p}} \le \exp(\bar{p}\eta V(u_0) - \bar{p}C_{\Pi}) \,, \tag{5.3}$$

for some constant C_{Π} such that $C_{\Pi} - C_J > 2\kappa C_L$ where $\kappa = \eta/(1-\eta')$. The constants η and $\bar{p}$ appearing in this bound are the same as the ones appearing in Assumption B.3, the constant η' is the same as the one appearing in Assumption B.2, and the projection Π is the same as the one appearing in Assumption B.1.

Remark 5.2 The condition $C_{\Pi} - C_J > 2\kappa C_L$ may seem particularly unmotivated. In the next section, we try to give some insight into its meaning.

Remark 5.3 We will see in the proof of Theorem 8.1 below that if we assume that the linear operator L has compact resolvent, then Assumption B.4 can always be satisfied by taking for Π the projection onto a sufficiently large number of eigenvectors of L.

Remark 5.4 Notice that if $\text{Range}(\Pi) \subset \text{span}\{g_1, \ldots, g_d\}$, then in light of the last representation in (4.11) it is reasonable to expect (5.1) to hold as long as one has some control over moments the modulus of continuity of $s \mapsto K_{s,t}$. (This is made more precise in Lemma 6.18.) We refer to such an assumption on the range as the "essentially elliptic" setting since all of the directions whose (pathwise) dynamics are not controlled by Assumption B.4 are directly forced.

Under these assumptions we have the following result which is the fundamental "smoothing" estimate of this paper. It is the linchpin on which all of the ergodic results rest.

Theorem 5.5 *Let Assumptions A.1 and B.1-B.4 hold. Then for any $\zeta \in [0, (C_{\Pi} - C_J)/2 - \kappa C_L)$ there a exist positive constants C such that for all $n \in \mathbf{N}$ and measurable $\varphi \colon \mathcal{H} \to \mathbf{R}$*

$$\|\mathrm{D}(\mathcal{P}_{2n}\varphi)(u)\| \le e^{4\kappa V(u_0)}\Big(CU^2(u_0)\sqrt{(\mathcal{P}_{2n}\varphi^2)(u)} + \gamma^{2n}\sqrt{(\mathcal{P}_{2n}\|\mathrm{D}\varphi\|^2)(u)}\Big) \tag{5.4}$$

where $\gamma = \exp(-\zeta)$.

Remark 5.6 By $(\mathcal{P}_t\|\mathrm{D}\varphi\|^2)(u)$, we simply mean $\mathbf{E}_u(\sup_{\|\xi\|=1}|(\mathrm{D}\varphi)(u_t)\xi|^2)$.

Remark 5.7 If $\|\varphi\|_\infty$ or $\|D\varphi\|_\infty$ are bounded by one then the corresponding terms under the square root are bounded by one. Furthermore, in light of Assumption B.2, if $\varphi(u)^2 \le \exp(V(u))$, then

$$\sqrt{\mathcal{P}_{2n}\varphi^2(u)} \le \|\varphi\|_\infty \sqrt{\mathbf{E}\exp(V(u_{2n}))} \le \|\varphi\|_\infty \exp(\eta' V(u_0)/2 + C_L/(2 - 2\eta')) \;.$$

Of course, the same bound for holds for $\sqrt{(\mathcal{P}_{2n}\|D\varphi\|^2)(u)}$, provided that one has an estimate of the type $\|D\varphi\|^2(u) \le \exp(V(u))$.

5.1 Motivating discussion

We now discuss in what sense (5.4) implies smoothing. When the term "smoothing" is used in the mathematics literature to describe a linear operator T, it usually means that $T\varphi$ belongs to a smoother function space than φ. This usually means that $T\varphi$ is "more differentiable" then φ. A convenient way to express this fact analytically would be an estimate of the form

$$\|D(T\varphi)(u)\| \le C(u)\|\varphi\|_\infty \;. \tag{5.5}$$

(Of course the "smoothing" property may improve the smoothness by less than a whole derivative, or one may consider functions φ that are not bounded, but let us consider (5.5) just for the sake of the argument.) This shows in a quantitative way that $T\varphi$ is differentiable while φ need not be. In light of Remark 5.7, this is in line with the first term on the right hand side of (5.4).

The second term on the right hand side of (5.4) embodies smoothing of a different type. Suppose that T satisfies the estimate

$$\|D(T\varphi)\|_\infty \le C\|\varphi\|_\infty + \gamma\|D\varphi\|_\infty \tag{5.6}$$

for some positive C and some $\gamma \in (0,1)$. (Note that this is a variation of what is usually referred to as the Lasota-Yorke inequality [LY73, Liv03] or the Ionescu-Tulcea-Marinescu inequality [ITM50].) Though (5.6) does not imply that $T\varphi$ belongs to a smoother function space then φ, it does imply that the gradients of $T\varphi$ are smaller then those of φ, at least as long as the gradients of φ are sufficiently steep. This is in line with a more colloquial idea of smoothing, though not in line with the traditional mathematical definition used.

5.1.1 Strongly dissipative setting

Where does the assumption $C_\Pi > C_J + 2\kappa C_L$ come from? This is easy to understand if we consider the "trivial" case $\Pi = 0$. In this case, Assumption B.1 is empty and the projection $\Pi^\perp$ is the identity. Therefore, the left hand sides from Assumptions B.3 and B.4 coincide, so that one has $C_J = -C_\Pi$ and our restriction becomes $C_J + \kappa C_L < 0$.

This turns out to be precisely the right condition to impose if one wishes to show that $\mathbf{E}\|J_{0,n}\| \to 0$ at an exponential rate:

Proposition 5.8 *Let Assumptions B.2 and B.3 hold. Then, for any $p \in [0, \bar{p}/2]$, one has the bound*

$$\mathbf{E}\|J_{0,n}\|^p \le \exp(p\kappa V(u_0) + pC_T n) \;,$$

with $\kappa = \eta/(1 - \eta')$ and $C_T = C_J + \kappa C_L$.

Proof. Using the fact that $\|J_{0,n}\| \leq \|J_{n-1,n}\|\|J_{0,n-1}\|$, we have the following recursion relation:

$$\mathbf{E}\big(\exp(p\kappa V(u_n))\|J_{0,n}\|^p\big) \leq \mathbf{E}\Big(\mathbf{E}(\exp(p\kappa V(u_n))\|J_{n-1,n}\|^p \mid \mathcal{F}_{n-1})\|J_{0,n-1}\|^p\Big)$$

$$\leq \mathbf{E}\Big(\big(\mathbf{E}(\|J_{n-1,n}\|^{\bar{p}} \mid \mathcal{F}_{n-1})\big)^{\frac{p}{\bar{p}}}\Big(\mathbf{E}\Big(\exp\Big(\frac{p\bar{p}}{\bar{p}-p}\kappa V(u_n)\Big) \,\Big|\, \mathcal{F}_{n-1}\Big)\Big)^{\frac{\bar{p}-p}{\bar{p}}}\|J_{0,n-1}\|^p\Big)$$

$$\leq e^{pC_T}\mathbf{E}(\exp(p\kappa V(u_{n-1}))\|J_{0,n-1}\|^p)\,,$$

where we made use of Assumptions B.2 and B.3 in the second inequality. It now suffices to apply this n times and to use the fact that $\|J_{0,0}\| = 1$. The assumptions $\bar{p}\kappa < 1$ and $p \leq \bar{p}/2$ ensure that $p\bar{p} \leq \bar{p} - p$ so that the bound (5.2) can be used. $\qquad\square$

We now use this estimate to prove a version of Theorem 5.5 when the system is strongly dissipative:

Proposition 5.9 *Let Assumptions B.2 and B.3 hold and set $C_T = C_J + \kappa C_L$ with $\kappa = \eta/(1-\eta')$ as before. Then, for any $\varphi : \mathcal{H} \to \mathbf{R}$ and $n \in \mathbf{N}$ one has*

$$\|\mathrm{D}(\mathcal{P}_n\varphi)(u)\| \leq \gamma^n e^{\kappa V(u)}\sqrt{\mathcal{P}_n\|\mathrm{D}\varphi\|^2(u)}\,.$$

with $\gamma = e^{C_T}$. In particular, the semigroup $\mathcal{P}_t$ has the asymptotic strong Feller property whenever $C_T < 0$.

Proof. Fixing any $\xi \in \mathcal{H}$ with $\|\xi\| = 1$, observe that

$$\mathrm{D}(\mathcal{P}_t\varphi)(u)\xi = \mathbf{E}_u(\mathrm{D}\varphi)(u_t)J_{0,t}\xi \leq \sqrt{\mathbf{E}\|J_{0,t}\|^2}\sqrt{\mathbf{E}\|\mathrm{D}\varphi\|^2(u_t)}\,.$$

Applying Proposition 5.8 completes the proof. $\qquad\square$

Comparing this result to the bound (5.4) stated in Theorem 5.5 shows that, the combination of the smoothing Assumption B.4 with Assumption B.1 on the Malliavin matrix allows us to consider the system as if its Jacobian was contracting at an average rate $(C_\Pi - C_J)/2$ instead of expanding at a rate C_J. This is precisely the rate that one would obtain by projecting the Jacobian with $\Pi^\perp$ at every second step. The additional term containing $\mathcal{P}_{2n}\varphi^2$ appearing in the right hand side of (5.4) should then be interpreted as the probabilistic "cost" of performing that projection. Since this "projection" will be performed by using an approximate inverse to the Malliavin matrix, it makes sense that the larger the lower bound on $\mathcal{M}_t$ is, the lower the corresponding probabilistic cost.

Remark 5.10 It is worth mentioning, that nothing in this section required that the number of Wiener process be finite. Hence one is free to take $d = \infty$, as long as all of the solutions and linearization are well defined (which places conditions on the g_k).

5.2 Transfer of variation

Having analyzed the strongly dissipative setting, we now turn to the general setting. We would like to mimic the calculation used in Proposition 5.9, but we do not want to require the system to be "contractive" in the sense of being strongly dissipative. However, in settings where one can prove (5.5) there is usually no requirement of strong dissipativity but rather an assumption of hypoellipticity. This is because the variation

in the initial condition is transferred to a variation in the Wiener space. Mirroring the discussion in [Mat08, HM06] (where more details can be found), we begin sketching a proof of (5.5) and then show how to modify it to obtain (5.6). The central idea is to compensate as much as possible the effect of an infinitesimal perturbation in the initial condition to an infinitesimal variation in the driving Wiener process. In short, to transfer one type of variation to another.

Denoting by $\mathcal{S} = \{\xi \in \mathcal{H} : \|\xi\| = 1\}$ the set of possible directions in $\mathcal{H}$, let there be given a map from $\mathcal{S} \times \mathcal{C}([0, \infty), \mathbf{R}^d) \to \mathcal{CM}'$ denoted by $(\xi, W) \mapsto h^\xi(W)$, mapping variations in the initial condition u to variations in the Wiener path W. We will worry about constructing a suitable map in the next sections; for the moment we just explore which properties of h^ξ might be useful. Fixing t, let us begin by assuming that the following identity holds:

$$\mathrm{D}_\xi \Phi_t^u(W) = \langle \mathrm{D}\Phi_t^u(W), \xi \rangle = \langle \mathcal{D}\Phi_t^u(W), h^\xi(W) \rangle = \mathcal{D}_{h^\xi} \Phi_t^u(W) \,. \qquad (5.7)$$

(The first and last equalities are just changes in notation.) Here, D_ξ denotes derivative with respect to the initial condition in the direction $\xi \in \mathcal{H}$, while $\mathcal{D}$ denotes the (Malliavin) derivative with respect to the noise. In words, the middle equality states that the variation in $u_t(W)$ caused by an infinitesimal shift in the initial condition in the direction ξ is equal to the variation in u_t caused by an infinitesimal shift of the Wiener process W in the direction $h^\xi(W)$. This is the basic reasoning behind smoothness estimates proved by Malliavin calculus. We begin as in the proof of Proposition 5.9. For any $\xi \in \mathcal{S}$, one has that

$$\mathrm{D}_\xi(\varphi(\Phi_t^u)) = (\mathrm{D}\varphi)(\Phi_t^u)\mathrm{D}_\xi \Phi_t^u = (\mathrm{D}\varphi)(\Phi_t^u)\mathcal{D}_{h^\xi} \Phi_t^u = \mathcal{D}_{h^\xi}(\varphi(\Phi_t^u)) \,. \qquad (5.8)$$

Taking expectations and using the Malliavin integration by parts formula (4.6) to obtain the last equality yields

$$\mathrm{D}_\xi \mathcal{P}_t \varphi(u) = \mathbf{E}\mathrm{D}_\xi(\varphi(\Phi_t^u)) = \mathbf{E}\mathcal{D}_{h^\xi}(\varphi(\Phi_t^u)) = \mathbf{E}_u \varphi(\Phi_t^u) \int_0^t h_s^\xi \cdot dW(s) \,.$$

Applying the Cauchy-Schwartz inequality to the last term produces a term of the form of the first term on the right-hand side of (5.4) provided $\mathbf{E}|\int_0^t h_s^\xi \cdot dW(s)|^2 < \infty$. Taken alone, provided one can find a mapping $(\xi, W) \mapsto h^\xi(W)$ satisfying (5.7) with $\mathbf{E}|\int_0^t h_s^\xi \cdot dW(s)| < \infty$, we have proven an inequality of the form (5.5).

In the infinite-dimensional SPDE setting of this paper, finding a map $(\xi, W) \mapsto h^\xi(W)$ satisfying (5.7) seems hopeless, unless the noise is infinite-dimensional itself and acts in a very non-degenerate way on the equation, see [Mas89, DPEZ95, EH01] or the monograph [DPZ96] for some results in this direction. Instead, we only "approximately compensate" for the variation due to differentiating in the initial direction ξ with a shift in the Wiener process. As such, given an mapping $(\xi, W) \mapsto h^\xi(W)$, we replace the requirement in (5.7) with the definition

$$\rho_t(W) = \mathrm{D}_\xi \Phi_t^u(W) - \mathcal{D}_{h^\xi} \Phi_t^u(W) \qquad (5.9)$$

and hope that we can choose h^ξ in such a way that $\rho_t \to 0$ as $t \to \infty$. As before, we postpone choosing a mapping $(\xi, W) \mapsto h^\xi(W)$ until the next section. For the moment we are content to explore the implications of finding such a mapping with desirable properties.

Returning to (5.8) but using (5.9), we now have

$$D_\xi(\varphi(\Phi_t^u)) = (D\varphi)(\Phi_t^u)D_\xi\Phi_t^u = (D\varphi)(\Phi_t^u)\mathcal{D}_{h^\xi}\Phi_t^u + (D\varphi)(\Phi_t^u)\rho_t$$
$$= \mathcal{D}_{h^\xi}(\varphi(\Phi_t^u)) + (D\varphi)(\Phi_t^u)\rho_t \ . \tag{5.10}$$

Taking expectations of both sides and applying the Malliavin integration by parts the first term on the right-hand side produces

$$D_\xi\mathcal{P}_t\varphi(u) = \mathbf{E}D_\xi(\varphi(\Phi_t^u)) = \mathbf{E}\varphi(\Phi_t^u)\int_0^t h_s^\xi \cdot dW(s) + \mathbf{E}(D\varphi)(\Phi_t^u)\rho_t$$

which in turn, after application of the Cauchy-Schwartz inequality twice, yields

$$\|D\mathcal{P}_t\varphi(u)\| \le C(t)\sqrt{(\mathcal{P}_t\varphi^2)(u)} + \Gamma(t)\sqrt{(\mathcal{P}_t\|D\varphi\|^2)(u)} \tag{5.11}$$

with $C(t) = \sqrt{\mathbf{E}|\int_0^t h_s^\xi \cdot dW(s)|^2}$ and $\Gamma(t) = \sqrt{\mathbf{E}|\rho_t|^2}$. Observe that provided that

$$\limsup_{n\in\mathbf{N}} C(n) < \infty \quad \text{and} \quad \limsup_{n\in\mathbf{N}} \Gamma(n)\gamma^{-n} < \infty \tag{5.12}$$

for some $\gamma \in (0,1)$ we will have proved Theorem 5.5. Choosing a mapping $(\xi, W) \mapsto h^\xi(W)$ so that these two conditions hold is the topic of the next four sections.

5.3 Choosing a variation h_t^ξ

As discussed in [HM06] and at length in [Mat08], if one looks for the variation h^ξ such that (5.7) holds and $\int_0^t |h_s^\xi|^2 ds$ is minimized, then the answer is $h_s^\xi = (\mathcal{A}_t^*\mathcal{M}_t^{-1}J_t\xi)(s)$ which by the observation in (4.5) is simply $h_s^\xi = G^*K_{s,t}\mathcal{M}_t^{-1}J_t\xi$. While this is not quite the correct optimisation problem to solve since its solution h^ξ is not adapted to W and hence $\mathbf{E}|\int_0^t h_s^\xi \cdot dW(s)|^2 \ne \int_0^t \mathbf{E}|u_s|^2 ds$, it is in general a good enough choice.

A bigger problem is that the space on which $\mathcal{M}_t$ can be inverted is far from evident. If the range of G was dense in $\mathcal{H}$ (which requires infinitely many driving Wiener processes), then there is some chance that $\text{Range}(J_t) \subset \text{Range}(\mathcal{M}_t)$ and the above formula for h_t could be used. This is in fact the case where the Bismut-Elworthy-Li formula is often used and which might be refereed to as "truly elliptic." It this case the system *is* in fact strong Feller. We are precisely interested in the case when only a finite number of directions are forced (or the variance decays so fast that this is effectively true). One of the fundamental ideas used in this article is that we need only effective control of the system on a finite dimensional subspace since the dynamic pathwise control embodied in Assumption B.4 can control the remaining degrees of freedom.

While Theorem 6.7 of the next section gives conditions that ensure that $\mathcal{M}_t$ is almost surely non-degenerate, it does not give much insight into the structure of the range since it only deals with finite dimensional projections. However, Assumption B.1 ensures that it is unlikely the eigenvectors with sizable projection in $\Pi\mathcal{H}$ have small eigenvalues. As long as this is true, the "regularised inverse" $(\mathcal{M}_t + \beta)^{-1}$, which always exists since $\mathcal{M}_t$ is positive definite, will be a "good inverse" for $\mathcal{M}_t$, at least on $\Pi\mathcal{H}$. This suggests that we make the choice $h_s^\xi = G^*K_{s,t}(\mathcal{M}_t + \beta)^{-1}J_t\xi$ for some very small $\beta > 0$. Observe that

$$D_\xi u_t - \mathcal{D}_{h^\xi}u_t = J_t\xi - \mathcal{M}_t(\mathcal{M}_t + \beta)^{-1}J_t\xi = \beta(\mathcal{M}_t + \beta)^{-1}J_t\xi \ , \tag{5.13}$$

which will be expected to be small as long as $J_t\xi$ has small projection (relative to the size of β) in $\Pi^\perp\mathcal{H}$. But in any case, the norm of the right hand side in (5.13) will never

exceed the norm of $J_t\xi$, so that for small values of β, $\|D_\xi u_t - \mathcal{D}_{h^\xi} u_t\|$ is expected to behave like $\|\Pi^\perp J_t\xi\|$.

Assumption B.4 precisely states that if one projects the Jacobian onto $\Pi^\perp \mathcal{H}$, then the system behaves as if it was "strongly dissipative" as in Section 5.1.1. All together, this motivates alternating between choosing $h^\xi = \mathcal{A}_{n,n+1}^* (\mathcal{M}_{n,n+1} + \beta_n)^{-1} J_{n,n+1}\rho_n$ for even n and $h^\xi \equiv 0$ on $[n, n+1]$ for odd n.

Since we will split time into intervals of length one, we introduce the following notations:
$$J_n = J_{n,n+1}\,, \quad \mathcal{A}_n = \mathcal{A}_{n,n+1}\,, \quad \mathcal{M}_n = \mathcal{M}_{n,n+1}\,.$$

We then define the map $(\xi, W) \mapsto h^\xi(W)$ recursively by

$$h_s^\xi = \begin{cases} (\mathcal{A}_{2n}^* (\beta_{2n} + \mathcal{M}_{2n})^{-1} J_t \rho_{2n})(s) & \text{for } s \in [2n, 2n+1) \text{ and } n \in \mathbf{N}\,, \\ 0 & \text{for } s \in [2n-1, 2n) \text{ and } n \in \mathbf{N}\,. \end{cases} \quad (5.14)$$

Here, as before, $\rho_0 = \xi$, $\rho_t = J_{0,t}\xi - \mathcal{A}_{0,t} h_s^\xi = D_\xi u_t - \mathcal{D}_{h^\xi} u_t$, and β_n is a sequence of positive random numbers measurable with respect to $\mathcal{F}_n$ which will be chosen later.

Observe that these definitions are not circular since the construction of h_s^ξ for $s \in [n, n+1)$ only requires the knowledge of ρ_n, which in turn depends only on h_s^ξ for $s \in [0, n)$. The remainder of this section is devoted to showing that this particular choice of h^ξ is "good" in the sense that it allows to satisfy (5.12). We are going to assume throughout this section that Assumptions A.1 and B.1-B.4 hold, so that we are in the setting of Theorem 5.5, and that h^ξ is defined as in (5.14).

5.4 Preliminary bounds and definitions

We start by a stating a few straightforward consequences of Assumption B.2:

Proposition 5.11 *For any $\alpha \leq 1$, one has the bound*

$$\mathbf{E}\exp(\alpha V(u_1)) \leq \exp(\alpha\eta' V(u_0) + \alpha C_L)\,.$$

Furthermore, for $\eta > 0$ and $p > 0$ such that $\eta p \leq 1$, one has

$$\mathbf{E}\exp(\eta p V(u_n)) \leq \exp(p\eta(\eta')^n V(u_0) + p\kappa C_L)\,.$$

Finally, setting $\kappa = \eta/(1 - \eta')$ as before, one has the bound

$$\mathbf{E}\exp\left(\eta p \sum_{k=0}^n V(u_k)\right) \leq \exp(p\kappa V(u_0) + p\kappa C_L n)\,,$$

provided that $\kappa p \leq 1$.

Proof. The first bound follows immediately from Jensen's inequality. The second and third inequalities are shown by rewriting the estimate from Assumption B.2 as

$$\mathbf{E}\big(\exp(\eta p V(u_n))|\mathcal{F}_{n-1}\big) \leq \exp\big(\eta p\eta' V(u_{n-1}) + \eta p C_L\big)\,,$$

and iterating it. $\qquad\square$

Similarly, we obtain a bound on the Jacobian and on the Malliavin derivative $\mathcal{A}_n$ of the solution flow between times n and $n+1$:

Proposition 5.12 *For any $p \in [0, \bar{p}]$, one has*

$$\sup_{n \leq s < t \leq n+1} \mathbf{E}\|J_{s,t}\|^p \leq \exp(p(\eta')^n \eta V(u_0) + pC_J + p\kappa C_L) \tag{5.15}$$

$$\mathbf{E}\|\mathcal{A}_n\|^p \leq \|G\|^p \exp(p\eta(\eta')^n V(u_0) + p\kappa C_L + pC_J) . \tag{5.16}$$

Furthermore, (5.15) also holds for $J_{s,t}^{(2)}$ with C_J replaced by $C_J^{(2)}$.

Proof. We only need to show the bound for $p = \bar{p}$, since lower values follow again from Jensen's inequality. The bound (5.15) is an immediate consequence of Assumption B.2 and Proposition 5.11. The second bound follows by writing

$$\|\mathcal{A}_n h\|^p = \left\| \int_n^{n+1} J_{r,n+1} G h_r dr \right\|^p$$

$$\leq \|G\|^p \Big(\int_n^{n+1} \|J_{r,n+1}\|^2 dr \Big)^{\frac{p}{2}} \Big(\int_n^{n+1} |h_r|^2 dr \Big)^{\frac{p}{2}}$$

$$\leq \|G\|^p \Big(\int_n^{n+1} \|J_{r,n+1}\|^p dr \Big) \|h\|_n^p ,$$

and then applying the first bound. $\qquad\square$

In addition to these first Malliavin derivatives, we will need the control of the derivative of various objects involving the Malliavin derivative. The following lemma gives control over two objects related to the second Malliavin derivative:

Lemma 5.13 *For all $p \in [0, \bar{p}/2]$, one has the bounds*

$$\sup_{s,r \in [n,n+1]} \mathbf{E}\|\mathcal{D}_s^i J_{r,n+1}\|^p \leq \exp(2p\eta(\eta')^n V(u_0) + 2p\kappa C_L + pC_J + pC_J^{(2)}) ,$$

$$\sup_{s \in [n,n+1]} \mathbf{E}\|\mathcal{D}_s^i \mathcal{A}_n\|^p \leq \|G\|^p \exp(2p\eta(\eta')^n V(u_0) + 2p\kappa C_L + pC_J + pC_J^{(2)}) .$$

Proof. For this, we note that by (4.9) one has the identities

$$\mathcal{D}_s^i J_{r,n+1}\xi = \begin{cases} J_{s,n+1}^{(2)}(J_{r,s}\xi, g_i) & \text{for } r \leq s, \\ J_{r,n+1}^{(2)}(J_{s,r}g_i, \xi) & \text{for } s \leq r, \end{cases}$$

$$\mathcal{D}_s^i \mathcal{A}_n v = \int_n^{n+1} \mathcal{D}_s^i J_{r,n+1} G v_r \, dr .$$

Hence if $p \in [0, \bar{p}/2]$ (which by the way also ensures that $2p\kappa < 1$) it follows from Proposition 5.12 that

$$\mathbf{E}\|\mathcal{D}_s^i J_{r,n+1}\|^p \leq (\mathbf{E}\|J_{s,n+1}^{(2)}\|^{2p} \, \mathbf{E}\|J_{r,s}\|^{2p})^{\frac{1}{2}} \leq \mathbf{E}\exp(2p\eta V(u_n) + pC_J + pC_J^{(2)})$$

$$\leq \exp(2p\eta(\eta')^n V(u_0) + 2p\kappa C_L + pC_J + pC_J^{(2)})$$

for $r \leq s$ and similarly for $s \leq r$. Since, for $p \geq 1$, we can write

$$\mathbf{E}\|\mathcal{D}_s^i \mathcal{A}_n\|^p \leq \|G\|^p \int_n^{n+1} \mathbf{E}\|\mathcal{D}_s^i J_{r,n+1}\|^p dr ,$$

the second estimate then follows from the first one. $\qquad\square$

5.5 Controlling the error term ρ_t

The purpose of this section is to show that the "error term" $\rho_t = D_\xi u_t - \mathcal{D}_{h^\xi} u_t$ goes the zero as $t \to \infty$, provided that the "control" h^ξ is chosen as explained in Section 5.3. We begin by observing that for even integer times, ρ_n is given recursively by

$$\rho_{2n+2} = J_{2n+1}\rho_{2n+1} = J_{2n+1}\mathcal{R}_{2n}^{\beta_{2n}} J_{2n}\rho_{2n} \,, \tag{5.17}$$

where $\mathcal{R}_k^\beta$ is the operator

$$\mathcal{R}_k^\beta \stackrel{\text{def}}{=} 1 - \mathcal{M}_k(\beta + \mathcal{M}_k)^{-1} = \beta(\beta + \mathcal{M}_k)^{-1} \,.$$

Observe that $\mathcal{R}_k^\beta$ measures the error between $\mathcal{M}_k(\beta + \mathcal{M}_k)^{-1}$ and the identity, which we will see is small for β very small. This recursion is of the form $\rho_{2n+2} = \Xi_{2n+2}\rho_{2n}$, with the (random) operator $\Xi_{2n+2}\colon \mathcal{H} \to \mathcal{H}$ defined by $\Xi_{2n+2} = J_{2n+1}\mathcal{R}_{2n}^{\beta_{2n}} J_{2n}$. Notice that Ξ_{2n} is $\mathcal{F}_{2n}$-measurable and that Ξ_k is defined only for even integers k. Define the n-fold product of the Ξ_{2k} by

$$\Xi^{(2n)} = \prod_{k=1}^{n} \Xi_{2k} \,,$$

so that $\rho_{2n} = \Xi^{(2n)}\xi$.

It is our aim to show that it is possible under the assumptions of Section 5 to choose the sequence β_n in an adapted way such that for a sufficiently small constant $\bar\eta$ and $p \in [0, \bar{p}/2]$ one has

$$\mathbf{E}\|\rho_{2n}\|^p \leq \mathbf{E}(\|\Xi^{(2n)}\|^p)\|\rho_0\|^p \leq \exp(p\bar\eta V(u_0) - pn\tilde\kappa)\|\rho_0\|^p \,. \tag{5.18}$$

for some $\tilde\kappa > 0$. This will give the needed control over the last term in (5.11).

By Assumption B.1, we have a bound on the Malliavin covariance matrix of the form

$$\mathbf{P}\Big(\inf_{\|\Pi\varphi\|\geq\alpha\|\varphi\|} \langle\varphi, \mathcal{M}_k\varphi\rangle \leq \varepsilon\|\varphi\|^2 \,\Big|\, \mathcal{F}_k\Big) \leq C(\alpha, p)\, U^p(u_k)\, \varepsilon^p \,. \tag{5.19}$$

Here, by the Markov property, the quantities ε and α do not necessarily need to be constant, but are allowed to be $\mathcal{F}_k$-measurable random variables.

In order to obtain (5.18), the idea is to decompose Ξ_{2n+2} as

$$\begin{aligned}
\Xi_{2n+2} &= J_{2n+1}\mathcal{R}_{2n}^{\beta_{2n}} J_{2n} = (J_{2n+1}\Pi^\perp)\mathcal{R}_{2n}^{\beta_{2n}} J_{2n} + J_{2n+1}(\Pi\mathcal{R}_{2n}^{\beta_{2n}})J_{2n} \\
&\stackrel{\text{def}}{=} I_{2n+2,1} + I_{2n+2,2} \,.
\end{aligned} \tag{5.20}$$

The crux of the matter is controlling the term $\Pi\mathcal{R}_{2n}^{\beta_{2n}}$ since $J_{2n+1}\Pi^\perp$ is controlled by Assumption B.4 and we know that $\|\mathcal{R}_{2n}^{\beta_{2n}}\| \leq 1$. To understand and control the $I_{2n+2,2}$ term, we explore the properties of a general operator of the form of $\mathcal{R}_{2n}^\beta$.

Lemma 5.14 *Let Π be an orthogonal projection on $\mathcal{H}$ and M be a self-adjoint, positive linear operator on $\mathcal{H}$ satisfying for some $\gamma > 0$ and $\delta \in (0, 1]$*

$$\inf_{\xi \in \Lambda_\delta} \frac{\langle M\xi, \xi\rangle}{\|\xi\|^2} \geq \gamma \,, \tag{5.21}$$

where $\Lambda_\delta = \{\xi : \|\Pi\xi\| \geq \delta\|\xi\|\}$. Then, defining $R = 1 - M(\beta + M)^{-1} = \beta(\beta + M)^{-1}$ for some $\beta > 0$, one has $\|\Pi R\| \leq \delta \vee \sqrt{\beta/\gamma}$.

Proof. Since $\|R\| \leq 1$, for $R\xi \in \Lambda_\delta^c$ one has

$$\frac{\|\Pi R\xi\|}{\|\xi\|} \leq \frac{\|\Pi R\xi\|}{\|R\xi\|} \leq \delta \ .$$

Now for $R\xi \in \Lambda_\delta$, we have by assumption (5.21)

$$\gamma \frac{\|\Pi R\xi\|^2}{\|\xi\|^2} \leq \gamma \frac{\|R\xi\|^2}{\|\xi\|^2} \leq \frac{\langle M R\xi, R\xi \rangle}{\|\xi\|^2} \leq \frac{\langle (M+\beta)R\xi, R\xi \rangle}{\|\xi\|^2} = \beta \frac{\langle \xi, R\xi \rangle}{\|\xi\|^2} \leq \beta \ .$$

Combining both estimates gives the required bound. $\qquad\square$

This result can be applied almost directly to our setting in the following way:

Corollary 5.15 *Let $M(\omega)$ be a random operator satisfying the conditions of Lemma 5.14 almost surely for some random variable γ. If we choose β such that, for some (deterministic) $\delta \in (0,1)$, $p \geq 1$ and $C > 0$, one has the bound $\mathbf{P}(\beta \geq \delta^2 \gamma) \leq C\delta^p$, then $\mathbf{E}\|\Pi R\|^p \leq (1+C)\delta^p$.*

In particular, for any $\delta \in (0,1)$, setting

$$\beta_k = \frac{\delta^3}{U(u_k) C(\delta, \bar{p})^{\frac{1}{p}}} \ , \tag{5.22}$$

where C is the constant from (5.19), produces the bound $\mathbf{E}(\|\Pi \mathcal{R}_{2n}^{\beta_{2n}}\|^p | \mathcal{F}_{2n}) \leq 2\delta^p$, valid for every $p \leq \bar{p}$.

Proof. To see the first part define $\Omega_0 = \{\omega : \beta(\omega) \leq \delta^2 \gamma(\omega)\}$. It the follows from Lemma 5.14, the fact that $\|\Pi R\| \leq 1$ and the assumption $\mathbf{P}(\Omega_0^c) \leq C\delta^p$, that

$$\mathbf{E}\|\Pi R\|^p \leq \mathbf{E}\left(\left(\delta \vee \sqrt{\frac{\beta}{\gamma}} \right)^p \mathbf{1}_{\Omega_0} + \mathbf{1}_{\Omega_0^c} \right) \leq \delta^p + \mathbf{P}(\Omega_0^c) \leq (1+C)\delta^p \ , \tag{5.23}$$

as required.

To obtain the second statement, it is sufficient to consider (5.19) with $\varepsilon = \beta_{2n}/\delta^2$, so that one can take for γ the random variable equal to ε on the set for which the bound (5.19) holds and 0 on its complement. It then follows from the choice (5.22) for β_{2n} that the assumption for the first part are satisfied with $C = 1$ and $p = \bar{p}$, so that the statement follows. $\qquad\square$

We now introduce a "compensator"

$$\chi_{2n+2} = \exp\left(\eta V(u_{2n+1}) + \eta V(u_{2n}) \right) \ ,$$

and, in analogy to before, we set $\chi^{(2n)} = \prod_{k=1}^n \chi_{2k}$. Proposition 5.11 implies that for any $p \in [0, \bar{p}]$

$$\mathbf{E}(\chi^{(2n)})^p \leq \exp(p\kappa V(u_0) + p\kappa C_L 2n) \ , \tag{5.24}$$

where $\kappa = \eta/(1-\eta')$. Note that Assumption B.3 made sure that η is sufficiently small so that $\kappa\bar{p} \leq 1$. With these preliminaries complete, we now return to the analysis of (5.20).

Lemma 5.16 *For any $\varepsilon > 0$ and $p \in [0, \bar{p}/2]$, there exists a $\delta > 0$ sufficiently small so that if one chooses β_n as in Corollary 5.15 and η such that $\kappa \bar{p} \le 1$, one has*

$$\mathbf{E}(\|\Xi_{2n+2}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n}) \le \exp(pC_J - pC_\Pi + \varepsilon p) .$$

Proof. Since for every $\varepsilon > 0$ there exists a constant C_ε such that $|x+y|^p \le e^{p\varepsilon/2}|x|^p + C_\varepsilon^p |y|^p$, recalling the definition of $I_{2n+2,1}$ and $I_{2n+2,2}$ from (5.20) we have that

$$\mathbf{E}(\|\Xi_{2n+2}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n}) \le e^{\varepsilon p/2} \mathbf{E}(\|I_{2n+2,1}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n})$$
$$+ C_\varepsilon^p \mathbf{E}(\|I_{2n+2,2}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n}) .$$

We begin with the first term since it is the most straightforward one. Using the fact that $\|\mathcal{R}_{2n}^{\beta_{2n}}\| \le 1$ and that $\bar{p}\eta < 1 - \eta'$ by the assumption on η, we obtain from Assumptions B.2 and B.3 that

$$\mathbf{E}(\|I_{2n+2,1}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n}) \le \exp(-p\eta V(u_{2n})) \mathbf{E}\Big(\mathbf{E}(\|J_{2n+1}\Pi^\perp\|^p | \mathcal{F}_{2n+1})$$
$$\times \exp(-p\eta V(u_{2n+1})) \|J_{2n}\|^p \Big| \mathcal{F}_{2n} \Big)$$
$$\le \exp(pC_J - pC_\Pi) .$$

Turning to the second term, we obtain for any $\delta \in (0, 1)$ the bound

$$\mathbf{E}(\|I_{2n+2,2}\|^p \chi_{2n+2}^{-p} | \mathcal{F}_{2n}) \le \exp(-p\eta V(u_{2n})) \sqrt{\mathbf{E}(\|\Pi \mathcal{R}_{2n}^{\beta_{2n}}\|^{2p} | \mathcal{F}_{2n})}$$
$$\times \sqrt{\mathbf{E}\Big(\mathbf{E}(\|J_{2n+1}\|^{2p} | \mathcal{F}_{2n+1}) \exp(-2p\eta V(u_{2n+1})) \|J_{2n}\|^{2p} \Big| \mathcal{F}_{2n} \Big)}$$
$$\le \exp(p2C_J) \delta^p \sqrt{2} ,$$

provided that we choose β_n as in Corollary 5.15. Choosing now δ sufficiently small (it suffices to choose it such that $\delta^p \le \frac{\varepsilon p}{2\sqrt{2}} C_\varepsilon^{-p} e^{-pC_J - pC_\Pi}$ for every $p \le \bar{p}/2$) we obtain the desired bound. □

Combining Lemma 5.16 with (5.24), we obtain the needed result which ensures that the "error term" ρ_t from (5.11) goes to zero.

Lemma 5.17 *For any $p \in [0, \bar{p}/4]$ and $\tilde{\kappa} \in [0, C_\Pi - C_J - 2\kappa C_L)$ there exists a choice of the β_n of the form (5.22) so that*

$$\mathbf{E}\|\Xi^{(2n+2)}\|^p \le \exp\left(p\kappa V(u_0) - p\tilde{\kappa}n\right) ,$$

for all $u_0 \in \mathcal{H}$.

Proof. Since

$$\mathbf{E}\|\Xi^{(2n+2)}\|^p \le \left(\mathbf{E}\|\Xi^{(2n+2)}\|^{2p}(\chi^{(2n+2)})^{-2p}\right)^{\frac{1}{2}} \left(\mathbf{E}(\chi^{(2n+2)})^{2p}\right)^{\frac{1}{2}} ,$$

the result follows by combining Lemma 5.16 with (5.24). □

5.6 Controlling the size of the variation h_t^ξ

We now turn to controlling the size of

$$\mathbf{E}\left|\int_0^n \langle h_s^\xi, dW_s\rangle\right|^2 , \tag{5.25}$$

uniformly as $n \to \infty$. We assume throughout this section that h_t^ξ was constructed as in Section 5.3 with β_n as in (5.22).

Since our choice of h_s^ξ is not adapted to the W_s, this does *not* follow from a simple application of Itô's isometry. However, the situation is not as bad as it could be, since the control is "block adapted." By this we mean that h_n is adapted to $\mathcal{F}_n$ for every integer value of n. For non-integer values $t \in (n, n+1]$, h_t has no reason to be $\mathcal{F}_t$-measurable in general, but it is nevertheless $\mathcal{F}_{n+1}$-measurable. The stochastic integral in (5.25) is accordingly not an Itô integral, but a Skorokhod integral. Hence to estimate (5.25) we must use its generalization given in (4.7) which produces

$$\mathbf{E}\left|\int_0^{2n} \langle h_s^\xi, dW_s\rangle\right|^2 \leq \mathbf{E}\|h^\xi\|_{[0,2n]}^2 + \sum_{k=0}^{n-1} \int_{2k}^{2k+1} \int_{2k}^{2k+1} \mathbf{E}\|\mathcal{D}_s h_t\|_{\mathrm{HS}}^2 \, ds \, dt \tag{5.26}$$

where $\|f\|_I^2 = \int_I |f(s)|^2 \, ds$ and $\|M\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm on linear operators from $\mathbf{R}^d$ to $\mathbf{R}^d$. We see the importance of the "block adapted" structure of h_s. If not for this structure, the integrand appearing in the second term above would need to decay both in s and t to be finite.

The main result of this section is

Proposition 5.18 *Let Assumptions B.1–B.4 hold. Then, if one chooses β_n as in (5.22), there exists a constant $C > 0$ such that*

$$\lim_{n\to\infty} \mathbf{E}\left|\int_0^n \langle h_s^\xi, dW_s\rangle\right|^2 \leq C \exp((8\eta + 2\kappa)V(u_0))U^2(u_0)\|\xi\|^2 .$$

Proof of Proposition 5.18. In the interest of brevity we will set $\widetilde{\mathcal{M}}_n = \mathcal{M}_n + \beta_n$ and $I_n = [n, n+1]$. We will also write $\|h\|_I$ for the norm on $L^2(I, \mathbf{R}^d)$ viewed as a subset of $\mathcal{CM}'$ and we will use $\|\cdot\|$ and $\|\cdot\|_I$ to denote respectively the induced operator norm on linear maps from $\mathcal{H}$ to $\mathcal{H}$ and $\mathcal{CM}'$ to $\mathcal{H}$. Hopefully without too much confusion, we will also use $\|\cdot\|_I$ to denote the induced operator norm on linear maps from $\mathcal{H}$ to $\mathcal{CM}'$. In all cases, we will further abbreviate $\|h\|_{I_n}$ to $\|h\|_n$.

Observe now that the definitions of $\widetilde{\mathcal{M}}_n$ and $\mathcal{A}_n$ imply the following almost sure bounds:

$$\|\widetilde{\mathcal{M}}_n^{-1/2}\mathcal{A}_n\|_n \leq 1 , \qquad \|\mathcal{A}_n^*\widetilde{\mathcal{M}}_n^{-1/2}\|_n \leq 1 , \qquad \|\widetilde{\mathcal{M}}_n^{-1/2}\| \leq \beta_n^{-1/2} . \tag{5.27}$$

We start by bounding the first term on the right hand-side of (5.26). Observe that

$$\|h\|_{[0,2n]}^2 = \sum_{k=0}^n \|h\|_{2k}^2 . \tag{5.28}$$

Using the bound on $\mathcal{A}_k^*\widetilde{\mathcal{M}}_k^{-1}$ from (5.27), we obtain

$$\|h\|_{2k} = \|\mathcal{A}_{2k}^*\widetilde{\mathcal{M}}_{2k}^{-1}J_{2k}\rho_{2k}\|_{2k} \leq \beta_{2k}^{-1/2}\|J_{2k}\|\|\rho_{2k}\| .$$

By our assumption that $10/\bar{p} + 2/\bar{q} \leq 1$ we can find $1/q + 1/r + 1/p = 1$ with $q \leq \bar{q}$, $2r \leq \bar{p}$ and $2p \leq \bar{p}$. By the Hölder inequality we thus have

$$\mathbf{E}\|h\|_{2k}^2 \leq (\mathbf{E}\beta_{2k}^{-q})^{1/q}(\mathbf{E}\|J_{2k}\|^{2r})^{1/r}(\mathbf{E}\|\rho_{2k}\|^{2p})^{1/p} \,.$$

From Proposition 5.12, Assumption B.1 and Lemma 5.17, we obtain the existence of a positive constant C (depending only on the choice made for $\tilde{\kappa}$ and on the bounds given by our standing assumptions) such that one has the bounds

$$
\begin{aligned}
(\mathbf{E}\|J_{2k}\|^{2r})^{1/r} &\leq \exp(2\eta(\eta')^{2k}V(u_0) + 2C_J + 2\kappa C_L) \,, \\
(\mathbf{E}\beta_{2k}^{-q})^{1/q} &\leq CU(u_0) \,, \\
(\mathbf{E}\|\rho_{2k}\|^{2p})^{1/p} &\leq \exp(2\kappa V(u_0) - 2\tilde{\kappa}k)\|\xi\|^2 \,.
\end{aligned}
\tag{5.29}
$$

combining these bounds and summing over k yields

$$\mathbf{E}\|h\|_{[0,2n]}^2 \leq CU(u_0)\exp(2(\eta + \kappa)V(u_0))\|\xi\|^2 \,, \tag{5.30}$$

uniformly in $n \geq 0$.

We now turn to bound the second term on the right hand side of (5.26). Since the columns of the matrix representation of the integrand are just $\mathcal{D}_s^i$, the ith component of the Malliavin derivative, we have

$$\int_{2k}^{2k+1} \int_{2k}^{2k+1} \|\mathcal{D}_s h_t\|_{\mathrm{HS}}^2 \, ds \, dt = \sum_{i=1}^m \int_{2k}^{2k+1} \|\mathcal{D}_s^i h\|_{2k}^2 \, ds \,. \tag{5.31}$$

From the definition of h_t, Lemma 5.13, the relation $\widetilde{\mathcal{M}}_{2k} = \mathcal{A}_{2k}\mathcal{A}_{2k}^* + \beta_{2k}$, and the fact that both ρ_{2k} and β_{2k} are $\mathcal{F}_{2k}$-measurable, we have that for fixed $s \in I_{2k}$, $\mathcal{D}_s^i h$ is an element of $L^2(I_{2k}, \mathbf{R}) \subset \mathcal{CM}'$ with:

$$
\begin{aligned}
\mathcal{D}_s^i h &= (\mathcal{D}_s^i \mathcal{A}_{2k}^*)\widetilde{\mathcal{M}}_{2k}^{-1} J_{2k}\rho_{2k} + \mathcal{A}_{2k}^*\widetilde{\mathcal{M}}_{2k}^{-1}(\mathcal{D}_s^i J_{2k})\rho_{2k} \\
&\quad - \mathcal{A}_{2k}^*\widetilde{\mathcal{M}}_{2k}^{-1}((\mathcal{D}_s^i \mathcal{A}_{2k})\mathcal{A}_{2k}^* + \mathcal{A}_{2k}(\mathcal{D}_s^i \mathcal{A}_{2k}^*))\widetilde{\mathcal{M}}_{2k}^{-1} J_{2k}\rho_{2k} \,.
\end{aligned}
\tag{5.32}
$$

For brevity we suppress the subscripts k on the operators and norms for a moment. It then follows from (5.27) that one has the almost sure bounds

$$
\begin{aligned}
\|\widetilde{\mathcal{M}}^{-1}\mathcal{A}\| &\leq \|\widetilde{\mathcal{M}}^{-1/2}\|\|\widetilde{\mathcal{M}}^{-1/2}\mathcal{A}\| \leq \beta^{-1/2} \,, \\
\|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}\| &\leq \|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1/2}\|\|\widetilde{\mathcal{M}}^{-1/2}\| \leq \beta^{-1/2} \,, \\
\|(\mathcal{D}_s^i\mathcal{A}^*)\widetilde{\mathcal{M}}^{-1}J\| &\leq \|\mathcal{D}_s^i\mathcal{A}^*\|\|\widetilde{\mathcal{M}}^{-1}\|\|J\| \leq \beta^{-1}\|\mathcal{D}_s^i\mathcal{A}\|\|J\| \,, \\
\|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}(\mathcal{D}_s^i J)\| &\leq \|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}\|\|\mathcal{D}_s^i J\| \leq \beta^{-1/2}\|\mathcal{D}_s^i J\| \,.
\end{aligned}
$$

In particular, this yields the bounds

$$
\begin{aligned}
\|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}(\mathcal{D}_s^i\mathcal{A})\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}J\| &\leq \|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}\|^2\|\mathcal{D}_s^i\mathcal{A}\|\|J\| \leq \beta^{-1}\|\mathcal{D}_s^i\mathcal{A}\|\|J\| \\
\|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1}\mathcal{A}(\mathcal{D}_s^i\mathcal{A}^*)\widetilde{\mathcal{M}}^{-1}J\| &\leq \|\mathcal{A}^*\widetilde{\mathcal{M}}^{-1/2}\|^2\|\mathcal{D}_s^i\mathcal{A}^*\|\|\widetilde{\mathcal{M}}^{-1}\|\|J\| \\
&\leq \beta^{-1}\|\mathcal{D}_s^i\mathcal{A}\|\|J\| \,.
\end{aligned}
$$

Applying all of these estimates to (5.32) we obtain the bound

$$\|\mathcal{D}_s^i h\|_{2k} \leq 3\beta_{2k}^{-1}\|\mathcal{D}_s^i\mathcal{A}_{2k}\|_{2k}\|J_{2k}\|\|\rho_{2k}\| + \beta_{2k}^{-1/2}\|\mathcal{D}_s^i J_{2k}\|\|\rho_{2k}\| \,.$$

The assumption that $10/\bar{p} + 2/\bar{q} \leq 1$ ensures that we can find $q \leq \bar{q}/2$, $r \leq \bar{p}/2$ and $p \leq \bar{p}/4$ with $1/r + 2/p + 1/q = 1$. Applying Hölder's inequality to the preceding products yields:

$$\mathbf{E}\|\mathcal{D}_s^i h\|_{2k}^2 \leq 18(\mathbf{E}\beta_{2k}^{-2q})^{1/q}(\mathbf{E}\|\mathcal{D}_s^i \mathcal{A}_{2k}\|_{2k}^{2p}\mathbf{E}\|\rho_{2k}\|^{2p})^{1/p}(\mathbf{E}\|J_{2k}\|^{2r})^{1/r}$$
$$+ 2(\mathbf{E}\beta_{2k}^{-q})^{1/q}(\mathbf{E}\|\mathcal{D}_s^i J_{2k}\|^{2p}\mathbf{E}\|\rho_{2k}\|^{2p})^{1/p} .$$

We now use previous estimates to control each term. From Lemma 5.13 and Proposition 5.12, we have the bounds

$$(\mathbf{E}\|\mathcal{D}_s^i J_{2k}\|^{2p})^{1/p} \leq \exp(4\eta(\eta')^{2k}V(u_0) + 4\kappa C_L + 2C_J + 2C_J^{(2)}) ,$$
$$(\mathbf{E}\|\mathcal{D}_s^i \mathcal{A}_{2k}\|_{2k}^{2p})^{1/p} \leq \|G\|^2 \exp(4\eta(\eta')^{2k}V(u_0) + 4\kappa C_L + 2C_J + 2C_J^{(2)}) .$$

Recall furthermore the bounds on ρ_{2k} and J_{2k} already mentioned in (5.29). Lastly, from Assumption B.1 we have that, similarly as before, there exists a positive constant C such that

$$(\mathbf{E}\beta_{2k}^{-q})^{1/q} \leq (\mathbf{E}\beta_{2k}^{-2q})^{1/q} \leq CU^2(u_0) .$$

Combining all of these estimates produces

$$\sum_{i=1}^m \int_{2k}^{2k+1} \|\mathcal{D}_s^i h\|_{2k}^2 \, ds \leq C \exp((8\eta + 2\kappa)V(u_0))U^2(u_0) ,$$

for some different constant C depending only on $C_J, C_J^{(2)}, C_L, \eta, \kappa, \tilde{\kappa}$ and the choice of δ in (5.22). Combining this estimate with (5.30) and (5.26) concludes the proof. □

6   Spectral properties of the Malliavin matrix

The results in this section build on the ideas and techniques from [MP06] and [BM07]. In the first, the specific case of the 2D-Navier Stokes equation was studied using similar ideas. The time reversed representation of the Malliavin matrix used there is also the basis of our analysis here (see also [Oco88]). In the context for the 2D-Navier Stokes equations, a result analogous to Theorem 6.7 was proven. As here, one of the key results needed is a connection between the typical size of a non-adapted Wiener polynomial and the typical size of its coefficients. In [MP06], since the non-linearity was quadratic, only Wiener polynomials of degree one were considered and the calculations and formulation were made a coordinate dependent fashion. In [BM07], the calculations were reformulated in a basis free fashion which both made possible the extension to more complicated non-linearities and the inclusion of forcing which was not diagonal in the chosen basis. Furthermore in [BM07], a result close to Theorem 6.7 was claimed. Unfortunately, the auxiliary Lemma 9.12 in that article contains a mistake, which left the proof of this result incomplete.

That being said, the techniques and presentation used in this and the next section build on and refine those from [BM07]. One technical, but important, distinction between Theorem 6.7 and the preceding versions is that Theorem 6.7 allows for rougher test functions. This is accomplished by allowing $K_{t,T}$ to have a singularity in a certain interpolation norm as $t \to T$. See equation (6.3a) for the precise form. This extension is important in correcting an error in [HM06] which requires control of the Malliavin matrix of a type given by Theorem 6.7, that is with test functions rougher than those

allowed in [MP06]. Indeed, the second inequality in equation (4.25) of [HM06] is not justified, since the operator M_0 is only selfadjoint in L^2 and not in H^1. Theorem 6.7 rectifies the situation by dropping the requirement to work with H^1 completely.

6.1 Bounds on the dynamic

As the previous sections have shown, it is sufficient to have control on the moments of u and J in $\mathcal{H}$ to control their moments in many stronger norms. This motivates the next assumption. For the entirety of this section we fix a $T_0 > 0$.

Assumption C.1 *There exists a continuous function $\Psi_0 \colon \mathcal{H} \to [1, \infty)$ such that, for every $T \in (0, T_0]$ and every $p \geq 1$ there exists a constant C such that*

$$\mathbf{E} \sup_{T \leq t \leq 2T} \|u_t\|^p \leq C\Psi_0^p(u_0) \,,$$

$$\mathbf{E} \sup_{T \leq s < t \leq 2T} \|J_{s,t}\|^p \leq C\Psi_0^p(u_0) \,,$$

for every $u_0 \in \mathcal{H}$. Here, $\|J\|$ denotes the operator norm of J from $\mathcal{H}$ to $\mathcal{H}$.

Under this assumption, we immediately obtain control over the adjoint $K_{s,t}$.

Proposition 6.1 *Under Assumption C.1 for every $T \in (0, T_0]$ and every $p \geq 1$ there exists a constant C such that*

$$\mathbf{E} \sup_{T \leq s < t \leq 2T} \|K_{s,t}\|^p \leq C\Psi_0^p(u_0) \,,$$

for every $u_0 \in \mathcal{H}$.

Proof. By Proposition 3.10 we know that $K_{s,t}$ is the adjoint of $J_{s,t}$ in $\mathcal{H}$. Combined with Assumption C.1 this implies the result. $\square$

In the remainder of this section, we will study the solution to (3.1) away from $t = 0$ and up to some terminal time T which we fix from now on. We also introduce the interval $I_\delta = [\frac{T}{2}, T - \delta]$ for some $\delta \in (0, \frac{T}{4}]$ to be determined later. Given u_t a solution to (3.1), we also define a process v_t by $v_t = u_t - GW(t)$, which is more regular in time. Using Assumption C.1 and the a priori estimates from the previous sections, we obtain:

Proposition 6.2 *Let Assumption C.1 hold and Ψ_0 be the function introduced there. For any fixed $\gamma < \gamma_\star$ and $\beta < \beta_\star$ there exists a positive q so that if $\Psi = \Psi_0^q$ then the solutions to (3.1) satisfy the following bounds for every initial condition $u_0 \in \mathcal{H}$:*

$$\mathbf{E} \sup_{t \in I_\delta} \|u_t\|_{\gamma+1}^p \leq C_p \Psi^p(u_0) \,, \tag{6.1a}$$

$$\mathbf{E} \sup_{t \in I_\delta} \|\partial_t v_t\|_\gamma^p \leq C_p \Psi^p(u_0) \,. \tag{6.1b}$$

Furthermore, its linearization $J_{0,t}$ is bounded by

$$\mathbf{E} \sup_{t \in I_\delta} \sup_{\|\varphi\| \leq 1} \|J_{0,t}\varphi\|_{\gamma+1}^p \leq C_p \Psi^p(u_0) \,, \tag{6.2a}$$

$$\mathbf{E} \sup_{t \in I_\delta} \sup_{\|\varphi\| \leq 1} \|\partial_t J_{0,t}\varphi\|_\gamma^p \leq C_p \Psi^p(u_0) \,. \tag{6.2b}$$

Finally, the adjoint $K_{t,T}$ to the linearization satisfies the bounds

$$\mathbf{E} \sup_{t \in I_\delta} \sup_{\|\varphi\| \le 1} \|K_{t,T}\varphi\|_{\beta+1}^p \le \frac{C_p \Psi^p(u_0)}{\delta^{\bar{p}_\beta p}} , \tag{6.3a}$$

$$\mathbf{E} \sup_{t \in I_\delta} \sup_{\|\varphi\| \le 1} \|\partial_t K_{t,T}\varphi\|_{\beta}^p \le \frac{C_p \Psi^p(u_0)}{\delta^{\bar{p}_\beta p}} , \tag{6.3b}$$

where $\bar{p}_\beta$ is as in Proposition 3.9. In all these bounds, C_p is a constant depending only on p and on the details of the equation (3.1).

Remark 6.3 One can assume without loss of generality, and we will do so from now on, that the exponent q defining Ψ is greater or equal to n, the degree of the nonlinearity. This will be useful in the proof of Lemma 6.16 below.

Proof. It follows immediately from Assumption C.1 that

$$\mathbf{E} \sup_{t \in [T/4,T]} \|u_t\|^p \le C \Psi_0^p(u_0) .$$

Combining this with Proposition 3.6 yields the first of the desired bounds with $q = p_\gamma$. Here, Ψ_0 is as in Assumption C.1 and p_γ is as in Proposition 3.6.

Turning to the bound on $\partial_t v_t$, observe that v satisfies the random PDE

$$\partial_t v_t = F(v_t + GW(t)) = F(u_t) , \quad v_0 = u_0 .$$

It follows at once from Proposition 3.6 and Assumption A.1.2 that the quoted estimate holds with $q = p_{\gamma+1}$. More precisely, it follows from Proposition 3.6 that $u_t \in \mathcal{H}_\alpha$ for every $\alpha < \gamma_\star + 1$. Therefore, $Lu_t \in \mathcal{H}_\gamma$ for $\gamma < \gamma_\star$. Furthermore, $N \in \text{Poly}(\mathcal{H}_{\gamma+1}, \mathcal{H}_\gamma)$ by Assumption A.1.2, so that $N(u_t) \in \mathcal{H}_\gamma$ as well. The claim then follows from the *a priori* bounds obtained in Proposition 3.6.

Concerning the bound (6.2a) on the linearization $J_{0,t}$, Proposition 3.7 combined with Assumption C.1 proves the result with $q = \bar{q}_\gamma + 1$. The line of reasoning used to bound $\|\partial_t v_t\|_\gamma$ also controls $\|\partial_t J_{s,t}\|_\gamma$ for $s < t$ and $s,t \in I_\delta$, since $\partial_t J_{s,t} = -L J_{s,t} + DN(u_t) J_{s,t}$.

Since Proposition 6.1 give an completely analogous bound for $K_{s,t}$ in $\mathcal{H}$ as for $J_{s,t}$ the results on K follow from the *a priori* bounds in Proposition 3.9. $\qquad\square$

6.2 A Hörmander-like theorem in infinite dimensions

In this section, we are going to formulate a lower bound on the Malliavin covariance matrix $\mathcal{M}_t$ under a condition that is very strongly reminiscent of the bracket condition in Hörmanders celebrated "sums of squares" theorem [Hör85, Hör67]. The proof of the result presented in this section will be postponed until Section 6.3 and constitutes the main technical result of this work.

Throughout all of this section and Section 6.3, we are going to make use of the bounds outlined in Proposition 6.2. We therefore now fix once and for all some choice of constants

$$\gamma \in [-a, \gamma_\star) \quad \text{and} \quad \beta \in [-a, \beta_\star) \quad \text{satisfying} \quad \gamma + \beta \ge -1 . \tag{6.4}$$

From now on, we will only ever use Proposition 6.2 with this fixed choice for γ and β. This is purely a convenience for expositional clarity since we will need these bounds

only finitely many times. As a side remark, note that one should think of these constants as being arbitrarily close to $\gamma_\star$ and $\beta_\star$ respectively.

With γ and β fixed as in (6.4), we introduce the set

$$\mathrm{Poly}(\gamma, \beta) \stackrel{\text{def}}{=} \mathrm{Poly}(\mathcal{H}_\gamma, \mathcal{H}_{-\beta-1}) \cap \mathrm{Poly}(\mathcal{H}_{\gamma+1}, \mathcal{H}_{-\beta}) \tag{6.5}$$

for notational convenience. (For integer m, $\mathrm{Poly}^m(\gamma, \beta)$ is defined analogously.) A polynomial $Q \in \mathrm{Poly}(\gamma, \beta)$ is said to be *admissible* if

$$[Q_\alpha, F_\sigma] \in \mathrm{Poly}(\gamma, \beta) \,,$$

for *every* pair of multi-indices α, σ. Here, Q_α and F_σ are defined as in (3.4) and F is the drift term of the SPDE (3.1) defined in (3.2).

This definition allows us to define a family of increasing subsets $\mathrm{A}_i \subset \mathrm{Poly}(\gamma, \beta)$ by the following recursion:

$$\mathrm{A}_1 = \{g_k \,, \quad k = 1, \ldots, d\} \subset \mathcal{H}_{\gamma_\star + 1} \approx \mathrm{Poly}^0(\mathcal{H}_{\gamma_\star + 1}) \subset \mathrm{Poly}(\gamma, \beta) \,,$$
$$\mathrm{A}_{i+1} = \mathrm{A}_i \cup \{Q_\alpha, \, [F_\sigma, Q_\alpha] \, : \, Q \in \mathrm{A}_i, \, Q \text{ admissible, and } \alpha, \sigma \text{ multi-indices}\} \,.$$

Remark 6.4 Recall from (3.5) that Q_α is proportional to the iterated "Lie bracket" of Q with $g_{\alpha_1}, g_{\alpha_2}$ and so forth. Similarly, $[F_\sigma, Q_\alpha]$ is the Lie bracket between two different iterated Lie brackets. As such, except for the issue of admissibility, the set of brackets considered here is exactly the same as in the traditional statement of Hörmander's theorem, only the order in which they appear is slightly different.

To each A_N we associate a positive symmetric quadratic form-valued function $\mathcal{Q}_N$ by

$$\langle \varphi, \mathcal{Q}_N(u)\varphi \rangle = \sum_{Q \in \mathrm{A}_N} \langle \varphi, Q(u) \rangle^2 \,.$$

Lastly for $\alpha \in (0,1)$, and for a given orthogonal projection $\Pi \colon \mathcal{H} \to \mathcal{H}$, we define $\mathcal{S}_\alpha \subset \mathcal{H}$ by

$$\mathcal{S}_\alpha = \{\varphi \in \mathcal{H} \setminus \{0\} \, : \, \|\Pi\varphi\| \geq \alpha\|\varphi\|\} \,. \tag{6.6}$$

With this notation, we make the following non-degeneracy assumption:

Assumption C.2 *For every $\alpha > 0$, there exists $N > 0$ and a function $\Lambda_\alpha \colon \mathcal{H} \to [0, \infty)$ such that*

$$\inf_{\varphi \in \mathcal{S}_\alpha} \frac{\langle \varphi, \mathcal{Q}_N(u)\varphi \rangle}{\|\Pi\varphi\|^2} \geq \Lambda_\alpha^2(u) \,,$$

for every $u \in \mathcal{H}_a$. Furthermore, for every $p \geq 1$, $t > 0$ and every $\alpha \in (0,1)$, there exists C such that $\mathbf{E}\,\Lambda_\alpha^{-p}(u_t) \leq C\Psi^p(u_0)$ for every initial condition $u_0 \in \mathcal{H}$.

Remark 6.5 Assumption C.2 is in some sense weaker than the usual non-degeneracy condition of Hörmander's theorem, since it only requires $\mathcal{Q}_N$ to be sufficiently non-degenerate on the range of Π. In particular, if $\Pi = 0$, then Assumption C.2 is void and always holds with $\Lambda_\alpha = 1$, say. This is the reason why, by choosing for Π a projector onto some finite-dimensional subspace of $\mathcal{H}$, one can expect Assumption C.2 to hold for a finite value of N, even in our situation where A_N only contains finitely many elements.

Remark 6.6 As will be seen in Section 8, it is often possible to choose Λ_α to be a constant, so that the second part of Assumption C.2 is automatically satisfied.

When Assumption C.2 holds, we have the following result whose proof is given in Section 6.3.

Theorem 6.7 *Consider an SPDE of the type (3.1) such that Assumptions A.1 and C.1 hold. Let furthermore the Malliavin matrix $\mathcal{M}_t$ be defined as in (4.10) and $\mathcal{S}_\alpha$ as in (6.6). Let Π be a finite rank orthogonal projection satisfying Assumption C.2. Then, there exists $\theta > 0$ such that, for every $\alpha \in (0, 1)$, every $p \geq 1$ and every $t > 0$ there exists a constant C such that the bound*

$$\mathbf{P}\Big(\inf_{\varphi \in \mathcal{S}_\alpha} \frac{\langle \varphi, \mathcal{M}_t \varphi \rangle}{\|\varphi\|^2} \leq \varepsilon \Big) \leq C \Psi^{\theta p}(u_0) \varepsilon^p \, ,$$

holds for every $u_0 \in \mathcal{H}$ and every $\varepsilon \leq 1$.

Remark 6.8 If Π is a finite rank orthogonal projection satisfying Assumption C.2 then Theorem 6.7 provides the critically ingredient to prove the smoothness of the density of $(\mathcal{P}_t^* \delta_x) \Pi^{-1}$ with respect to Lebesgue measure. Though [BM07] contains a few unfortunate errors, it still provides the framework needed to deduce smoothness of these densities from Theorem 6.7. In particular, one needs to prove that Πu_t is infinitely Malliavin differentiable. Section 5.1 of [BM07] shows how to accomplish this in a setting close to ours, see also [MP06].

6.3 Proof of Theorem 6.7

While the aim of this section is to prove Theorem 6.7, we begin with some preliminary definitions which will simplify its presentation. Many of the arguments used will rely on the construction of "exceptional sets" of small probability outside of which certain intuitive implications hold. This justifies the introduction of the following notational shortcut:

Definition 6.9 Given a collection $H = \{H^\varepsilon\}_{\varepsilon \leq 1}$ of subsets of the ambient probability space Ω, we will say that "H is a family of negligible events" if, for every $p \geq 1$ there exists a constant C_p such that $\mathbf{P}(H^\varepsilon) \leq C_p \varepsilon^p$ for every $\varepsilon \leq 1$.

Given such a family H and a statement Φ_ε depending on a parameter $\varepsilon > 0$, we will say that "Φ_ε holds modulo H" if, for every $\varepsilon \leq 1$, the statement Φ_ε holds on the complement of H^ε.

We will say that the family H is "universal" if it does not depend on the problem at hand. Otherwise, we will indicate which parameters it depends on.

Given two families H_1 and H_2 of negligible sets, we write $H = H_1 \cup H_2$ as a shortcut for the sentence "$H^\varepsilon = H_1^\varepsilon \cup H_2^\varepsilon$ for every $\varepsilon \leq 1$." Let us state the following useful fact, the proof of which is immediate:

Lemma 6.10 *Let H_n^ε be a collection of events with $n \in \{1, \ldots, C\varepsilon^{-\kappa}\}$ for some arbitrary but fixed constants C and κ and assume that $\mathbf{P}(H_k^\varepsilon) = \mathbf{P}(H_\ell^\varepsilon)$ for any pair (k, ℓ). Then, if the family $\{H_1^\varepsilon\}$ is negligible, the family $\{H^\varepsilon\}$ defined by $H^\varepsilon = \bigcup_n H_n^\varepsilon$ is also negligible.*

Remark 6.11 The same statement also holds of course if the equality between probabilities of events is replaced by two-sided bounds with multiplicative constants that do not depend on k, ℓ, and ε.

An important particular case is when the family H depends on the initial condition u_0 to (3.1). We will then say that H is "Ψ-controlled" if the constant C_p can be bounded by $\tilde{C}_p \Psi^p(u_0)$, where $\tilde{C}_p$ is independent of u_0.

In this language, the conclusion of Theorem 6.7 can be restated as saying that there exists $\theta > 0$ such that, for every $\alpha > 0$, the event

$$\inf_{\varphi \in \mathcal{S}_\alpha} \langle \varphi, \mathcal{M}_T \varphi \rangle \leq \varepsilon \|\varphi\|^2$$

is a Ψ^θ-controlled family of negligible events. Recall that the terminal time T was fixed once and for all and that the function Ψ was defined in Proposition 6.2. We further restate this as an implication in the following theorem which is easily seen to be equivalent to Theorem 6.7:

Theorem 6.12 *Let Π be a finite rank orthogonal projection satisfying Assumption C.2. Then, there exists $\theta > 0$ such that for every $\alpha \in (0, 1)$, the implication*

$$\varphi \in \mathcal{S}_\alpha \qquad \Longrightarrow \qquad \langle \varphi, \mathcal{M}_T \varphi \rangle > \varepsilon \|\varphi\|^2$$

holds modulo a Ψ^θ-controlled family of negligible events.

6.4   Basic structure and idea of proof of Theorem 6.12

We begin with an overly simplified version of the argument which neglects some technical difficulties. The basic idea of the proof is to argue that if $\langle \mathcal{M}_T \varphi, \varphi \rangle$ is small then $\langle \mathcal{Q}_k(u_T)\varphi, \varphi \rangle$ must also be small (with high probability) for every $k > 0$. This is proved inductively, beginning with the directions which are directly forced, namely those belonging to A_1. Assumption C.2 then guarantees in turn that $\|\Pi \varphi\|$ must be small with high probability. On the other hand, since $\varphi \in \mathcal{S}_\alpha$, we know for a fact that $\|\Pi \varphi\| \geq \alpha \|\varphi\|$ which is not small. Hence one of the highly improbable events must have occurred.

This sketch of proof is essentially the same as that of Hörmander's theorem in finite dimensions, see [Mal78, KS84, KS85a, Nor86, Nua95]. Trying to adapt this argument to the infinite-dimensional case, one is rapidly faced with two major hurdles. First, processes of the form $t \mapsto \langle J_{t,T} g, \varphi \rangle$ appearing in the definition of $\mathcal{M}_T$ are not adapted to the filtration generated by the driving noise. In finite dimensions, this difficulty is overcome by noting that

$$\mathcal{M}_t = J_{0,t} \hat{\mathcal{M}}_t J_{0,t}^* , \qquad \hat{\mathcal{M}}_t = \int_0^t J_{0,s}^{-1} GG^* (J_{0,s}^*)^{-1} \, ds ,$$

and then working with $\hat{\mathcal{M}}_t$ instead of $\mathcal{M}_t$. ($\hat{\mathcal{M}}_t$ is often called the reduced Malliavin covariance matrix.) The processes $t \mapsto \langle J_{0,t}^{-1} g, \varphi \rangle$ appearing there are now perfectly nice semimartingales and one can use Norris' lemma [Nor86], which is a quantitative version of the Doob-Meyer decomposition theorem, to show inductively that if $\langle \varphi, \mathcal{M}_T \varphi \rangle$ is small, then $t \mapsto \langle J_{0,t}^{-1} Q(u_t), \varphi \rangle$ must be small for every vector field $Q \in A_k$. In our setting, unlike in some previous results for infinite-dimensional systems [BT05], the Jacobian $J_{0,t}$ is not invertible. This is a basic feature of dissipative

PDEs with a smoothing linear term which is the dominating term on the right hand side. Such dynamical systems only generate semi-flows as opposed to invertible flows.

Even worse, there appears to be no good theory characterising a large enough subset belonging to its range. The only other situations to our knowledge where this has been overcome previously are the linear case [Oco88], as well as the particular case of the two-dimensional Navier-Stokes equations on the torus [MP06] and in [BM07] for a setting close to ours. As in those settings, we do not attempt to define something like the operator $\hat{\mathcal{M}}_t$ mentioned above but instead we work directly with $\mathcal{M}_t$, so that we do not have Norris' lemma at our disposal. It will be replaced by the result from Section 7 on "Wiener polynomials." This result states that if one considers a polynomial where the variables are independent Wiener processes and the coefficients are arbitrary (possibly non-adapted) Lipschitz continuous stochastic processes, then the polynomial being small implies that with high probability each individual monomial is small. It will be shown in this section how it is possible to exploit the polynomial structure of our nonlinearity in order to replace Norris' lemma by such a statement.

Another slightly less serious drawback of working in an infinite-dimensional setting is that we encounter singularities at $t = 0$ and at $t = T$ (for the operator $J_{t,T}$). Recall the definition of the time interval $I_\delta = [\frac{T}{2}, T - \delta]$ from Section 3. We will work on this interval which is strictly included in $[0, T]$ to avoid these singularities. There will be a trade-off between larger values of δ that make it easy to avoid the singularity and smaller values of δ that make it easier to infer bounds for $\langle \mathcal{Q}_k(u_T)\varphi, \varphi \rangle$.

When dealing with non-adapted processes, it is typical to replace certain standard arguments which hinge on adaptivity by arguments which use local time-regularity properties instead. This was also the approach used in [MP06, BM07]. To this end we introduce the following Hölder norms. For $\theta \in (0, 1]$, we define the Hölder norm for functions $f : I_\delta \to \mathcal{H}$ by

$$\|f\|_\theta = \sup_{s,t \in I_\delta} \frac{\|f(s) - f(t)\|}{|t - s|^\theta} \, , \tag{6.7}$$

and similarly if f is real-valued. (Note that even though we use the same notation as for the norm in the Cameron-Martin space in the previous section, these have nothing to do with each other. Since on the other hand the Cameron-Martin norm is never used in the present section, we hope that this does not cause too much confusion.) We furthermore set

$$\|f\|_{\theta,\gamma} = \sup_{s,t \in I_\delta} \frac{\|f(s) - f(t)\|_\gamma}{|t - s|^\theta} \, ,$$

where $\| \cdot \|_\gamma$ denotes the γth interpolation norm defined in Assumption A.1. Finally, we are from now on going to assume that δ is a function of ε through a scaling relation of the type

$$\delta = \frac{T}{4}\varepsilon^r \tag{6.8}$$

for some (very small) value of r to be determined later.

6.5 Some preliminary calculations

We begin with two preliminary calculations. The first translates a given growth of the moments of a family of random variables into a statement saying that the variables are "small," modulo a negligible family of events. As such, it is essentially a translation of Chebyshev's inequality into our language. The second is an interpolation result which

controls the supremum of a function's derivative by the supremum of the function and the size of some Hölder coefficient.

Lemma 6.13 *Let δ be as in (6.8) with $r > 0$, let $\Psi \colon \mathcal{H} \to [1, \infty)$ be an arbitrary function, and let X_δ be a δ-dependent family of random variables such that there exists $b \in$ $\mathbf{R}$ (b is allowed to be negative) such that, for every $p \geq 1$, $\mathbf{E}|X_\delta|^p \leq C_p \Psi^p(u_0)\delta^{-bp}$. Then, for any $q > br$ and any $c > 0$, the family of events*

$$\left\{|X_\delta| > \frac{\varepsilon^{-q}}{c}\right\}$$

is $\Psi^{\frac{1}{q-br}}$-dominated negligible.

Proof. It follows from Chebychev's inequality that

$$\mathbf{P}\left(|X_\delta| > \frac{\varepsilon^{-q}}{c}\right) \leq C_p c^p \Psi^p \delta^{-bp} \varepsilon^{qp} = \bar{C}_\ell \big(\Psi^{\frac{1}{q-br}}\big)^\ell \varepsilon^\ell \,,$$

where $\bar{C}_\ell$ is equal to $C_p c^p$ with $\ell = p(q - br)$. Provided that $q - br > 0$, this holds for every $\ell > 0$ and the claim follows. $\square$

Lemma 6.14 *Let $f \colon [0, T] \to \mathbf{R}$ be continuously differentiable and let $\alpha \in (0, 1]$. Then, the bound*

$$\|\partial_t f\|_{L^\infty} = \|f\|_1 \leq 4\|f\|_{L^\infty} \max\left\{\frac{1}{T}, \|f\|_{L^\infty}^{-\frac{1}{1+\alpha}} \|\partial_t f\|_\alpha^{\frac{1}{1+\alpha}}\right\}$$

holds, where $\|f\|_\alpha$ denotes the best α-Hölder constant for f.

Proof. Denote by x_0 a point such that $|\partial_t f(x_0)| = \|\partial_t f\|_{L^\infty}$. It follows from the definition of the α-Hölder constant $\|\partial_t f\|_{\mathcal{C}^\alpha}$ that $|\partial_t f(x)| \geq \frac{1}{2}\|\partial_t f\|_{L^\infty}$ for every x such that $|x - x_0| \leq (\|\partial_t f\|_{L^\infty}/2\|\partial_t f\|_{\mathcal{C}^\alpha})^{1/\alpha}$. The claim then follows from the fact that if f is continuously differentiable and $|\partial_t f(x)| \geq A$ over an interval I, then there exists a point x_1 in the interval such that $|f(x_1)| \geq A|I|/2$. $\square$

6.6 Transferring properties of φ back from the terminal time

We now prove a result which shows that if $\varphi \in \mathcal{S}_\alpha$ then with high probability both $\|\Pi K_{T-\delta,T}\varphi\|$ and the ratio $\|\Pi K_{T-\delta,T}\varphi\|/\|K_{T-\delta,T}\varphi\|$ can not change dramatically for small enough δ. This allows us to step back from the terminal time T to the right end point of the time interval I_δ. As mentioned at the start of this section, this is needed to allow the rougher test functions used in Theorem 6.7.

Lemma 6.15 *Let* (6.9b) *hold and fix any orthogonal projection Π of $\mathcal{H}$ onto a finite dimensional subspace of $\mathcal{H}$ spanned by elements of $\mathcal{H}_1$. Recall furthermore the relation* (6.8) *between δ and ε. There exists a constant $c \in (0, 1)$ such that, for every $r > 0$ and every $\alpha > 0$, the implication*

$$\varphi \in \mathcal{S}_\alpha \quad \Longrightarrow \quad K_{T-\delta,T}\varphi \in \mathcal{S}_{c\alpha} \quad and \quad \|\Pi K_{T-\delta,T}\varphi\| \geq \frac{\alpha}{2}\|\varphi\| \,,$$

holds modulo a $\Psi^{1/r}$-controlled family of negligible events.

To prove this Lemma, we will need the following axillary lemma whose proof is given at the end of the section.

Lemma 6.16 *For any $\delta \in (0, T/2]$, one has the bound*

$$\mathbf{E} \sup_{\|\varphi\| \leq 1} \|K_{T-\delta,T}\varphi - e^{-\delta L}\varphi\|^p \leq C_p \Psi^{np}(u_0)\delta^{(1-a)p} , \qquad (6.9a)$$

$$\mathbf{E} \sup_{\|\varphi\| \leq 1} \|K_{T-\delta,T}\varphi - \varphi\|_{-1}^p \leq C_p \Psi^{np}(u_0)\delta^{(1-a)p} , \qquad (6.9b)$$

for every $p \geq 1$ and every $u_0 \in \mathcal{H}$. Here, n is the degree of the nonlinearity N.

Proof of Lemma 6.15. We begin by showing that, modulo some $\Psi^{1/r}$-dominated family of negligible events,

$$\|\Pi\varphi\| \geq \alpha\|\varphi\| \quad \Longrightarrow \quad \|\Pi K_{T-\delta,T}\varphi\| \geq \frac{\alpha}{2}\|\varphi\| .$$

By the assumption on Π, we can find a collection $\{v_k\}_{k=1}^N$ in $\mathcal{H}_1$ with $\|v_k\| = 1$ such that $\Pi\varphi = \sum_k v_k\langle v_k, \varphi\rangle$. Therefore, there exists a constant $C_1 = \sup_k \|v_k\|_1$ so that $\|\Pi\varphi\| \leq C_1\|\varphi\|_{-1}$. Combining Lemma 6.13 with Lemma 6.16, we see that

$$\sup_{\varphi \in \mathcal{H} : \|\varphi\|=1} \|K_{T-\delta,T}\varphi - \varphi\|_{-1} \leq \frac{\alpha}{2C_1} , \qquad (6.10)$$

modulo a $\Psi^{\frac{n}{(1-a)r}}$-dominated family of negligible events. Hence, modulo the same family of events,

$$\|\Pi K_{T-\delta,T}\varphi\| \geq \|\Pi\varphi\| - C_1\|K_{T-\delta,T}\varphi - \varphi\|_{-1}$$
$$\geq \alpha\|\varphi\| - \frac{\alpha}{2}\|\varphi\| = \frac{\alpha}{2}\|\varphi\| .$$

Combining now Lemma 6.13 with (6.9a), we see that

$$\|K_{T-\delta,T}\varphi\| \leq \|\varphi\| + \|e^{-\delta L}\varphi\| \leq C\|\varphi\| ,$$

modulo a $\Psi^{\frac{n}{(1-a)r}}$-dominated family of negligible events, thus showing that $K_{T-\delta,T}\varphi$ belongs to $\mathcal{S}_{c\alpha}$ with $c = 1/(2C)$ and concluding the proof. □

We now give the proof of the auxiliary lemma used in the proof of Lemma 6.15.

Proof of Lemma 6.16. It follows from (3.13) and the variation of constants formula that

$$K_{T-\delta,T}\varphi - e^{-\delta L}\varphi = \int_{T-\delta}^T e^{-(T-s)L} DN^*(u_s)K_{s,T}\varphi\, ds .$$

It now follows from Assumption A.1, point 3 that there exists $\gamma_0 \in [0, \gamma_\star + 1)$ such that $DN^*(u)$ is a bounded linear map from $\mathcal{H}$ to $\mathcal{H}_{-a}$ for every $u \in \gamma_0$ and that its norm is bounded by $C\|u\|_{\gamma_0}^{n-1}$ for some constant C. The first bound then follows by combining Proposition 6.2 with the fact that e^{-Lt} is bounded by Ct^{-a} as an operator from $\mathcal{H}_{-a}$ to $\mathcal{H}$ as a consequence of standard analytic semigroup theory [Kat80].

In order to obtain the second bound, we write

$$\|K_{T-\delta,T}\varphi - \varphi\|_{-1} \leq \|K_{T-\delta,T}\varphi - e^{-L\delta}\varphi\|_{-1} + \|e^{-L\delta}\varphi - \varphi\|_{-1}$$
$$\leq \|K_{T-\delta,T}\varphi - e^{-L\delta}\varphi\| + C\delta ,$$

where the last inequality is again a consequence of standard analytic semigroup theory. The claim then follows from (6.9a). □

6.7 The smallness of $\mathcal{M}_T$ implies the smallness of $\mathcal{Q}_N(u_{T-\delta})$

In this section, we show that if $\langle \mathcal{M}_T\varphi, \varphi \rangle$ is small then $\langle \mathcal{Q}_N(u_t)K_{t,T}\varphi, K_{t,T}\varphi \rangle$ must also be small with high probability for every $t \in I_\delta$. The precise statement is given by the following result:

Lemma 6.17 *Let the Malliavin matrix $\mathcal{M}_T$ be defined as in (4.10) and assume that Assumptions A.1 and C.1 are satisfied. Then, for every $N > 0$, there exist $r_N > 0$, $p_N > 0$ and $q_N > 0$ such that, provided that $r \leq r_N$, the implication*

$$\langle \varphi, \mathcal{M}_T\varphi \rangle \leq \varepsilon \|\varphi\|^2 \quad \Longrightarrow \quad \sup_{Q \in \mathsf{A}_N} \sup_{t \in I_\delta} |\langle K_{t,T}\varphi, Q(u_t) \rangle| \leq \varepsilon^{p_N} \|\varphi\| \,,$$

holds modulo some Ψ^{q_N}-dominated negligible family of events.

Proof. The proof proceeds by induction on N and the steps of this induction are the content of the next two subsections. Since $\mathsf{A}_1 = \{g_1, \ldots, g_d\}$, the case $N = 1$ is implied by Lemma 6.18 below, with $p_1 = 1/4$, $q_1 = 8$, and $r_1 = 1/(8\bar{p}_\beta)$.
    The inductive step is then given by combining Lemmas 6.21 and 6.24 below. At each step, the values of p_n and r_n decrease while q_n increases, but all remain strictly positive and finite after finitely many steps. $\qquad\square$

6.8 The first step in the iteration

The "priming step" in the inductive proof of Lemma 6.17 follows from the fact that the directions which are directly forced by the Wiener processes are not too small with high probability.

Lemma 6.18 *Let the Malliavin matrix $\mathcal{M}$ be defined as in (4.10) and assume that Assumptions A.1 and C.1 are satisfied. Then, provided that $r \leq 1/(8\bar{p}_\beta)$, the implication*

$$\langle \varphi, \mathcal{M}_T\varphi \rangle \leq \varepsilon \|\varphi\|^2 \quad \Longrightarrow \quad \sup_{k=1\ldots d} \sup_{t \in I_\delta} |\langle K_{t,T}\varphi, g_k \rangle| \leq \varepsilon^{1/4} \|\varphi\| \,,$$

holds modulo some Ψ^8-dominated negligible family of events. Here, $\bar{p}_\beta$ is as in (6.3b) and β was fixed in (6.4).

Proof. For notational compactness, we scale φ to have norm one by replacing φ with $\varphi/\|\varphi\|$. We will still refer to this new unit vector as φ. Now assume that $\langle \varphi, \mathcal{M}_T\varphi \rangle \leq \varepsilon$. It then follows from (4.10) that

$$\sup_{k=1\ldots d} \int_{I_\delta} \langle g_k, K_{t,T}\varphi \rangle^2 \, dt \leq \varepsilon \,.$$

Applying Lemma 6.14 with $f(t) = \int_{T/2}^t |\langle g_k, K_{s,T}\varphi \rangle| \, ds$ and $\alpha = 1$, it follows that there exists a constant $C > 0$ such that, for every $k = 1 \ldots d$, either

$$\sup_{t \in I_\delta} |\langle g_k, K_{t,T}\varphi \rangle| \leq \varepsilon^{1/4} \,,$$

or

$$\||\langle g_k, K_{\cdot,T}\varphi \rangle\|_1 \geq C\varepsilon^{-1/4} \,. \tag{6.11}$$

Therefore, to complete the proof, we need only to show that the latter events form a Ψ^4-dominated negligible family for every k. Since $\|\langle g_k, K_{\cdot,T}\varphi\rangle\|_1 \leq \|g_k\|_{-\beta}\|K_{\cdot,T}\varphi\|_{1,\beta}$, the bound (6.11) implies that

$$\sup_{\varphi\in\mathcal{H}\,:\,\|\varphi\|=1} \|K_{t,T}\varphi\|_{1,\beta} \geq \frac{C\varepsilon^{-1/4}}{g^*} \, , \tag{6.12}$$

where $g^* = \max_k \|g_k\|_{-\beta}$ (which is finite since we have by assumption that $-\beta \leq \gamma + 1 < \gamma_\star + 1$ and since $g_k \in \mathcal{H}_{\gamma_\star+1}$ for every k) . This event depends only on the initial condition u_0 and on the model under consideration. In particular, it is independent of φ.

The claim now follows from the *a priori* bound (6.3b) and Lemma 6.13 with $q = \frac{1}{4}$ and $b = \bar{p}_\beta$. □

6.9 The iteration step

Recall that we consider evolution equations of the type

$$du_t = F(u_t)\, dt + \sum_{k=1}^d g_k dW_k(t) \, , \tag{6.13}$$

where F is a "polynomial" of degree n. The aim of this section is to implement the following recursion: if, for any given polynomial Q, the expression $\langle Q(u_t), K_{t,T}\varphi\rangle$ is "small" in the supremum norm, then both the expression $\langle [Q,F](u_t), K_{t,T}\varphi\rangle$ and $\langle [Q, g_k](u_t), K_{t,T}\varphi\rangle$ must be small in the supremum norm as well.

The main technical tool used in this section will be the estimates on "Wiener polynomials" from Section 7. Using the notation

$$W_\alpha(t) \stackrel{\text{def}}{=} W_{\alpha_1}(t)\, W_{\alpha_2}(t)\cdots W_{\alpha_\ell}(t) \, ,$$

for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, this estimate states that if an expression of the type $\sum_{|\alpha|\leq m} A_\alpha(t)W_\alpha(t)$ is small, then, provided that the processes A_α are sufficiently regular in time, each of the A_α must be small. In other words, two distinct monomials in a Wiener polynomial cannot cancel each other out. Here, the processes A_α do not have to be adapted to the filtration generated by the W_k, so this gives us some kind of anticipative replacement of Norris' lemma. The main trick that we use in order to take advantage of such a result is to switch back and forth between considering the process u_t solution to (6.13) and the process v_t defined by

$$v_t \stackrel{\text{def}}{=} u_t - \sum_{k=1}^d g_k W_k(t) \, ,$$

which has more time-regularity than u_t. Recall furthermore that given a polynomial Q and a multi-index α, we denote by Q_α the corresponding term (3.5) appearing in the (finite) Taylor expansion of Q.

Recall the definition $\text{Poly}^m(\gamma, \beta) = \text{Poly}^m(\mathcal{H}_\gamma, \mathcal{H}_{-\beta-1}) \cap \text{Poly}^m(\mathcal{H}_{\gamma+1}, \mathcal{H}_{-\beta})$. We first show that if $Q \in \text{Poly}^m(\gamma, \beta)$ and $\langle Q(u_t), K_{t,T}\varphi\rangle$ is small, then the expression $\langle Q_\alpha(v_t), K_{t,T}\varphi\rangle$ (note the appearance of v_t rather than u_t) must be small as well for every multi-index α:

Lemma 6.19 *Let $Q \in \mathrm{Poly}^m(\gamma, \beta)$ for some $m \geq 0$ and for γ and β as chosen in (6.4). Let furthermore $q > 0$ an set $\bar{q} = q3^{-m}$. Then, the implication*

$$\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi \rangle| \leq \varepsilon^q \|\varphi\| \quad \Longrightarrow \quad \sup_\alpha \sup_{t \in I_\delta} |\langle Q_\alpha(v_t), K_{t,T}\varphi \rangle| \leq \varepsilon^{\bar{q}} \|\varphi\| ,$$

holds modulo some $\Psi^{6(m+1)/\bar{q}}$-dominated negligible family of events, provided that $r < \bar{q}/(6\bar{p}_\beta)$.

Proof. Note first that both inner products appearing in the implication are well-defined by Proposition 6.2 and the assumptions on Q. By homogeneity, we can assume that $\|\varphi\| = 1$. Since Q is a polynomial, (3.4) implies that

$$\langle Q(u_t), K_{t,T}\varphi \rangle = \sum_\alpha \langle Q_\alpha(v_t), K_{t,T}\varphi \rangle W_\alpha(t) .$$

Applying Theorem 7.1, we see that, modulo some negligible family of events Osc_W^m, $\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi \rangle| \leq \varepsilon^q$ implies that either

$$\sup_\alpha \sup_{t \in I_\delta} |\langle Q_\alpha(v_t), K_{t,T}\varphi \rangle| \leq \varepsilon^{\bar{q}} , \tag{6.14}$$

or there exists some α such that

$$\||\langle Q_\alpha(v_t), K_{\cdot,T}\varphi \rangle\||_1 \geq \varepsilon^{-\bar{q}/3} . \tag{6.15}$$

We begin by arguing that the second event is negligible. Since Q is of degree m, there exists a constant C such that

$$\||\langle Q_\alpha(v_t), K_{\cdot,T}\varphi \rangle\||_1 \leq \sup_{t \in I_\delta} \|K_{t,T}\varphi\|_{\beta+1} \||Q_\alpha(v_\cdot)\||_{1,-\beta-1} + \sup_{t \in I_\delta} \|Q_\alpha(v_t)\|_{-\beta} \||K_{\cdot,T}\varphi\||_{1,\beta}$$

$$\leq C \sup_{t \in I_\delta} \|K_{t,T}\varphi\|_{\beta+1} \sup_{t \in I_\delta} \|v_t\|_\gamma^{m-1} \||v\||_{1,\gamma} + C \sup_{t \in I_\delta} \|v_t\|_{\gamma+1}^m \||K_{\cdot,T}\varphi\||_{1,\beta} .$$

Here, we used the fact that $Q_\alpha \in \mathrm{Poly}^m(\mathcal{H}_\gamma, \mathcal{H}_{-1-\beta})$ to bound the first term and the fact that $Q_\alpha \in \mathrm{Poly}^m(\mathcal{H}_{\gamma+1}, \mathcal{H}_{-\beta})$ to bound the second term. The fact that Q_α belongs to these spaces is a consequence of $g_k \in \mathcal{H}_{\gamma_\star+1}$ and of the definition (3.4) of Q_α.

Therefore, (6.15) implies that either

$$X_\delta \stackrel{\mathrm{def}}{=} \sup_{\varphi \in \mathcal{H} : \|\varphi\|=1} \sup_{t \in I_\delta} \|K_{t,T}\varphi\|_{\beta+1} \sup_{t \in I_\delta} \|v_t\|_\gamma^{m-1} \||v\||_{1,\gamma} \geq \frac{1}{2C}\varepsilon^{-\bar{q}/3} \tag{6.16}$$

or

$$Y_\delta \stackrel{\mathrm{def}}{=} \sup_{\varphi \in \mathcal{H} : \|\varphi\|=1} \sup_{t \in I_\delta} \|v_t\|_{\gamma+1}^m \||K_{\cdot,T}\varphi\||_{1,\beta} \geq \frac{1}{2C}\varepsilon^{-\bar{q}/3} . \tag{6.17}$$

Combining the Cauchy-Schwarz inequality with (6.3b) of Proposition 6.2, we see that X_δ and Y_δ satisfy the assumptions of Lemma 6.13 with $\Phi = \Psi^{m+1}$ and $b = \bar{p}_\beta$, thus showing that the families of events (6.16) and (6.17) are both $\Psi^{6(m+1)/\bar{q}}$-dominated negligible, provided that $r < \bar{q}/(6\bar{p}_\beta)$. □

In the sequel, we will need the follow simple result which is, in some way, a converse to Theorem 7.1.

Lemma 6.20 *Given any integer $N > 0$ and any two exponents $0 < \bar{q} < q$, there exists a universal family of negligible events Sup_W^N such that the implication*

$$\sup_{\alpha} \sup_{t \in I_\delta} |A_\alpha(t)| < \varepsilon^q \quad \Longrightarrow \quad \sup_{t \in I_\delta} \left| \sum_{\alpha:|\alpha| \le N} A_\alpha(t) W_\alpha(t) \right| < \varepsilon^{\bar{q}}$$

holds modulo Sup_W^N for any collection of processes $\{A_\alpha(t) : |\alpha| \le N\}$.

Proof. Observe that

$$\sup_{t \in I_\delta} \left| \sum_{\alpha:|\alpha| \le N} A_\alpha(t) W_\alpha(t) \right| \le \left(\sup_{\alpha} \sup_{t \in I_\delta} |A_\alpha(t)| \right) \left(\sum_{\alpha:|\alpha| \le N} \sup_{t \in I_\delta} |W_\alpha| \right)$$

Since for any $p > 0$,

$$\sum_{\alpha:|\alpha| \le N} \sup_{t \in I_\delta} |W_\alpha| > \varepsilon^{-p}$$

is a negligible family of events, the claim follows at once. $\qquad\square$

As a corollary to Lemmas 6.19 and 6.20, we now obtain the key estimate for Lemma 6.17 in the particular case where the commutator is taken with one of the constant vector fields:

Lemma 6.21 *Let $Q \in \text{Poly}^m(\gamma, \beta)$ be a polynomial of degree m and let $q > 0$. Then, for $\bar{q} = q3^{-(m+1)}$, the implication*

$$\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi \rangle| \le \varepsilon^q \|\varphi\| \implies \sup_{\alpha} \sup_{t \in I_\delta} |\langle Q_\alpha(u_t), K_{t,T}\varphi \rangle| \le \varepsilon^{\bar{q}} \|\varphi\| ,$$

holds for all $\varphi \in \mathcal{H}$ modulo some $\Psi^{2(m+1)/\bar{q}}$-dominated negligible family of events, provided that $r < \bar{q}/(2\bar{p}_\beta)$.

Proof. Since it follows from (3.4) that $(Q_\alpha)_\beta = Q_{\alpha \cup \beta}$, we have the identity

$$Q_\alpha(u_t) = \sum_\beta (Q_\alpha)_\beta(v_t) W_\beta = \sum_\beta Q_{\alpha \cup \beta}(v_t) W_\beta .$$

Combining Lemma 6.19 and Lemma 6.20 with $N = m$ proves the claim. $\qquad\square$

In the next step, we show a similar result for the commutators between Q and F. We are going to use the fact that if a function f is differentiable with Hölder continuous derivative, then f being small implies that $\partial_t f$ is small as well, as made precise by Lemma 6.14. As previously, we start by showing a result that involves the process v_t instead of u_t:

Lemma 6.22 *Let Q be as in Lemma 6.19 and such that $[Q_\alpha, F_\sigma] \in \text{Poly}(\gamma, \beta)$ for any two multi-indices α, σ. Let furthermore $q > 0$ and set $\bar{q} = q3^{-2m}/8$. Then the implication*

$$\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi \rangle| \le \varepsilon^q \|\varphi\| \quad \Longrightarrow \quad \sup_{\alpha, \sigma} \sup_{t \in I_\delta} |\langle [Q_\alpha, F_\sigma](v_t), K_{t,T}\varphi \rangle| \le \varepsilon^{\bar{q}} \|\varphi\| ,$$

holds modulo some $\Psi^{6(m+1)/\bar{q}}$-dominated negligible family of events, provided that $r < \bar{q}/(6\bar{p}_\beta)$. (As before the empty multi-indices are included in the supremum.)

Proof. By homogeneity, we can assume that $\|\varphi\| = 1$. Combining Lemma 6.19 with Lemma 6.14 and defining $\hat{q} = q3^{-m}$, we obtain that $\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi\rangle| \leq \varepsilon^q$ implies for $f_{\alpha,\varphi}(t) \stackrel{\text{def}}{=} \partial_t \langle Q_\alpha(v_t), K_{t,T}\varphi\rangle$ the bound

$$\sup_{t \in I_\delta} |f_{\alpha,\varphi}(t)| \leq C \max\{\varepsilon^{\hat{q}}, \varepsilon^{\frac{\hat{q}}{4}} \|f_{\alpha,\varphi}\|_{1/3}^{3/4}\} , \tag{6.18}$$

modulo some $\Psi^{6(m+1)/\hat{q}}$-dominated negligible family of events, provided that $r \leq \hat{q}/(6\bar{p}_\beta)$. Note that this family is in particular independent of both α and φ. Here and in the sequel, we use the letter C to denote a generic constant depending on the details of the problem that may change from one expression to the next.

One can see that $\langle Q_\alpha(v_t), K_{t,T}\varphi\rangle$ is differentiable in t by combining Proposition 6.2 with the fact that $Q_\alpha \in \text{Poly}(\mathcal{H}_\gamma, \mathcal{H}_{-1-\beta}) \cap \text{Poly}(\mathcal{H}_{\gamma+1}, \mathcal{H}_{-\beta})$ as in the proof of Lemma 6.19. See [DL92] for a more detailed proof of a similar statement.

Computing the derivative explicitly, we obtain

$$f_{\alpha,\varphi}(t) = \langle DQ_\alpha(v_t)F(u_t) - DF(u_t)Q_\alpha(v_t), K_{t,T}\varphi\rangle \stackrel{\text{def}}{=} \langle B_\alpha(t), K_{t,T}\varphi\rangle .$$

The function B_α can be further expanded to

$$B_\alpha(t) = \sum_\sigma (DQ_\alpha(v_t)F_\sigma(v_t) - DF_\sigma(v_t)Q_\alpha(v_t)) W_\sigma(t) = \sum_\sigma [Q_\alpha, F_\sigma](v_t)W_\sigma(t) .$$

Notice that, by the assumption that $[Q_\alpha, F_\sigma] \in \text{Poly}(\gamma, \beta)$, one has

$$\|[Q_\alpha, F_\sigma](v.)W_\sigma(\cdot)\|_{\frac{1}{3}, -1-\beta} \leq C(1 + \sup_{t \in I_\delta} \|v_t\|_\gamma)^{n+m-2-|\alpha|-|\sigma|}\|\partial_t v_t\|_\gamma \sup_{t \in I_\delta} |W_\sigma(t)|$$
$$+ C\|W_\sigma\|_{\frac{1}{3}}(1 + \sup_{t \in I_\delta} \|v_t\|_\gamma)^{n+m-1-|\alpha|-|\sigma|} ,$$

$$\|[Q_\alpha, F_\sigma](v_t)W_\sigma(t)\|_{-\beta} \leq C(1 + \|v_t\|_{\gamma+1})^{n+m-1-|\alpha|-|\sigma|}|W_\sigma(t)| .$$

(Here it is understood that if one of the exponents of the norm of v_t is negative, the term in question actually vanishes.) It therefore follows from Proposition 6.2 that

$$\mathbf{E}\|B_\alpha\|_{\frac{1}{3}, -1-\beta}^p \leq C_p \Psi^{(n+m-1)p}(u_0) , \qquad \mathbf{E} \sup_{t \in I_\delta} \|B_\alpha(t)\|_{-\beta}^p \leq C_p \Psi^{(n+m-1)p}(u_0) ,$$

for every $p \geq 1$ and some constants C_p.

Since the Hölder norm of $f_{\alpha,\varphi}$ is bounded by

$$\|\langle B_\alpha(\cdot), K_{\cdot,T}\varphi\rangle\|_{\frac{1}{3}} \leq \|B_\alpha\|_{\frac{1}{3}, -1-\beta} \sup_{t \in I_\delta} \|K_{t,T}\|_{\beta+1} + \|K_{\cdot,T}\|_{\frac{1}{3}, \beta} \sup_{t \in I_\delta} \|B_\alpha(t)\|_{-\beta} ,$$

we can use the bounds on B_α just obtained, the Cauchy-Schwarz inequality, Proposition 6.2, and Lemma 6.13, to obtain

$$\sup_\alpha \sup_{\|\varphi\| \leq 1} \|f_{\alpha,\varphi}\|_{1/3}^{3/4} \leq \varepsilon^{-\frac{\hat{q}}{8}} , \tag{6.19}$$

modulo some $\Psi^{12(n+m)/\hat{q}}$-dominated negligible family of events, provided that $r \leq \min\{\hat{q}/12, \hat{q}/(6\bar{p}_\beta)\}$. As a consequence, modulo this family, we obtain from (6.18) the bound $\sup_\alpha \sup_{t \in I_\delta} |f_{\alpha,\varphi}(t)| \leq C\varepsilon^{\frac{\hat{q}}{8}}$ which can be rewritten as

$$\sup_\alpha \sup_{t \in I_\delta} \left| \sum_\sigma \langle [Q_\alpha, F_\sigma](v_t), K_{t,T}\varphi\rangle W_\sigma(t) \right| \leq C\varepsilon^{\frac{\hat{q}}{8}} . \tag{6.20}$$

Since $[Q_\alpha, F_\sigma] \in \mathrm{Poly}(\gamma, \beta)$ the same reasoning as in Lemma 6.19 combined with Theorem 7.1 on Wiener polynomials implies that modulo some negligible family of events Osc_W^m, the estimate (6.20) implies that either

$$\sup_{\alpha, \sigma} \sup_{t \in I_\delta} |\langle [Q_\alpha, F_\sigma](v_t), K_{t,T}\varphi \rangle| \leq \varepsilon^{\bar{q}} \,, \tag{6.21}$$

or there exists some α and σ such that

$$\| \langle [Q_\alpha, F_\sigma](v_t), K_{\cdot,T}\varphi \rangle \|_1 \geq \varepsilon^{-\bar{q}/3} \,. \tag{6.22}$$

Again following the same logic as Lemma 6.19, we see that the family of events in (6.22) is $\Phi^{6(m+1)/\bar{q}}$-dominated negligible provided that $r < \bar{q}/(6\bar{p}_\beta)$. $\qquad\square$

In order to turn this result into a result involving the process u_t, we need the following expansion:

Lemma 6.23 *Given any two multi-indices α and σ (including the empty indices), there exist an integer N and a collection of multi-indices $\{\alpha_i, \sigma_i, \zeta_i \,:\, i = 1 \ldots N\}$ and constants $\{c_i : i = 1 \ldots N\}$ so that*

$$[Q_\alpha, F_\sigma](u_t) = \sum_{i=1}^{N} c_i [Q_{\alpha_i}, F_{\sigma_i}](v_t) W_{\zeta_i}(t)$$

Proof. First observe that

$$[Q_\alpha, F_\sigma](u_t) = \sum_{\zeta} [Q_\alpha, F_\sigma]_\zeta(v_t) W_\zeta(t) \,.$$

The Jacobi identity for Lie bracket states that

$$\mathrm{D}_{g_k}[Q_\alpha, F_\sigma] = [g_k, [Q_\alpha, F_\sigma]] = [[g_k, Q_\alpha], F_\sigma] + [Q_\alpha, [g_k, F_\sigma]]$$
$$= (|\alpha| + 1)[Q_{\alpha \cup (k)}, F_\sigma] + (|\sigma| + 1)[Q_\alpha, F_{\sigma \cup (k)}] \,.$$

By iterating this calculation, we see that for any multi-index ζ, $[Q_\alpha, F_\sigma]_\zeta$ is equal to some linear combination of a finite number of terms of the form $[Q_{\alpha_i}, F_{\sigma_i}]$ for some multi-indices α_i and σ_i. $\qquad\square$

In very much the same way as before, it then follows that:

Corollary 6.24 *Let Q be as in Lemma 6.19 and such that $[Q_\alpha, F_\sigma] \in \mathrm{Poly}(\gamma, \beta)$ for any two multi-indices α, σ. Let furthermore $q > 0$ and set $\bar{q} = q3^{-2(m+1)}/8$. Then the implication*

$$\sup_{t \in I_\delta} |\langle Q(u_t), K_{t,T}\varphi \rangle| \leq \varepsilon^q \|\varphi\| \quad \Longrightarrow \quad \sup_{\alpha, \sigma} \sup_{t \in I_\delta} |\langle [Q_\alpha, F_\sigma](u_t), K_{t,T}\varphi \rangle| \leq \varepsilon^{\bar{q}} \|\varphi\| \,,$$

holds modulo some $\Psi^{2(m+1)/(3\bar{q})}$-dominated negligible family of events, provided that $r < 3\bar{q}/(2\bar{p}_\beta)$.

Proof. It follows from Lemma 6.23 that

$$\langle [Q_\alpha, F_\sigma](u_t), K_{T,t}\varphi \rangle = \sum_{i=1}^{N} c_i \langle [Q_{\alpha_i}, F_{\sigma_i}](v_t), K_{T,t}\varphi \rangle W_{\gamma_i}(t) \,.$$

Combining the control of the $\langle [Q_{\alpha_i}, F_{\sigma_i}](v_t), K_{T,t}\varphi \rangle$ obtained in Lemma 6.22 with Lemma 6.20 gives the quoted result. $\qquad\square$

6.10 Putting it all together: proof of Theorem 6.12

We now finally combine all of the results we have just accumulated to give the proof of the main theorem of these sections.

Proof of Theorem 6.12. We are going to prove the statement by showing that there exists $\theta > 0$ and, for every $\alpha > 0$, a Ψ^θ-dominated family of negligible events such that, modulo this family, the assumption $\inf_{\varphi \in \mathcal{S}_\alpha} \langle \varphi, \mathcal{M}_T \varphi \rangle \le \varepsilon \|\varphi\|^2$ leads to a contradiction for all ε sufficiently small.

From now on, fix N as in Assumption C.2. By Lemmas 6.15 and 6.17, we see that there exist constants $\theta, q, r_0 > 0$ such that, modulo some Ψ^θ-dominated family of negligible events, one has the implication

$$\left. \begin{array}{r} \varphi \in \mathcal{S}_\alpha \\ \langle \varphi, \mathcal{M}_T \varphi \rangle \le \varepsilon \|\varphi\|^2 \end{array} \right\} \implies \left\{ \begin{array}{l} K_{T-\delta,T}\varphi \in \mathcal{S}_{c\alpha} \text{ and } \|\Pi K_{T-\delta,T}\varphi\| \ge \frac{\alpha}{2}\|\varphi\| \\ \langle K_{T-\delta,T}\varphi, \mathcal{Q}_N(u_{T-\delta}) K_{T-\delta,T}\varphi \rangle \le \varepsilon^q \|\varphi\|^2 \,, \end{array} \right.$$

provided that we choose $r \le r_0$ in the definition (6.8) of δ. By Assumption C.2, this in turn implies (modulo the same family of negligible events)

$$\cdots \implies \frac{\alpha}{2}\|\varphi\| \le \|\Pi K_{T-\delta,T}\varphi\| \le \Lambda_{c\alpha}^{-1}(u_0)\varepsilon^{\frac{q}{2}}\|\varphi\| \,.$$

On the other hand, it follows from Lemma 6.13 and the assumption on the inverse moments of $\Lambda_{c\alpha}$ that, modulo some $\Psi^{\frac{4}{q}}$-dominated family of negligible events, one has the bound

$$\Lambda_{c\alpha}^{-1}(u_0) \le \varepsilon^{-\frac{q}{4}} \,.$$

Possibly making θ smaller, it follows that, modulo some Ψ^θ-dominated family of negligible events, one has the implication

$$\left. \begin{array}{r} \varphi \in \mathcal{S}_\alpha \\ \langle \varphi, \mathcal{M}_T \varphi \rangle \le \varepsilon \|\varphi\|^2 \end{array} \right\} \implies \frac{\alpha}{2} \le \varepsilon^{\frac{q}{4}} \,,$$

which cannot hold for ε small enough, thus concluding the proof of Theorem 6.12 $\qquad\square$

7 Bounds on Wiener polynomials

We will use the terminology of "negligible sets" introduced in Definition 6.9. We will always work on the time interval $[0, 1]$, but all the results are independent (modulo change of constants) of the time interval, provided that its length is bounded from above and from below by two positive constants independent of ε. This is seen easily from the scaling properties of the Wiener process.

The results of this section are descendents of similar results obtained in [MP06, BM07] by related techniques. In [BM07] it was proven that if a Wiener polynomial, with continuous, bounded variation coefficients, is identically zero on an interval then so are its coefficients. This is enough to prove the almost sure invertibility of projections of the Malliavin matrix, which in turn implies the existence of a density for the projections of the transition probabilities. To prove smoothness of the densities or the ergodic results of this paper, more quantitative control is needed. In [BM07], a result close to (7.1) is claimed. However an error in Lemma 9.12 of that article leaves the proof incomplete. Arguing along similar, though slightly different lines, we prove the needed result below. We build upon the presentation in [BM07] but simplify it significantly. (The presentation in [BM07] was already a significant simplification over that in [MP06].)

Theorem 7.1 *Let $\{W_k\}_{k=1}^d$ be a family of i.i.d. standard Wiener processes and, for every multi-index $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, define $W_\alpha = W_{\alpha_1} \ldots W_{\alpha_\ell}$ with the convention that $W_\alpha = 1$ if $\alpha = \phi$. Let furthermore A_α be a family of (not necessarily adapted) stochastic processes with the property that there exists $m \geq 0$ such that $A_\alpha = 0$ whenever $|\alpha| > m$ and set $Z_A(t) = \sum_\alpha A_\alpha(t) W_\alpha(t)$.*

Then, there exists a universal family of negligible events Osc_W^m depending only on m such that the implication

$$\|Z_A\|_{L^\infty} \leq \varepsilon \quad \Longrightarrow \quad \begin{cases} either & \sup_\alpha \|A_\alpha\|_{L^\infty} \leq \varepsilon^{3^{-m}} \\ or & \sup_\alpha \|A_\alpha\|_{\mathrm{Lip}} \geq \varepsilon^{-3^{-(m+1)}} \end{cases} \tag{7.1}$$

holds modulo Osc_W^m. (The supremum norms are taken on the interval $[0,1]$.)

Remark 7.2 Informally, we can read the statement of Theorem 7.1 as "if Z_A is small, then either all of the coefficients A_α are small, or at least one of them oscillates very fast." The exponents appearing in the statement of Theorem 7.1 are somewhat arbitrary. By going through the proof more carefully, we can see that for any $\kappa > 2$, it is possible to find a constant $C_\kappa > 0$ such that the exponents in (7.1) can be replaced by κ^{-m} and $-C_\kappa \kappa^{-m}$ respectively. Here, the coefficient C_κ tends to 0 as $\kappa \to 2$. While the precise values of the exponents in (7.1) arising from our proof are unlikely to be sharp, they are not far from it, as can be seen by looking at processes of the form $Z(t) = \varepsilon^{1-\frac{\theta}{2}}(W_\theta(t) - W(t))$, where W_θ is the linear interpolation of the Wiener process W over intervals of size ε^θ.

Remark 7.3 The reason why the family of negligible sets appearing in this statement is called Osc_W^m is that it relies on the fact that the Wiener processes typically fluctuate sufficiently fast on every small time interval so that their effects can be distinguished from those of the multiplicators A_α which fluctuate over much longer timescales. It is important to note that Osc_W^m depends on the processes A_α only through the value of m.

Before we start with the proof, we show the following result, which is essentially the particular case of Theorem 7.1 where $m = 1$ and where the coefficients A_α do not depend on time. Here, $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbf{R}^d$.

Lemma 7.4 *Let $\{W_k\}_{k=1}^d$ be a collection of i.i.d. standard Wiener processes. Then, for any exponent $\kappa > 0$, there exists a universal family Osc_W of negligible events such that the bound*

$$\sup_{t \in [0,1]} |\langle A, W(t) \rangle| \geq \varepsilon^\kappa |A|, \tag{7.2}$$

holds modulo Osc_W for any choice of coefficients $A \in \mathbf{R}^d$.

Remark 7.5 We would like to stress again the fact that the family of events Osc_W is *independent* of the choice of coefficients A and depends only on the realisation of the W_k's.

Proof. Fix $\kappa > 0$ and define a family of events B by $B^\varepsilon = \{\sup_{t \in [0,1]} |W(t)| \geq \varepsilon^{-\kappa}\}$. It follows immediately from the fact that the supremum of a Wiener process has Gaussian tails that the family B is negligible. Consider now the unit sphere S^d in $\mathbf{R}^d$. For every $A \in S^d$, the process $W_A(t) = \langle A, W(t) \rangle$ is a standard Wiener process and

so $\mathbf{P}(\sup_{t\in[0,1]} |W_A(t)| \leq 2\varepsilon^\kappa) \leq C_1 \exp(-C_2\varepsilon^{-2\kappa})$ for some constants C_1 and C_2 that are independent of A. Denote this event by H_A^ε.

Choose now a collection $\{A_k\}$ of points in S^d such that $\sup_{A\in S^d} \inf_k |A - A_k| \leq \varepsilon^{2\kappa}$ and define $H^\varepsilon = \bigcup_k H_{A_k}^\varepsilon$. Since this can be achieved with $\mathcal{O}(\varepsilon^{-2\kappa(d-1)})$ points, the family H is negligible by Lemma 6.10. We now define $\mathrm{Osc}_W = H \cup B$ and we note that, modulo Osc_W, one has for every $\bar{A} \in \mathbf{R}^d$ the bound

$$\sup_{t\in[0,1]} |\langle \bar{A}, W(t)\rangle| \geq |\bar{A}| \inf_{A\in S^d} \sup_{t\in[0,1]} |\langle A, W(t)\rangle|$$

$$\geq |\bar{A}| \Big(\inf_k \sup_{t\in[0,1]} |\langle A_k, W(t)\rangle| - \varepsilon^\kappa \Big) \geq |\bar{A}|\varepsilon^\kappa ,$$

as required. $\qquad\qquad\square$

We now turn to the

Proof of Theorem 7.1. The proof proceeds by induction on the parameter m. For $m = 0$, the statement is trivial since in this case one has $Z_A(t) = A_\phi(t)$, so that one can take $\mathrm{Osc}_W^0 = \phi$.

Fix now a value $m \geq 1$ and assume that, for some ε, both inequalities

$$\|Z_A\|_{L^\infty} \leq \varepsilon , \tag{7.3a}$$

$$\sup_{|\alpha| \leq m} \|A_\alpha\|_{\mathrm{Lip}} \leq \varepsilon^{-3^{-(m+1)}} \tag{7.3b}$$

hold. Our aim is to find a (universal) family of negligible sets Osc_W^m such that, modulo Osc_W^m, these two bounds imply the bound $\sup_\alpha \|A_\alpha\|_{L^\infty} \leq \varepsilon^{3^{-m}}$. Before we proceed, we localise our argument to Wiener processes that do not behave too "wildly." Using the fact that the Hölder norm of a Wiener process has Gaussian tails for every Hölder exponent smaller than $1/2$, we see that the bounds

$$\sup_{t\in[0,1]} \sup_{|\alpha|\leq m} |W_\alpha(t)| \leq \varepsilon^{-1/10} , \quad \sup_{s\neq t} \sup_{|\alpha|\leq m} \frac{|W_\alpha(t) - W_\alpha(s)|}{|t-s|^{2/5}} \leq \varepsilon^{-1/30} , \quad (7.4)$$

both hold modulo some universal family Wien of negligible events. The reason for these particular choices of exponents will become clearer later on, but any two negative exponents would have been admissible.

Choose an exponent κ to be determined later and define a sequence of times $t_\ell = \ell\varepsilon^\kappa$ for $\ell = 0, \ldots, \varepsilon^{-\kappa}$, so that the interval $[0,1]$ gets divided into $\varepsilon^{-\kappa}$ subintervals of the form $[t_\ell, t_{\ell+1}]$. We define $A_\alpha^\ell = A_\alpha(t_\ell)$ and similarly for W_α^ℓ. We also define the Wiener increments $\bar{W}_i^\ell(t) = W_i(t) - W_i(t_\ell)$ and their products $\bar{W}_\alpha^\ell = \Pi_{j\in\alpha} \bar{W}_j^\ell$. With these notations, one has for $t \in [t_\ell, t_{\ell+1}]$ the equality

$$Z_A(t) = Z_A(t_\ell) + \sum_{\alpha\neq\phi} A_\alpha^\ell(W_\alpha(t) - W_\alpha^\ell) + \sum_\alpha (A_\alpha(t) - A_\alpha^\ell)W_\alpha(t) \tag{7.5}$$

$$= Z_A(t_\ell) + \sum_{\alpha\neq\phi} \sum_{\substack{\sigma\subset\alpha \\ \sigma\neq\phi}} A_\alpha^\ell W_{\alpha\setminus\sigma}^\ell \bar{W}_\sigma^\ell(t) + \sum_\alpha (A_\alpha(t) - A_\alpha^\ell)W_\alpha(t)$$

$$= Z_A(t_\ell) + \sum_\nu \sum_{\sigma\neq\phi} C_{\nu,\sigma} A_{\nu\cup\sigma}^\ell W_\nu^\ell \bar{W}_\sigma^\ell(t) + \sum_\alpha (A_\alpha(t) - A_\alpha^\ell)W_\alpha(t)$$

$$\equiv Z_A(t_\ell) + \sum_\nu \sum_{j=1}^d C_{\nu,(j)} A_{\nu\cup(j)}^\ell W_\nu^\ell \bar{W}_j^\ell(t) + E_\ell(t) ,$$

for some "error term" E_ℓ that will be analysed later. Here, the combinatorial factor $C_{\alpha,\sigma}$ counts the number of ways in which the multi-index σ can appear in the multi-index $\alpha \cup \sigma$ (for example $C_{(i,j),(j)}$ is equal to 2 if $i \neq j$ and 3 if $i = j$). Using the Brownian scaling and the fact that the supremum of a Wiener process has Gaussian tails, we see that for every $\kappa' < \kappa$, the bound

$$\sup_{\ell \leq \varepsilon^{-\kappa}} \sup_{t \in [0,\varepsilon^\kappa]} \sup_{j \in \{1,\ldots,d\}} |\bar{W}_j^\ell(t)| \leq \varepsilon^{\kappa'/2} , \tag{7.6}$$

holds modulo some universal family $\text{Wien}_{\kappa',m}$ of negligible events.

Note now that all the terms appearing in E_ℓ are (up to combinatorial factors) either of the form $A_{\alpha \cup \sigma}^\ell W_\alpha^\ell \bar{W}_\sigma(t)$ with $|\sigma| \geq 2$, or of the form $(A_\alpha(t) - A_\alpha^\ell)W_\alpha(t)$. Together with (7.6) and the first bound in (7.4), this shows that there exists a constant C depending only on m such that (7.3b) implies

$$\sup_{\ell \leq \varepsilon^{-\kappa}} \sup_{t \in [t_\ell,t_{\ell+1}]} |E_\ell(t)| \leq C(\varepsilon^{\kappa'-1/27-1/10} + \varepsilon^{\kappa-1/9-1/10}) , \tag{7.7}$$

modulo $\text{Wien}_{\kappa',m}$. Here we used the fact that (7.3b) implies in particular that the bound $\|A_\alpha\|_{L^\infty} \leq \varepsilon^{-1/27}$ holds for every α with $|\alpha| \geq 2$ (note that these terms are non-zero only if $m \geq 2$) and that $\|A_\alpha\|_{\text{Lip}} \leq \varepsilon^{-1/9}$, since we assumed $m \geq 1$. At this point, we fix $\kappa = \frac{5}{4}$ and $\kappa' = \frac{6}{5}$, so that in particular both exponents appearing in (7.7) are greater than 1. We then define $\text{Wien}' = \text{Wien} \cup \text{Wien}_{\kappa',m}$ so that, modulo Wien', (7.3a) and (7.5) imply

$$\sup_{t \in [t_\ell,t_{\ell+1}]} \left| \sum_\alpha \sum_{j=1}^d C_{\alpha,(j)} A_{\alpha \cup (j)}^\ell W_\alpha^\ell \bar{W}_j(t) \right| \leq 2\varepsilon + \sup_{\ell \leq \varepsilon^{-\kappa}} \sup_{t \in [t_\ell,t_{\ell+1}]} |E_\ell(t)| \leq C\varepsilon . \tag{7.8}$$

The left hand side of this expression motivates the introduction of operators M_j acting on the set of families of stochastic processes by

$$(M_j A)_\alpha = C_{\alpha,(j)} A_{\alpha \cup (j)} .$$

Note that M_j lowers the "degree" of A by one in the sense that if $A_\alpha = 0$ for every $|\alpha| \geq m$, then $(M_j A)_\alpha = 0$ for every $|\alpha| \geq m - 1$.

With this notation, we can rewrite (7.8) as

$$\sup_{t \in [t_\ell,t_{\ell+1}]} \left| \sum_{j=1}^d Z_{M_j A}(t_\ell) \bar{W}_j(t) \right| \leq C\varepsilon . \tag{7.9}$$

Using the Brownian scaling and applying Lemma 7.4, combined with Lemma 6.10, shows the existence of a family Osc_W of negligible events such that (7.9) implies

$$|Z_{M_j A}(t_\ell)| \leq \varepsilon^{7/20} , \quad \forall \ell \leq \varepsilon^{-5/4} .$$

Here, we used the fact that our choice of κ implies that $1 - \kappa/2 > 7/20$. This shows that the statements (7.3) imply

$$\|Z_{M_j A}\|_{L^\infty} \leq \varepsilon^{7/20} + C_m \sup_\alpha (\varepsilon^\kappa \|A_\alpha\|_{\text{Lip}} \|W_\alpha\|_{L^\infty} + \varepsilon^{2\kappa/5} \|A_\alpha\|_{L^\infty} \|W_\alpha\|_{C^{2/5}})$$

$$\leq \varepsilon^{7/20} + C_m(\varepsilon^{\kappa-1/9-1/0} + \varepsilon^{1/2-1/9-1/30}) \leq C_m \varepsilon^{7/20} , \tag{7.10}$$

modulo $\mathrm{Wien}' \cup \mathrm{Osc}_W$. Here, the constant $C_m > 1$ depends only on m.

We now finally arrived at the stage where we are able to apply our induction hypothesis to each of the processes $Z_{M_j A}$. Note that since $7/20 > 1/3$, (7.3b) implies that

$$\sup_{\alpha,j} \|(M_j A)_\alpha\|_{\mathrm{Lip}} \leq (C_m \varepsilon^{7/20})^{-3^{-m}} \; ,$$

for all sufficiently small ε. Therefore, outside of the event $(\mathrm{Osc}_W^{m-1})_{C_m \varepsilon^{7/20}}$, one has the implication

$$\Big\{ \sup_j \|Z_{M_j A}\|_{L^\infty} \leq C_m \varepsilon^{7/20} \Big\} \, \& \, \Big\{ \sup_\alpha \|A_\alpha\|_{\mathrm{Lip}} \leq \varepsilon^{-3^{-(m+1)}} \Big\}$$
$$\implies \Big\{ \sup_{\alpha,j} \|(M_j A)_\alpha\|_{L^\infty} \leq C'_m \varepsilon^{\frac{7}{20} 3^{-(m-1)}} \Big\} \; , \tag{7.11}$$

for some different constant C'_m depending also only on m. Since $7/20 > 1/3$ and since $\|(M_j A)_\alpha\|_{L^\infty} \geq \|A_{\alpha \cup (j)}\|_{L^\infty}$, this implies in particular that $\|A_\alpha\|_{L^\infty} \leq \varepsilon^{3^{-m}}$ for every $\alpha \neq \phi$.

In order to conclude the proof of the theorem, it therefore only remains to obtain a similar bound on $\|A_\phi\|_{L^\infty}$. We define a family of negligible events Wien''_m so that $\mathrm{Wien}' \subset \mathrm{Wien}''_m$ and such that the bound

$$\sup_{t \in [0,1]} \sup_{|\alpha| \leq m} |W_\alpha(t)| \leq \varepsilon^{-\frac{1}{70} 3^{-(m-1)}} \; , \tag{7.12}$$

holds modulo Wien''_m. We claim that if we define recursively

$$(\mathrm{Osc}_W^m)_\varepsilon = (\mathrm{Osc}_W^{m-1})_{C \varepsilon^{7/20}} \cup (\mathrm{Wien}''_m)_\varepsilon \; ,$$

the family Osc_W^m has the requested properties. It follows indeed from (7.3a), (7.12) and the definition of Z_A that, modulo Osc_W^m, (7.3) imply the bound

$$\|A_\phi\|_{L^\infty} \leq \varepsilon + \sum_{\alpha \neq \phi} \|A_\alpha\|_{L^\infty} \|W_\alpha\|_{L^\infty} \leq \varepsilon + C'_m \varepsilon^{(7/20 - 1/70)3^{-(m-1)}} \; . \tag{7.13}$$

Since we choose the bound (7.12) in such a way that $7/20 - 1/70 > 1/3$, we obtain $\|A_\phi\|_{L^\infty} \leq \varepsilon^{1/3}$ for sufficiently small ε. Together with the remark following (7.11), this concludes the proof of Theorem 7.1. $\qquad\square$

8   Examples

In this section, we apply the abstract framework developed in this article to two concrete examples: the stochastic Navier-Stokes equations on a sphere and a class of stochastic reaction-diffusion equations. The examples are chosen in order to highlight the techniques that can be used to verify the assumptions of our results and to get some idea of their scope of applicability. In particular, the Navier-Stokes equations provide an example where bounds on the Jacobian are not very uniform, so that an initial condition dependent control is required in Assumption C.1. The stochastic reaction-diffusion system on the other hand satisfies very strong a priori bounds, but Assumption A.1 is not verified with the usual choice $\mathcal{H} = L^2$, so that one has to work a bit more to fit the equations into the framework presented here. Our strategy is as follows: in a first section, we provide a simplified version of our results. We tried to find a formulation that strikes a balance between powerful results and easily verifiable assumptions. This general formulation will then be used by both of the examples mentioned above.

8.1 A general formulation

The 'general purpose' theorem formulated in this section allows to obtain the asymptotic strong Feller property for a large class of semilinear SPDEs under a Hörmander-type bracket condition. Our first assumption ensures that all the stability conditions of the previous sections can be verified.

Assumption D.1 *The operator L has compact resolvent. Furthermore, there exists a measurable function $V: \mathcal{H} \to \mathbf{R}_+$ such that there exist constants $c > 0$ and $\alpha > 0$ such that the bound*

$$V(u) \geq c\|u\|^{\alpha} \,,$$

holds for all $u \in \mathcal{H}$ and such that the following bounds hold:
   There exists a constant $C > 0$ and $\eta' \in [0, 1)$ such that

$$\mathbf{E} \exp(V(u_1)) \leq C \exp(\eta' V(u_0)) \,. \tag{8.1}$$

We also require the following bounds on the Jacobian, as well as the second variation on the dynamic. For every $p > 0$ and every $\delta > 0$, there exists a constant C such that the bounds

$$\sup_{t \in [0,1]} \mathbf{E}\|u_t\|^p \leq C \exp(\delta V(u_0)) \,, \tag{8.2a}$$

$$\mathbf{E} \sup_{s,t \in [0,1]} \|J_{s,t}\|^p \leq C \exp(\delta V(u_0)) \,, \tag{8.2b}$$

$$\sup_{s,t \in [0,1]} \mathbf{E}\|J_{s,t}^{(2)}\|^p \leq C \exp(\delta V(u_0)) \,, \tag{8.2c}$$

hold for every $u_0 \in \mathcal{H}$.

Our next assumption is simply a restatement of the Hörmander bracket condition (considering only constant 'vector fields'), with the additional condition that the g_i belong to $\mathcal{H}_\infty$. This ensures that all the relevant brackets are in $\mathcal{H}_\infty$ and hence admissible in the sense of Section 6.2.

Assumption D.2 *The forcing directions g_i belong to $\mathcal{H}_\infty$. Furthermore, define a sequence of subsets of $\mathcal{H}$ recursively by $\mathrm{A}_0 = \{g_j \, : \, j = 1, \ldots, d\}$ and*

$$\mathrm{A}_{k+1} \stackrel{\text{def}}{=} \mathrm{A}_k \cup \{N_m(h_1, \ldots, h_m) \, : \, h_j \in \mathrm{A}_k\} \,.$$

Then, the linear span of $\mathrm{A}_\infty \stackrel{\text{def}}{=} \bigcup_{n>0} \mathrm{A}_n$ is dense in $\mathcal{H}$.

With these assumptions in hand, a simplified, yet sufficiently powerful for many uses, formulation of our main results is as follows:

Theorem 8.1 *Consider the setting of equation (1.1) and assume that Assumptions A.1, D.1, and D.2 hold. Then, there exist constants $C, \kappa > 0$ and $\gamma \in (0, 1)$ such that the Markov semigroup $\mathcal{P}_t$ generated by (1.1) satisfies the bound*

$$\|\mathrm{D}(\mathcal{P}_{2n}\varphi)(u)\| \leq Ce^{\kappa V(u_0)}\left(\sqrt{(\mathcal{P}_{2n}\varphi^2)(u)} + \gamma^n \sqrt{(\mathcal{P}_{2n}\|\mathrm{D}\varphi\|^2)(u)}\right) \,, \tag{8.3}$$

for every integer $n > 0$. In particular, it satisfies the asymptotic strong Feller property.
   Furthermore, if $\beta_\star > a - 1$, then for every $m > 0$, every $u \in \mathcal{H}$, and every linear map $\mathcal{T}: \mathcal{H} \to \mathbf{R}^m$, the projections of the time-2 transition probabilities $\mathcal{T}^*\mathcal{P}_2(u, \cdot)$ have $\mathcal{C}^\infty$ densities with respect to Lebesgue measure on $\mathbf{R}^m$.

Remark 8.2 The final times 1 and 2 appearing in the statement are somewhat arbitrary since it suffices to rescale the equation in time, which does not change any of our assumptions. We chose to keep them in this way in order to avoid awkward notations in the proof.

In this result, the Hörmander-type assumption, Assumption C.2 is verified by using constant vector fields only. Before we turn to the proof of Theorem 8.1, we therefore present the following useful little lemma:

Lemma 8.3 *Let $\mathcal{H}$ be a separable Hilbert space and $\{g_i\}_{i=1}^{\infty} \subset \mathcal{H}$ a collection of elements such that its span is dense in $\mathcal{H}$. Define a family of symmetric bilinear forms $\mathcal{Q}_n$ on $\mathcal{H}$ by $\langle h, \mathcal{Q}_n h \rangle = \sum_{i=1}^{n} \langle g_i, h \rangle^2$. Let $\Pi \colon \mathcal{H} \to \mathcal{H}$ be any orthogonal projection on a finite-dimensional subspace of $\mathcal{H}$. Then, there exists $N > 0$ and, for every $\alpha > 0$ there exists $c_\alpha > 0$ such that $\langle h, \mathcal{Q}_n h \rangle \geq c_\alpha \|\Pi h\|^2$ for every $h \in \mathcal{H}$ with $\|\Pi h\| \geq \alpha \|h\|$ and every $n \geq N$.*

Proof. Assume by contradiction that the statement does not hold. Then, there exists $\alpha > 0$ and a sequence h_n in $\mathcal{H}$ such that $\|\Pi h_n\| = 1$, $\|h_n\| \leq \alpha^{-1}$, and such that $\lim_{n \to 0} \langle h_n, \mathcal{Q}_n h_n \rangle \to 0$. Since $\|h_n\| \leq \alpha^{-1}$ is bounded, we can assume (modulo extracting a subsequence) that there exists $h \in \mathcal{H}$ such that $h_n \to h$ in the weak topology. Since Π has finite rank, one has $\|\Pi h\| = 1$. Furthermore, since the maps $h \mapsto \langle h, Q_n h \rangle$ are continuous in the weak topology and since $n \mapsto \langle h, Q_n h \rangle$ is increasing for every n, one has

$$\langle h, Q_n h \rangle = \lim_{m \to \infty} \langle h_m, Q_n h_m \rangle \leq \lim_{m \to \infty} \langle h_m, Q_m h_m \rangle = 0 ,$$

so that $\langle h, g_i \rangle = 0$ for every $i > 0$. This contradicts the fact that the span of the g_i is dense in $\mathcal{H}$. $\qquad\square$

We are now in a position to turn to the proof of our general result.

Proof of Theorem 8.1. We show first that the supremum in (8.2a) can easily be pulled under the expectation. Indeed, it follows from the variation of constants formula that we have the bound

$$\|u_t\| \leq \|S(t)u_0\| + C \int_0^s (t-s)^{-a} \|N(u_s)\|_{-a}\, ds + \|W_L(t)\| ,$$

where W_L is the stochastic convolution of $G\,W$ with the semigroup S generated by L. It follows immediately from Hölder's inequality that there exists a constant C and an exponent $p > 0$ such that

$$\sup_{t \leq 1} \|u_t\| \leq \|u_0\| + C \Big(\int_0^1 (1 + \|u_s\|^{np})\, ds \Big)^{1/p} + \sup_{t \leq 1} \|W_L(t)\| .$$

Combining this with (8.2a), we conclude immediately that for every $p > 0$ and every $\delta > 0$ there exists $C > 0$ such that

$$\mathbf{E} \sup_{t \in [0,1]} \|u_t\|^p \leq C \exp(\delta V(u_0)) . \tag{8.4}$$

We now verify that Assumptions C.1 and C.2 are satisfied for our problem. It follows from (8.4) and (8.2b) that for every $\delta > 0$, Assumption C.1 holds with the choice $\Psi_0(u) = \exp(\delta V(u))$.

Furthermore, Assumption C.2 holds for every finite-rank orthogonal projection $\Pi\colon \mathcal{H} \to \mathcal{H}$ by Assumption D.2 and Lemma 8.3. Note that the function Λ_α is then constant, so that the condition on its moments is trivially satisfied. We can therefore apply Theorem 6.7 which states that for every $\alpha \in (0,1)$, every $\delta > 0$, every finite-rank projection Π, and every $p \geq 1$ there exists a constant C such that the bound

$$\mathbf{P}\Big(\inf_{\varphi \in \mathcal{S}_\alpha} \frac{\langle \varphi, \mathcal{M}_1 \varphi \rangle}{\|\varphi\|^2} \leq \varepsilon \Big) \leq C \exp(\delta V(u_0)) \varepsilon^p , \tag{8.5}$$

holds for every $u_0 \in \mathcal{H}$ and every $\varepsilon \leq 1$.

Combining this statement with (8.1), we see that Assumption B.1 is satisfied with $\bar{q} = 8$ (for example) and $U(u) = \exp(\delta V(u))$ with every $\delta \leq \frac{1}{8}$.

The bound (8.1) is nothing but a restatement of Assumption B.2. Since we assume that (8.2b) and (8.2c) hold for every $\delta > 0$, we infer that Assumption B.3 holds with $\bar{p} = 20$ and η sufficiently small. It remains to verify that, for every $C_\Pi > 0$ there exists a finite-rank projection Π such that (5.3) is satisfied. This ensures that the required relation $C_\Pi > C_J + 2\eta C_L/(1 - \eta')$ can be satisfied by a suitable choice of Π.

Because L has compact resolvent by assumption, it has a complete system of eigenvectors with the corresponding eigenvalues $\{\lambda_n\}$ satisfying $\lim_{n\to\infty} \lambda_n = \infty$. Therefore, if we denote by Π_N the projection onto the subspace of $\mathcal{H}$ spanned by the first N eigenfunctions, we have the identity

$$\|e^{-Lt}\Pi_N^\perp\| = e^{-\lambda_{N+1}t} .$$

This allows us to get a bound on $J_{0,1}\Pi^\perp$ as follows. It follows from (3.10) and the variation of constants formula that

$$\|J_{0,t}\Pi^\perp\| \leq \|e^{-Lt}\Pi^\perp\| + \int_0^t C s^{-a} \|DN(u_s)\|_{-a} \, ds$$
$$\leq e^{-\lambda_{N+1}t} + C t^{1-a} \sup_{s \leq t} \|u_s\|^k ,$$

so that, for every $\delta > 0$, we have by (8.4) the bound

$$\mathbf{E}\|J_{0,t}\Pi^\perp\|^p \leq C_{\delta,p}(e^{-\lambda_{N+1}t} + t^{1-a}) \exp(\delta V(u_0)) ,$$

for some family of constants $C_{\delta,p}$ independent of $t \in [0,1]$. Since $a < 1$, it follows that for every $\varepsilon, \delta > 0$ and $p > 0$, we can find N sufficiently large and t sufficiently small such that

$$\mathbf{E}\|J_{0,t}\Pi^\perp\|^p \leq \varepsilon \exp(\delta V(u_0))) .$$

Combining this with (8.2b) and the fact that $\|J_{0,1}\Pi^\perp\| \leq \|J_{t,1}\| \|J_{0,t}\Pi^\perp\|$, we obtain

$$\mathbf{E}\|J_{0,1}\Pi^\perp\|^{\bar{p}} \leq \Big(\mathbf{E}\|J_{0,t}\|^{2\bar{p}} \mathbf{E}\|J_{t,1}\|^{2\bar{p}}\Big)^{\frac{1}{2}} \leq C\varepsilon \exp(2\delta V(u_0)) ,$$

provided that N is sufficiently large. By choosing δ sufficiently small, it follows that Assumption B.4 (with arbitrary values for $\bar{p}$ and C_Π) can always be satisfied by choosing for Π the projection onto the first N eigenvectors of L for some large enough value of N. The bound (8.3) now follows from a simple application of Theorem 5.5.

It remains to prove the statement about the smoothness of $\mathcal{T}^*\mathcal{P}_2(u,\cdot)$, which will be a consequence of (8.5) by [Nua95, Cor. 2.1.2]. The reason why we consider the process at time 2 is that, in order to avoid the singularity at the origin, we consider the

solution u_2 as an element of the probability space with Gaussian structure given by the increments of W over the interval $[1, 2]$. The increments of W over $[0, 1]$ are then considered as some "redundant" randomness, which is irrelevant by [Nua95, Ch. 1]. With this slightly tweaked Gaussian structure, the Malliavin matrix of Πu_2 is given almost surely by $\Pi \mathcal{M}_1(u_1) \Pi$, where $\mathcal{M}_1$ is defined as before, but over the interval $[1, 2]$. The claim now follows from (8.5) and (8.1), provided that the random variable Πu_2 belongs to the space $\mathcal{D}^\infty$ of random variables whose Malliavin derivatives of all orders have moments of all orders.

Recall now (see for example [BM07, Section 5.1]) that for any n-tuple of elements $h_1, \ldots, h_n \in L^2([1, 2], \mathbf{R}^d)$, the nth Malliavin derivative of u_2 in the directions $h_1, \ldots, h_n$ is given by

$$\mathcal{D}^n u_2(h) = \int_{1 \le s_1 < \cdots < s_n \le 2} J_{s,1}^{(n)} G h_s \, ds \, . \tag{8.6}$$

Applying (6.1a) in Proposition 6.2 we see that, for every $u_0 \in \mathcal{H}$, every $\gamma < \gamma_\star + 1$, and every $p > 0$, one has the bound

$$\mathbf{E} \sup_{t \in [1,2]} \|u_t\|_\gamma^p < \infty \, .$$

We conclude from Proposition 3.11 that

$$\mathbf{E} \sup_{1 \le s_1 < \cdots < s_n \le 2} \sup_{\|\varphi_j\| \le 1} \|J_{s,t}^{(k)}(\varphi_1, \ldots, \varphi_k)\|^p \le \infty \, ,$$

so that, by (8.6), u_2 does indeed have Malliavin derivatives of all orders with bounded moments of all orders. This concludes the proof. $\qquad \square$

8.2 The 2D Navier-Stokes equations on a sphere

Consider the stochastically forced two-dimensional Navier-Stokes equations on the two-dimensional sphere S^2:

$$du = \nu \Delta u \, dt + \nu \operatorname{Ric} u \, dt - \nabla_u u \, dt - \nabla p \, dt + Q \, dW(t) \, , \qquad \operatorname{div} u = 0 \, . \tag{8.7}$$

Here, the velocity field u is an element of $H^1(S^2, TS^2)$, $\nabla_u u$ denotes the covariant differentiation of u along itself with respect to the Levi-Civita connection on S^2, $\Delta = -\nabla^* \nabla$ is the (negative of the) Bochner Laplacian on S^2, and Ric denotes the Ricci operator from TS^2 into itself. In the case of the sphere, the latter is just the multiplication with the scalar 1. See also [Tay92, TW93, Nag97] for more details on the Navier-Stokes equations on manifolds.

As in the flat case, it is possible to represent u uniquely by a scalar "vorticity" field w given by

$$w = \operatorname{curl} u \stackrel{\text{def}}{=} -\operatorname{div}(n \wedge u) \, , \tag{8.8}$$

where n denotes the unit vector in $\mathbf{R}^3$ normal to the surface of the sphere (so that $n \wedge u$ defines again a vector field on the sphere). With this notation, one can rewrite (8.7) as

$$dw = \nu \Delta w \, dt - \operatorname{div}(w \, Kw) \, dt + G \, dW(t) \, . \tag{8.9}$$

Here, we denoted by K the operator that reconstructs a velocity field from its vorticity field, that is

$$u = Kw = -\operatorname{curl} \Delta^{-1} w \stackrel{\text{def}}{=} n \wedge \nabla \Delta^{-1} w \, ,$$

and Δ denotes the Laplace-Beltrami operator on the sphere. See [TW93] for a more detailed derivation of these equations. In order to fit the framework developed in this article, we assume that the operator G is of finite rank and that its image consists of smooth functions, so that the noise term can be written as

$$G\,dW(t) = \sum_{i=1}^{n} g_i\,dW_i(t)\,, \qquad g_i \in H^{\infty}(S^2, \mathbf{R})\,.$$

We choose to work in the space $\mathcal{H} = L^2(S^2, \mathbf{R})$ for the equation (8.9) in vorticity formulation, so that the interpolation spaces $\mathcal{H}_\alpha$ coincide with the fractional Sobolev spaces $H^{2\alpha}(S^2, \mathbf{R})$, see [Tri86]. In particular, elements $w \in \mathcal{H}_\alpha$ are characterised by the fact that the functions $x \mapsto \varphi(x)w(\psi(x))$ belong to $H^{2\alpha}(\mathbf{R}^2)$ for any compactly supported smooth function φ and any function $\psi\colon \mathbf{R}^2 \to S^2$ which is smooth on an open set containing the support of φ. Since the sphere is compact, this implies that the usual Sobolev embeddings for the torus also hold true in this case.

Define now $\mathrm{A}_0 = \{g_i\,:\,i = 1, \ldots, n\}$ and set recursively

$$\mathrm{A}_{n+1} = \mathrm{A}_n \cup \big\{B(v,w)\,:\,v, w \in \mathrm{A}_n\big\}\,,$$

where we made use of the symmetrised nonlinearity

$$B(v,w) = \tfrac{1}{2}\big(\mathrm{div}(w\,Kv) + \mathrm{div}(v\,Kw)\big)\,.$$

We then have the following result:

Theorem 8.4 *If the closure of the linear span of $\mathrm{A}_\infty = \bigcup_{n \geq 0} \mathrm{A}_n$ is equal to all of $L^2(S^2, \mathbf{R})$, then the equations (8.7) have a unique invariant measure.*

Remark 8.5 Sufficient conditions for density of A_∞ and for approximate controllability are given in [AS08]. In particular, the authors there give an example of A_0 containing five spherical harmonics that satisfies our condition. Note however that controllability is not required for our result to hold, since we only use the fact that the origin belongs to the topological support of every invariant measure. On the other hand, as shown in [MP06], controllability allows to obtain positivity of the projected densities of transition probabilities.

Proof. The main step in the proof is to check that we can apply Theorem 5.5 to conclude that the Markov semigroup generated by the solutions to (8.7) has the asymptotic strong Feller property. Let us first check that the Navier-Stokes nonlinearity on the sphere does indeed satisfy Assumption A.1 for some $a \in [0, 1)$. It is clear that the nonlinearity N, defined by $N(w) = B(w, w)$, is continuous from $\mathcal{H}_\infty$ to $\mathcal{H}_\infty$ (which coincides with the space of infinitely differentiable functions on the sphere), so in order to show point 2, it remains to show that N maps $\mathcal{H}_\gamma$ into $\mathcal{H}_{\gamma - a}$ for a range of values $\gamma \geq 0$ and some $a \in [0, 1)$.

Setting $\hat{B}(w, w') = \mathrm{div}(w\,Kw')$ so that $N(w) = B(w, w) = \hat{B}(w, w)$, one can show exactly as in [CF88] that, for any triplet (s_1, s_2, s_3) with $s_i \geq 0$, $\sum_i s_i > 1$, one has bounds of the type

$$\int_{S^2} v(x)\,\hat{B}(w, w')(x)\,dx \leq C\|v\|_{H^{s_1}}\|w\|_{H^{1+s_2}}\|w'\|_{H^{s_3 - 1}}\,,$$

$$\int_{S^2} v(x)\,\hat{B}(w, w')(x)\,dx \leq C\|v\|_{H^{1+s_1}}\|w\|_{H^{s_2}}\|w'\|_{H^{s_3 - 1}}\,,$$

for some constant C depending on the choice of the s_i. In particular, $\hat{B}$ can be interpreted as a continuous linear map from $\mathcal{H} \otimes \mathcal{H}$ into $\mathcal{H}_{-\frac{3}{4}}$ (for example) and from $\mathcal{H}_{\frac{1}{2}} \otimes \mathcal{H}_{\frac{1}{2}}$ into $\mathcal{H}$ (using the usual identification of bilinear maps with linear maps between tensor products). It thus follows from the Calderón-Lions interpolation theorem as in Remark 3.3 that $\hat{B}$ is a continuous linear map from $\mathcal{H}_\alpha \otimes \mathcal{H}_\alpha$ into $\mathcal{H}_\beta$ for $\beta = \frac{3\alpha}{2} - \frac{3}{4}$ and $\alpha \in [0, \frac{1}{2}]$. For $\alpha > \frac{1}{2}$, we use the fact that $\mathcal{H}_\alpha$ is an algebra [Tri92] to deduce that $\hat{B}$ is continuous from $\mathcal{H}_\alpha \otimes \mathcal{H}_\alpha$ into $\mathcal{H}_{\alpha - \frac{1}{2}}$. This shows that point 2 of Assumption A.1 is satisfied with $a = \frac{3}{4}$ (any exponent strictly larger than $\frac{1}{2}$ would do, actually) and $\gamma_\star = +\infty$.

Turning to point 3 of Assumption A.1, it suffices to show that, for v sufficiently smooth, the map $w \mapsto \hat{B}(v, w)$ is bounded from $\mathcal{H}_{-\beta}$ into $\mathcal{H}_{-\beta - a}$. It is well-known on the other hand that if $v \in \mathcal{C}^k$ then the multiplication operator $w \mapsto vw$ is continuous in H^s for all $|s| \leq k$. It follows immediately that $DN^*(v)$ is continuous from $\mathcal{H}_{-\beta}$ into $\mathcal{H}_{-\beta - \frac{1}{2}}$, provided that $v \in \mathcal{C}^k$ for $k \geq 2\beta$. Point 3 then follows with $\beta_\star = \infty$.

For any fixed $\eta > 0$, it follows exactly as in [HM06, Lemma 4.10] that Assumption D.1 is verified with $V(w) = \eta \|w\|^2$ for η sufficiently small. This concludes the verification of the assumptions of Theorem 8.1 and the claim follows. □

Remark 8.6 Just as in [HM06], this result is optimal in the following sense. The closure $\bar{A}_\infty$ of the linear span of A_∞ in L^2 is always an invariant subspace for (8.9) and the invariant measure for the Markov process restricted to $\bar{A}_\infty$ is unique. However, if $\bar{A}_\infty \neq L^2$, then one expects in general the presence of more than one invariant probability measure in L^2 at low values of the viscosity ν.

8.3 Stochastic reaction-diffusion equations

In this section, we consider a general class of reaction-diffusion equations on a "nice" domain D. The dimension m of the ambient space is chosen smaller or equal to 3 for technical reasons. However, the number ℓ of components in the reaction is arbitrary. The domain D is assumed to be either of

- A compact smooth m-dimensional Riemannian manifold.
- A bounded open domain of $\mathbf{R}^m$ with smooth boundary.
- A hypercube in $\mathbf{R}^m$.

We furthermore denote by Δ the Laplace (resp. Laplace-Beltrami) operator on D, endowed with either Neumann or Dirichlet boundary conditions. With these notations in place, the equations that we consider are

$$du = \Delta u\, dt + f \circ u\, dt + \sum_{i=1}^d g_i\, dW_i(t)\,, \qquad (8.10)$$

with $u(t)\colon D \to \mathbf{R}^\ell$ and $f\colon \mathbf{R}^\ell \to \mathbf{R}^\ell$ a polynomial of arbitrary degree n with $n \geq 3$ an odd integer. (We exclude the case $n = 1$ since this gives rise to a linear equation and is trivial to analyse.) The functions g_i describing the stochastic component of the equations are assumed to belong to $\mathcal{H}_\infty$, the intersection of the domains of Δ^α in $L^2(D)$ for all $\alpha > 0$. It is a straightforward exercise to check that (8.10) has unique local solutions in $\mathcal{E} = \mathcal{C}(D, \mathbf{R}^\ell)$ for every initial condition in $\mathcal{C}(D, \mathbf{R}^\ell)$ (replace $\mathcal{C}$ by $\mathcal{C}_0$ in the case of Dirichlet boundary conditions). In order to obtain global solutions, we make the following assumption on the nonlinearity:

Assumption RD.1 *Writing $f = \sum_{k=0}^{n} f_k$ for f with f_k being k-linear maps from $\mathbf{R}^\ell$ to itself, we assume that n is odd and that*

$$\langle f_n(u, \ldots, u, v), v \rangle < 0 \,,$$

for every $u, v \in \mathbf{R}^\ell \setminus \{0\}$.

Remark 8.7 Provided that Assumption RD.1 holds, one can check that there exist positive constants c and C such that the inequality

$$\langle f(u + v), u \rangle \leq C(1 + \|v\|^{n+1}) - c\|u\|^{n+1} \,, \tag{8.11}$$

holds for every $u, v \in \mathbf{R}^\ell$.

Essentially, Assumption RD.1 makes sure that the function $u \mapsto |u|^2$ is a Lyapunov function for the "reaction" part $\dot{u} = f(u)$ of (8.10). In the interest of brevity, we define $\mathrm{Sup}_{t,\infty}(v) = 1 + \sup_{s \leq t} \|v(s)\|_{\mathcal{E}}$ for any function $v \in L^\infty([0, t], \mathcal{E})$ and $\mathrm{Sup}_{t,r}(v) = 1 + \sup_{s \leq t} \|v(s)\|_{H^r}$ for $v \in L^\infty([0, t], H^r(D))$, As a consequence of Assumption RD.1, we obtain the following *a priori* bound on the solutions to (8.10):

Proposition 8.8 *Under Assumption RD.1, there exist constants c and C such that the bound*

$$\|u(t)\|_{L^\infty} \leq C\Big(\frac{\|u_0\|_{L^\infty}}{(1 + t\|u_0\|_{L^\infty}^n)^{1/n}} + \mathrm{Sup}_{t,\infty}(W_\Delta)\Big) \,,$$

holds almost surely for every $u_0 \in \mathcal{E}$, where $\mathcal{E}$ is either $\mathcal{C}(D, \mathbf{R}^\ell)$ or $\mathcal{C}_0(D, \mathbf{R}^\ell)$, depending on the boundary conditions of Δ. In particular, for every $t_0 > 0$ there exists a constant C such that one has the almost sure bound

$$\|u(t)\|_{L^\infty} \leq C\mathrm{Sup}_{t,\infty}(W_\Delta) \,, \tag{8.12}$$

independently of the initial condition, provided that $t \geq t_0$.

Proof. The proof is straightforward and detailed calculations for a variant of it can be found for example in [Hai08]. Setting $v = u - W_\Delta(t)$ where W_Δ is the "stochastic convolution" solving the linearised equation (8.10) with $f \equiv 0$, and defining $V(v) = \|v\|_{L^\infty}^2$, we obtain from (8.11) the almost sure bound

$$\frac{d}{dt} V(v(t)) \leq C\mathrm{Sup}_{t,\infty}^{n+1}(W_\Delta) - cV^{(n+1)/2}(v(t)) \,.$$

In particular, there exist possibly different constants such that

$$\frac{d}{dt} V(v(t)) \leq -cV^{(n+1)/2}(v(t))$$

for all v such that $V(v(t)) \geq C\mathrm{Sup}_{t,\infty}^2(W_\Delta))$. Since we assumed that $n \geq 3$, a simple comparison theorem for ODEs then implies that

$$V(v(t)) \leq C\frac{\|u_0\|^2}{(1 + t\|u_0\|^{2/\alpha})^\alpha} \wedge \mathrm{Sup}_{t,\infty}^2(W_\Delta) \,,$$

where we set $\alpha = 2/n$. The requested bound then follows at once. The second bound is an immediate consequence of the first one. $\qquad\square$

Remark 8.9 The function $t \mapsto V(v(t))$ is of course not differentiable in time in general. The left hand side in (8.12) should therefore be interpreted as the right upper Dini derivative $\limsup_{h \to 0+} h^{-1}(V(v(t+h)) - V(v(t)))$.

In order to fit the framework developed in this article, we cannot take L^2 as our base space, since the nonlinearity will not in general map L^2 into any Sobolev space of negative order. However, provided that $k > m/2$, the Sobolev spaces H^k form an algebra, so that the nonlinearity $u \mapsto N(u) \stackrel{\text{def}}{=} f \circ u$ is continuous from H^k to H^k in this case. It is therefore natural to choose $\mathcal{H} = H^k$ for some $k > m/2$. In this case, for $\alpha > 0$, the interpolation spaces $\mathcal{H}_\alpha$ coincide with the Sobolev spaces $H^{k+2\alpha}$, so that one has $N \in \text{Poly}(\mathcal{H}_\alpha, \mathcal{H}_\alpha)$ for every $\alpha > 0$. This shows that Assumption A.1 is satisfied with $a = 0$, $\gamma_\star = \infty$ and $\beta_\star = \infty$. It turns out that it is relatively easy to obtain bounds in the Sobolev space H^2. From now on, we do therefore assume that the following holds:

Assumption RD.2 *The space dimension m is smaller or equal to* 3.

This will allow us to work in $\mathcal{H} = H^2$. Before we state the main theorem of this section, we obtain a number of *a priori* bounds that will allow us to verify that the assumptions from the previous parts of this article do indeed apply to the problem at hand.

By using a bootstrapping argument similar to Proposition 3.6, we can obtain the following *a priori* estimate:

Proposition 8.10 *Assume that Assumptions RD.1 and RD.2 hold. If u is the solution to (8.10) with initial condition $u_0 \in H^2$ then there exists a constant C such that the bounds*

$$\|u(t)\|_{H^2} \leq C\text{Sup}_{t,\infty}^{2n}(u)(\|u_0\|_{H^2} + \text{Sup}_{t,2}(W_\Delta)) \,,$$

$$\|u(t)\|_{H^2} \leq C\text{Sup}_{t,\infty}^{2n}(u)(\frac{1}{t}\|u_0\|_{L^2} + \text{Sup}_{t,2}(W_\Delta)) \,,$$

hold for all $t \leq 1$ almost surely.

Proof. From Duhamel's formula, we obtain the bound

$$\|u(t)\|_{H^1} \leq \|u_0\|_{H^1} + \int_0^t \frac{C}{\sqrt{t-s}}\|f \circ u(s)\|_{L^2}\,ds + \text{Sup}_{t,1}(W_\Delta)$$

$$\leq \|u_0\|_{H^2} + C\sqrt{t}\text{Sup}_{t,\infty}^n(u) + \text{Sup}_{t,1}(W_\Delta) \,.$$

At this stage, we use that since f is a polynomial of degree n, there exists a constant C such that

$$\|f \circ u\|_{H^1} \leq C(1 + \|u\|_{L^\infty}^n + \|u\|_{L^\infty}^{n-1}\|u\|_{H^1}) \,. \tag{8.13}$$

Using Duhamel's formula again, this yields

$$\|u(t)\|_{H^2} \leq \|u_0\|_{H^2} + \int_0^t \frac{C}{\sqrt{t-s}}\|f \circ u(s)\|_{H^1}\,ds + \text{Sup}_{t,2}(W_\Delta)$$

$$\leq \|u_0\|_{H^2} + \int_0^t \frac{C}{\sqrt{t-s}}(1 + \|u(s)\|_{L^\infty}^n + \|u(s)\|_{L^\infty}^{n-1}\|u(s)\|_{H^1})\,ds$$

$$+ \text{Sup}_{t,2}(W_\Delta)$$

$$\leq \|u_0\|_{H^2} + \int_0^t \frac{C}{\sqrt{t-s}} \Big(\mathrm{Sup}_{t,\infty}^n(u) + \mathrm{Sup}_{t,\infty}^{n-1}(u)(\|u_0\|_{H^2}$$
$$+ \sqrt{s}\mathrm{Sup}_{t,\infty}^n(u) + \mathrm{Sup}_{t,1}(W_\Delta)) \Big) \, ds + \mathrm{Sup}_{t,2}(W_\Delta) \,.$$

Integrating the last term yields the first bound. The second bound can be obtained in exactly the same way, using the smoothing properties of the semigroup generated by the Laplacian. $\qquad\square$

As a consequence, we obtain the following bound on the exponential moments in H^2 of the solution starting from any initial condition:

Proposition 8.11 *For every $T > 0$, there exists a constant $C > 0$ such that*

$$\mathbf{E}\exp(\|u(T)\|_{H^2}^{1/n}) \leq C \,,$$

for every initial condition $u_0 \in H^2$.

Proof. Without loss of generality, we set $T = 1$. Combining Proposition 8.10 and the Markov property, we see that there exists a constant $C > 0$ such that

$$\|u(1)\|_{H^2} \leq C \Big(\sup_{\frac{1}{2} \leq s \leq 1} \|u(s)\|_{L^\infty}^n \Big) \Big(\|u(\tfrac{1}{2})\|_{L^2} + \mathrm{Sup}_{1,2}(W_\Delta) \Big)$$
$$\leq C \Big(\sup_{\frac{1}{2} \leq s \leq 1} \|u(s)\|_{L^\infty}^n \Big) \mathrm{Sup}_{1,2}(W_\Delta) \,.$$

The requested bound then follows from (8.12) and the fact that $\mathrm{Sup}_{1,2}(W_\Delta)$ has Gaussian tails by Fernique's theorem. $\qquad\square$

We now turn to bounds on the Jacobian J for (8.10). Recall from (3.10) that, given any "tangent vector" ξ, the Jacobian $J_{s,t}\xi$ satisfies the random PDE

$$\frac{d}{dt} J_{s,t}\xi = \Delta J_{s,t}\xi + (Df \circ u)(t)J_{s,t}\xi \,,$$

where Df denotes the derivative of the map f. Our main tool is the fact that, from Assumption RD.1, we obtain the existence of a constant $C > 0$ such that

$$\langle Df(u)v, v \rangle \leq C|v|^2 \,,$$

for every $u, v \in \mathbf{R}^\ell$. In particular, we obtain the *a priori* L^2 estimate:

$$\frac{d}{dt}\|J_{s,t}\xi\|_{L^2}^2 = -2\|\nabla J_{s,t}\xi\|_{L^2}^2 + 2\langle J_{s,t}\xi, (Df \circ u)(t)J_{s,t}\xi \rangle \leq 2C\|J_{s,t}\xi\|_{L^2}^2 \,, \quad (8.14)$$

so that $\|J_{s,t}\xi\|_{L^2} \leq e^{C(t-s)}\|\xi\|_{L^2}$ almost surely. We now us similar reasoning to obtain a sequence of similar estimates in smoother spaces.

Proposition 8.12 *For any $u_0 \in H^2$, the Jacobian satisfies the operator bounds*

$$\|J_{s,t}\|_{L^2 \to L^2} \leq C$$
$$\|J_{s,t}\|_{L^2 \to H^1} \leq C \Big(\frac{1}{\sqrt{t-s}} + \mathrm{Sup}_{t,\infty}^n(u) \Big) \,,$$
$$\|J_{s,t}\|_{H^1 \to H^1} \leq C\mathrm{Sup}_{t,\infty}^n(u) \,,$$
$$\|J_{s,t}\|_{H^2 \to H^2} \leq C\mathrm{Sup}_{t,\infty}^{4n}(u)(\|u_0\|_{H^2} + \mathrm{Sup}_{t,2}(W_\Delta))$$
$$\|J_{s,t}\|_{H^1 \to H^2} \leq C\mathrm{Sup}_{t,\infty}^{4n}(u)(\|u_0\|_{H^2} + \mathrm{Sup}_{t,2}(W_\Delta)) + \frac{C}{\sqrt{t-s}}$$

for $0 \leq s < t \leq 1$ with $\mathrm{Sup}_{t,\infty}$ defined just before Proposition 8.10.

Proof. The first estimate is just a rewriting of the calculation before the Proposition. As in the proof of the *a priori* bounds for the solution, we are going to use a bootstrapping argument, starting from the bound (8.14). Applying Duhamel's formula and using the notation $\mathrm{Sup}_{t,\infty}$ as before, we obtain

$$\|J_{s,t}\xi\|_{H^1} \leq \|\xi\|_{H^1} + \int_s^t \frac{C}{\sqrt{t-r}} \|(Df \circ u)(r)J_{s,r}\xi\|_{L^2}\, dr$$

$$\leq \|\xi\|_{H^1}\left(1 + \mathrm{Sup}_{t,\infty}^{n-1}(u)\int_s^t \frac{C}{\sqrt{t-r}} e^{C(r-s)}\, dr\right)$$

$$\leq \|\xi\|_{H^1}\left(1 + C\mathrm{Sup}_{t,\infty}^{n-1}(u)e^{C|t-s|}\sqrt{t-s}\right)$$

$$\leq \|\xi\|_{H^1} C\mathrm{Sup}_{t,\infty}^{n-1}(u)e^{C|t-s|} \ .$$

(And similarly for the second bound.) Regarding the H^2 norm of the Jacobian, we use the fact that there is a constant C such that the bound

$$\|Df(u)v\|_{H^1} \leq C(\|u\|_{L^\infty}^{n-1}\|v\|_{H^1} + \|u\|_{L^\infty}^{n-2}\|\nabla u\|_{L^4}\|v\|_{L^4})$$

$$\leq C\|u\|_{L^\infty}^{n-2}\|u\|_{H^2}\|v\|_{H^1}$$

holds. Hence we get similarly to before

$$\|J_{s,t}\xi\|_{H^2} \leq \|\xi\|_{H^2} + \int_s^t \frac{C}{\sqrt{t-r}} \|(Df \circ u)(r)J_{s,r}\xi\|_{H^1}\, dr$$

$$\leq \|\xi\|_{H^2} + \int_s^t \frac{C}{\sqrt{t-r}} \|u(r)\|_{L^\infty}^{n-1}\|u(r)\|_{H^2}\|J_{s,r}\xi\|_{H^1}\, dr$$

$$\leq C\mathrm{Sup}_{t,\infty}^{4n}(u)(\|u_0\|_{H^2} + \mathrm{Sup}_{t,2}(W_\Delta))\|\xi\|_{H^2} \ ,$$

which is the requested bound. To obtain the last bound, one proceeds identically except that one used $\|e^{L(t-s)}\xi\|_{H^2} \leq C\|\xi\|_{H^1}/\sqrt{t-s}$. $\qquad\square$

We now turn to the second variation.

Proposition 8.13 *For any $u_0 \in H^2$, the second variation $J^{(2)}$ of the solution to (8.10) satisfies*

$$\|J_{s,t}^{(2)}\|_{H^2 \otimes H^2 \to H^2} \leq C\mathrm{Sup}_{t,\infty}^{13n}(u)(\|u_0\|_{H^2} + \mathrm{Sup}_{t,2}(W_\Delta))^4$$

for $0 \leq s < t \leq 1$ with $\mathrm{Sup}_{t,\infty}$ defined just before Proposition 8.10.

Proof. Again using Duhamel's formula, we have

$$J_{s,t}^{(2)}(\varphi,\psi) = \int_s^t J_{r,t}D^2F(u_r)(J_{s,r}\varphi, J_{s,r}\psi)dr \ . \tag{8.15}$$

To control the H^2 norm we will need the following estimate:

$$\|\nabla^2 D^2 F(u)(\varphi, \psi)\|_{L^2} \leq C(1 + \|u\|_\infty^{n-2})(\|(\nabla^2 u)\varphi\psi\|_{L^2} + \|(\nabla u)^2 \varphi\psi\|_{L^2}$$
$$+ \|(\nabla^2 \varphi)\psi\|_{L^2} + \|\varphi(\nabla^2 \psi)\|_{L^2} + \|(\nabla\varphi)(\nabla\psi)\|_{L^2}$$
$$+ \|(\nabla u)(\nabla\psi)\varphi\|_{L^2} + \|(\nabla u)(\nabla\varphi)\psi\|_{L^2})$$
$$\leq C(1 + \|u\|_\infty^{n-2})(\|u\|_{H^2}\|\varphi\|_{L^\infty}\|\psi\|_{L^\infty} + \|u\|_{H^2}^2\|\varphi\|_{L^\infty}\|\psi\|_{L^\infty}$$
$$+ \|\varphi\|_{H^2}\|\psi\|_{L^\infty} + \|\varphi\|_{L^\infty}\|\psi\|_{H^2} + \|\varphi\|_{H^2}\|\psi\|_{H^2}$$
$$+ \|u\|_{H^2}\|\varphi\|_{L^\infty}\|\psi\|_{H^2} + \|u\|_{H^2}\|\varphi\|_{H^2}\|\psi\|_{L^\infty})$$
$$\leq C(1 + \|u\|_\infty^{n-2})(1 + \|u\|_{H^2}^2)\|\varphi\|_{H^2}\|\psi\|_{H^2} .$$

In this estimate, we have used repeatedly the fact that $\|v\|_{L^4} \leq C\|v\|_{H^1}$ and $\|v\|_{L^\infty} \leq C\|v\|_{H^2}$. Using this estimate in (8.15), we obtain

$$\|J_{s,t}^{(2)}(\varphi, \psi)\|_{H^2} \leq C\mathrm{Sup}_{t,\infty}^{n-2}(u)\int_s^t (1 + \|u_r\|_{H^2}^2)\|J_{r,t}\|_{H^2 \to H^2}\|J_{s,r}\varphi\|_{H^2}\|J_{s,r}\psi\|_{H^2} dr$$
$$\leq C\mathrm{Sup}_{t,\infty}^{17n}(u)(\|u_0\|_{H^2} + \mathrm{Sup}_{t,2}(W_\Delta))^5\|\psi\|_{H^2}\|\varphi\|_{H^2} ,$$

which completes the proof. $\qquad\square$

We now set the stage to prove the analogue of Theorem 5.5 for equation (8.10). We begin by collecting a number for relevant results implied by the preceding calculations. We will work in H^2 since this will be the base space for what follows.

Proposition 8.14 *Define $V(u) = \|u\|_{H^2}^{1/n}$. Then, for every $p > 0$ there exists a constant C_p and, for every $\eta > 0$ and $p > 0$ there exists a constant $C_{\eta,p}$ so that the bounds*

$$\mathbf{E}\sup_{1/2 \leq t \leq 1}\|u(t)\|_{H^2}^p \leq C_p$$
$$\mathbf{E}\sup_{1/2 \leq s < t \leq 1}\|J_{s,t}\|_{H^2 \to H^2}^p \leq C_p$$
$$\sup_{0 \leq s < t \leq 1}\mathbf{E}\|J_{s,t}\|_{H^2 \to H^2}^p \leq \exp(\eta p V(u) + pC_{\eta,p})$$
$$\sup_{0 \leq s < t \leq 1}\mathbf{E}\|J_{s,t}^{(2)}\|_{H^2 \otimes H^2 \to H^2}^p \leq \exp(\eta p V(u) + pC_{\eta,p})$$

hold for all $u_0 \in H^2$.

Proof. The first two bounds are a consequence of Propositions 8.11, 8.12 and 8.8.
In order to get the second two bounds, note that

$$\sup_{0 \leq s < t \leq 1}\mathbf{E}\|J_{s,t}^{(2)}\|_{H^2 \otimes H^2 \to H^2}^p \leq C_p(1 + \|u_0\|_{H^2})^{(17n+5)p}$$
$$= \exp\Big((17n + 5)p\log(1 + \|u_0\|_{H^2}) + \log C_p\Big) ,$$

as a consequence of Propositions 8.13, 8.11 and 8.8. A similar bound holds for J. Since, for any positive q, r, η, K there exists a $C_{q,r,\eta,K}$ so that $q\log(1 + x) + \log(K) \leq \eta x^r + C_{q,r,\eta,K}$ for all $x \geq 0$, the quoted bound holds. $\qquad\square$

We assume from now on that the g_k used in the definition of the forcing all belong to H^4. We now construct a particular subset of the A_n defined in Section 6.2 using on

the highest degree nonlinear term. By doing so we obtain only constant vector fields, thus trivializing Assumption C.2 in light of Lemma 8.3. Setting $\widetilde{A}_1 = \{g_1, \cdots, g_d\}$, we define recursively $\widetilde{A}_{k+1} = \widetilde{A}_k \cup \{F_n(h_1, \cdots, h_n) : h_j \in \widetilde{A}_k\}$ and $\widetilde{A}_\infty = \bigcup \widetilde{A}_k$. Notice that since $g_k \in H^4$, we know that all of the $\widetilde{A}_n \subset H^4$ since H^4 is a multiplicative algebra in our setting.

Proposition 8.15 *If $span(\widetilde{A}_\infty)$ is dense in H^2 then given any H^2-orthogonal projection Π onto a finite dimensional subspace, there exists $\theta > 0$ such that Assumption B.1 holds with $U(u) = \Psi_0^\theta$.*

Proof. Proposition 8.14 guarantees that all of the assumptions of Theorem 6.7 hold except Assumption C.2. However since by construction all of the vector fields in $\widetilde{A}_n$ are constant Assumption C.2 clearly holds with Λ a constant if Π is an orthogonal projection onto a subspace of $span(\widetilde{A}_n)$. Lemma 8.3 furthermore shows that it actually holds for any finite rank orthogonal projection. $\square$

Let Π_M be the projection on the eigenfunctions of the Laplacian with eigenvalues smaller than M^2. We will now restrict ourselves to such a projection since it allows for easy verification of the pathwise smoothing/contracting properties needed for Assumption B.4. We have indeed the following bound:

Proposition 8.16 *Given any positive η, r and p, there exists a $C_{\eta,r,p}$ so that the bound*

$$\mathbf{E}\|J_{0,1}\Pi_M^\perp\|_{H^2 \to H^2}^p \leq \exp(p\eta\|u_0\|_{H^2}^r - p\log(M) + pC_{\eta,r,p})$$

holds for all $u_0 \in H^2$, and all $M \in \mathbf{N}$.

Proof. First observe that

$$\begin{aligned}
\|J_{0,1}\Pi_M^\perp\|_{H^2 \to H^2} &\leq \|J_{0,1}\|_{H^1 \to H^2}\|\Pi_M^\perp\|_{H^2 \to H^1} \leq M^{-1}\|J_{0,1}\|_{H^1 \to H^2} \\
&\leq CM^{-1}\mathrm{Sup}_{1,\infty}^{4n}(\|u_0\|_{H^2} + \mathrm{Sup}_{1,2}(W_\Delta)) \\
&\leq CM^{-1}\mathrm{Sup}_{1,2}(W_\Delta)^{4n+1}(\|u_0\|_{H^2} + 1)^{4n+1}
\end{aligned}$$

Raising both sides to the power p, taking expectations, and using the fact that the law of $\mathrm{Sup}_{1,2}(W_\Delta)$ has Gaussian tails, we obtain

$$\mathbf{E}\|J_{0,1}\Pi_M^\perp\|_{H^2 \to H^2}^p \leq \exp(p(4n+1)\log(1 + \|u_0\|) - p\log(M) + pC_p) .$$

The claim now follows from the fact that, for any $\eta > 0$ and $r > 0$, there exist a $C_{\eta,r}$ with $(4n+1)\log(1+x) \leq \eta x^r + C_{\eta,r}$ for all $x \geq 0$. $\square$

Theorem 8.17 *Let $\mathcal{P}_t$ be the Markov semigroup on H^2 generated by (8.10). If the linear span of $\widetilde{A}_\infty$ is dense in H^2 then, for every orthogonal finite rank projection $\Pi \colon H^2 \to H^2$, for every $p > 0$, and for every $\alpha > 0$, there exists a constant $C(\alpha, p, \Pi)$ such that the bound (5.1) on the Malliavin matrix holds with $U = 1$.*

Proof. The result follows from Theorem 6.7. One can check that Assumption A.1 holds with $\mathcal{H} = H^2$, $a = 0$, $\gamma_\star = \beta_\star = \infty$ since H^ℓ is a multiplicative algebra for every $\ell \geq 2$ (this is true because we restricted ourselves to dimension $m \leq 3$). Since

the most involved part is the assumption on the adjoint, part 3, we give the details for that one. One can verify that the adjoint of $DN(u)$ in $\mathcal{H}$ acts on elements v in $\mathcal{H}^\infty$ as

$$DN^*(u)v = \Delta^{-2}f'(u)\Delta^2 v \ .$$

(This is because $\mathcal{H}$ is the Sobolev space H^2 and not the space L^2.) The claim then follows from the fact that the multiplication by a smooth enough function is a bounded operator in every Sobolev space H^ℓ with $\ell \in \mathbf{R}$.

Since Assumption C.2 (with Λ_α a constant depending on Π) can be verified by using Lemma 8.3, it remains to verify Assumption C.1 with $\Psi_0 = 1$. This in turn is an immediate consequence of Proposition 8.14. $\qquad\square$

Combining all of these results, we finally obtain the following result on the asymptotic strong Feller property of a general reaction-diffusion equation:

Theorem 8.18 *Let $\mathcal{P}_t$ be the Markov semigroup on H^2 generated by* (8.10) *and let Assumptions RD.1 and RD.2 hold. If the linear span of $\widetilde{A}_\infty$ is dense in H^2 then, for any $\zeta > 0$, there exists a positive constant C so that for every $u \in H^2$, and $\varphi : H^2 \to \mathbf{R}$ on has*

$$\|\mathrm{D}\mathcal{P}_t\varphi(u)\|_{L^2\to\mathbf{R}} \le C(\|\varphi\|_{L^\infty} + e^{-\zeta t}\sup_{v\in H^2}\|\mathrm{D}\varphi(v)\|_{H^2\to\mathbf{R}}) \ . \tag{8.16}$$

In particular, $\mathcal{P}_t$ has the asymptotic strong Feller property in H^2.

Remark 8.19 It is easy to infer from the *a priori* bounds given in Propositions 8.10, 8.11, 8.12 and 8.13 that the assumptions of our 'all purpose' Theorem 8.1 hold with $V(u) = \|u\|^\alpha$ for a sufficiently small exponent α. However, the bound (8.3) is slightly weaker than the bound (8.16). This shows that it may be worth under some circumstances to make the effort to apply the more general Theorem 5.5.

Remark 8.20 As a corollary, we see that for the semigroup on $\mathcal{E}$, one has

$$\|\mathrm{D}\mathcal{P}_t\varphi(u)\| \le C(\|\varphi\|_{L^\infty} + e^{-\zeta t}\|\mathrm{D}\varphi\|_{L^\infty}) \ ,$$

where all the derivatives a Fréchet derivatives of functions from $\mathcal{E}$ to $\mathbf{R}$.

In particular, in space dimension $m = 1$, the same bound is obtained in the space H^1 since one then has $H^1 \subset \mathcal{E}$.

Proof. The result follows from Theorem 5.5. Fix $\Pi = \Pi_M$, the projection onto the eigenfunctions of Δ with eigenvalues of modulus less than M^2. The constant M is going to be determined later on. Assumption B.2 with $V(u) = \|u\|_{\mathcal{H}}^{1/n}$ and $\eta' = 0$ follows immediately from Proposition 8.11. Fix any $\bar{p} > 10$ and any positive $\eta < 1/\bar{p}$. Assumption B.3 then follows from Proposition 8.14. It then follows from Proposition 8.16 that we can choose the value of M in the definition of Π sufficiently large so that Assumption B.4 holds and such that $(C_\Pi - C_J)/2 - \eta C_L > \zeta$. Since, in view of Theorem 8.17, Assumption B.1 holds with $U = 1$, we thus obtain from Theorem 5.5 the bound

$$\|\mathrm{D}\mathcal{P}_t\varphi(u)\|_{H^2\to\mathbf{R}} \le Ce^{\eta\|u\|^{1/n}}(\|\varphi\|_{L^\infty} + e^{-\zeta t}\sup_{v\in H^2}\|\mathrm{D}\varphi(v)\|_{H^2\to\mathbf{R}}) \ . \tag{8.17}$$

In order to obtain (8.16), we note that one has

$$\mathbf{E}\|J_{0,2}\|_{L^2\to H^2}^2 \leq \mathbf{E}\|J_{0,1}\|_{L^2\to L^2}^2\|J_{1,2}\|_{L^2\to H^2}^2 \leq C\mathbf{E}\|J_{1,2}\|_{L^2\to H^2}^2 \leq C \ , \quad (8.18)$$

where C is a universal constant independent of the initial condition. Here, we combined the bounds of Proposition 8.12 with Proposition 8.8 in order to obtain the last bound. We thus have

$$\begin{aligned}
\|\mathrm{D}\mathcal{P}_t\varphi(u)\|_{L^2\to\mathbf{R}} &= \|\mathrm{D}\mathcal{P}_2\mathcal{P}_{t-2}\varphi(u)\|_{L^2\to H^2} \\
&\leq \mathbf{E}\|\mathrm{D}\mathcal{P}_{t-2}\varphi(u_2)\|_{H^2\to\mathbf{R}}\|J_{0,2}\|_{L^2\to H^2} \\
&\leq C(\|\varphi\|_{L^\infty} + e^{-\zeta t}\sup_v\|\mathrm{D}\varphi(v)\|_{H^2\to\mathbf{R}})\mathbf{E}(e^{2\eta\|u_2\|^{1/n}}\|J_{0,2}\|_{L^2\to H^2}) \ ,
\end{aligned}$$

where we made use of (8.17) to obtain the last inequality. The requested bound now follows from (8.18) and Proposition 8.11. $\qquad\square$

8.4 Unique ergodicity of the stochastic Ginzburg-Landau equation

In this section, we show under very weak conditions on the driving noise that the stochastic real Ginzburg-Landau equation has a unique invariant measure. Recall that this equation is given by

$$du(x,t) = \nu\partial_x^2 u(x,t)\,dt + \eta u(x,t)\,dt - u^3(x,t)\,dt + \sum_{j=1}^d g_j(x)\,dW_j(t) \ , \quad (8.19)$$

where the spatial variable x takes values on the circle $x \in S^1$ and the driving functions g_j belong to $\mathcal{C}^\infty(S^1, \mathbf{R})$. The two positive parameters ν and η are assumed to be fixed throughout this section. This is a particularly simple case of the type of equation considered above, so that Theorem 8.18 applies. The aim of this section is to show one possible technique for obtaining the uniqueness of the invariant measure for such a parabolic SPDE. It relies on Corollary 2.2 and yields:

Theorem 8.21 *Consider (8.19) and suppose that*

1. *there exists a linear combination g of the g_j that has only finitely many simple zeroes,*
2. *the smallest vector space containing all the g_j and closed under the operation $(f, g, h) \mapsto fgh$ is dense in $H^1(S^1)$.*

Then (8.19) has exactly one invariant probability measure.

Remark 8.22 The second assumption is satisfied for example if $d \geq 3$ and $g_1(x) = 1$, $g_2(x) = \sin x$ and $g_3(x) = \cos x$.

Remark 8.23 We believe that the first condition in Theorem 8.21 is not needed, since in finite dimensions such a Lie bracket condition implies global controllability for polynomial systems of odd degree. See for example [Jur97].

Remark 8.24 Actually, we could have relaxed the regularity assumption on the g_j's. If we choose $\mathcal{H} = H^1$, $\gamma_\star = 2\varepsilon - 1$, $a = 1 - \varepsilon$, and $\beta_\star = 1$, we can check that Assumption A.1 is satisfied as soon as $g_j \in H^{1+4\varepsilon}$. Furthermore, in this case, all the relevant Lie brackets for assumption 2 in Theorem 8.21 are admissible, so that its conclusion still holds.

Looking at Corollary 2.2, the two main ingredients needed to prove Theorem 8.21 are the establishment of the estimate in (1.2) and the needed form of irreducibility. The first will follows almost instantly from the second assumption of Theorem 8.21 which ensures that $\mathrm{span}(\widetilde{A}_\infty)$ is dense in H^1. The irreducibility is given by the following proposition whose proof is postponed to the end of this section.

Proposition 8.25 *Consider* (8.19) *under the second condition in Theorem 8.21. Then there exists a positive K so that for any $\epsilon > 0$ there is a v with $\|v\|_{H^1} \leq K$ and a $T > 0$ so that $\mathcal{P}_T(u_0, \mathcal{B}_\epsilon(v)) > 0$ for all $u_0 \in H^1$. Here $\mathcal{B}_\epsilon(v)$ is the ϵ ball in the H^1–norm.*

Proof of Theorem 8.21. The existence of an invariant probability measure for (8.19) is standard, see for example [Cer99]. Furthermore, since we are working in space dimension 1, H^1 is already a multiplicative algebra and one can retrace the proof of Theorem 8.18 for $\mathcal{H} = H^1$. This shows that assumption 2. implies that the semigroup generated by (8.19) satisfies (1.2) on the Hilbert space $\mathcal{H} = H^1(S^1)$. It therefore remains to show that assumption 1. implies the assumption of Corollary 2.2. □

In fact we have established much more than just uniqueness of the invariant measure. We now use the results from [HM08] to establish a spectral gap. For any Fréchet differentiable functions from $\varphi : H^1 \to \mathbf{R}$ define the norm $\|\varphi\|_{\mathrm{Lip}} = \sup_u(|\varphi(u)| + \|\mathrm{D}\varphi(u)\|_{H^1 \to \mathbf{R}})$. In turn we define a metric on probability measures μ, ν on H^1 by $d(\mu, \nu) = \sup\{\int \varphi d\mu - \int \varphi d\nu : \|\varphi\|_{\mathrm{Lip}} \leq 1\}$. Combining (8.16), Proposition 8.25 and [HM08, Theorem 2.5] yields the following corollary to Theorem 8.21.

Corollary 8.26 *Under the assumption of Theorem 8.21, there exist positive constants C and γ so that $d(\mathcal{P}_t^* \mu, \mathcal{P}_t^* \nu) \leq Ce^{-\gamma t} d(\mu, \nu)$ for any two probability measures μ and ν on H^1 and $t \geq 1$.*

Proof of Proposition 8.25. Fix an arbitrary initial condition u_0 and some $\varepsilon > 0$. Our aim is to find a target v, bounded controls $V_j(t)$, and a terminal time $T > 0$, so that the solution to the controlled problem

$$\partial_t u(x,t) = \nu \partial_x^2 u(x,t) + \eta u(x,t) - u^3(x,t) + f(x,t) , \quad f(x,t) \overset{\text{def}}{=} \sum_{j=1}^d g_j(x) V_j(t) ,$$
(8.20)

satisfies $\|u(T) - v\|_{H^1} \leq \varepsilon$. Furthermore, we want to be able to choose v such that $\|v\|_{H^1} \leq K$ for some constant K independent of ε. The claim on the topological supports of transition probabilities then follows immediately from the fact that the Itô map $(u_0, W) \mapsto u_t$ is continuous in the second argument in our case.

The idea is to choose f of the form

$$f(x,t) = \begin{cases} \varepsilon^{-\gamma} g(x) & \text{for } 1 \leq t \leq 2, \\ 0 & \text{otherwise,} \end{cases}$$

and to set $T = 3$. We furthermore set v to be the solution at time 1 for the uncontrolled equation (that is (8.20) with $f = 0$) with an initial condition v_0 satisfying

$$\nu \partial_x^2 v_0(x) + \eta v_0(x) - v_0^3(x) + \varepsilon^{-\gamma} g(x) = 0 , \tag{8.21}$$

for some exponent $\gamma > 0$ to be determined. Such a v_0 always exists since the coercive "energy functional"

$$E(v) = \int_{S^1} \left(\frac{\nu}{2} |\partial_x v(x)|^2 - \frac{\eta}{2} |v(x)|^2 + \frac{1}{4} |v(x)|^4 - \varepsilon^{-\gamma} g(x) v(x) \right) dx$$

has at least one critical point. Even though v_0 is in general very large (see however Lemma 8.27 below), it follows from (8.12) that the target v constructed in this way is bounded independently of ε.

The remaining ingredient of the proof are Lemmas 8.28 and 8.27 below. To show that this is sufficient, note first that (8.12) implies the existence of a constant C such that $\|u(1)\|_{L^2} \leq C$ independently of u_0. It then follows from Lemmas 8.28 and 8.27 that (choosing for example $\beta = \gamma/14$) there exists a constant C such that one has the bound

$$\|u(2) - v_0\|_{L^2} \leq C \varepsilon^{\frac{\gamma}{6}} .$$

Since the uncontrolled equation expands at rate at most η, this immediately yields $\|u(T) - v\|_{L^2} \leq C\varepsilon^{\frac{\gamma}{6}}$. On the other hand, we know from Proposition 8.10 that there exists a constant C such that $\|u(T) - v\|_{H^2} \leq C$, so that

$$\|u(T) - v\|_{H^1} \leq \left(\|u(T) - v\|_{L^2} \|u(T) - v\|_{H^2} \right)^{1/2} \leq C\varepsilon^{\gamma/12} ,$$

and the claim follows by choosing $\gamma > 12$. $\qquad\square$

Lemma 8.27 *There exists a constant C_v independent of $\varepsilon < 1$ such that the bound $\|v_0(x)\|_{L^\infty} \leq C_v \varepsilon^{-\gamma/3}$ holds.*

Proof. It follows immediately from (8.21), using the fact that $\partial_x^2 v_0 \leq 0$ at the maximum and $\partial_x^2 v_0 \geq 0$ at the minimum. $\qquad\square$

Lemma 8.28 *For every exponent $\beta \in [0, \gamma/4]$ there exists a constant C such that the bound*

$$\int_{S^1} (u - v_0)(u^3 - v_0^3) \, dx \geq C\varepsilon^{-2\beta} \int_{S^1} (u - v_0)^2 \, dx - C\varepsilon^{\frac{\gamma - 13\beta}{3}}$$

holds for every $\varepsilon \leq 1$ and every $u \in L^2(S^1)$.

Proof. The proof is based on the fact that since g has only isolated zeroes, the function v_0 necessarily has the property that it is large at most points. More precisely, consider some exponent $\beta \in [0, \gamma/3]$ and define the set $A = \{x \in S^1 : |v_0(x)| > \varepsilon^{-\beta}\}$. We claim that there then exists a constant C such that the Lebesgue measure of A is bounded by $|A| \leq C\varepsilon^\alpha$ for $\alpha = \min\{\gamma - 3\beta, \frac{\gamma - \beta}{3}\}$. Indeed, consider the set $\tilde{A}$ of points such that $|g(x)| \leq 2\varepsilon^\alpha$. Since g is assumed to be smooth and have simple zeroes, $|\tilde{A}| \leq C\varepsilon^\alpha$ and the complement of $\tilde{A}$ consists of finitely many intervals on which g has a definite sign.

Consider one such interval I on which $g(x) > 9\varepsilon^\alpha$, so that the definition of v_0 yields the estimate $v_0'' < -9\varepsilon^{\alpha-\gamma} - v_0 + v_0^3$. It follows that, for every $x \in I$, one either has $v_0''(x) < -\varepsilon^{\alpha-\gamma}$, or one has $v_0(x) > 2\varepsilon^{\frac{\alpha-\gamma}{3}} \geq 2\varepsilon^{-\beta}$ (since we set $\alpha \leq \gamma - 3\beta$). We conclude that $I \cap A$ consists of at most two intervals and that $v_0(x) > \varepsilon^{\frac{\alpha-\gamma}{3}}$ for every $x \in I \cap A$, so that $|I \cap A| \leq C\varepsilon^{\frac{\gamma-\alpha-\beta}{2}}$ and the bound follows. (The same reasoning but with opposite signs applies to those intervals on which $g(x) < -9\varepsilon^\alpha$.)

This yields the sequence of bounds

$$2 \int_{S^1} (u - v_0)(u^3 - v_0^3) \, dx \geq \int_{S^1} (u - v_0)^2 (u^2 + v_0^2) \, dx$$

$$\geq \varepsilon^{-2\beta} \int_A (u - v_0)^2 \, dx + \int_{A^c} (u - v_0)^2 (u^2 + v_0^2) \, dx$$

$$\geq \varepsilon^{-2\beta} \int_A (u - v_0)^2 \, dx + \frac{1}{4} \int_{A^c} (u - v_0)^4 \, dx$$

$$\geq \varepsilon^{-2\beta} \int_A (u - v_0)^2 \, dx + \frac{1}{4|A^c|} \left(\int_{A^c} (u - v_0)^2 \, dx \right)^2$$

$$\geq \varepsilon^{-2\beta} \int_A (u - v_0)^2 \, dx + \frac{C}{\varepsilon^\alpha} \left(\varepsilon^{\alpha - 2\beta} \int_{A^c} (u - v_0)^2 \, dx - \varepsilon^{2\alpha - 4\beta} \right)$$

$$\geq C \varepsilon^{-2\beta} \int_{S^1} (u - v_0)^2 \, dx - C \varepsilon^{\alpha - 4\beta} \,,$$

which is the required estimate. $\qquad \square$

References

[AKSS07] A. AGRACHEV, S. KUKSIN, A. SARYCHEV, and A. SHIRIKYAN. On finite-dimensional projections of distributions for solutions of randomly forced 2D Navier-Stokes equations. *Ann. Inst. H. Poincaré Probab. Statist.* **43**, no. 4, (2007), 399–415.

[AS04] A. A. AGRACHEV and Y. L. SACHKOV. *Control theory from the geometric viewpoint*, vol. 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[AS05] A. A. AGRACHEV and A. V. SARYCHEV. Navier-Stokes equations: controllability by means of low modes forcing. *J. Math. Fluid Mech.* **7**, no. 1, (2005), 108–152.

[AS08] A. AGRACHEV and A. SARYCHEV. Solid controllability in fluid dynamics. In *Instability in models connected with fluid flows. I*, vol. 6 of *Int. Math. Ser. (N. Y.)*, 1–35. Springer, New York, 2008.

[BH07] F. BAUDOIN and M. HAIRER. A version of Hörmander's theorem for the fractional Brownian motion. *Probab. Theory Related Fields* **139**, no. 3-4, (2007), 373–395.

[BKL01] J. BRICMONT, A. KUPIAINEN, and R. LEFEVERE. Ergodicity of the 2D Navier-Stokes equations with random forcing. *Comm. Math. Phys.* **224**, no. 1, (2001), 65–81. Dedicated to Joel L. Lebowitz.

[BM05] Y. BAKHTIN and J. C. MATTINGLY. Stationary solutions of stochastic differential equations with memory and stochastic partial differential equations. *Commun. Contemp. Math.* **7**, no. 5, (2005), 553–582.

[BM07] Y. BAKHTIN and J. C. MATTINGLY. Malliavin calculus for infinite-dimensional systems with additive noise. *J. Funct. Anal.* **249**, no. 2, (2007), 307–353.

[BT05] F. BAUDOIN and J. TEICHMANN. Hypoellipticity in infinite dimensions and an application in interest rate theory. *Ann. Appl. Probab.* **15**, no. 3, (2005), 1765–1777.

[Cer99] S. CERRAI. Ergodicity for stochastic reaction-diffusion systems with polynomial coefficients. *Stochastics Stochastics Rep.* **67**, no. 1-2, (1999), 17–51.

[CF88] P. CONSTANTIN and C. FOIAS. *Navier-Stokes equations*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1988.

[DL92] R. DAUTRAY and J.-L. LIONS. *Mathematical analysis and numerical methods for science and technology. Vol. 5*. Springer-Verlag, Berlin, 1992. Evolution problems.

I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.

[DPEZ95] G. DA PRATO, K. D. ELWORTHY, and J. ZABCZYK. Strong Feller property for stochastic semilinear equations. *Stochastic Anal. Appl.* **13**, no. 1, (1995), 35–45.

[DPZ92] G. DA PRATO and J. ZABCZYK. *Stochastic Equations in Infinite Dimensions*. University Press, Cambridge, 1992.

[DPZ96] G. DA PRATO and J. ZABCZYK. *Ergodicity for Infinite Dimensional Systems*, vol. 229 of *London Mathematical Society Lecture Note Series*. University Press, Cambridge, 1996.

[EH01] J.-P. ECKMANN and M. HAIRER. Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Commun. Math. Phys.* **219**, no. 3, (2001), 523–565.

[EMS01] W. E, J. C. MATTINGLY, and Y. SINAI. Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation. *Comm. Math. Phys.* **224**, no. 1, (2001), 83–106. Dedicated to Joel L. Lebowitz.

[Fla95] F. FLANDOLI. *Regularity theory and stochastic flows for parabolic SPDEs*, vol. 9 of *Stochastics Monographs*. Gordon and Breach Science Publishers, Yverdon, 1995.

[FM95] F. FLANDOLI and B. MASLOWSKI. Ergodicity of the 2-d Navier-Stokes equation under random perturbations. *Commun. Math. Phys.* **172**, no. 1, (1995), 119–141.

[GM06] B. GOLDYS and B. MASLOWSKI. Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's. *Ann. Probab.* **34**, no. 4, (2006), 1451–1496.

[Hai02] M. HAIRER. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probab. Theory Related Fields* **124**, no. 3, (2002), 345–380.

[Hai08] M. HAIRER. An introduction to stochastic PDEs, 2008. URL http://www.hairer.org/Teaching.html. Unpublished lecture notes.

[HM06] M. HAIRER and J. C. MATTINGLY. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)* **164**, no. 3, (2006), 993–1032.

[HM08] M. HAIRER and J. C. MATTINGLY. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.* **36**, no. 6, (2008), 2050–2091.

[HM10] M. HAIRER and A. J. MAJDA. A simple framework to justify linear response theory. *Nonlinearity* **23**, (2010), 909–922.

[HMS10] M. HAIRER, J. MATTINGLY, and M. SCHEUTZOW. Asymptotic coupling and a general form of harris theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **??**, no. ??, (2010), ???

[Hör67] L. HÖRMANDER. Hypoelliptic second order differential equations. *Acta Math.* **119**, (1967), 147–171.

[Hör85] L. HÖRMANDER. *The Analysis of Linear Partial Differential Operators I–IV*. Springer, New York, 1985.

[ITM50] C. T. IONESCU TULCEA and G. MARINESCU. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)* **52**, (1950), 140–147.

[Jur97] V. JURDJEVIC. *Geometric control theory*, vol. 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[Kat80] T. KATO. *Perturbation Theory for Linear Operators*. Springer, New York, 1980.

[KS84] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. I. In *Stochastic analysis (Katata/Kyoto, 1982)*, vol. 32 of *North-Holland Math. Library*, 271–306. North-Holland, Amsterdam, 1984.

[KS85a] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**, no. 1, (1985), 1–76.

[KS85b] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**, no. 1, (1985), 1–76.

[KS00] S. B. KUKSIN and A. SHIRIKYAN. Stochastic dissipative PDE's and Gibbs measures. *Commun. Math. Phys.* **213**, (2000), 291–330.

[Liv03] C. LIVERANI. Invariant measures and their properties. A functional analytic point of view. In *Dynamical systems. Part II*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, 185–237. Scuola Norm. Sup., Pisa, 2003.

[LY73] A. LASOTA and J. A. YORKE. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186**, (1973), 481–488 (1974).

[Mal78] P. MALLIAVIN. Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, 195–263. Wiley, New York, 1978.

[Mal97] P. MALLIAVIN. *Stochastic analysis*, vol. 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997.

[Mas89] B. MASLOWSKI. Strong Feller property for semilinear stochastic evolution equations and applications. In *Stochastic systems and optimization (Warsaw, 1988)*, vol. 136 of *Lecture Notes in Control and Inform. Sci.*, 210–224. Springer, Berlin, 1989.

[Mat02] J. C. MATTINGLY. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Commun. Math. Phys.* **230**, no. 3, (2002), 421–462.

[Mat08] J. C. MATTINGLY. *Saint Flour Lectures*. Springer, 2008. In preparation.

[MP06] J. C. MATTINGLY and É. PARDOUX. Malliavin calculus for the stochastic 2D Navier-Stokes equation. *Comm. Pure Appl. Math.* **59**, no. 12, (2006), 1742–1790.

[MSVE07] J. C. MATTINGLY, T. SUIDAN, and E. VANDEN-EIJNDEN. Simple systems with anomalous dissipation and energy cascade. *Comm. Math. Phys.* **276**, no. 1, (2007), 189–220.

[MT93] S. P. MEYN and R. L. TWEEDIE. *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993.

[MY02] N. MASMOUDI and L.-S. YOUNG. Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs. *Comm. Math. Phys.* **227**, no. 3, (2002), 461–481.

[Nag97] T. NAGASAWA. Navier-Stokes flow on Riemannian manifolds. In *Proceedings of the Second World Congress of Nonlinear Analysts, Part 2 (Athens, 1996)*, vol. 30, 825–832. 1997.

[Nor86] J. NORRIS. *Simplified Malliavin Calculus*, vol. 1204 of *Lecture Notes in Mathematics*. Springer, New York, 1986.

[Nua95] D. NUALART. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[Oco88] D. OCONE. Stochastic calculus of variations for stochastic partial differential equations. *J. Funct. Anal.* **79**, no. 2, (1988), 288–331.

[RR04] M. RENARDY and R. C. ROGERS. *An introduction to partial differential equations*, vol. 13 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second ed., 2004.

[RS80] M. REED and B. SIMON. *Methods of Modern Mathematical Physics I–IV*. Academic Press, San Diego, California, 1980.

[Str83]   D. W. STROOCK. Some applications of stochastic calculus to partial differential equations. In *Eleventh Saint Flour probability summer school—1981 (Saint Flour, 1981)*, vol. 976 of *Lecture Notes in Math.*, 267–382. Springer, Berlin, 1983.

[Tay92]   M. E. TAYLOR. Analysis on Morrey spaces and applications to Navier-Stokes and other evolution equations. *Comm. Partial Differential Equations* **17**, no. 9-10, (1992), 1407–1456.

[Tri86]   H. TRIEBEL. Spaces of Besov-Hardy-Sobolev type on complete Riemannian manifolds. *Ark. Mat.* **24**, no. 2, (1986), 299–337.

[Tri92]   H. TRIEBEL. *Theory of function spaces. II*, vol. 84 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992.

[TW93]   R. TEMAM and S. H. WANG. Inertial forms of Navier-Stokes equations on the sphere. *J. Funct. Anal.* **117**, no. 1, (1993), 215–242.