ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

The strong Feller property for singular stochastic PDEs

M. Hairer[a] and J. Mattingly[b]

[a]*Mathematics Research Centre, University of Warwick, UK. E-mail: m.hairer@warwick.ac.uk*
[b]*Departments of Mathematics and Statistical Science, Duke University, USA. E-mail: jonm@math.duke.edu*

Abstract. We show that the Markov semigroups generated by a large class of singular stochastic PDEs satisfy the strong Feller property. These include for example the KPZ equation and the dynamical Φ_3^4 model. As a corollary, we prove that the Brownian bridge measure is the unique invariant measure for the KPZ equation with periodic boundary conditions.

Résumé. Nous montrons que les semi-groupes de Markov engendrés par une classe large d'EDPs stochastiques singulières satisfont la propriété forte de Feller. Cette classe inclut par exemple l'équation KPZ et le modèle Φ_3^4. Nous montrons comme corollaire que la distribution du pont Brownien est l'unique mesure invariante pour l'équation KPZ avec conditions frontières périodiques.

1. Introduction

Recall that a Markov operator $\mathfrak{P}$ on some separable metric space $\mathcal{X}$ is said to satisfy the *strong Feller property* if it maps all bounded measurable functions $\mathcal{X} \to \mathrm{R}$ into bounded continuous functions. Intuitively, the strong Feller property allows us to link measure theoretic properties of a process to the corresponding topological properties. As a result, it is a very useful ingredient when trying to establish the ergodicity of a given Markov process, see for example [44]. When considering finite-dimensional diffusions with sufficiently smooth coefficients, the strong Feller property is a consequence of the parabolic Hörmander condition [33,41,47], and this condition is essentially sharp.

In infinite dimensions, no such sharp condition exists, despite considerable progress on notions of hypoellipticity in that case, see for example [5,16,30,31,43]. This should of course not come as a surprise since measures on infinite-dimensional spaces have a rather annoying tendency of being mutually singular. The strong Feller property on the other hand implies that the total variation distance between transition probabilities starting from nearby points is small [24,49], thus ruling out mutual singularity. However when considering parabolic stochastic PDEs in bounded domains driven by noise that is "sufficiently non-degenerate and not overly smooth spatially," it is well-known that the strong Feller property holds, see for example [12–14,16,42].

All these results do however rely very strongly on the well-posedness of the equations under consideration, as well as on good a priori control on their derivative with respect to initial conditions. This typically enforces some conditions on the driving noise requiring it to be sufficiently regular for the standard solution theory [15] to apply. The aim of this article is to extend these results to singular stochastic PDEs like the dynamical $P(\Phi)_2$ model [1,11,38], the KPZ equation [27,39], and the dynamical Φ_3^4 model [28].

Both of these examples contain noise which is spatially rough enough to make the existence and uniqueness theory extremely non-trivial. However, when proving the strong Feller property, rougher noise only makes the core of the proof easier, the only difficulties being making sense of all of the objects being manipulated. This is in contrast to

the case of spatially smooth or degenerate noise where the existence and uniqueness of solutions and the correctness of all manipulations is straightforward while the proof of any property like the strong Feller property which links the long time behaviour of nearby points and illuminates the ergodic properties of the system are more complicated, see for example [30,31].

The main novelty of our approach is that while it is close in spirit to proofs based on the Bismut–Elworthy–Li formula [12,17], we do not require any a priori control on the solutions: they are in principle allowed to blow up, even with positive probability. In particular, this allows to strengthen well-posedness results for almost every initial condition as in [1,11,32] to every initial condition in the topological support of the invariant measure. It also yields as a corollary the uniqueness of the Brownian bridge measure (modulo height shifts) for the KPZ equation with periodic boundary conditions, which had not been established before.

Both our approach and that used in the proof based on the Bismut–Elworthy–Li formula rely on transferring the variation caused by shifting the initial condition infinitesimally to an infinitesimal shift in the noise. In versions of the argument closest to ours, one then integrates by parts against the Gaussian measure, moving the infinitesimal variation in the noise to the Wiener measure and finally averaging over the realisations of the noise. This in particular requires the solution to be well posed almost surely. In our approach we accumulate these infinitesimal shifts to build a macroscopic shift in the noise corresponding to a macroscopic shift in the initial condition. This allows us to work in a more pathwise manner and consider equations which might explode with positive probability.

Remark 1.1. It is very natural to ask whether the solutions to the stochastic Navier–Stokes equations driven by space-time white noise in dimensions 2 and 3 as constructed in [10,52] also satisfy the strong Feller property. While we believe this to be true, our results as they are presented in Section 4 do not cover this case because of the presence of the Leray projection. We do however expect the results of Section 3 to be applicable to that situation as well by slightly modifying the argument of Section 4. We feel that this is largely a technical issue but do not explore it here.

Let us also mention that some related results have recently been obtained. In [48], the authors show that the Φ_2^4 measure (as constructed in [46]) is indeed the unique invariant measure for the dynamical Φ_2^4 model. The proof of this fact however does not make use of the strong Feller property but instead relies on an asymptotic coupling argument. While this is sufficient to prove many ergodic properties, it does not establish the local regularity in the total variation topology of the transition density. More recently, in [51], the authors obtained not only the strong Feller property for the dynamical $P(\Phi)_2$ model, but also the exponential ergodicity of the dynamical Φ_2^4 model. While these results are much stronger than ours, they are restricted to one particular model and rely strongly on good a priori bounds on the solutions which are not available in all the cases we treat.

The structure of this article goes as follows. In Sections 2 and 3, we set up an abstract framework and give sufficient conditions for a Markovian continuous random dynamical system to satisfy the strong Feller property. This framework is very general and we expect it to be useful also in other contexts. It is designed so that, as shown in Section 4, it covers a very large class of semilinear stochastic PDEs, provided that we build their solutions via the theory of regularity structures. We do however expect that constructions using paracontrolled calculus as in [6,22,23] can also be fitted into our framework. In our last section, we then finally show that many interesting examples of singular stochastic PDEs are covered by our results.

2. Abstract setup

In this section, we set up the abstract framework needed to make precise the idea that an infinitesimal variation in the initial condition can be equated with an infinitesimal variation in the noise. This is the content of equation (2.4) in Assumption 5 below. It is in establishing (2.4) for an adapted shift that we require noise which is non-degenerate and with sufficient spatial roughness, see also Assumption 11 below.

Over the next two sections, we will consider a general random dynamic in some Banach space U, defined over a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We will always consider Gaussian probability spaces endowed with a filtration. More precisely, we assume that Ω is a separable Banach space and there exists a Hilbert space H_0 such that $\mathcal{H} = L^2(\mathbf{R}, H_0) \subset \Omega$ is the Cameron–Martin space for the Gaussian measure $\mathbf{P}$. The canonical random variable ω drawn from Ω according to $\mathbf{P}$ induces the two-sided continuous filtration $\{\mathcal{F}_{s,t}, s < t\}$ by $\mathcal{F}_{s,t} = \sigma\{h^*(\omega) : h \in L^2([s,t], H_0)\}$, with the canonical inclusion $L^2([s,t], H_0) \subset \mathcal{H}$, where we use the canonical identification $h \leftrightarrow h^*$ between the Cameron–Martin

space and measurable linear functionals on Ω, see [3,25]. We say that a stochastic process over $(\Omega, \mathcal{F}, \mathbf{P})$ is adapted if it is adapted to the filtration $\mathcal{F}_t \overset{\text{def}}{=} \mathcal{F}_{-\infty,t}$.

We also assume that we are given a complete separable metric space $\mathcal{M}$ as well as a random variable

$$\boldsymbol{\xi} \colon \Omega \to \mathcal{M}.$$

In the applications we have in mind, $\mathcal{M}$ is the space of admissible models for a given regularity structure as in [28] and $\boldsymbol{\xi}$ is the map constructing a suitable admissible model from the underlying driving noise. We will generally write $\boldsymbol{\xi}$ for the $\mathcal{M}$-valued random variable given by $\boldsymbol{\xi} = \boldsymbol{\xi}(\omega)$ where ω is drawn according to $\mathbf{P}$ and ξ for a generic element of $\mathcal{M}$. For any measurable map $X \colon \Omega \to \mathcal{X}$ with $\mathcal{X}$ a Polish space, and for every $h \in \mathcal{H}$, we furthermore use without further ado the notation $\omega \mapsto X(\omega + h)$ as a shortcut for the (unique up to null sets) random variable Y such that $\mathbf{E} F(Y) = \mathbf{E} F(X) \exp(h^*(\omega) - \|h\|_{\mathcal{H}}^2/2)$ for every bounded measurable $F \colon \mathcal{X} \to \mathbf{R}$.

The type of random dynamics that we are interested in are then defined on a Banach space U and described by a family of maps

$$\Phi_{s,t} \colon \overline{U} \times \mathcal{M} \to \overline{U}, \quad 0 \le s \le t \le 1,$$

where we have extended the state space U to include a "death state" by defining $\overline{U} = U \cup \{\mathbf{\maltese}\}$. The space $\overline{U}$ is again a separable metric space by setting for example $d(\mathbf{\maltese}, u)^2 = 1 + \|u\|^2$ for all $u \in U$. These maps are assumed to be consistent in the sense that, for any $t \in [s, r]$ with $s \le r$, one has the identity

$$\Phi_{t,r}\big(\Phi_{s,t}(u, \xi), \xi\big) = \Phi_{s,r}(u, \xi), \quad \forall (u, \xi) \in \overline{U} \times \mathcal{M}. \tag{2.1}$$

The event $\Phi_{s,t}(u, \boldsymbol{\xi}) = \mathbf{\maltese}$ should be understood as the system "blowing up". After blow-up, the dynamics can no longer be defined. Consistent with this, we assume that $\Phi_{s,t}(\mathbf{\maltese}, \xi) = \mathbf{\maltese}$ for all ξ and all $0 \le s \le t \le 1$, which embodies the belief that "resurrection is impossible." We also use the shorthands $\Phi_t = \Phi_{0,t}$ and $\Phi = \Phi_{0,1}$.

Remark 2.1. The property (2.1) may look strange to someone used to random dynamical systems since we do not perform any time-shift of the element ξ. Since we never look at times beyond time 1, this turns out to be a much more convenient convention in our setting.

Throughout this article, we make the following rather strong regularity assumption on the maps $\Phi_{s,t}$.

Assumption 1. The preimage of U under the map

$$\Phi \colon (s, t, u, \xi) \mapsto \Phi_{s,t}(u, \xi),$$

is open and Φ is jointly continuous on $\Phi^{-1}(U)$. Furthermore, Φ is Fréchet differentiable in u at every point of $\Phi^{-1}(U)$.

In view of this assumption, we define the sets

$$\mathcal{N}_t = \big\{(u, \xi) \colon \Phi_t(u, \xi) \ne \mathbf{\maltese}\big\}, \quad \mathcal{N} \overset{\text{def}}{=} \mathcal{N}_1.$$

In particular, our assumption guarantees that the sets $\mathcal{N}_t$ are all open in $U \times \mathcal{M}$. We will denote the Fréchet derivative of Φ_t in the direction $v \in U$ by $D\Phi_t(u, \xi)v$, with the understanding that $D\Phi_t$ is only defined on $\mathcal{N}_t$. We also assume that we are given a somewhat more quantitative way of measuring how close a given point (u, ξ) is to being outside of $\mathcal{N}_t$. This is encapsulated in the following assumption.

Assumption 2. We are given a lower semi-continuous map $r \colon [0, 1] \times U \times \mathcal{M} \to [0, \infty]$ with the following properties.

1. For every $u \in U$ and every $t \in [0, 1]$, the map $\omega \mapsto r_t(u, \boldsymbol{\xi}(\omega))$ is $\mathcal{F}_t$-measurable.
2. For every $(u, \xi) \in U \times \mathcal{M}$, one has $r_0(u, \xi) < \infty$ and the map $t \mapsto r_t(u, \xi)$ is continuous and increasing.

3. One has $\{(u, \xi) : \Phi_t(u, \xi) = \skull\} = \{(u, \xi) : r_t(u, \xi) = \infty\}$.
4. For every $t \in (0, 1]$, the map r_t is locally Lipschitz continuous on $\mathcal{N}_t$.

The continuity of the map $t \mapsto r_t$ combined with the above assumptions imply that $\mathcal{N}_s \supset \mathcal{N}_t$ for all $0 \leq s \leq t \leq 1$ and that the following continuity property holds

$$\mathcal{N}_t = \bigcup_{a \in (t, 1]} \mathcal{N}_a \quad \text{for all } 0 \leq t \leq 1.$$

In particular, this implies that for any $u \in U$

$$\lim_{t \to 0^+} \mathbf{P}\big((u, \boldsymbol{\xi}) \in \mathcal{N}_t\big) = 1.$$

So far, we have described a rather generic class of random dynamical systems. The only part of our setup that is somewhat non-standard is the factorisation of our map into a fixed measurable map $\boldsymbol{\xi}$ which has a nice Gaussian probability space as its domain and a map Φ which has nice continuity properties. This factorisation is essential to the pathwise approach to stochastic analysis, be it via rough paths [18,40] or regularity structures [28]. Indeed, for the examples we have in mind, it is in general not possible to simplify the above setup by assuming that $\mathcal{M}$ is a topological vector space and the map $\boldsymbol{\xi}$ is linear.

In the next subsection, we introduce the additional structural assumptions that are crucial in guaranteeing the non-degeneracy of the noise required for the strong Feller property to hold.

2.1. *The shift map and differentiation with respect to the noise*

We will be interested in transformations of the dynamics induced by shifting the noise by translation. The assumptions we present in this section are twofold. On the one hand, we want to exhibit a space of Cameron–Martin shifts which are sufficiently "nice" to be compatible with the topology of the space $\mathcal{M}$. On the other hand, we want to guarantee that this space is sufficiently large to contain shifts allowing to compensate for arbitrary shifts in the initial condition.

Regarding the first type of assumption, we consider a space $E = L^p([0, 1], X_0) \subset \mathcal{H}$ of directions, where X_0 is some separable Banach subspace of H_0 and $p \in (2, \infty)$ is some fixed value. (In the sequel, we will typically choose p very large. The main reason not to use $p = \infty$ is the lack of separability.) It will sometimes be convenient to also consider shifts in $E_s = L^p([0, s], X_0)$ for some $s \in (0, 1]$. We then view E_s as a subspace of $E = E_1$ by identifying its elements with the corresponding function that vanishes on $(s, 1]$. For $\bar{s} \leq s$, we also write $\pi_{\bar{s}} \colon E_s \to E_{\bar{s}}$ for the restriction operator. Our assumption regarding the compatibility of E with the topology of $\mathcal{M}$ is then as follows.

Assumption 3. We are given a *continuous* action $\tau \colon E \times \mathcal{M} \to \mathcal{M}$ of E onto $\mathcal{M}$ such that, for every $\xi \in \mathcal{M}$, the map $h \mapsto \tau(h, \xi)$ is locally Lipschitz continuous and such that, for every $h \in E$, the identity

$$\boldsymbol{\xi}(\omega + h) = \tau\big(h, \boldsymbol{\xi}(\omega)\big), \tag{2.2}$$

holds $\mathbf{P}$-almost surely. Furthermore, the action τ is compatible with the maps $\Phi_{s,t}$ in the sense that if h is such that $h(r) = 0$ for $r \in [s, t]$, then

$$\Phi_{s,t}\big(u, \tau(h, \xi)\big) = \Phi_{s,t}(u, \xi), \tag{2.3}$$

for every $u \in \bar{U}$.

Since we now have a notion of a shift on $\mathcal{M}$, it is reasonable to explore differentiation of $\Phi(u, \xi)$ with respect to ξ. We do not want to assume full Malliavin differentiability, which would essentially amount to defining a directional derivative for all directions in $\mathcal{H}$, although in many cases of interest one does expect this to be achievable, albeit at the cost of significant additional technical difficulties, see for example [7]. Instead, we will only assume that the shift is well behaved in the directions given by $E \subset \mathcal{H}$ as follows.

Assumption 4. For every $(u, \xi) \in \mathcal{N}_t$, the map

$$h \mapsto \Phi_t\big(u, \tau(h, \xi)\big),$$

is Fréchet differentiable (as a map between the Banach spaces E and U) at $h = 0$.

We denote this Fréchet derivative by $\mathcal{D}\Phi_t(u, \xi)$. Lastly, and this is the main assumption guaranteeing that the strong Feller property holds, we would like to assume that the range of $\mathcal{D}\Phi_t(u, \xi)$ contains that of $D\Phi_t(u, \xi)$. More precisely, we would like to find a linear operator $A_t(u, \xi)$ transferring a variation $v \in U$ in the initial condition u to a variation $h \in E_t$ in the driving noise.

However, since we furthermore want to use adapted shift maps, the precise formulation of this condition requires some care. The formulation we choose to go with is as follows.

Assumption 5. For every $s \leq t$ with $t \in (0, 1]$, we are given a map

$$A_t^{(s)} : \mathcal{N}_s \to L(U, E_s), \,^{[1]}$$

and these maps are compatible in the sense that, for any $0 < s < r \leq t$, any $(u, \xi) \in \mathcal{N}_r$, and any $v \in U$, one has

$$A_t^{(r)}(u, \xi)v\!\restriction\![0, s] = A_t^{(s)}(u, \xi)v.^{[2]}$$

Furthermore, for every $u \in U$ and $s \leq t$, the map

$$\omega \mapsto A_t^{(s)}\big(u, \boldsymbol{\xi}(\omega)\big),$$

is $\mathcal{F}_s$-measurable and one has the identity

$$D\Phi_t(u, \xi)v + \mathcal{D}\Phi_t(u, \xi)\big(A_t^{(t)}(u, \xi)v\big) = 0, \tag{2.4}$$

for all $v \in U$ and all $(u, \xi) \in \mathcal{N}_t$. Furthermore, for every $0 < s \leq t \leq 1$, the map $A_t^{(s)}$ is locally Lipschitz continuous from $\mathcal{N}_s$ to $L(U, E_s)$ and bounded on every set of the form $\{(u, \xi) : r_s(u, \xi) \leq R\}$ with $R > 0$.

Remark 2.2. This formulation seems somewhat awkward: why not simply define a map A_t on $\mathcal{N}_t$ satisfying (2.4) and be done with it? The problem is that the event $(u, \boldsymbol{\xi}(\omega)) \in \mathcal{N}_t$ is in general not $\mathcal{F}_s$-measurable for $s < t$, so that our measurability condition would fail in this case. On the other hand, Assumption 2 does guarantee that the event $(u, \boldsymbol{\xi}(\omega)) \in \mathcal{N}_t$ is $\mathcal{F}_t$-measurable, so that the condition as stated has at least some chance of being fulfillable. We will however sometimes use the notational convention $A_t = A_t^{(t)}$ for simplicity.

Remark 2.3. We say that a map is "locally Lipschitz" if, for every x in its domain, there exists a neighbourhood of x on which it is Lipschitz continuous.

3. Strong Feller property

We now show how one can "integrate up" the infinitesimal shift given in (2.4) to produce a macroscopic shift which drives the solution starting from one initial condition to the same location at the terminal time as the solution starting from a nearby initial condition. This pathwise coupling will then be used to establish the strong Feller property. Since we will allow the existence of trajectories which blow up, we will only be able to prove the existence of this "compensating shift" on a set of positive (though not necessarily full) measure of noise trajectories.

---

[1] Here, $L(U, E)$ denotes the space of bounded linear operators $U \to E$.
[2] As usual, we write $F\!\restriction\!A$ for the restriction of a map F to a subset A of its domain.

Throughout this section, we assume that we are given Φ and $\boldsymbol{\xi}$ as above and that Assumptions 1–5 are satisfied. We will henceforth simply refer to this as "the context of Section 2". For any bounded measurable map $\Psi : \bar{U} \to \mathbb{R}$, we set $\mathfrak{P}\Psi(u) = \mathbf{E}\Psi(\Phi(u, \boldsymbol{\xi}(\omega)))$, which defines a Markov operator on $\bar{U}$. The main abstract result of this article is that in the context of the previous section, the Markov operator $\mathfrak{P}$ enjoys the strong Feller property, namely $\mathfrak{P}\Psi$ is continuous for all Ψ that are only bounded and measurable. Our proof actually gives the apparently stronger conclusion that the transition probabilities of $\mathfrak{P}$ are continuous in the total variation distance, but this is known [9,24,49] to be essentially equivalent.

3.1. *Existence of a compensating shift*

We now prove the existence of a "compensating shift" or rather a shift in the noise so that the solution starting from one initial condition coincides with the solution starting from another initial condition at time t. This is the content of equation (3.1) in the following theorem.

Theorem 3.1. *In the context of Section 2, for every $u \in U$, every M sufficiently large and every $t \in (0, 1)$, there exists a map $(\bar{u}, \xi) \mapsto h^{(t,M)}(u, \bar{u}, \xi) \in E_t$ such that*

1. *For every $(u, \xi) \in \mathcal{N}_t$, there exist $M_+, \gamma_+ > 0$ such that the identity*

$$\Phi_t(u, \xi) = \Phi_t\big(\bar{u}, \tau\big(h^{(t,M)}(u, \bar{u}, \xi), \xi\big)\big), \tag{3.1}$$

   holds for every $M \geq M_+$ and for every $\bar{u}$ with $\|\bar{u} - u\| \leq \gamma_+$.
2. *For every $u, \bar{u}$ in U and every $s \in (0, t)$, the map $\omega \mapsto \pi_s h^{(t,M)}(u, \bar{u}, \boldsymbol{\xi}(\omega))$ is $\mathcal{F}_s$-measurable.*
3. *One has the bound $\|h^{(t,M)}(u, \bar{u}, \xi)\|_{\mathcal{H}} \leq M \|\bar{u} - u\|_U$.*

Proof. We begin by defining a "truncated" version of the maps A_t. For this, we choose a smooth function $\chi : \mathbb{R}_+ \to [0, 1]$ such that $\chi(r) = 1$ for $r \leq 1$ and $\chi(r) = 0$ for $r \geq 2$, and we set

$$\big(\tilde{A}_t(u, h, \xi)v\big)(s) = \big(A_t^{(s)}(u, \tau(h, \xi))v\big)(s),$$

$$\big(Z^R(u, h, \xi)\big)(s) = \chi\big(r_s(u, \tau(h, \xi))/R\big),$$

for all $s \in [0, t]$. As a consequence of Assumption 5, for every $s \in (0, t)$, the restriction of the map $\tilde{A}_t(u, h, \xi)$ to $[0, s]$ is locally Lipschitz continuous as a function of h and ξ on the set $\{(h, \xi) : (u, \tau(h, \xi)) \in \mathcal{N}_s\}$. For any direction $v \in U$ with $\|v\| = 1$ and for any $R > 0$, we then consider the ODE in $U \times E$ given by

$$\partial_\gamma u_\gamma = v, \quad u_0 = u,$$
$$\partial_\gamma h_\gamma = \tilde{A}_t(u_\gamma, h_\gamma, \xi)v \inf_{\bar{\gamma} \leq \gamma} Z^R(u_{\bar{\gamma}}, h_{\bar{\gamma}}, \xi), \quad h_0 = 0. \tag{3.2}$$

We claim that this ODE has unique global solutions. Indeed, consider s such that $(u, \xi) \in \mathcal{N}_s$. Then, it follows from Assumptions 3 and 5 that the ODE in $U \times E_s$ given by (3.2), but with the right-hand side interpreted as taking values in E_s, admits unique solutions $(u_\gamma, h_\gamma^{(s)})$ up to the first value of γ such that $r_s(u_\gamma, \tau(h_\gamma^{(s)}, \xi)) = 2R$. Let us denote this value by $\gamma_s^\star$ (which also depends on (u, ξ) of course). Note that since $s \mapsto r_s$ is increasing, the map $s \mapsto \gamma_s^\star$ is decreasing. Note also that for $\bar{s} \leq s$ and for $\gamma \leq \gamma_s^\star$, one has $\pi_{\bar{s}} h_\gamma^{(s)} = h_\gamma^{(\bar{s})}$. The desired solution to (3.2) is then given by

$$h_\gamma(s) = h_{\gamma \wedge \gamma_s^\star}^{(s)}(s).$$

We now show that the choice $h^{(t,M)}(u, \bar{u}, \xi) = h_\gamma$ indeed satisfies (3.1) for $\bar{u} = u + \gamma v$ with $\|v\| = 1$, provided that γ is small enough and that R is such that $\tilde{A}_t$ is bounded by M on the set $\{(u, h, \xi) : r_t(u, \tau(h, \xi)) \leq R\}$.

Since $(u, \xi) \in \mathcal{N}_t$, we can find an R large enough so $r_t(u, \xi) \leq R/2$. Once this value of R is fixed, we can find $\gamma_+ > 0$ such that $r_t(u, \tau(h, \xi)) \leq R$ for every h with $\sup_{s \in [0,t]} \|h(s)\| \leq \gamma_+$. Since $\tilde{A}_t$ is bounded by M, we can

furthermore choose γ_+ sufficiently small so that $\tilde{A}_t \le 1/\gamma_+$. This choice of γ_+ guarantees *a priori* that, for $\gamma \le \gamma_+$, the solution to (3.2) satisfies

$$\partial_\gamma h_\gamma = A_t^{(t)}\big(u_\gamma, \tau(h_\gamma, \xi)\big) v. \tag{3.3}$$

Differentiating the quantity $\Phi_t(u_\gamma, \tau(h_\gamma, \xi))$ with respect to γ, we then obtain from the definition of $\mathcal{D}\Phi_t$, the chain rule, (3.3), and (2.4) that

$$\frac{d}{d\gamma}\Phi_t\big(u_\gamma, \tau(h_\gamma, \xi)\big) = D\Phi_t\big(u_\gamma, \tau(h_\gamma, \xi)\big)v + \mathcal{D}\Phi_t\big(u_\gamma, \tau(h_\gamma, \xi)\big) A_t^{(t)}\big(u_\gamma, \tau(h_\gamma, \xi)\big)v$$

$$= 0,$$

which shows that the first property of h does indeed hold.

The measurability of $\pi_s h^{(t)}$ with respect to $\mathcal{F}_s$ is then an immediate consequence of the fact that both $A_t^{(s)}$ and r_s are $\mathcal{F}_s$-measurable, and the last property holds by definition since $\|\partial_\gamma h_\gamma(s)\| \le M$ for $\gamma \le \gamma_s^\star$ and $\partial_\gamma h_\gamma(s) = 0$ afterwards. $\qquad\square$

3.2. *The Markov operator $\mathfrak{P}$ is strong Feller*

We now use the estimate given in (3.1) to prove that $\mathfrak{P}$ satisfies the strong Feller property. The existence of a compensating shift as in (3.1) can be combined with Girsanov's theorem to show that nearby points have transition measures which are close in total variation.

Theorem 3.2. *In the context of Section 2, the Markov operator $\mathfrak{P}$ satisfies the strong Feller property.*

Proof. Without loss of generality, we assume that $\Psi(\text{☻}) = 0$ and that $\sup_{u \in \bar{U}} |\Psi(u)| \le 1$. It is sufficient to show that, for any fixed $u \in U$, one has

$$\lim_{\gamma \to 0} \sup_{\|\bar{u} - u\| \le \gamma} \big|\mathfrak{P}\Psi(u) - \mathfrak{P}\Psi(\bar{u})\big| = 0. \tag{3.4}$$

For this, we proceed as follows. Viewing the initial condition u and any time $t \in (0, 1)$ as fixed, we choose a direction $v \in U$ and write $h_{\gamma,t}^{(M)}(\xi, v)$ for the corresponding shift obtained in Theorem 3.1, namely

$$h_{\gamma,t}^{(M)}(\xi, v) = h^{(t,M)}(u, u + \gamma v, \xi).$$

We also fix some $\varepsilon > 0$.

It follows from Theorem 3.1 that we can find $M_+, \gamma_+ > 0$ (possibly depending on t and ξ) such that the identity

$$\Phi_t(u, \xi) = \Phi_t\big(u + \gamma v, \tau\big(h_{\gamma,t}^{(M)}(\xi, v), \xi\big)\big),$$

and the bound $\|h_{\gamma,t}^{(M)}(\xi, v)\|_{\mathcal{H}} \le \gamma M$ hold for all $\gamma \le \gamma_+(\xi, t)$, all $M \ge M_+(\xi, t)$, and all v with $\|v\| = 1$. In particular, as a consequence of (2.1) and (2.3), one has

$$\Phi(u, \xi) = \Phi\big(u + \gamma v, \tau\big(h_{\gamma,t}^{(M)}(\xi, v), \xi\big)\big),$$

since if the trajectories coincide at time t then they coincide for all $T > t$.

Fixing a $u \in U$, $v \in U$ with $\|v\| = 1$, we write $\bar{u} = u + \gamma v$ for $\gamma > 0$. We also define the sets

$$\overline{\Omega}_{\gamma,M,t} = \big\{\omega : \gamma \le \gamma_+\big(\boldsymbol{\xi}(\omega), t\big) M \ge M_+\big(\boldsymbol{\xi}(\omega), t\big)\big\},$$

as well as

$$\Omega_{\gamma,M,t} = \big\{\omega : \omega + h_{\gamma,t}^{(M)}\big(\boldsymbol{\xi}(\omega), v\big) \in \overline{\Omega}_{\gamma,M,t}\big\}.$$

(The latter actually also depends on v, but this will be irrelevant for our proof since all of our estimates are uniform in v.) Observe that since both γ_+ and M_+ are finite for all $(u, \xi) \in \mathcal{N}_t$, one has that $\mathbf{P}(\overline{\Omega}_{\gamma, M, t}) \to \mathbf{P}((u, \xi) \in \mathcal{N}_t)$ as $\gamma \to 0$ and $M \to \infty$. Furthermore, $\mathbf{P}((u, \xi) \in \mathcal{N}_t) \to 1$ as $t \to 0$. Combining these facts implies for any $\varepsilon > 0$ there exists a t_0 and, for every $t \le t_0$, there exist γ_0 and M_0 so that

$$\mathbf{P}(\overline{\Omega}_{\gamma, M, t}^c) \le \varepsilon \quad \text{for any } t \in (0, t_0], \gamma \le \gamma_0 \text{ and } M \ge M_0, \tag{3.5}$$

where we wrote B^c for the complement of an event $B \subset \Omega$. We will henceforth always assume that γ is sufficiently small so that $\gamma M \le 1$.

Next, observe that we can then write

$$\mathfrak{P}\Psi(u) - \mathfrak{P}\Psi(\bar{u}) = \mathbf{E}\big(\Psi\big(\Phi(u, \xi)\big) - \Psi\big(\Phi(\bar{u}, \xi)\big)\big)$$

$$= \mathbf{E}\big(\Psi\big(\Phi(u, \xi)\big)(\mathbf{1}_{\Omega_{\gamma, M, t}} + \mathbf{1}_{\Omega_{\gamma, M, t}^c}) - \Psi\big(\Phi(\bar{u}, \xi)\big)(\mathbf{1}_{\overline{\Omega}_{\gamma, M, t}} + \mathbf{1}_{\overline{\Omega}_{\gamma, M, t}^c})\big). \tag{3.6}$$

Using first Girsanov's theorem, then the definition of $\Omega_{\gamma, M, t}$ combined with Proposition 3.6 below and finally (3.1), we obtain the identity

$$\mathbf{E}\Psi\big(\Phi(\bar{u}, \xi)\big)\mathbf{1}_{\overline{\Omega}_{\gamma, M, t}(\xi)} = \mathbf{E}\Psi\big(\Phi\big(\bar{u}, \xi\big(\omega + h_{\gamma, t}^{(M)}(\xi)\big)\big)\big)\mathcal{E}_\gamma(\omega, v)\mathbf{1}_{\overline{\Omega}_{\gamma, M, t}}\big(\omega + h_{\gamma, t}^{(M)}(\xi)\big)$$

$$= \mathbf{E}\Psi\big(\Phi\big(\bar{u}, \tau\big(h_{\gamma, t}^{(M)}(\xi), \xi\big)\big)\big)\mathcal{E}_\gamma(\omega, v)\mathbf{1}_{\Omega_{\gamma, M, t}}(\omega)$$

$$= \mathbf{E}\Psi\big(\Phi(u, \xi)\big)\mathcal{E}_\gamma(\omega, v)\mathbf{1}_{\Omega_{\gamma, M, t}}(\omega)$$

$$= \mathbf{E}\Psi\big(\Phi(u, \xi)\big)\mathbf{1}_{\Omega_{\gamma, M, t}}(\omega) + \mathbf{E}\Psi\big(\Phi(u, \xi)\big)\big(\mathcal{E}_\gamma(\omega, v) - 1\big)\mathbf{1}_{\Omega_{\gamma, M, t}}(\omega),$$

where $\mathcal{E}_\gamma(\omega, v)$ is the exponential martingale given by Girsanov's theorem. Rearranging this equality and using the bound $|\Psi| \le 1$ produces

$$\big|\mathbf{E}\Psi\big(\Phi(\bar{u}, \xi)\big)\mathbf{1}_{\overline{\Omega}_{\gamma, M, t}}(\xi) - \mathbf{E}\Psi\big(\Phi(u, \xi)\big)\mathbf{1}_{\Omega_{\gamma, M, t}}(\omega)\big| \le \sqrt{\mathbf{E}\big(\mathcal{E}_\gamma(\omega, v) - 1\big)^2}.$$

Since $\|h_{\gamma, t}^{(M)}(\xi, v)\|_{\mathcal{H}} \le 1$ by Theorem 3.1 and the assumption $\gamma M \le 1$, one has

$$\mathbf{E}\big(\mathcal{E}_\gamma(\omega, v) - 1\big)^2 = \mathbf{E}\mathcal{E}_\gamma(\omega, v)^2 - 1 = \mathbf{E}\exp\big(\big\|h_{\gamma, t}^{(M)}(\xi, v)\big\|_{\mathcal{H}}^2\big) - 1 \le e\mathbf{E}\big\|h_{\gamma, t}^{(M)}(\xi, v)\big\|_{\mathcal{H}}^2.$$

Similarly, using again Girsanov's theorem, if $\mathcal{A} = \{\omega : (u, \xi) \in \mathcal{N}\}$ and $\bar{\mathcal{A}} = \{\omega : (\bar{u}, \xi) \in \mathcal{N}\}$ we also obtain the bound

$$\mathbf{P}\big(\Omega_{\gamma, M, t}^c\big) = \mathbf{E}(\mathcal{E}_\gamma \mathbf{1}_{\Omega_{\gamma, M, t}^c}) - \mathbf{E}\big((\mathcal{E}_\gamma - 1)\mathbf{1}_{\Omega_{\gamma, M, t}^c}\big)$$

$$= \mathbf{P}\big(\overline{\Omega}_{\gamma, M, t}^c\big) - \mathbf{E}\big((\mathcal{E}_\gamma - 1)\mathbf{1}_{\Omega_{\gamma, M, t}^c}\big)$$

$$\le \mathbf{P}\big(\overline{\Omega}_{\gamma, M, t}^c\big) + \big(\mathbf{E}(\mathcal{E}_\gamma - 1)^2\big)^{1/2}.$$

Combining all of these bounds with (3.6) and $|\Psi| \le 1$, we finally obtain that

$$\big|\mathfrak{P}\Psi(u) - \mathfrak{P}\Psi(\bar{u})\big| \le \big|\mathbf{E}\Psi\big(\Phi(u, \xi)\big)\mathbf{1}_{\Omega_{\gamma, M, t}} - \mathbf{E}\Psi\big(\Phi(\bar{u}, \xi)\big)\mathbf{1}_{\overline{\Omega}_{\gamma, M, t}}\big| + \mathbf{P}\big(\Omega_{\gamma, M, t}^c\big) + \mathbf{P}\big(\overline{\Omega}_{\gamma, M, t}^c\big)$$

$$\le 2\mathbf{P}\big(\overline{\Omega}_{\gamma, M, t}^c\big) + 2e\sqrt{\mathbf{E}\big\|h_{\gamma, t}^{(M)}(\xi, v)\big\|_{\mathcal{H}}^2}. \tag{3.7}$$

Fixing any $\varepsilon > 0$ from (3.5) we can find a t small enough so that the first term above is less than $\varepsilon/2$ for M sufficiently large and γ sufficiently small. Then, possibly choosing γ even smaller so that $2e\gamma \le \varepsilon/2$ and $\gamma M \le 1$, the second term will also be less that $\varepsilon/2$. Hence we conclude that for all γ sufficiently small one has the bound

$$\sup_{\|\bar{u} - u\| \le \gamma} \big|\mathfrak{P}\Psi(u) - \mathfrak{P}\Psi(\bar{u})\big| \le \varepsilon.$$

As ε was arbitrarily, we conclude that (3.4) holds and the proof is complete. $\qquad \square$

Remark 3.3. If one knows that the system is not explosive then one can often obtain a quantitative bound in (3.7) and show that $\mathfrak{P}\Psi$ is Hölder continuous whenever Ψ is bounded. In particular, if $\mathbf{P}(\boldsymbol{\xi} \in \mathcal{N}_t) = 1$ for some $t \in (0, 1]$ then one is free to fix that t in (3.7). Having fixed t, one knows from Theorem 3.1 that

$$\sqrt{\mathbf{E}\big\|h_{\gamma,t}^{(M)}(\boldsymbol{\xi}, v)\big\|_{\mathcal{H}}^2} \le C\gamma = C\|u - \bar{u}\|.$$

Combining this with any estimate on $|\mathbf{P}(\overline{\Omega}_{\gamma,M,t}) - 1|$ as $\gamma \to 0$, such as $\mathbf{P}(\overline{\Omega}_{\gamma,M,t}) \le K\gamma^p$ for some $p > 0$ will produce a quantitative estimate. However, as such a quantitative bound is not needed to prove unique ergodicity of the system, we do not dwell on this point.

Remark 3.4. While we have emphasised the use of (3.1) at points which do not blow up, our result also shows that the probability of blowing up varies continuously with the initial condition: just apply $\mathfrak{P}$ to the test function Ψ given by $\Psi(\skull) = 1$ and $\Psi(u) = 0$ for $u \in U$. The intuition is clear: if there is a compensating shift which shifts the trajectory from a nearby initial condition to coincide with a trajectory which is en route to exploding, then the nearby point also blows up with positive probability.

3.3. *Robustness of shift maps*

In the proof given in the previous subsection, we used the fact that (2.2) also holds for *random* shifts h as long as they are adapted. The aim of this section is to show that this is indeed automatically the case, at least under some Novikov-type condition. (This could probably be relaxed, but it is trivially satisfied in our application since our cut-off procedure guarantees that we only consider shifts that are bounded by a deterministic constant.) We will make use of the following fact.

Lemma 3.5. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces, $\mathbf{P}$ a probability measure on $\mathcal{X}$, and X, Y two measurable maps $\mathcal{X} \to \mathcal{Y}$. Then $X = Y$ $\mathbf{P}$-almost surely if and only if the identity*

$$\mathbf{E}f\big(X(x)\big)g(x) = \mathbf{E}f\big(Y(x)\big)g(x),$$

holds for any two bounded uniformly continuous functions $f : \mathcal{Y} \to \mathrm{R}$ and $g : \mathcal{X} \to \mathrm{R}$.

The main result of this subsection is the following. Here, for any separable Banach space $\mathcal{X}$, $L_{\mathrm{pr}}^p(\Omega \times [0, 1], \mathcal{X})$ denotes the set of L^p functions from $\Omega \times [0, 1]$ into $\mathcal{X}$ that are measurable with respect to the predictable σ-algebra on $\Omega \times [0, 1]$ (with respect to the filtration $\{\mathcal{F}_t\}_t$). In the particular case $\mathcal{X} = X_0$, one has of course $L_{\mathrm{pr}}^p(\Omega \times [0, 1], \mathcal{X}) \subset L^p(\Omega \times [0, 1], \mathcal{X}) \approx L^p(\Omega, E)$, so that we alternatively view these elements as E-valued random variables.

Proposition 3.6. *Let $\boldsymbol{\xi} \colon \Omega \to \mathcal{M}$ and let τ be a continuous action of E on $\mathcal{M}$ as above. If (2.2) holds for every deterministic $h \in E$, then it also holds for every predictable $h \colon \Omega \to E$ satisfying $\mathbf{E}\exp(6\|h\|_{\mathcal{H}}^2) < \infty$.*

Proof. Let $h_n \in L_{\mathrm{pr}}^p(\Omega \times [0, 1], X_0)$ be a sequence of random variables such that each h_n takes only countably many values in E, and such that the h_n converge to h both in $L_{\mathrm{pr}}^2(\Omega \times [0, 1], H_0)$ and in probability in E. We choose the h_n furthermore such that $\sup_n \mathbf{E}\exp(6\|h_n\|_{\mathcal{H}}^2) < \infty$. This is always possible by the separability of X_0 and the density of simple processes in $L_{\mathrm{pr}}^p(\Omega \times [0, 1], X_0)$.

Since h_n only takes countably many values, the identity

$$\boldsymbol{\xi}\big(\omega + h_n(\omega)\big) = \tau\big(h_n(\omega), \boldsymbol{\xi}(\omega)\big),$$

holds almost surely for every n. Here, assuming that h_n takes the value $h_n^{(k)}$ on the set $A_k \subset \Omega$ forming a partition of Ω, the left-hand side should be interpreted as the random variable which equals $\boldsymbol{\xi}(\omega + h_n^{(k)})$ on A_k for every k. In particular, for every n, we have the identity

$$\mathbf{E}f\big(\boldsymbol{\xi}\big(\omega + h_n(\omega)\big)\big)g\big(\omega + h_n(\omega)\big)\mathcal{E}(h_n) = \mathbf{E}f\big(\tau\big(h_n(\omega), \boldsymbol{\xi}(\omega)\big)\big)g\big(\omega + h_n(\omega)\big)\mathcal{E}(h_n), \tag{3.8}$$

where we used the notation

$$\mathcal{E}(h) \stackrel{\mathrm{def}}{=} \exp\left(-\int_0^1 \langle h(s), d\omega(s)\rangle_{\mathcal{H}} - \frac{1}{2}\int_0^1 \|h(s)\|_{\mathcal{H}}^2 \, ds\right).$$

It follows from the first and second properties that $\mathcal{E}(h_n)$ converges toward $\mathcal{E}(h)$ in $L^1(\Omega, \mathbf{P})$: writing $\mathcal{E}(h) = \exp(\mathcal{I}(h))$, it is easy to verify that

$$\mathbf{E}|\mathcal{I}(h) - \mathcal{I}(h_n)|^2 \leq \mathbf{E}\|h - h_n\|_{\mathcal{H}}^2 \left(1 + \|h + h_n\|_{\mathcal{H}}^2\right)$$
$$\lesssim \left(\mathbf{E}\|h - h_n\|_{\mathcal{H}}^2 \mathbf{E}\left(1 + \|h\|_{\mathcal{H}} + \|h_n\|_{\mathcal{H}}\right)^6\right)^{1/2} \lesssim \left(\mathbf{E}\|h - h_n\|_{\mathcal{H}}^2\right)^{1/2}.$$

We conclude that

$$\mathbf{E}|\mathcal{E}(h) - \mathcal{E}(h_n)| \leq \mathbf{E}|\mathcal{I}(h) - \mathcal{I}(h_n)|\left(\mathcal{E}(h) + \mathcal{E}(h_n)\right)$$
$$\leq \left(2\mathbf{E}|\mathcal{I}(h) - \mathcal{I}(h_n)|^2 \mathbf{E}\left(\mathcal{E}(h)^2 + \mathcal{E}(h_n)^2\right)\right)^{1/2} \lesssim \left(\mathbf{E}|\mathcal{I}(h) - \mathcal{I}(h_n)|^2\right)^{1/4},$$

which does indeed converge to 0 by assumption.

As a consequence of the uniform continuity of g and τ and the convergence of h_n to h, the right-hand side of (3.8) converges as $n \to \infty$ to

$$\mathbf{E}f\left(\tau\left(h(\omega), \boldsymbol{\xi}(\omega)\right)\right)g\left(\omega + h(\omega)\right)\mathcal{E}(h). \tag{3.9}$$

On the other hand, for any fixed n, we can use Girsanov's theorem to show that the left-hand side equals

$$\mathbf{E}f\left(\boldsymbol{\xi}(\omega)\right)g(\omega),$$

independently of n. Applying Girsanov's theorem in the opposite direction and comparing this to (3.9), we conclude that

$$\mathbf{E}f\left(\boldsymbol{\xi}\left(\omega + h(\omega)\right)\right)g\left(\omega + h(\omega)\right)\mathcal{E}(h) = \mathbf{E}f\left(\tau\left(h(\omega), \boldsymbol{\xi}(\omega)\right)\right)g\left(\omega + h(\omega)\right)\mathcal{E}(h).$$

Writing $\tilde{\mathbf{E}}$ for expectations under the probability measure $\tilde{\mathbf{P}} = \mathcal{E}(h)\mathbf{P}$ and writing σ_h for the σ-algebra generated by the shift map T_h, it follows from Lemma 3.5 that the identity

$$f\left(\boldsymbol{\xi}\left(\omega + h(\omega)\right)\right) = \tilde{\mathbf{E}}\left(f\left(\tau\left(h(\omega), \boldsymbol{\xi}(\omega)\right)\right)|\tilde{\sigma}\right),$$

holds $\tilde{\mathbf{P}}$-almost surely. Since this holds for every bounded uniformly continuous function f, it also holds for its square, so that

$$\tilde{\mathbf{E}}\left(f\left(\tau\left(h(\omega), \boldsymbol{\xi}(\omega)\right)\right)^2|\tilde{\sigma}\right) = \left(\tilde{\mathbf{E}}\left(f\left(\tau\left(h(\omega), \boldsymbol{\xi}(\omega)\right)\right)|\tilde{\sigma}\right)\right)^2.$$

This implies that $\omega \mapsto f(\tau(h(\omega), \boldsymbol{\xi}(\omega)))$ is $\tilde{\sigma}$-measurable, thus concluding the proof. □

3.4. *A few consequences*

Recall that every invariant measure for a Markov operator $\mathfrak{P}$ on a Polish space can be written as a superposition of ergodic invariant measures. In our case, the Dirac mass located on ☠ is always invariant because of the property $\Phi(☠, \xi) = ☠$, but one might be interested in asking whether $\mathfrak{P}$ also has a unique invariant measure located on U. The main ingredient for such a statement is the fact [14] that the strong Feller property yields the following strengthening of the fact that distinct ergodic invariant measures are mutually singular.

Proposition 3.7. *Let $\mathfrak{P}$ be a strong Feller Markov operator on a Polish space and let $\mu \neq \nu$ be ergodic invariant probability measures for $\mathfrak{P}$. Then, the topological supports of μ and ν are disjoint.*

We immediately obtain the following two corollaries.

Corollary 3.8. *In the context of Section 2, assume that, for every open set $A \subset U$ and every $u \in U$, one has*

$$\mathbf{P}\big(\{\omega : \Phi(u, \xi(\omega)) \in A\}\big) > 0. \tag{3.10}$$

Then, $\mathfrak{P}$ admits at most one invariant measure μ with $\mu(U) = 1$.

Proof. The condition (3.10) implies that every invariant measure located on U has full support, so that the claim follows from Proposition 3.7. □

The next corollary is useful in situations where an invariant measure is already known, as is the case for the KPZ equation [20], the 2D Navier–Stokes equations driven by space-time white noise [10], the dynamical $P(\Phi)_2$ model [1,11], as well as the dynamical Φ_3^4 model [32].

Corollary 3.9. *In the context of Section 2, assume that $\mathfrak{P}$ admits an invariant measure μ with $\operatorname{supp} \mu = U$. Then μ is the only invariant measure concentrated on U.*

Proof. Since this result does not appear to be easy to find in the literature, we give a self-contained proof. We first show that $\mathfrak{P}$ can have at most countably many ergodic invariant measures as a consequence of the strong Feller property. Recall that the strong Feller property implies the existence of a probability measure λ such that the transition probabilities $\mathfrak{P}(u, \cdot)$ are absolutely continuous with respect to λ for every $u \in U$. (See for example [24, Lem. 1.6.4].) In particular, by Prokhorov's theorem, there exists a σ-compact (i.e. countable union of compact) set $K \subset U$ such that $\mathfrak{P}(u, K) = 1$ for every $u \in U$.

The strong Feller property also implies that for every $u \in U$ there exists a neighbourhood N_u of u such that there exists at most one ergodic invariant measure μ_u for with $\operatorname{supp} \mu \cap N_u \neq \varnothing$. By the σ-compactness of K, we can find a sequence $\{u_n\}_{n \in \mathbb{N}}$ so that $K \subset \bigcup_{n \in \mathbb{N}} N_{u_n}$. Since every invariant measure is concentrated on K, the claim follows.

Assume now by contradiction that μ is not ergodic. Then, by the above, we can write $\mu = \sum_n p_n \mu_n$ for at most countably many distinct ergodic invariant measures μ_n. Furthermore, by Proposition 3.7, the supports of the μ_n are all disjoint. Since the supports are furthermore closed, their union has to be all of U by our assumption on μ, and U is connected, we immediately conclude that the sum can have only one term. □

4. Application to singular SPDEs

In this section, we show how to apply the previous abstract result to solutions to singular stochastic PDEs of the type studied in [28]. We henceforth assume that we are in the setting of [28, Section 7.3] which we summarise in the next subsection.

4.1. *A very general setting*

We are working with an ambient regularity structure $(\mathcal{T}, \mathcal{G})$ which contains the usual polynomial regularity structure $(\bar{\mathcal{T}}, \bar{\mathcal{G}})$ on $\mathbf{R}^{d+1}$ with scaling $\mathfrak{s} = (2q, \bar{\mathfrak{s}})$ for some scaling $\bar{\mathfrak{s}}$ of $\mathbf{R}^d$. In particular, we have a distinguished basis vector $\mathbf{1} \in \bar{\mathcal{T}}$ of degree 0 which represents the constant functions. We assume that it is also endowed with an abstract integration map $\mathcal{I}$ of order $2q$ as well as a convolution operator $\mathcal{P}$ (associated to $\mathcal{I}$ as in [28, Section 4]), obtained from the linear evolution problem associated to a homogeneous (with respect to the scaling $\bar{\mathfrak{s}}$) elliptic differential operator $\mathcal{L}$ of order $2q$ with constant coefficients on $\mathbf{R}^d$. In particular, one has

$$\mathcal{R}\mathcal{P}\mathbf{\Phi} = P \star \mathcal{R}\mathbf{\Phi},$$

for any $\mathbf{\Phi} \in \mathcal{D}^\gamma$ (taking values in the domain of $\mathcal{I}$) with $\gamma > 0$ and for P the Green's function of $\partial_t - \mathcal{L}$. A very important property of the kernel P we will use in the sequel is that it is non-anticipative in the sense that $P(t, x) = 0$ for $t < 0$. Henceforth, we will also assume that we are given a lattice L on $\mathbf{R}^d$ with compact fundamental domain and

that all quantities of interest are periodic with respect to L. For example, we simply write $\mathcal{C}^\alpha$ for the space of Hölder continuous functions of regularity α that are L-periodic, without specifically mentioning L. The same is true for our spaces $\mathcal{D}^{\gamma,\eta}$ of modelled distributions.

As usual, we write $\mathcal{M}$ for the space of all admissible models for $(\mathcal{T}, \mathcal{G})$ and the scaling $\mathfrak{s}$ that are furthermore periodic (in the sense of [28, Def. 3.33]) with respect to the lattice L. Here, a model (Π, Γ) is said to be admissible if

- it agrees with the polynomial model on $\bar{\mathcal{T}}$;
- it realises K for $\mathcal{I}$ (in the sense of [28, Def. 5.9]) for a kernel K such that $P = K + R$ where R is smooth and K satisfies [28, Ass. 5.1 & 5.4].

Such a decomposition for P exists by [28, Lem. 5.5].

Remark 4.1. Note that the operator $\mathcal{P}$ here is equal to the operator $\mathcal{K}_{\bar{\gamma}} + R_\gamma \mathcal{R}$ in [28, Thm. 7.8].

Remark 4.2. In general, one can allow for several operators $\mathcal{I}^{(k)}$ with k taking values in some finite index set, associated to integral kernels $P^{(k)}$. All of the results in this section carry over to this more general setting without any substantial change, but we prefer to stick to the scalar-valued case to simplify notations.

We also assume that we are given two finite-dimensional sectors V and $\bar{V}$ such that

$$\bar{\mathcal{T}} \subset V \subset \bar{\mathcal{T}} \oplus \mathcal{T}_{\geq \zeta}, \qquad \bar{\mathcal{T}} \subset \bar{V} \subset \bar{\mathcal{T}} \oplus \mathcal{T}_{\geq \bar{\zeta}}, \tag{4.1}$$

for some $0 < \zeta \leq \bar{\zeta} + 2q$. We also fix from now on exponents $\gamma, \bar{\gamma}$ such that $\bar{\gamma} + 2q > \gamma \geq \bar{\gamma} > 0$, such that $\bar{\gamma} \leq \zeta$, and such that $\mathcal{Q}_{<\gamma} \mathcal{I} \bar{V}_{<\bar{\gamma}} \subset V_{<\gamma}$, where $\mathcal{Q}_{<\gamma}$ denotes projection onto $\mathcal{T}_{<\gamma}$.

To describe the nonlinearity of our evolution problem, we assume that we are given a distinguished element $N \in \mathcal{T}$ such that $\Gamma N = N$ for every Γ belonging to the structure group $\mathcal{G}$. In particular, as a consequence of the fact that N is invariant under $\mathcal{G}$ and V contains $\bar{\mathcal{T}}$, the subspace $\hat{V} \subset \mathcal{T}$ given by

$$\hat{V} = V + R\mathcal{I}(N),$$

is again a sector. In some situations, the choice $N = 0$ is possible, but we will see in Section 5.1 below that one sometimes needs to make a different choice in order to remain compatible with the condition $\zeta > 0$ with ζ appearing in the properties of V. Given this sector $\hat{V}$, we make the following structural assumption on both our models and our nonlinearity.

Assumption 6. We are given a continuous "canonical lift map" $\mathcal{L} : \mathcal{C}^\infty(\mathbf{R} \times \mathbf{R}^d) \to \mathcal{M}$ as well as a finite-dimensional Lie group $\mathfrak{R}$, together with a continuous action M of $\mathfrak{R}$ onto $\mathcal{M}$.

We are also given a map $F : \hat{V}_{<\gamma} \to \bar{V}_{<\bar{\gamma}}$ such that the corresponding composition operator is strongly locally Lipschitz continuous (in the sense of [28, Section 7.3]) from $\mathcal{D}^{\gamma,\eta}$ to $\mathcal{D}^{\bar{\gamma},\bar{\eta}}$, locally uniformly over the underlying admissible model, and for some exponents $\eta, \bar{\eta} > -2q$ such that $(\bar{\eta} \wedge \bar{\zeta}) + 2q > (\eta \vee 0)$ and such that $\eta \leq \zeta$.

The first part of this assumption leads us to the following definition of a "nice" model.

Definition 4.3. An admissible model $\mathbf{\Pi} = (\Pi, \Gamma) \in \mathcal{M}$ for $(\mathcal{T}, \mathcal{G})$ is *nice* if there exist $f_n \in \mathcal{C}^\infty(\mathbf{R} \times \mathbf{R}^d)$ and $g_n \in \mathfrak{R}$ such that $\mathbf{\Pi} = \lim_{n \to \infty} M_{g_n} \mathcal{L}(f_n)$ and furthermore the distribution $P \star \Pi_z N$ belongs to $\mathcal{C}(\mathbf{R}, \mathcal{C}^\eta)$.

Write $\mathcal{M}$ for the closure of all smooth and nice models in the space of nice admissible models for the regularity structure $(\mathcal{T}, \mathcal{G})$. Setting $I_0 = [-2, 3]$, we endow $\mathcal{M}$ with the "seminorm" given by

$$\|(\Pi, \Gamma)\| = \sup_{z, \varphi, \lambda, \alpha, \tau} \lambda^{-\alpha} \left| (\Pi_z \tau)(\varphi_z^\lambda) \right| + \sup_{z, \bar{z}, \alpha, \beta, \tau} |z - \bar{z}|_{\mathfrak{s}}^{-(\alpha - \beta)} \|\Gamma_{z\bar{z}} \tau\|_{\mathcal{T}_\beta}$$

$$+ \|P \star \Pi_0 N\|_{\mathcal{C}(I_0, \mathcal{C}^\eta)},$$

as well as the corresponding distance function as in [28, Eq. 2.17]. Here, the supremum over z, as well as the one over $z, \bar{z}$ run over $I_0 \times R^d$, while the suprema over all other variables are as in [28, Def. 2.17]. We also identify elements that are at distance 0, thus turning $\mathcal{M}$ into a separable metric space. Although, strictly speaking, an element of $\mathcal{M}$ is now an equivalence class of models, all of the operations we will ever make use of only ever concern modelled distributions defined on $[0, 1] \times R^d$ (which are canonically extended to be 0 outside), so that both $\mathcal{P}$ and the reconstruction operator are well-defined and do not depend on the representative in $\mathcal{M}$.

Writing $(\Omega, \mathbf{P})$ for a filtered Gaussian probability space as in Section 2, we furthermore assume without loss of generality that Ω is of the form $\mathcal{C}^{-1/3}(\mathrm{R}, H)$ for some separable Hilbert space H containing H_0 as a dense subspace and such that both H and H_0 are canonically identified with some space of (L-periodic) functions / distributions on R^d in the sense that one has continuous and dense embeddings

$$\mathcal{C}^\infty(R^d) \subset H_0 \subset H \subset \mathcal{D}'(R^d).$$

In particular, elements of Ω can be viewed as distributions on R^{d+1}. We then make the following assumption.

Assumption 7. For every compactly supported mollifier $\varrho \in \mathcal{C}_0^\infty(R^{d+1})$ there exists a sequence $g_\varepsilon \in \mathfrak{R}$ such that the sequence of random models $M_{g_\varepsilon} \mathcal{L}(\varrho^\varepsilon \star \omega)$ converges in probability in $\mathcal{M}$ to a limiting random model $\boldsymbol{\xi}(\omega)$.

Furthermore, the model $\boldsymbol{\xi} = (\Pi, \Gamma)$ is such that, for every $\bar{t} \le t \in [0, 1]$, the random variables $(\Pi_z \tau)(\varphi_z)$ and $\Gamma_{z\bar{z}}$ are $\mathcal{F}_t$-measurable for every $\varphi \in \mathcal{C}_0^\infty$ with support in $R_- \times R^d$. Here, the time coordinates of $z, \bar{z}$ are given by $t, \bar{t}$.

Remark 4.4. In all the examples we have in mind, the sequence g_ε can be chosen in such a way that the limiting random model $\boldsymbol{\xi}$ does not depend on the choice of mollifier ϱ. This property however is not essential for our analysis.

It then follows from a combination of [28, Thm 7.8] and the argument given in [28, Section 9] that, for every periodic initial condition $\Phi_0 \in \mathcal{C}^\eta$ and every nice periodic admissible model, one has a maximal solution $\boldsymbol{\Phi}$ to the equation

$$\boldsymbol{\Phi} = \mathcal{P}\mathbf{1}_+\big(F(\boldsymbol{\Phi}) + N\big) + P\Phi_0, \tag{4.2}$$

where $\mathbf{1}_+$ is the multiplication operator by the indicator function of the set $\{(t, x) : t > 0\}$ and $P\Phi_0$ denotes the solution to the linearised problem, viewed via its truncated Taylor expansion as an element in $\mathcal{D}^{\gamma,\eta}$ with values in the usual Taylor polynomials $\bar{\mathcal{T}}$.

In other words, there exists $T \in (0, 2]$ (depending on the initial condition Φ_0 and on the underlying model Π) such that, for every $t < T$, there exists a unique element $\boldsymbol{\Phi} \in \mathcal{D}^{\gamma,\eta}((-\infty, t] \times R^d, \hat{V}_{<\gamma})$ vanishing on $R_- \times R^d$ and such that the identity (4.2) holds on $(-\infty, t] \times R^d$. As a consequence of the model being nice and the condition $\zeta > 0$, the solution is furthermore such that $\mathcal{R}\boldsymbol{\Phi}$ takes values in $\mathcal{C}((0, T), \mathcal{C}^\eta)$ and, if $T \le 1$, then $\lim_{t \to T} \|(\mathcal{R}\boldsymbol{\Phi})(t, \cdot)\|_{\mathcal{C}^\eta} = \infty$.

Remark 4.5. The upper bound 2 we impose on T is of course completely arbitrary, as long as it is larger than the final time 1 we are really interested in. However, since our models are undefined outside the time interval I_0, we have to put some threshold otherwise the equation makes no longer sense.

In particular, provided that we set $U = \mathcal{C}^\eta$ (or rather the closure of smooth functions in the $\mathcal{C}^\eta$-norm so that U is separable) and we similarly consider solutions starting with an initial condition at time s instead of time 0, this does indeed yield a collection of flow maps $\Phi_{s,t} : \bar{U} \times \mathcal{M} \to \bar{U}$ which is consistent in the sense of (2.1) as a consequence of [28, Prop. 7.11]. Here, we set $\Phi_{s,t}(\Phi_s, \Pi) = \text{💀}$ if and only if the maximal existence time T for the solution to (4.2) with initial condition Φ_s at time s and model $\Pi \in \mathcal{M}$ satisfies $T \le t$.

4.2. *A nice class of nonlinearities*

We first show that, under suitable assumptions on F, the solution to (4.2) also satisfies the remaining assumptions of Section 2. Our main assumption is that solutions to (4.2) driven by nice models are limits of solutions to stochastic PDEs with local nonlinearities. Given a smooth function u on R^{d+1} we write $G(\partial^* u)$ for any expression of the form $G(u, \partial_i u, \partial_{ij}^2 u, \ldots)$ which depends on *finitely many* spatial derivatives of u. In our context, this will only ever include derivatives of order strictly less than the order $2q$ of the operator $\mathcal{L}$. We then make the following assumption on F:

Assumption 8. For every $g \in \mathfrak{R}$ there exist locally Lipschitz continuous functions $F_g^{(1)}$ and $F_g^{(2)}$ such that, for every model of the form $\mathbf{\Pi} = M_g \mathcal{L}(f)$ with $f \in \mathcal{C}^\infty(\mathrm{R}^{d+1})$, the solution $\mathbf{\Phi}$ to (4.2) is such that $u = \mathcal{R} \mathbf{\Phi}$ solves the PDE

$$\partial_t u = \mathcal{L} u + F_g^{(1)}(\partial^* u) + F_g^{(2)}(\partial^* u) f, \qquad u_0 = \Phi_0.$$

We will furthermore impose the following regularity assumption, where the various objects and exponents are as introduced above.

Assumption 9. The map F is twice differentiable as a map between the finite-dimensional vector spaces $\hat{V}_{<\gamma}$ and $\bar{V}_{<\bar{\gamma}}$. Its derivative DF is such that if $\mathbf{\Phi}$ and J belong to $\mathcal{D}^{\gamma,\eta}$ with values in $\hat{V}_{<\gamma}$ and $V_{<\gamma}$ respectively, then the map

$$z \mapsto (DF(\mathbf{\Phi}(z))) J(z)$$

belongs to $\mathcal{D}^{\bar{\gamma},\bar{\eta}}$ and the map $(\mathbf{\Phi}, J) \mapsto DF(\mathbf{\Phi})J$ is strongly locally Lipschitz between these spaces.

See [28, Section 7.3] for the definition of strongly locally Lipschitz in this context. We will also assume henceforth that we are given a continuous action τ of a suitable space of shifts onto the space $\mathcal{M}$ of nice admissible models for $(\mathcal{T}, \mathcal{G})$. This is formulated more precisely in the following structural assumption on our equation, where $\zeta > 0$ denotes the regularity of the sector V as above.

Assumption 10. There exist $p \in (2, \infty)$ as well as a separable Banach space X_0 with dense and continuous inclusions $\mathcal{C}^\infty(\mathrm{R}^d) \subset X_0 \subset H_0$ such that

- There is a continuous action τ of $E = L^p([0,1], X_0)$ on $\mathcal{M}$ such that, for all smooth functions h, ω and every $g \in \mathfrak{R}$, one has the identity $\tau(h, M_g \mathcal{L}(f)) = M_g \mathcal{L}(f + h)$.
- One has $E \subset \mathcal{C}^{\gamma-2}$, extending functions on $[0,1]$ to vanish outside.
- The space X_0 admits $\mathcal{C}^\zeta$ as a multiplier and, for every $\varrho \in \mathcal{C}_0^\infty(\mathrm{R}^d)$ integrating to 1, the operators $h \mapsto \varrho^\varepsilon \star h$ are uniformly bounded as $\varepsilon \to 0$ and converge pointwise to the identity.

Furthermore, there exists a $\mathcal{C}^2$ map $G \colon \mathrm{R} \to \mathrm{R}$ such that, for any $h \in E$ supported in $(0,1) \times \mathrm{R}^d$ and any model $\mathbf{\Pi} \in \mathcal{M}$, the solution $\mathbf{\Phi}^h$ to (4.2) with model $\mathbf{\Pi}^h = \tau(h, \mathbf{\Pi})$ is related to the solution $\hat{\mathbf{\Phi}}^h$ to

$$\hat{\mathbf{\Phi}}^h = \mathcal{P} \mathbf{1}_+ \big(F(\hat{\mathbf{\Phi}}^h) + N \big) + P \star \big(G(\langle \mathbf{1}, \hat{\mathbf{\Phi}}^h \rangle) h \big) + P \Phi_0, \tag{4.3}$$

with model $\mathbf{\Pi}$ by $\mathcal{R}^h \mathbf{\Phi}^h = \mathcal{R} \hat{\mathbf{\Phi}}^h$. Here, $\mathcal{R}^h$ denotes the reconstruction operator for the model $\mathbf{\Pi}^h$, $\mathcal{R}$ denotes the reconstruction operator for $\mathbf{\Pi}$, and $\langle \mathbf{1}, \cdot \rangle \colon \hat{V}_{<\gamma} \to \mathrm{R}$ is the canonical projection onto the component $\mathbf{1}$. Finally, we assume that either G is constant or $\hat{V} = V$.

Remark 4.6. In (4.3), the term $G(\langle \mathbf{1}, \hat{\mathbf{\Phi}}^h \rangle) h$ is interpreted as an element of E, thanks to the assumption on G and the fact that $\mathcal{C}^\zeta$ is a multiplier on X_0. (The fact that $\langle \mathbf{1}, \hat{\mathbf{\Phi}}^h \rangle$ belongs to $\mathcal{C}^\zeta$ follows from [28, Prop. 3.28].) Since we assume that h is supported away from 0, the singularity at the origin appearing in $G(\langle \mathbf{1}, \hat{\mathbf{\Phi}}^h \rangle)$ plays no role. It then follows from the second assumption on X_0 that one has $P \star (G(\langle \mathbf{1}, \hat{\mathbf{\Phi}}^h \rangle) h) \in \mathcal{C}^\gamma$, which we interpret as an element of $\mathcal{D}^\gamma$ via its canonical lift to the Taylor polynomials $\bar{\mathcal{T}}$. In particular, this shows that (4.3) again satisfies the assumptions of [28, Section 7], so that it does indeed admit unique local solutions. The identity (4.3) should then of course be interpreted as holding before the possible explosion time, and that this explosion time is the same for both equations.

The framework considered here covers all current examples (on bounded domains with periodic boundary conditions) of parabolic singular stochastic PDEs for which the theory of regularity structures applies. In the cases of the KPZ equation and the dynamical Φ_d^4 equation with $d \in \{2, 3\}$ as treated in [27,28] one can take $N = \Xi$, where $\Xi \in \mathcal{T}$ is the basis vector representing the driving noise. In the case of the one-dimensional stochastic heat equation with multiplicative noise covered in [34], one can simply take $N = 0$.

The following proposition shows that this framework is compatible with our abstract result.

Proposition 4.7. *In the setting of this section, Assumptions 6–10 guarantee that Assumptions 1–4 of Section 2 are satisfied.*

Proof. Recall that we have set $U = \mathcal{C}^\eta$ and that, for t, Φ_0 and Π such that the explosion time of (4.2) is greater than t, one has $\Phi_t(\Phi_0, \Pi) = (\mathcal{R}\Phi)(t, \cdot)$, where Φ is the solution in $\mathcal{D}^{\gamma,\eta}$ to (4.2) and $\mathcal{R}$ is the reconstruction operator associated to the nice model Π. The only parts of Assumption 1 which do not follow immediately from the well-posedness results of [28, Section 7] are the continuity in time, since one may have $\eta < 0$, and the Fréchet differentiability with respect to the initial condition. Time continuity follows from the fact that we assumed $\zeta > 0$, combined with the condition $P \star \Pi_z N \in \mathcal{C}(\mathbf{R}, \mathcal{C}^\eta)$, which guarantee that $\mathcal{R}\Phi \in \mathcal{C}(\mathbf{R}, \mathcal{C}^\eta)$ for every $\Phi \in \mathcal{D}^{\gamma,\eta}(V)$. To show Fréchet differentiability, we note that as a consequence of Assumption 9, the map

$$(\tau, J) \mapsto \big(F(\tau), \big(DF(\tau)\big)J\big),$$

from $V_{<\gamma} \oplus V_{<\gamma}$ into $\bar{V}_{<\bar{\gamma}} \oplus \bar{V}_{<\bar{\gamma}}$ satisfies the assumptions of [28, Thm 7.8]. As a consequence, we have unique maximal solutions to the fixed point problem

$$\begin{aligned}
\Phi &= \mathcal{P}\mathbf{1}_+\big(F(\Phi) + N\big) + P\Phi_0, \\
J &= \mathcal{P}\mathbf{1}_+\big(DF(\Phi)J\big) + PJ_0.
\end{aligned} \tag{4.4}$$

As a consequence of the implicit function theorem and the fact that $\Phi \mapsto \mathcal{P}\mathbf{1}_+ F(\Phi)$ is a contraction on $\mathcal{D}^{\gamma,\eta}([0, T], V)$ for small enough T, it follows that the solution Φ_t is indeed differentiable in the initial condition and that its derivative in the direction $J_0 \in \mathcal{C}^\eta$ is given by J. A standard patching argument shows that this is true not only for small t, but all the way up to the explosion time of Φ.

We now turn to Assumption 2. If $\Phi_t(u, \xi) = \text{☠}$, then we set $r_t(u, \xi) = +\infty$, otherwise, we set

$$r_t(u, \xi) = \|\!|\Phi\|\!|_{\gamma, \eta; [0, t]},$$

where Φ is the solution to (4.2) with initial condition u and underlying model ξ, and the norm $\|\!|\cdot\|\!|_{\gamma, \eta; [0, t]}$ is as in [28, Def. 6.2] with $\mathfrak{K} = [0, t] \times \mathbf{R}^d$. (The fact that this set is not compact is irrelevant since we only consider periodic functions.) The $\mathcal{F}_t$-measurability of r_t is then an immediate consequence of the non-anticipativity of P, combined with the second part of Assumption 7, which guarantee that the random variable $\Phi(s, x)$ is $\mathcal{F}_s$-measurable for all $s \in [0, 1]$ and all $x \in \mathbf{R}^d$. The map is also increasing by definition and it is continuous in t, except at the explosion time when it has to diverge to $+\infty$ as a consequence of the fact that the $\mathcal{C}^\eta$-norm of the solution is dominated by r_t and has to blow up at the explosion time. The third required property of r is again satisfied by definition, while the last property follows from the fact that the fixed point map defining our solutions is locally Lipschitz continuous with values in $\mathcal{D}^{\gamma,\eta}$ as a function of both the initial condition and the underlying model.

To check that Assumption 3 is verified, we exploit the identity $\tau(h, M_g \mathcal{L}(f)) = M_g \mathcal{L}(f + h)$ given by Assumption 10. Fix a mollifier ϱ, then we have for every $\varepsilon > 0$ and every $\omega \in \Omega$ the identity

$$\tau\big(h_\varepsilon, M_{g_\varepsilon} \mathcal{L}(\omega_\varepsilon)\big) = M_{g_\varepsilon} \mathcal{L}(\omega_\varepsilon + h_\varepsilon),$$

where we set $h_\varepsilon = h \star \varrho^\varepsilon$ and similarly for ω_ε. For every bounded continuous $X \colon \mathcal{M} \to \mathbf{R}$, we therefore have, by Girsanov's theorem,

$$\mathbf{E}X\big(\tau\big(h_\varepsilon, M_{g_\varepsilon} \mathcal{L}(\omega_\varepsilon)\big)\big) = \mathbf{E}X\big(M_{g_\varepsilon} \mathcal{L}(\omega_\varepsilon)\big) \exp\big(h_\varepsilon^*(\omega_\varepsilon) - \|h_\varepsilon\|_{\mathcal{H}}^2/2\big).$$

As a consequence of the third property of X_0, one has $h_\varepsilon \to h$ in E and therefore also in $\mathcal{H}$ since $X_0 \subset H_0$. Furthermore $h_\varepsilon^*(\omega_\varepsilon) = \tilde{h}_\varepsilon^*(\omega)$ with $\tilde{h}_\varepsilon = h \star \varrho^\varepsilon \star \varrho^\varepsilon$, which also converges to h in $\mathcal{H}$, so that $h_\varepsilon^*(\omega_\varepsilon) \to h^*(\omega)$ in probability. Taking limits on both sides and exploiting the fact that $M_{g_\varepsilon} \mathcal{L}(\omega_\varepsilon) \to \xi$ in probability in $\mathcal{M}$ and τ is jointly continuous, the claim (2.2) follows. The second claim of that assumption immediately follows from Assumption 8, in particular from the locality of the maps $F_g^{(i)}$.

It remains to show that Assumption 4 holds, namely that the solution with model $\tau(h, \Pi)$ is Fréchet differentiable in the direction of $h \in E$ at $h = 0$. Thanks to Assumption 10 however, we can fix the model Π and it suffices to

show that the solution $\hat{\boldsymbol{\Phi}}^h$ to (4.3) is differentiable at $h = 0$ for every $t \in (0, 1]$ and every $(\Phi_0, \boldsymbol{\Pi}) \in \mathcal{N}_t$. This however follows again immediately from the implicit functions theorem. □

Of course we cannot expect that Assumption 5 is satisfied without further condition since we could for example encode a situation where $F_g^{(2)} = 0$, so that our solutions are not random at all. The next subsection introduces a condition under which this last assumption is also satisfied, thus yielding the strong Feller property for the corresponding Markov process.

4.3. *A non-degeneracy assumption*

In order to have a chance for Assumption 5 to hold, we need to know that the Cameron–Martin space $\mathcal{H}$ is sufficiently large to contain enough shifts to be able to compensate for any shift in the initial condition. One possible assumption guaranteeing that this is indeed the case is the following.

Assumption 11. One has $\mathcal{C}^\zeta \subset X_0 \subset \mathcal{C}^\eta$, and the map G appearing in Assumption 10 satisfies $G(u) > 0$ for every $u \in \hat{V}_{<\gamma}$.

This is then the final ingredient in the proof of the strong Feller property.

Theorem 4.8. *In the setting of this section, Assumptions 6–11 guarantee that Assumptions 1–5 of Section 2 are satisfied. In particular, the Markov operator $\mathfrak{P}$ satisfies the strong Feller property.*

Proof. We first note that, similarly to the proof of the first part of Proposition 4.7, our assumptions guarantee that, for $(\Phi_0, \boldsymbol{\Pi}) \in \mathcal{N}_t$ and $0 \le s \le t$, the map $\Phi_{s,t}$ is Fréchet differentiable with respect to its first argument at the point $\Phi_s(\Phi_0, \boldsymbol{\Pi})$. Denote its derivative in the direction u by $J_{s,t}u$. Since the second component of the solution to the fixed point problem (4.4) belongs to $\mathcal{D}^{\gamma,\eta}(V)$ with $V \subset \bar{\mathcal{T}} \oplus \mathcal{T}_{\ge\zeta}$, it follows that for $t > s$, $J_{s,t}$ is a bounded linear operator from $\mathcal{C}^\eta$ to $\mathcal{C}^\zeta$. (Since it is the identity for $s = t$, its norm blows up as $t \to s$ and one can quantify this blow-up, but its details are not important to us.)

As a consequence of the variation of constants formula given in Corollary A.3 of the Appendix, the derivative of Φ_t with respect to $h \in E$ is given by

$$\mathcal{D}\Phi_t(\Phi_0, \boldsymbol{\Pi})h = \int_0^t J_{s,t} G(\langle \mathbf{1}, \boldsymbol{\Phi}(s, \cdot) \rangle) h(s) \, ds, \tag{4.5}$$

as long as $(\Phi_0, \boldsymbol{\Pi}) \in \mathcal{N}_t$. Here, $\boldsymbol{\Phi}$ denotes the solution to (4.2), while $\Phi_t = (\mathcal{R}\boldsymbol{\Phi})(t, \cdot)$.

This suggests the following definition. First, fix an arbitrary bounded function $\chi : [0, 1] \to \mathbb{R}$ such that $\chi(s) = 0$ for $s \le 1/4$ and such that $\int_0^1 \chi(s) \, ds = 1$. For $s \le t$ and using the shorthand $J_r = J_{0,r}$ we then define $A_t^{(s)} : \mathcal{N}_s \to L(U, E_s)$ by

$$\left(A_t^{(s)}(\Phi_0, \boldsymbol{\Pi}) v \right)(r) = \frac{1}{t} \chi(r/t) G^{-1}(\langle \mathbf{1}, \boldsymbol{\Phi}(r, \cdot) \rangle) J_r v, \quad r \in [0, s]. \tag{4.6}$$

Here, the Jacobian J_s does of course depend on $(\Phi_0, \boldsymbol{\Pi})$, but we have suppressed this dependency in our notations. We also used the notation G^{-1} for the map $u \mapsto 1/G(u)$. Note that the right-hand side depends on s only through the range of the variable r. In particular, one has the identity

$$\left(A_t^{(s)}(\Phi_0, \boldsymbol{\Pi}) v \right)(r) = \left(A_t^{(\bar{s})}(\Phi_0, \boldsymbol{\Pi}) v \right)(r),$$

provided that $(\Phi_0, \boldsymbol{\Pi}) \in \mathcal{N}_s \cap \mathcal{N}_{\bar{s}}$. We claim that this definition satisfies all the properties of Assumption 5. Indeed, $\mathcal{F}_s$-measurability follows again from the fact that our solutions $\boldsymbol{\Phi}(s, \cdot)$ are $\mathcal{F}_s$-measurable. The identity (2.4) follows immediately from inserting (4.6) into (4.5) and using the fact that $J_{r,t} J_r = J_t$ for every $r \in [0, t]$ and every $(\Phi_0, \boldsymbol{\Pi}) \in \mathcal{N}_t$. The fact that $A_t^{(s)}(\Phi_0, \boldsymbol{\Pi}) v$ belongs to E for every $v \in U$ follows from the fact that J_r maps U into $\mathcal{C}^\zeta$ as already mentioned above, combined with the fact that χ is supported away from 0 and the map $x \mapsto G(\langle \mathbf{1}, \boldsymbol{\Phi}(r, x) \rangle)$ belongs to $\mathcal{C}^\zeta$ by Assumption 10, so that the same is true for $x \mapsto 1/G(\langle \mathbf{1}, \boldsymbol{\Phi}(r, x) \rangle)$ by the positivity of G. It follows that one actually has the stronger statement that $A_t^{(s)}(\Phi_0, \boldsymbol{\Pi}) v \in L^\infty([0, s], \mathcal{C}^\zeta)$. □

5. Application to concrete examples

We now show that the abstract results obtained in this article can be applied to the concrete examples mentioned in the Introduction. We will treat the case of the dynamical Φ^4_3 model in some detail and then only focus on the differences with the other examples.

5.1. *Dynamical Φ^4_3 model*

Recall that the dynamical Φ^4_3 model is the stochastic PDE formally given by

$$\partial_t \Phi = \Delta \Phi - \Phi^3 + \xi, \tag{5.1}$$

where ξ denotes space-time white noise, see [28]. The space variable is assumed to take values in some bounded three-dimensional torus. This is of course only a formal notation, one way of constructing Φ is as limits of solutions to the equation

$$\partial_t \Phi_\varepsilon = \Delta \Phi_\varepsilon + C_\varepsilon \Phi_\varepsilon - \Phi^3_\varepsilon + \xi_\varepsilon,$$

where $\xi_\varepsilon = \xi \star \varrho^\varepsilon$ for some compactly supported space-time mollifier ϱ and for a suitable choice of (diverging) constants C_ε. Using regularity structures, it was shown in [28] that such limits exist and are "unique" in the sense that modulo a suitable choice of C_ε they are independent of the choice of mollifier ϱ.

Theorem 5.1. *The Markov semigroup generated by the solutions to the dynamical Φ^4_3 model on $\mathcal{C}^{-5/8}$ satisfies the strong Feller property.*

Proof. In view of Theorems 4.8 and 3.2, it suffices to show that Assumptions 6–11 are satisfied. Recall the construction [28, Sections 9–10] of the regularity structure associated to (5.1). Consider first the space $\bar{\mathcal{T}}$ of polynomials in four indeterminates $X_0, \ldots, X_3$, representing the usual Taylor polynomials. For any multiindex $k = (k_0, \ldots, k_3)$, we write X^k for the monomial $X_0^{k_0} \cdots X_3^{k_3}$. We also write $\mathbf{1}$ instead of X^0.

We then build a collection $\mathcal{U}$ of formal expressions as the smallest collection containing all the X^k and $\mathcal{I}(\Xi)$, and such that

$$\tau_1, \tau_2, \tau_3 \in \mathcal{U} \quad \Rightarrow \quad \mathcal{I}(\tau_1 \tau_2 \tau_3) \in \mathcal{U}, \tag{5.2}$$

where it is understood that $\mathcal{I}(X^k) = 0$ for every multiindex k. Here, it is understood that we make the necessary identifications so that the product appearing on the right-hand side is commutative and associative, with unit $\mathbf{1}$. We then set

$$\mathcal{W} = \{\Xi\} \cup \{\tau_1 \tau_2 \tau_3 : \tau_i \in \mathcal{U}\}, \tag{5.3}$$

and we define our space $\mathcal{T}$ as the set of all linear combinations of elements in $\mathcal{W}$. (Note that since $\mathbf{1} \in \mathcal{U}$, one does in particular have $\mathcal{U} \subset \mathcal{W}$.) Each formal expression in $\mathcal{W}$ is assigned a degree by setting $\deg \Xi = -\frac{5}{2} - \kappa$ for some (sufficiently small) $\kappa > 0$, $\deg \mathbf{1} = 0$, $\deg X_0 = 2$, $\deg X_i = 1$ for $i \in \{1, \ldots, 3\}$, and then extending this to all of $\mathcal{W}$ by postulating that

$$\deg(\tau \bar{\tau}) = \deg \tau + \deg \bar{\tau},$$
$$\deg \mathcal{I}(\tau) = \deg \tau + 2.$$

Naturally, $\mathcal{T}_\alpha$ is then the subspace of $\mathcal{T}$ spanned by those elements in $\mathcal{W}$ that are of degree α. With this definition, provided that κ is small enough ($\kappa < \frac{1}{2}$ suffices, but one needs κ to be even smaller later), each space $\mathcal{T}_{<\gamma}$ for $\gamma \in \mathbf{R}$ is finite-dimensional. Using the same graphical notations as in [28,37] (bullets for instances of Ξ, lines for instances of $\mathcal{I}$, so that for example $\overset{\curvearrowright}{\mathcal{V}} = \mathcal{I}(\Xi)^2 \mathcal{I}(\mathcal{I}(\Xi)^2))$, $\mathcal{T}_{<0}$ is for example spanned by

$$\{\Xi, \uparrow, \vee, \psi, \overset{\curvearrowright}{\psi}, \overset{\curvearrowright}{\mathcal{V}}, \overset{\curvearrowright}{\psi}\}. \tag{5.4}$$

The structure group $\mathcal{G}$ built in [28, Def. 8.20] then consists of all linear maps $\Gamma\colon \mathcal{T} \to \mathcal{T}$ such that

- One has $\Gamma\Xi = \Xi$ and there exist $x_i \in \mathbb{R}$ such that $\Gamma X_i = X_i - x_i\mathbf{1}$.
- For every $\tau \in \mathcal{T}_\alpha$ one has $\Gamma\tau - \tau \in \mathcal{T}_{<\alpha}$.
- For every $\tau, \bar\tau$ in $\mathcal{W}$ such that $\tau\bar\tau \in \mathcal{W}$, one has $\Gamma(\tau\bar\tau) = (\Gamma\tau)(\Gamma\bar\tau)$.
- For every $\tau \in \mathcal{W}$ such that $\mathcal{I}(\tau) \in \mathcal{W}$, one has $\Gamma\mathcal{I}(\tau) - \mathcal{I}(\Gamma\tau) \in \bar{\mathcal{T}}$, where $\mathcal{I}$ is extended from $\mathcal{W}$ to $\mathcal{T}$ by linearity.

These properties are consistent (in the sense that the last two properties never refer to any formal expression not contained in $\mathcal{W}$) and they do form a group.

We then define $V \subset \mathcal{T}$ as the subspace spanned by the X^k, as well as all expressions of the form $\mathcal{I}(\tau)$ for some τ, *except* for the expression $\mathcal{I}(\Xi)$. We do however set $N = \Xi$, so that $\hat{V}$ contains all $\mathcal{I}(\tau)$, and we simply set $\bar{V} = \mathcal{T}$. If κ is sufficiently small, one can verify that these choices of V and $\bar{V}$ satisfy the assumptions of our setting with exponents

$$\zeta = \frac{1}{2} - 3\kappa, \qquad \bar{\zeta} = -\frac{3}{2} - 3\kappa.$$

Possible choices for γ, $\bar\gamma$, η and $\bar\eta$ (again provided that κ is small enough) are given by

$$\gamma = \frac{7}{4}, \qquad \bar\gamma = \frac{1}{4}, \qquad \eta = -\frac{5}{8}, \qquad \bar\eta = -\frac{15}{8}.$$

(The point for the η exponents is that $\eta \in (-2/3, -1/2)$ and $\bar\eta = 3\eta$ which are imposed by the irregularity of the noise and the form of the nonlinearity.) These choices plainly satisfy the various inequalities and properties we imposed in Section 4 with $q = 1$. The nonlinearity $F\colon V_{<\gamma} \to \bar{V}_{<\bar\gamma}$ is simply given by $F(\boldsymbol{\Phi}) = -\mathcal{Q}_{<\bar\gamma}\boldsymbol{\Phi}^3$, where $\mathcal{Q}_{<\bar\gamma}$ is the projection onto terms of degree less than $\bar\gamma$. The fact that F satisfies the strong local Lipschitz property was shown in [28, Lem. 9.7]. In order to verify Assumption 6, it thus only remains to exhibit $\mathcal{L}$ and $\mathfrak{R}$. The canonical lift $\mathcal{L}$ is the same as in [28, Section 8.2]. The group $\mathfrak{R}$ in this case is simply given by $\mathbb{R}^2$ with addition, acting on the space of admissible models as in [28, Section 9.2]. It acts on $\mathcal{T}$ in a natural way by essentially performing the substitutions $\vee \mapsto C_1\mathbf{1}$ and $\psi \mapsto C_2\mathbf{1}$ in the sense that an arbitrary symbol is mapped to the sum over all possible ways of contracting (possibly multiple) occurrences of $\vee$ and ψ, multiplied by the corresponding powers of C_i. For example

$$M_g \psi = \psi + C_1 \mathsf{Y} + 3C_1 \mathsf{V} + 3C_1^2 \mathsf{\backslash} + 3C_2 \mathsf{\uparrow}. \tag{5.5}$$

This action can be lifted to an action on the space $\mathcal{M}$ of admissible models via the construction of [28, Section 8]. (See also [2] for a more systematic justification of this.)

The convergence of the mollified and renormalised models to a unique limiting model $\boldsymbol{\xi}$ is the content of [28, Section 10.5]. The fact that the limiting model is adapted to the filtration $\mathcal{F}_t$ in the required way is shown in the same way as in [34], so that Assumption 7 is also satisfied. Assumption 8 is the content of [28, Prop. 9.10], which shows in particular that one has

$$F_g^{(1)}(\partial^* u) = c_g u - u^3, \qquad F_g^{(2)}(\partial^* u) = 1,$$

for a constant c_g given by a suitable linear combination of the two components of g. The regularity assumption in Assumption 9 follows immediately from the fact that F is polynomial. The fact that the derivative map $(\boldsymbol{\Phi}, J) \mapsto -3\boldsymbol{\Phi}^2 J$ is strongly locally Lipschitz between the spaces $\mathcal{D}^{\gamma,\eta}$ and $\mathcal{D}^{\bar\gamma,\bar\eta}$ follows from [28, Prop. 6.12].

We now turn to Assumption 10, which is the one that requires the largest amount of work. Regarding a suitable space of shifts which generate a continuous action on our space $\mathcal{M}$ of nice models, we claim that the space $\mathcal{C}^{-\kappa}(\mathbb{R} \times \mathbb{R}^d)$ has the required properties provided that κ is sufficiently small, so that we can the choose for X_0 space of the form $\mathcal{C}^\zeta$ for ζ so that $\zeta \geq -\kappa$ and $X_0 \subset H_0$. It is relatively easy to show this "by hand", but we want to have a more systematic proof which also carries over to the other examples. We introduce an auxiliary regularity structure $(\hat{\mathcal{T}}, \hat{\mathcal{G}})$ defined as follows, similarly to what was done in [7, Section 3.2]. The space $\hat{\mathcal{T}}$ is constructed in exactly the same way as $\mathcal{T}$, with the exception that in the first step of the construction we start with a collection $\hat{\mathcal{U}}$ that contains X^k and $\mathcal{I}(\Xi)$, as well as an additional expression $\mathcal{I}(\hat{\Xi})$. We then also add $\hat{\Xi}$ to the collection $\hat{\mathcal{W}}$ otherwise defined from $\hat{\mathcal{U}}$ as

above, and taken as a basis for the space $\hat{\mathcal{T}}$. Graphically, if we denote $\hat{\Xi}$ by a circle, the symbols appearing in $\hat{\mathcal{T}}$ are the same as those appearing in $\mathcal{T}$, but with any occurrence of a bullet possibly replaced by a circle, so for example $\mathcal{I}(\Xi)\mathcal{I}(\hat{\Xi}) = \vee$.

Regarding the grading of $\hat{\mathcal{T}}$, we set $\deg \hat{\Xi} = -\kappa$, the degrees of the remaining basis vectors being obtained by using the same rules as above. The structure group $\hat{\mathcal{G}}$ is also defined as above by additionally imposing that $\Gamma \hat{\Xi} = 0$ for every $\Gamma \in \hat{\mathcal{G}}$. The same group $\mathfrak{R}$ acts naturally on $\hat{\mathcal{T}}$ in exactly the same way as above. Denoting this action by $\hat{M}$, one has for example, for $g = (C_1, C_2)$

$$\hat{M}_g \Psi = \Psi + C_1 \Psi + C_1 \Psi + C_1^2 \rangle + C_2.$$

Compare this to (5.5) where the constants are slightly different since there are more inequivalent occurrences of $\vee$ and Ψ appearing in that symbol. Thanks to [2] this can again be lifted naturally to an action on the corresponding spaces of admissible models, so that we also have a space of "nice models" $\hat{\mathcal{M}}$ for this larger regularity structures. Note also that one has $(\mathcal{T}, \mathcal{G}) \subset (\hat{\mathcal{T}}, \hat{\mathcal{G}})$ in the sense of inclusions of regularity structures as in [28, Section 2.1].

A crucial remark is that if κ is sufficiently small (in our case one needs $\kappa < 1$ which guarantees that $\deg \vee > 0$), all of the elements of $\hat{\mathcal{W}} \setminus \mathcal{W}$ are of strictly positive degree, except for $\hat{\Xi}$ itself. As a consequence, by repeatedly applying [28, Prop. 3.31] and [28, Thm 5.14] and writing $\hat{\mathcal{M}}$ for the space of admissible models for $(\hat{\mathcal{T}}, \hat{\mathcal{G}})$, there exists a *unique* locally Lipschitz continuous map $\mathcal{Y} \colon \mathcal{C}^{-\kappa} \times \mathcal{M} \to \hat{\mathcal{M}}$ such that

- For every $h \in \mathcal{C}^{-\kappa}$ and $\mathbf{\Pi} \in \mathcal{M}$, the model $\hat{\mathbf{\Pi}}^h = \mathcal{Y}(h, \mathbf{\Pi}) \in \hat{\mathcal{M}}$ agrees with $\mathbf{\Pi}$ on $\mathcal{T} \subset \hat{\mathcal{T}}$ (or in other words $\hat{\mathbf{\Pi}}^h$ extends $\mathbf{\Pi}$) as in [28, Def. 2.22].
- Writing $(\hat{\Pi}, \hat{\Gamma}) = \mathcal{Y}(h, \mathbf{\Pi})$, one has $\hat{\Pi}_z \hat{\Xi} = h$ for every z.

It is easy to show that the action of $\mathfrak{R}$ commutes with $\mathcal{Y}$ in the sense that, for every $g \in \mathfrak{R}$, every $\mathbf{\Pi} \in \mathcal{M}$ and every $h \in \mathcal{C}^{-\kappa}$, one has $\hat{M}_g \mathcal{Y}(h, \mathbf{\Pi}) = \mathcal{Y}(h, M_g \mathbf{\Pi})$. Indeed, both $\hat{M}_g \mathcal{Y}(h, \mathbf{\Pi})$ and $\mathcal{Y}(h, M_g \mathbf{\Pi})$ agree on $\mathcal{T} \oplus \langle \hat{\Xi} \rangle$ and any admissible model is uniquely determined by this, as mentioned above.

It remains now to build a (locally Lipschitz continuous) map $\mathcal{Z} \colon \hat{\mathcal{M}} \to \mathcal{M}$, also commuting with the action of $\mathfrak{R}$, so that we can then define $\tau = \mathcal{Z} \circ \mathcal{Y}$. For this, given a model $\hat{\mathbf{\Pi}} \in \hat{\mathcal{M}}$, we introduce the following notion of a "$\hat{\mathbf{\Pi}}$-polynomial":

Definition 5.2. A $\hat{\mathbf{\Pi}}$-polynomial f is a map $f \colon \mathbb{R}^{d+1} \to \hat{\mathcal{T}}$ such that $f(z) = \Gamma_{z\bar{z}} f(\bar{z})$ for any $z, \bar{z}$. We say that f is of degree $\deg f = \gamma$ if the component of $f(z)$ in $\mathcal{T}_{>\gamma}$ vanishes and its component in $\mathcal{T}_\gamma$ is non-zero for some (and therefore all) z. In particular, one has $f \in \mathcal{D}^{\bar{\gamma}}$ for every $\bar{\gamma} > \deg f$.

It is immediate that the product of two $\hat{\mathbf{\Pi}}$-polynomials is again a $\hat{\mathbf{\Pi}}$-polynomial. Given an admissible model $\hat{\mathbf{\Pi}} = (\Pi, \Gamma) \in \hat{\mathcal{M}}$, we now define a collection of $\hat{\mathbf{\Pi}}$-polynomials $\{f_z^\tau\}$ for $\tau \in \mathcal{W}$ and $z \in \mathbb{R}^{d+1}$ recursively as follows. Recall first that our model $\mathbf{\Pi}$ defines operators $\mathcal{J}(z) \colon \hat{\mathcal{T}} \to \hat{\mathcal{T}}$ by [28, Eq. 5.11]. We also define an alternative notion $\overline{\deg}$ of degree on $\hat{\mathcal{W}}$ by setting $\overline{\deg} \tau = \deg \tau$ for $\tau \in \mathcal{W}$, but then setting $\overline{\deg} \hat{\Xi} = \deg \Xi$ and defining it on the rest of $\hat{\mathcal{W}}$ by using the same rules as for $\deg$. This allows us to define operators $\bar{\mathcal{J}}(z) \colon \hat{\mathcal{T}} \to \hat{\mathcal{T}}$ by setting

$$\bar{\mathcal{J}}(z)\tau = \mathcal{Q}_{<\overline{\deg}\tau} \mathcal{J}(z). \tag{5.6}$$

With these definitions at hand, we set

$$f_z^{X^k}(\bar{z}) = \Gamma_{\bar{z}z} X^k, \qquad f_z^{\Xi}(\bar{z}) = \Xi + \hat{\Xi}.$$

Then, we set recursively

$$f_z^{\tau\bar{\tau}}(\bar{z}) = f_z^\tau(\bar{z}) f_z^{\bar{\tau}}(\bar{z}), \qquad f_z^{\mathcal{I}(\tau)}(\bar{z}) = \left(\mathcal{I} + \mathcal{J}(\bar{z}) - \Gamma_{\bar{z}z}\bar{\mathcal{J}}(z)\Gamma_{z\bar{z}}\right) f_z^\tau(\bar{z}).$$

Using [28, Lem. 5.16] it is a simple exercise to verify that if f_z^τ is a $\hat{\mathbf{\Pi}}$-polynomial, then $f_z^{\mathcal{I}(\tau)}$ as defined above is indeed again a $\hat{\mathbf{\Pi}}$-polynomial. Furthermore, it follows by induction that, for every z and every τ, $f_z^\tau(\bar{z})$ is a linear

combination of terms of $\overline{\deg}$-degree equal to $\deg \tau$. Combining this with the definition of $\bar{\mathcal{J}}(z)$ and the fact that $\deg \hat{\Xi} > \deg \Xi$ allows to show inductively that, for every $\tau \in \mathcal{W}$, one has

$$\hat{\mathcal{Q}}_{<\deg \tau} f_z^\tau(z) = 0, \quad \forall z \in \mathrm{R}^d, \tag{5.7}$$

where $\hat{\mathcal{Q}}_{<\gamma}$ is the projection onto $\hat{\mathcal{T}}_{<\gamma}$.

Define now operators $\Lambda_{z\bar{z}} \colon \mathcal{T} \to \mathcal{T}$ by setting

$$\Lambda_{z\bar{z}} \Xi = \Xi, \qquad \Lambda_{z\bar{z}} X^k = \Gamma_{z\bar{z}} X^k,$$

and then recursively by

$$\Lambda_{z\bar{z}}(\tau \bar{\tau}) = (\Lambda_{z\bar{z}}\tau)(\Lambda_{z\bar{z}}\bar{\tau}),$$

as well as

$$\Lambda_{z\bar{z}} \mathcal{I}\tau = \mathcal{I}\Lambda_{z\bar{z}}\tau + \big(\bar{\mathcal{J}}(z)\Gamma_{z\bar{z}} - \Gamma_{z\bar{z}}\bar{\mathcal{J}}(\bar{z})\big) f_{\bar{z}}^\tau(\bar{z}).$$

We also extend the definition of f_z^τ to all of $\tau \in \mathcal{T}$ by linearity. It is then a straightforward exercise to verify by recursion that one has the identity

$$f_z^{\Lambda_{z\bar{z}}\tau} = f_{\bar{z}}^\tau. \tag{5.8}$$

The map $\mathcal{Z}$ is then defined as follows. Given a model $\hat{\mathbf{\Pi}} = (\Pi, \Gamma) \in \hat{\mathcal{M}}$, we define a new model $\mathcal{Z}\hat{\mathbf{\Pi}} = (\tilde{\Pi}, \tilde{\Gamma})$ by setting

$$\tilde{\Pi}_z \tau = \mathcal{R} f_z^\tau = \Pi_z f_z^\tau(z), \qquad \tilde{\Gamma}_{z\bar{z}} = \Lambda_{z\bar{z}}, \tag{5.9}$$

where $\mathcal{R}$ is the reconstruction operator associated to $\hat{\mathbf{\Pi}}$. Let us verify that this is indeed an admissible model. It follows from (5.8) and these definitions that one has the algebraic identity $\tilde{\Pi}_z \tilde{\Gamma}_{z\bar{z}} = \tilde{\Pi}_{\bar{z}}$ as required. Since the map $\tau \mapsto f_x^\tau$ is easily seen to be injective, (5.8) also implies that $\Lambda_{z\bar{z}}\Lambda_{\bar{z}\underline{z}} = \Lambda_{z\underline{z}}$, so that it remains to verify that the required analytical bounds hold. The bounds on $\tilde{\Pi}_z \tau$ follow at once from the corresponding bounds on $\mathbf{\Pi}$ and the fact that $\tilde{\Pi}_z \tau = \Pi_z f_z^\tau(z)$, combined with (5.7). The bounds on $\Lambda_{z\bar{z}}$ on the other hand follow inductively from the corresponding bounds on $\Gamma_{z\bar{z}}$, combined with [28, Lem. 5.21] and the fact that $f_z^\tau(z)$ is a linear combination of terms with $\overline{\deg}$-degree equal to $\deg \tau$. The fact that the new model is again admissible (in the sense that it realises K for $\mathcal{I}$) follows at once from the fact that since

$$\tilde{\Pi}_z \mathcal{I}(\tau) = \Pi_z \big(\mathcal{I} + \mathcal{J}(z) - \bar{\mathcal{J}}(z)\big) f_z^\tau(z),$$

and since

$$\Pi_z \big(\mathcal{I} + \mathcal{J}(z)\big) \bar{\tau} = K \star \Pi_z \bar{\tau}$$

for every z and every $\bar{\tau}$ in the domain of $\mathcal{I}$, it follows from the definitions (5.6) and (5.9) of $\tilde{\Pi}_z$ and $\bar{\mathcal{J}}(z)$ that $\tilde{\Pi}_z \mathcal{I}(\tau) - K \star \Pi_z \tau$ is a polynomial of degree τ. Since furthermore it satisfies the desired analytical bounds, it does indeed realise K for $\mathcal{I}$.

Setting $\tau(h, \mathbf{\Pi}) = \mathcal{Z}(\mathcal{Y}(h, \mathbf{\Pi}))$, we still need to show that τ is an action and that it commutes with the action of $\mathfrak{R}$. Regarding the canonical lift $\mathcal{L}$, it is immediate from the definitions that one has

$$\tau\big(h, \mathcal{L}(\xi)\big) = \mathcal{L}(\xi + h), \tag{5.10}$$

for every smooth ξ and h. If we can show that furthermore

$$\tau(h, M_g \mathbf{\Pi}) = M_g \tau(h, \mathbf{\Pi}), \tag{5.11}$$

for every smooth h, every smooth $\mathbf{\Pi} \in \mathcal{M}$ and every $g \in \mathfrak{R}$, then both the required identity for τ and the fact that it is an action follow by continuity. Since, as before, admissible models on $\mathcal{T}$ are uniquely determined by their action on those elements $\tau \in \mathcal{W}$ with $\deg \tau \leq 0$, the identity (5.11) only needs to be verified on the sector spanned by the elements in (5.4). This is non-trivial only for the last three elements, but can easily be verified from the explicit formulae of both the action of $\mathfrak{R}$ and that of τ.

Let us also sketch a more systematic way of verifying (5.11). For this, we note that the linear map $Z \colon \hat{\mathcal{T}} \to \hat{\mathcal{T}}$ obtained by substituting Ξ by $\Xi + \hat{\Xi}$ in every formal expression of $\hat{\mathcal{W}}$ belongs to the "renormalisation group" of the regularity structure $(\hat{\mathcal{T}}, \hat{\mathcal{G}})$ in the sense of [28, Def. 8.41] (see [35, Appendix B] for a proof of the fact that the second identity of [28, Eq. 8.39] actually follows from the first one), thus yielding an action $\hat{Z}$ on $\hat{\mathcal{M}}$. Furthermore, $\mathcal{Z}$ is precisely given by $\hat{\mathcal{Z}}$, followed by the canonical projection from $\hat{\mathcal{M}}$ to $\mathcal{M}$. The claim (5.11) is then an immediate consequence of the fact that the linear maps Z and M_g commute on $\hat{\mathcal{T}}$. This is indeed the case, with both $Z M_g$ and $M_g Z$ given by the map that maps a symbol τ to the sum over all ways of substituting instances of Ξ by $\hat{\Xi}$, as well as contracting occurrences of $\vee$ and $\dot{\vee}$ and replacing them by the corresponding renormalisation constants. Since $L^p([0, 1], \mathcal{C}) \subset \mathcal{C}^{-\kappa}$ for sufficiently large p, we can for example choose $X_0 = \mathcal{C}$, the space of L-periodic continuous functions. Furthermore, as a consequence of (5.10) and (5.11), it is immediate that (4.3) holds with $G = 1$, so that Assumption 10 holds. Finally, Assumption 11 is trivially verified since $\mathcal{C}^\zeta \subset \mathcal{C} \subset L^2 \cap \mathcal{C}^\eta$, and $G = 1$. □

5.2. *Multi-component KPZ equation*

Consider the system of coupled KPZ equations

$$\partial_t h^i = \partial_x^2 h^i + S^i_{jk} \partial_x h^j \partial_x h^k + \xi_i \tag{5.12}$$

(summation over repeated indices is implied) where the ξ_i are independent space-time white noises on $\mathbf{R} \times \mathbf{T}^1$ and the S^i_{jk} are constant coefficients. Such systems of equations arise naturally when considering the large-scale limits of systems with more than one locally conserved quantity, see [21,50]. We will sometimes assume that the coefficients S satisfy the symmetry conditions

$$S^i_{jk} = S^i_{kj}, \qquad S^i_{jk} = S^j_{ki}. \tag{5.13}$$

As shown formally in [21,50] and proven rigorously in [19], this guarantees that the invariant measure for (5.12) (modulo height shifts) is Gaussian and simply consist of independent Brownian bridges. As usual, (5.12) should be interpreted as the limit, as $\varepsilon \to 0$, of solutions to

$$\partial_t h^i = \partial_x^2 h^i + S^i_{jk} \partial_x h^j \partial_x h^k - C_i^{(\varepsilon)} + \xi_i^{(\varepsilon)}, \tag{5.14}$$

where the constants $C_i^{(\varepsilon)}$ are chosen in such a way that the bilinear term is given by a Wick product with respect to the Gaussian structure determined by the linearised equation. In the symmetric case (5.13), one chooses

$$C_i^{(\varepsilon)} = \frac{c}{\varepsilon} \sum_k S^i_{kk}, \tag{5.15}$$

where c is as in [36, Eq. 1.4]. The symmetry condition (5.13) is such that the additional logarithmically divergent renormalisation constants appearing for example in the analysis of [27] cancel out, otherwise one may have to add a logarithmically diverging term to (5.15). As shown in [18,27] in the one-component case, but the multi-component case does not add any difficulty whatsoever, the limit of (5.14) exists, up to the first time at which the $\mathcal{C}^\alpha$-norm of the solution blows up for some (and therefore all) $\alpha > 0$. It is now rather straightforward to verify that our assumptions are verified.

Theorem 5.3. *The solutions to the coupled KPZ equations* (5.14) *generate a Markov semigroup that satisfies the strong Feller property.*

Proof. In this case, we simply set $N = 0$, so that $V = \hat{V}$. The construction of the corresponding regularity structure works in a way analogous to what we mentioned above for the dynamical Φ_3^4 model and is exposed for example in [18]. The values of the exponents ζ and $\bar{\zeta}$ in this setting (assuming that the symbols Ξ_i representing the noises ξ_i have degree $-\frac{3}{2} - \kappa$) are given by

$$\zeta = \frac{1}{2} - \kappa, \qquad \bar{\zeta} = -1 - 2\kappa,$$

provided that κ is sufficiently small. Possible choices for γ, $\bar{\gamma}$, η and $\bar{\eta}$ (again for κ small enough) are given by

$$\gamma = \frac{7}{4}, \qquad \bar{\gamma} = \frac{1}{4}, \qquad \eta = \frac{1}{4}, \qquad \bar{\eta} = -\frac{3}{2}.$$

(This time the main constraint on the η exponents is $\eta \in (0, 1/2)$ and $\bar{\eta} = 2(\eta - 1)$ which are imposed by the irregularity of the noise and the form of the nonlinearity.) Again, these choices satisfy the various inequalities and properties we imposed in Section 4 with $q = 1$.

The rest of the proof is virtually identical to that given above for the dynamical Φ_3^4 model. □

Denote now by $\tilde{\mathcal{C}}^\alpha$ the space of $\mathcal{C}^\alpha$ functions, quotiented by constant functions and write μ for the probability measure on $\tilde{\mathcal{C}}^\alpha$ (for some $\alpha \in (0, \frac{1}{2})$) under which the h^i are independent Brownian bridges. We then have the following result.

Proposition 5.4. *Under condition* (5.13), *the measure μ is the unique invariant measure for the multi-component KPZ equation and its solutions are almost surely global in time for every initial condition $h_0 \in \mathcal{C}^\alpha$.*

Proof. As already mentioned, the fact that μ is invariant for (5.12) was recently obtained in [19]. The uniqueness of the invariant measure μ then follows immediately from Corollary 3.9 since the Brownian bridge measure has full support in the space $\tilde{\mathcal{C}}^\alpha$. Since furthermore μ has exponential moments, the argument of [4] (see also [32] for an application in a context very similar to here) yields global solutions for μ-almost every initial conditions.

Let $A \subset \tilde{\mathcal{C}}^\alpha$ be this set of full measure. Then A is dense in $\tilde{\mathcal{C}}^\alpha$ so that, for every $h \in \tilde{\mathcal{C}}^\alpha$, there exists a sequence $h_n \in A$ with $h_n \to h$. Using Theorem 3.1, it follows that for every $\varepsilon > 0$, there exist $t > 0$, $N > 0$ and a coupling between the solutions with initial conditions h and h_N such that $\mathbf{P}((h, \boldsymbol{\xi}(\omega)) \in \mathcal{N}_t) > 1 - \varepsilon$ and, under this coupling, $\mathbf{P}(\Phi_s(h(s), \boldsymbol{\xi}(\omega)) = \Phi_s(h_N(s), \boldsymbol{\xi}(\bar{\omega})) \, \forall s \geq t) > 1 - \varepsilon$. Since the solution with initial condition h_N exists for all times almost surely, it follows that the solution with initial condition h also exists for all time with probability at least $1 - 2\varepsilon$. Since ε was arbitrary, the claim follows. □

Remark 5.5. The class of equations considered in [26] can be treated in exactly the same way, with the same exponents appearing. In particular, using the fact that the measure μ in [26, Eq. 4.2] has full support, our result shows that it is the unique invariant measure for the process constructed in [26, Thm 4.1].

Remark 5.6. The class of equations formally given by

$$\partial_t u = \partial_x^2 u + H(u) + G(u)\xi,$$

with periodic boundary conditions considered in [34] also satisfies the assumptions of our theorem, provided that the function G appearing in [34, Eq. 1.2] is strictly positive. (In the case when G is a matrix, we need its singular values to stay away from 0.) This is the first example considered here where we need to consider a non-constant function G. In this particular example, the strong Feller property has long been known, see [14, Thm 7.1.1], albeit under rather strong boundedness conditions on the coefficients G and H.

One example with non-constant G to which our theory also applies is the natural "stochastic heat equation with values in a manifold" considered in [2,8,29]. In this case, the non-singularity of G is a consequence of the fact that the Riemannian metric tensor of the target manifold is strictly positive.

5.3. *The dynamical $P(\Phi)_2$ model*

This is the model formally given by

$$\partial_t \Phi = \Delta \Phi - P'(\Phi) + \xi, \tag{5.16}$$

where ξ denotes space-time white noise, P is an even polynomial with positive leading coefficient, and the space variable takes values in some bounded two-dimensional torus. The strong Feller property for this model can be obtained in the same way as for the dynamical Φ_3^4 model, but the arguments are a bit easier since there is much more "wriggle room". Since the invariant measure for (5.16) (interpreted in a suitable Wick-renormalised sense) is known [1,11] and has full support, this again allows one to obtain almost sure global solutions for every initial condition in $\mathcal{C}^{-\kappa}$ for suitable $\kappa > 0$. This result however has already been obtained by more PDE-oriented methods in [45,51] and the strong Feller property for (5.16) has very recently been obtained in [51], so we do not provide any more details here.

Appendix: Variation of constants formula

In this section, we derive a version of the variation of constants formula that is suitable for our needs. This allows us to relate the derivative of the solution map with respect to its initial condition to the derivative with respect to the driving noise by (4.5), which was used in a crucial way in our proof. Throughout this appendix we assume that we are in the setting described in Section 4 and that all of the assumptions we made there are satisfied, without further mentioning this in our statements.

To formulate our result, the following notation is useful. Let P_s be the hyperplane $\{(t, x) \in \mathbb{R}^{d+1} : t = s\}$ and write $\mathcal{D}_s^{\gamma, \eta} \stackrel{\text{def}}{=} \mathcal{D}_{P_s}^{\gamma, \eta}$ for the corresponding spaces with singularity at P_s as defined in [28, Def. 6.2]. Let also $W \subset \mathcal{T}$ be a sector of regularity α_0 of the regularity structure $(\mathcal{T}, \mathcal{G})$. Consider then a measurable (and L-periodic as usual) function $F : [0, 1] \times \mathbb{R}^{d+1} \to W_{<\gamma}$ with the property that $F_s(\cdot) \stackrel{\text{def}}{=} F(s, \cdot) \in \mathcal{D}_s^{\gamma, \eta}$ for every $s \in [0, 1]$ and write

$$\bar{F}(z) = \int_0^1 F_s(z) \, ds, \tag{A.1}$$

for every $z \in \mathbb{R}^{d+1}$. We then have the following result.

Lemma A.1. *Let $\gamma > 0$ and $\eta \in (\gamma - 2, \alpha_0)$, where α_0 is the regularity of the sector W above. Then, if F_s is bounded in $\mathcal{D}_s^{\gamma, \eta}$ uniformly over $s \in [0, 1]$, the function $\bar{F}$ given by (A.1) belongs to $\mathcal{D}^\gamma$.*

Proof. For any pair of points $(z, \bar{z})$ with $|z - \bar{z}| \leq 1$ and any $\beta < \gamma$, we estimate $\|\bar{F}(z) - \bar{F}(\bar{z})\|_\beta$ in the following way. Write $\delta = |z - \bar{z}|$ as a shorthand and write $I_\delta = [t - 4\delta^2, t + 4\delta^2]$, where t is such that $z = (t, x)$ for some $x \in \mathbf{T}^d$. We then write

$$\left\| \bar{F}(z) - \Gamma_{z\bar{z}} \bar{F}(\bar{z}) \right\|_\beta \leq \int_{[0,1] \setminus I_\delta} \left\| F_s(z) - \Gamma_{z\bar{z}} F_s(\bar{z}) \right\|_\beta ds$$

$$+ \int_{[0,1] \cap I_\delta} \left(\left\| F_s(z) \right\|_\beta + \left\| \Gamma_{z\bar{z}} F_s(\bar{z}) \right\|_\beta \right) ds. \tag{A.2}$$

For the first integral, we use the fact that, by the definition of the spaces $\mathcal{D}_s^{\gamma, \eta}$, one has the bound

$$\left\| F_s(z) - \Gamma_{z\bar{z}} F_s(\bar{z}) \right\|_\beta \lesssim |z - \bar{z}|^{\gamma - \beta} |t - s|^{\frac{\eta - \gamma}{2}}.$$

Since $\eta > \gamma - 2$ by assumption, this expression is integrable in s, thus leading to the required bound of order $|z - \bar{z}|^{\gamma - \beta}$. Regarding the second bound, it follows from the definition of the spaces $\mathcal{D}_s^{\gamma, \eta}$ and the fact that we chose $\eta \leq \alpha_0$

that

$$\left\| F_s(z) \right\|_\beta + \left\| \Gamma_{z\bar{z}} F_s(\bar{z}) \right\|_\beta \lesssim |t-s|^{\frac{\eta-\beta}{2}} + \sum_{\alpha \in [\beta, \gamma)} |\bar{t}-s|^{\frac{\eta-\alpha}{2}} |z-\bar{z}|^{\alpha-\beta},$$

where $\bar{t}$ is the time component of $\bar{z}$. Since $\eta > \gamma - 2$, so that in particular $\eta > \beta - 2$, these functions are all integrable in s and we obtain the bounds

$$\int_{[0,1] \cap I_\delta} |t-s|^{\frac{\eta-\beta}{2}} \, ds \lesssim \delta^{\eta-\beta+2}, \qquad \int_{[0,1] \cap I_\delta} |\bar{t}-s|^{\frac{\eta-\alpha}{2}} \, ds \lesssim \delta^{\eta-\alpha+2},$$

so that the corresponding term in (A.2) is bounded by $|z-\bar{z}|^{\eta-\beta+2}$. Since furthermore $|z-\bar{z}| \le 1$ and $\eta + 2 > \gamma$, this is in turn bounded by $|z-\bar{z}|^{\gamma-\beta}$ as desired. □

We now consider linear equations of the type

$$J = \mathcal{P}\mathbf{1}_+ \Psi J + P J_0, \tag{A.3}$$

where, for some $T > 0$, one has $\Psi \colon [0, T] \times \mathbb{R}^d \to L(V_\gamma, \bar{V}_{\bar{\gamma}})$ and we wrote $(\Psi J)(z) = \Psi(z) J(z)$. We make the following running assumption.

Assumption 12. The map $J \mapsto \Psi J$ with $(\Psi J)(z) = \Psi(z) J(z)$ maps $\mathcal{D}_s^{\gamma, \eta}$ into $\mathcal{D}_s^{\bar{\gamma}, \bar{\eta}}$ for every $s \ge 0$, with a bound of the type

$$\|\Psi J\|_{\bar{\gamma}, \bar{\eta}; [s, T]} \le C \|J\|_{\gamma, \eta; [s, T]},$$

holding uniformly over $s \in [0, T]$.

It follows immediately from [28, Thm 7.8] that under this assumption, (A.3) admits a unique global solution. Furthermore, by linearity of the equation, this solution is linear in the initial condition J_0. By our assumptions on the sector V, the solution J is such that, away from $t = 0$, $\mathcal{R}J$ is a Hölder continuous function of regularity ζ by (4.1) and [28, Prop. 3.28]. We then write $J^{(s)} \colon \mathcal{C}^\eta \to \mathcal{D}_s^{\gamma, \eta}$ for the solution to (A.3) with $\mathbf{1}_+$ replaced by the indicator function of the set $\{(t, x) : t \ge s\}$ and $P J_0$ replaced by the solution map to the linearised equation with initial condition J_s, but starting at time s. We also write $J_{s,t} \colon \mathcal{C}^\eta \to \mathcal{C}^\zeta$ for the linear maps such that, for every $f \in \mathcal{C}^\eta$, one has

$$\left(\mathcal{R}J^{(s)} f\right)(t, x) = (J_{s,t} J_s f)(x).$$

It follows from [28, Prop. 7.11] that these linear maps satisfy the identities $J_{t,u} \circ J_{s,t} = J_{s,u}$ for any $0 \le s \le t \le u$, where we use of course implicitly the canonical injection $\mathcal{C}^\zeta \hookrightarrow \mathcal{C}^\eta$. Consider now a function $f \in L^p([0, 1], \mathcal{C}^\eta)$ and let A be the solution to the fixed point equation

$$A = \mathcal{P}\mathbf{1}_+ \Psi A + P \star f. \tag{A.4}$$

As already mentioned in Remark 4.6, our assumptions guarantee that $P \star f \in \mathcal{C}^\gamma$, which we interpret as an element of $\mathcal{D}^\gamma$ by identifying it with its local Taylor expansion of order γ at each point. We then have the following version of the variation of constants formula.

Proposition A.2. *The equation* (A.4) *admits a unique local solution in* $\mathcal{D}^\gamma$, *given by the identity*

$$A = \int_0^\infty J^{(s)} f(s, \cdot) \, ds. \tag{A.5}$$

Proof. Defining A by (A.5), we only need to show that (A.4) is satisfied. First, note that A does indeed belong to $\mathcal{D}^\gamma$ as a consequence of Lemma A.1. We also note that, by the definition of $J^{(s)}$, one has for every $s \geq 0$ the identity

$$J^{(s)} f(s, \cdot) = \mathcal{P} \mathbf{1}_{t \geq s} \Psi J^{(s)} f(s, \cdot) + P_s f(s, \cdot).$$

Since furthermore $(J^{(s)} f(s, \cdot))(t, \cdot) = 0$ for $t \leq s$, one also has

$$J^{(s)} f(s, \cdot) = \mathcal{P} \Psi J^{(s)} f(s, \cdot) + P_s f(s, \cdot).$$

Integrating over s and exploiting the linearity of $\mathcal{P}$, together with the fact that, for $\bar{\mathcal{T}}$-valued elements $f \in \mathcal{D}^\alpha$, $\mathcal{P} \mathbf{1}_+ f$ coincides with the canonical lift of $P \star (\mathbf{1}_+ f)$, completes the proof. $\qquad\square$

Corollary A.3. *For every $t \in [0, 1]$ and every $(\Phi_0, \Pi) \in \mathcal{N}_t$, the identity (4.5) holds.*

Proof. Applying the implicit functions theorem to (4.3), it follows that

$$\mathcal{D}\Phi(\Phi_0, \Pi)h = \mathcal{R} U^h,$$

where $U^h \in \mathcal{D}^{\gamma, \eta}$ solves the fixed point problem

$$U^h = \mathcal{P} \mathbf{1}_+ \big(DF(\Phi)U^h\big) + P \star \big(G(\langle \mathbf{1}, \Phi \rangle)h\big),$$

with Φ the solution to (4.2) with initial condition Φ_0.

On the other hand, we already argued that the derivative $J = D\Phi$ of the solution with respect its initial condition Φ_0 solves (4.4). We are therefore almost exactly in the setting of Proposition A.2. The only problem is that the process $DF(\Phi)$ is only defined up to some possible blow-up time T, which is however guaranteed to satisfy $T > t$ by the definition of $\mathcal{N}_t$. This can easily be circumvented by simply multiplying $DF(\Phi)$ with a smooth cutoff function which leaves the equation unchanged before time t and makes it vanish before time T, so that it can be continued trivially for all times. We conclude that one does have the identity

$$U^h = \int_0^t J^{(s)} G\big(\langle \mathbf{1}, \Phi(s, \cdot) \rangle\big) h(s) \, ds.$$

Evaluating this at time t, the claim follows. $\qquad\square$

Acknowledgements

References

[1] S. Albeverio and M. Röckner. Stochastic differential equations in infinite dimensions: Solutions via Dirichlet forms. *Probab. Theory Related Fields* **89** (3) (1991) 347–386. MR1113223

[2] Y. Bruned, M. Hairer and L. Zambotti. Algebraic renormalisation of regularity structures, 2016. Available at arXiv:1610.08468.

[3] V. I. Bogachev. *Gaussian Measures. Mathematical Surveys and Monographs* **62**. American Mathematical Society, Providence, RI, 1998. MR1642391

[4] J. Bourgain. Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. *Comm. Math. Phys.* **176** (2) (1996) 421–445. MR1374420

[5] F. Baudoin and J. Teichmann. Hypoellipticity in infinite dimensions and an application in interest rate theory. *Ann. Appl. Probab.* **15** (3) (2005) 1765–1777. Available at arXiv:math/0508452. MR2152244

[6] R. Catellier and K. Chouk. Paracontrolled Distributions and the 3-dimensional Stochastic Quantization Equation, 2013. Available at arXiv:1310.6869.

[7] G. Cannizzaro, P. K. Friz and P. Gassiat. Malliavin calculus for regularity structures: The case of gPAM. *J. Funct. Anal.* **272** (1) (2017) 363–419. Available at arXiv:1511.08888. MR3567508

[8] A. Chandra and M. Hairer. An analytic BPHZ theorem for regularity structures, 2016. Available at arXiv:1612.08138.

[9] C. Dellacherie and P.-A. Meyer. *Probabilités et Potentiel. Chapitres IX à XI*, revised edition. Publications de l'Institut de Mathématiques de l'Université de Strasbourg, XVIII. Hermann, Paris, 1983. MR0727641

[10] G. Da Prato and A. Debussche. Two-dimensional Navier–Stokes equations driven by a space-time white noise. *J. Funct. Anal.* **196** (1) (2002) 180–210. MR1941997

[11] G. Da Prato and A. Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31** (4) (2003) 1900–1916. MR2016604

[12] G. Da Prato, K. D. Elworthy and J. Zabczyk. Strong Feller property for stochastic semilinear equations. *Stoch. Anal. Appl.* **13** (1) (1995) 35–45. MR1313205

[13] G. Da Prato and J. Zabczyk. Smoothing properties of transition semigroups in Hilbert spaces. *Stoch. Stoch. Rep.* **35** (2) (1991) 63–77. MR1110991

[14] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge University Press, Cambridge, 1996. MR1417491

[15] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, 2nd edition. *Encyclopedia of Mathematics and Its Applications*. **152**. Cambridge University Press, Cambridge, 2014. MR3236753

[16] J.-P. Eckmann and M. Hairer. Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Comm. Math. Phys.* **219** (3) (2001) 523–565. Available at arXiv:nlin/0009028. MR1838749

[17] K. D. Elworthy and X.-M. Li. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** (1) (1994) 252–286. MR1297021

[18] P. K. Friz and M. Hairer. *A Course on Rough Paths*. Springer International Publishing, Cham, 2014. MR3289027

[19] T. Funaki and M. Hoshino. A coupled KPZ equation, its two types of approximations and existence of global solutions, 2016. Available at arXiv:1611.00498. MR3653951

[20] T. Funaki and J. Quastel. KPZ equation, its renormalization and invariant measures. *Stoch. Partial Differ. Equ., Anal. Computat.* **3** (2) (2015) 159–220. Available at arXiv:1407.7310. MR3350451

[21] P. L. Ferrari, T. Sasamoto and H. Spohn. Coupled Kardar–Parisi–Zhang equations in one dimension. *J. Stat. Phys.* **153** (3) (2013) 377–399. Available at arXiv:1306.5643. MR3107649

[22] M. Gubinelli, P. Imkeller and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, e6 (2015). Available at arXiv:1210.2684. MR3406823

[23] M. Gubinelli and N. Perkowski. KPZ reloaded. *Comm. Math. Phys.* **349** (1) (2017) 165–269. Available at arXiv:1508.03877. MR3592748

[24] M. Hairer. Ergodic properties of a class of non-Markovian processes. In *Trends in Stochastic Analysis* 65–98. *London Math. Soc. Lecture Note Ser.* **353**. Cambridge Univ. Press, Cambridge, 2009. Available at arXiv:0708.3338. MR2562151

[25] M. Hairer. An introduction to stochastic PDEs, 2009. Available at arXiv:0907.4178.

[26] M. Hairer. Rough stochastic PDEs. *Comm. Pure Appl. Math.* **64** (11) (2011) 1547–1585. Available at arXiv:1008.1708.

[27] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)* **178** (2) (2013) 559–664. Available at arXiv:1109.6811. MR3071506

[28] M. Hairer. A theory of regularity structures. *Invent. Math.* **198** (2) (2014) 269–504. Available at arXiv:1303.5113. MR3274562

[29] M. Hairer. The motion of a random string, 2016. Available at arXiv:1605.02192.

[30] M. Hairer and J. C. Mattingly. Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)* **164** (3) (2006) 993–1032. Available at arXiv:math/0406087. MR2259251

[31] M. Hairer and J. C. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electron. J. Probab.* **16** (23) (2011) 658–738. Available at arXiv:0808.1361. MR2786645

[32] M. Hairer and K. Matetski. Discretisations of rough stochastic PDEs. *Ann. Probab.* (2015). Available at arXiv:1511.06937. MR2832168

[33] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.* **119** (1967) 147–171. MR0222474

[34] M. Hairer and É. Pardoux. A Wong–Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan* **67** (4) (2015) 1551–1604. Available at arXiv:1409.3138. MR3417505

[35] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ, 2015. Available at arXiv:1512.07845.

[36] M. Hairer and H. Shen. A central limit theorem for the KPZ equation. *Ann. Probab.* (2015). Available at arXiv:1507.01237.

[37] M. Hairer and H. Weber. Large deviations for white-noise driven, nonlinear stochastic PDEs in two and three dimensions. *Ann. Fac. Sci. Toulouse Math. (6)* **24** (1) (2015) 55–92. Available at arXiv:1404.5863. MR3325951

[38] G. Jona-Lasinio and P. K. Mitter. On the stochastic quantization of field theory. *Comm. Math. Phys.* **101** (3) (1985) 409–436. MR0815192

[39] M. Kardar, G. Parisi and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** (9) (1986) 889–892.

[40] T. Lyons and Z. Qian. *System Control and Rough Paths*. *Oxford Mathematical Monographs*. Oxford University Press, Oxford, 2002. Oxford Science Publications. MR2036784

[41] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations* 195–263. *Kyoto*, *1976*. *Res. Inst. Math. Sci., Kyoto Univ.* Wiley, New York–Chichester–Brisbane, 1978. MR0536013

[42] B. Maslowski. Strong Feller property for semilinear stochastic evolution equations and applications. In *Stochastic Systems and Optimization* 210–224. *Warsaw*, *1988*. *Lecture Notes in Control and Inform. Sci.* **136**. Springer, Berlin, 1989. MR1180781

[43] J. C. Mattingly and É. Pardoux. Malliavin calculus for the stochastic 2D Navier–Stokes equation. *Comm. Pure Appl. Math.* **59** (12) (2006) 1742–1790. Available at arXiv:math/0407215. MR2257860

[44] S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*, 2nd edition. Cambridge University Press, Cambridge, 2009. MR2509253

[45] J.-C. Mourrat and H. Weber. The dynamic Φ_3^4 model comes down from infinity. *Comm. Math. Phys.* (2016). Available at arXiv:1601.01234.

[46] E. Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles* 69–73. *Proc. Conf., Dedham, Mass., 1965*. M.I.T. Press, Cambridge, Mass, 1966. MR0210416

[47] J. Norris. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85* 101–130. *Lecture Notes in Math.* **1204**. Springer, Berlin, 1986. MR0942019

[48] M. Rockner, R. Zhu and X. Zhu. Ergodicity for the stochastic quantization problems on the 2D-torus. ArXiv e-prints, 2016. Available at arXiv:1606.02102. MR3631399

[49] J. Seidler. A note on the strong Feller property, 2001. Unpublished lecture notes.

[50] H. Spohn. Nonlinear fluctuating hydrodynamics for anharmonic chains. *J. Stat. Phys.* **154** (5) (2014) 1191–1227. Available at arXiv:1305.6412. MR3176405

[51] P. Tsatsoulis and H. Weber. Spectral Gap for the Stochastic Quantization Equation on the 2-dimensional Torus, 2016. Available at arXiv:1609.08447.

[52] R. Zhu and X. Zhu. Three-dimensional Navier–Stokes equations driven by space-time white noise. *J. Differential Equations* **259** (9) (2015) 4443–4508. Available at arXiv:1406.0047. MR3373412