Stochastics and Partial Differential Equations: Analysis and Computations
https://doi.org/10.1007/s40072-025-00365-0

On ergodic properties of stochastic PDEs

Le Chen[1] · Cheng Ouyang[2] · Samy Tindel[3] · Panqiu Xia[4]

Dedicated to Giuseppe (Beppe) Da Prato, in memoriam

Received: 5 December 2024 / Revised: 5 December 2024 / Accepted: 24 April 2025

Abstract
In this note we review several situations in which stochastic PDEs exhibit ergodic properties. We begin with the basic dissipative conditions, as stated by Da Prato and Zabczyk in their classical monograph. Then we describe the singular case of SPDEs with reflection. Next we move to some degenerate (and thus more demanding) settings. Namely we recall some results obtained around 2006, concerning stochastic Navier-Stokes equations with a very degenerate noise. We finish the article by handling some cases with degenerate coefficients. This includes a new result about the parabolic Anderson model in dimension $d \geq 3$, driven by a general class of noises and fairly general initial conditions. In this context, a phase transition is observed, expressed in terms of the noise intensity.

Keywords Invariant measure · Dalang's condition · Ergodicity · Stochastic heat equation · Parabolic Anderson model · Phase transition.

✉ Samy Tindel
stindel@purdue.edu

Le Chen
le.chen@auburn.edu

Cheng Ouyang
couyang@uic.edu

Panqiu Xia
xiap@cardiff.ac.uk

[1] Department of Mathematics and Statistics, Auburn University, Auburn, United States

[2] Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, Illinois, United States

[3] Department of Mathematics, Purdue University, West Lafayette, United States

[4] School of Mathematics, Cardiff University, Cardiff, United Kingdom

Published online: 25 May 2025

1 Introduction

Any person having attended a stochastic analysis conference in the 90's-00's will certainly remember Beppe, consistently beginning his talks with the famous singing words: "let A be an operator on a Hilbert space H, generating a C_0-semigroup $S(t)\dots$". However, behind what could sometimes be seen as a kind of ritual, lies an important fact: Beppe was an exceptional pioneer in the area of stochastic PDEs, who was completely passionate about his topic. Most importantly, he certainly was the undisputed leader of the field for decades. This outstanding legacy is well highlighted by the series of books [21–23] written in collaboration with Jerzy Zabczyk, which have been and still are an immense source of inspiration for the stochastic PDE community.

Among the influential contributions mentioned above, the current paper will single out [21]. This book brings together the worlds of general ergodic theory and stochastic models in infinite dimensions, a delicate task which is achieved by Da Prato and Zabczyk in an astonishingly clear and consistent way. Starting from the foundations, the book covers a wide range of applications to systems (dissipative equations, Navier-Stokes equations, spin systems) which are still the object of active research today.

Our contribution first proposes to review a few basic facts contained in [21]. Namely we will recall the fundamental notions allowing to get ergodic type results for infinite dimensional stochastic differential equations seen as dynamical systems. We will illustrate this technology by stating an ergodic result for a generic stochastic PDE. We shall then examine some cases departing from the standard Da Prato-Zabczyk setting. Without any pretension to be exhaustive, we wish to give an account on the following situations:

(i) *Case of reflected equations.* Following [53], we will see how reflections change the landscape in terms of invariant measures. This encompasses the identification of new invariant measures, as well as integration by parts formulae.

(ii) *Ergodicity in degenerate situations.* In the celebrated result [37], Hairer and Mattingly showed that a nondegenerate noise is not necessary for good mixing properties of a stochastic PDE. Their result focuses on a stochastic Navier-Stokes equation, for which only four active modes in the noise are requested.

(iii) *Phase transition in ergodicity.* Our own contribution will be related to another degenerate situation, namely the parabolic Anderson model (PAM). There we will see that for a spatial variable $x \in \mathbb{R}^d$ with $d \geq 3$, we have two cases: at high temperature the equation exhibits an ergodic behavior; In contrast, for low temperatures the moments of the solution blow up as $t \to \infty$. We shall give a self contained proof of the ergodic behavior based on contraction arguments, for a general spatial covariance of the noise. This result appears to be new and is inspired by [34, 36].

This constant dialogue with Da Prato and Zabczyk's books (seen as benchmarks) is a staple in the stochastic analysis literature. We will also mention possible new developments for polymer measures.

Our article is structured as follows: in Section 2 we recall the classical Da Prato-Zabczyk ergodic setting for stochastic PDEs. Section 3 is devoted to equations with reflection. In Section 4 we turn to the case of degenerate noises, focusing on the stochastic Navier-Stokes case. Eventually Section 5 deals with phase transitions in ergodicity for generalized PAMs. A Gronwall-type lemma is proved in Appendix A.

2 The standard stochastic PDE case

In this section, we describe the standard ergodic setting for stochastic PDEs in Da Prato and Zabczyk's language. We will spell out the general setting in Section 2.1 and then outline the main ergodic result in Section 2.2.

2.1 Abstract setting for stochastic PDEs

Let us briefly recall the abstract setting which is used in [21] in order to solve stochastic differential equations in infinite dimensions. It can be split in four ingredients (notice that for sake of conciseness we will not define every technical term below, we refer to Beppe's original book for a complete and self-contained exposition):

(a) As recalled in the introduction, everything starts with the infinitesimal generator A of a strongly continuous semigroup $\{S(t); t \geq 0\}$ defined on a separable Hilbert space H (notice that analyticity of $S(t)$ is further assumed in [21]).

(b) One considers a mapping $F : H \to H$ with linear growth and Lipschitz continuity. That is one has

$$F(x) \lesssim 1 + |x|, \quad \text{and} \quad |F(x) - F(y)| \lesssim |x - y|, \tag{2.1}$$

where $|\cdot|$ denotes the norm in H. The mapping F will be the drift of our equation.

(c) On another (possibly larger) Hilbert space U, let Q be a trace class operator. Then we can define a Q-Wiener process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This is a U-valued continuous process $\{W_t; t \geq 0\}$ with independent increments and such that for all $t \geq 0,\ h > 0$ we have

$$\frac{1}{h^{1/2}} (W_{t+h} - W_t) \sim \mathcal{N}(0, Q). \tag{2.2}$$

Notice that the space U is equipped with an inner product induced by Q:

$$\langle \xi, \psi \rangle_Q := \langle \xi, Q\psi \rangle_U . \tag{2.3}$$

(d) The diffusion coefficient B of our equation has to be compatible with the Q-Wiener process and the semigroup $S(t)$. Namely $B : H \to L(U; H)$ has to be a strongly continuous map such that the following growth conditions are fulfilled for every $t > 0$ and $x, y \in H$:

$$\|S(t)B(x)\|_{\text{HS}} \leq K(t)(1 + |x|), \tag{2.4}$$

$$\|S(t)B(x) - S(t)B(y)\|_{\text{HS}} \leq K(t)|x - y|, \tag{2.5}$$

where K is a kernel which sits in the space $L^2([0, T])$ for every $T > 0$. We also label the following assumption which is useful in order to get regularity properties of our stochastic differential equations: there exists $\alpha \in (0, 1/2)$ such that for every $T > 0$ we have

$$\int_0^T s^{-2\alpha} K^2(s) \mathrm{d}s < \infty. \tag{2.6}$$

As we will see in the next section, the above setting is sufficient to ensure existence, uniqueness and ergodic properties of a general class of stochastic PDEs.

2.2 Ergodic properties

With our setting of Section 2.1 in hand, the general type of H-valued equation considered in [21] can be written as

$$\mathrm{d}X_t = [AX_t + F(X_t)]\,\mathrm{d}t + B(X_t)\mathrm{d}W_t, \tag{2.7}$$

with an initial condition $X_0 = \xi$ for a given $\xi \in H$. Observe that equation (2.7) is solved in the so-called mild sense, which can be written as

$$X_t = S(t)\xi + \int_0^t S(t-s)F(X_s)\mathrm{d}s + \int_0^t S(t-s)B(X_s)\mathrm{d}W_s, \tag{2.8}$$

where the stochastic integral in (2.8) has to be interpreted in the Itô sense.

The conditions (a)– (d) in Section 2.1 ensure existence and uniqueness of the solution to (2.7), although a complete theory for this fact is better found in the first Da Prato-Zabczyk volume [22]. As far as ergodic properties are concerned, it should be noticed that ergodic behaviors often rely on damping terms in dynamic systems. In case of equation (2.7), this damping is provided by the operator A (whose spectrum is implicitly thought of as mostly negative). The main result in this direction is summarized in the following theorem.

Theorem 2.1 *Assume the setting of Section 2.1 holds true. In addition, suppose that the operators $S(t)$ are compact and that the following uniform bound on moments is satisfied: one can find $T_0 > 0$ and $p \geq 2$ such that*

$$\sup_{t \geq T_0,\, \xi \in H} \mathbb{E}_\xi \left[|X_t|^p \right] < \infty. \tag{2.9}$$

Then there exists an invariant measure for equation (2.7).

Proof The detailed proof is found in [21, Theorem 6.1.2]. We will content ourselves with an outline of the main ingredients. This will be split in several steps.
Step 1: Averaged measure. The solution $\{X_t; t \geq 0\}$ to equation (2.7) generates a Markov dynamics. We call $P_t(\xi, \cdot)$ the corresponding transition. Next for a Borel set

$\Gamma \in H$ and $T > 0$ we set

$$R_T(\xi, \Gamma) = \frac{1}{T} \int_0^T P_t(\xi, \Gamma) \mathrm{d}t. \tag{2.10}$$

According to a celebrated criteria from Krylov and Bogoliubov, there exists an invariant measure for X_t as long as the family $\{R_T; T > 0\}$ is tight. Let us recall that tightness of R_T means that for every $\epsilon \in (0, 1)$, one is able to find a compact set $K_\epsilon \subset H$ such that

$$R_T(\xi, K_\epsilon) \geq 1 - \epsilon, \quad \text{uniformly in } T. \tag{2.11}$$

Summarizing, relation (2.11) implies the existence of an invariant measure.

A criterion like (2.11) is easy to verify when $H = \mathbb{R}^n$. Indeed, in that case the compact K_ϵ can be taken as a (closed) ball $B(0, r_\epsilon)$ for a proper radius $r_\epsilon > 0$. In addition, a direct application of Markov's inequality entails that for any $p \geq 1$

$$\mathbb{P}\left(X_t \notin B(0, r_\epsilon)\right) \leq \frac{\mathbb{E}\left[|X_t|^p\right]}{r_\epsilon^p}, \tag{2.12}$$

so that (2.11) is easily implied by (2.9). This is not true anymore when H is an infinite dimensional functional space. Some additional ingredients are thus needed.

Step 2: Factorization method. In order to solve the problem raised in Step 1, Da Prato and Zabczyk resort to one of their most emblematic and clever trick. Namely under assumption (d), some elementary fractional integral arguments show that the stochastic integral in (2.8) can be expressed as

$$\int_0^t S(t-s) B(X_s) \mathrm{d}W_s = \frac{\sin(\alpha\pi)}{\pi} \int_0^t (t-s)^{\alpha-1} S(t-s) Y_s \mathrm{d}s, \tag{2.13}$$

where the process $Y = Y^{(\alpha)}$ is the element of $L^2([0, T]; H)$ defined by

$$Y_s = \int_0^s (s-r)^{-\alpha} S(s-r) B(X_r) \mathrm{d}W_r.$$

Formula (2.13) is a progress in the following sense: we have obtained a representation of stochastic convolutions as Lebesgue type integrals. This expression is more amenable to computations for regularity (and therefore compactness) estimates.

Step 3: Compactness. Recall that our goal is to establish relations (2.11)-(2.12) for a family of compact sets K_ϵ. This is achieved by writing the solution to (2.7) at time $t = 1$ in the following way (thanks to the factorization formula (2.13)):

$$X_1 = S(1)\xi + G_1 F(X) + G_\alpha Y, \tag{2.14}$$

where for $\alpha \in (0, 1]$ we introduce an operator $G_\alpha : L^p([0,1]; H) \to H$ by

$$G_\alpha f = \frac{\sin(\alpha\pi)}{\pi} \int_0^1 (1-s)^{\alpha-1} S(1-s) f(s) \mathrm{d}s.$$

An important point in the proof is that whenever $S(t)$ is compact for $t > 0$, G_α is also compact. Therefore whenever ξ, $F(X)$ and Y in (2.14) lye in proper balls in H, the random element X_1 belongs to a compact set. A full description of this family of sets $\{K_\epsilon; r > 0\}$ stems from the decomposition (2.14):

$$K_\epsilon = \Big\{ x \in H; \quad x = S(1)\xi + G_1 g + G_\alpha h, \quad \text{with} \\ |\xi| \le \frac{1}{\epsilon}, \quad |g|_{L^p([0,1];H)} \le \frac{1}{\epsilon}, \quad \text{and} \quad |h|_{L^p([0,1];H)} \le \frac{1}{\epsilon} \Big\}. \tag{2.15}$$

Step 4: Conclusion. The expression (2.15) allows to locate X_1 given by (2.14) in a compact set thanks to mere moment estimates. Furthermore, we have assumed in (2.9) that the moments of X_t are uniformly bounded. Hence for $\epsilon > 0$ and K_ϵ defined by (2.15), we end up with

$$\mathbb{P}_\xi(X_1 \in K_\epsilon) \ge 1 - C\epsilon^p (1 + |\xi|^p).$$

An easy Markov conditioning procedure allows then to conclude that (2.11) holds true whenever ϵ is small enough. This implies the existence of an invariant measure for X_t as explained in Step 1.

As the reader might see from the sketch above, Da Prato and Zabczyk's method is a very elaborate combination of deep functional analysis insight, stochastic calculus techniques and fundamental ergodic tools. It is difficult to overstate Da Prato's pioneering role in this type of development. Let us add a couple of remarks to close the section.

Remark 2.2 Existence of an invariant measure is the most basic result one can get in ergodic theory. The extra ingredients put forward in [21] in order to get uniqueness are essentially reduced to strong Feller properties for the transition P_t and irreducibility. We will go back to those issues in Section 4. However, let us mention at this point that uniqueness of the invariant measure is closely related to irreducibility of the Markovian transition $P_t(\xi, \cdot)$ alluded to in the proof of Theorem 2.1. This irreducibility is ensured by nondegeneracy properties of the noise W and the diffusion coefficient B in (2.8). A standard way to express this nondegeneracy when B is constant is the following:

$$\mathrm{Im}\,(S(t)) \subset \mathrm{Im}\big(Q_t^{1/2}\big), \quad \text{where} \quad Q_t = \int_0^t S(s) B\, S(s)^* \,\mathrm{d}s\,. \tag{2.16}$$

Remark 2.3 For dissipative systems, that is when $A + F$ in (2.7) can be considered as a damping coefficient, one can prove convergence of $\mathscr{L}(X_t)$ to an invariant measure by a direct analysis based on Itô formula. This type of consideration is obtained, e.g., in [21, Section 6.3 and 6.5]

Remark 2.4 As mentioned in the introduction, Da Prato and Zabczyk also investigate the ergodic behavior of specific stochastic systems of interest. Those include delayed systems, Burger's equation, spin systems, and the Navier-Stokes equation. Each case requires some serious modification of the general method.

3 A stochastic heat equation with reflection

We now start a series of deviations from the standard setting for ergodic stochastic PDEs (as introduced in Section 2), beginning with equations involving reflections. Observe that stochastic heat equations with reflection were firstly studied by Nualart and Pardoux [46] in a quasilinear setting, building upon a deterministic framework that had previously received considerable attention from various researchers; see [3] and references therein. It is worth noting that reflected stochastic heat equations can also arise as the scaling limit of microscopic models for random interfaces [31, 33]. As far as ergodic properties are concerned, reflections induce an extra singularity in the equation. In the current section, we will summarize an approximation procedure allowing us to handle this singularity. For further exploration of the topic, readers are referred to Zambotti's book [56] and survey paper [57, Section 6].

Stochastic PDEs with reflections can be handled in the abstract setting of Section 2. In this context, the underlying Hilbert space should be $H = L^2([0, 1])$. Nevertheless, since we wish to encompass Dirichlet boundary conditions, we will consider a state space of the form

$$H_0 := \{\xi \in H;\ \xi(0) = \xi(1) = 0\}, \tag{3.1}$$

equipped with the inner product inherited from H. In addition, the reflection measure will constrain solutions to live in a family of sets $\{K_\alpha;\ \alpha \geq 0\}$ defined by

$$K_\alpha := C([0, 1]) \cap \{\xi \in H_0;\ \xi \geq -\alpha\}\ . \tag{3.2}$$

With this preliminary notation in mind, we now consider the following equation driven by a Brownian sheet W:

$$\mathrm{d}X_t = [AX_t + F(X_t)]\,\mathrm{d}t + \mathrm{d}W_t + \eta(\mathrm{d}t). \tag{3.3}$$

In equation (3.3), the operator A is defined as $A = \frac{1}{2}\Delta$, and F is assumed to be a mapping similar to that considered in (2.1), which in particular might take the form

$$[F(\xi)]\,(x) = f(x, \xi(x)), \quad \text{for all } \xi \in H \text{ and } x \in [0, 1], \tag{3.4}$$

with a regular enough function $f : [0, 1] \times \mathbb{R} \to \mathbb{R}$. As mentioned above, W is a Brownian sheet. This means, in the language of Section 2, that W is a Q-Wiener process as in (2.2) with $Q = \mathrm{Id}_H$. In the literature, the time derivative of W (interpreted as a generalized function) is also referred to as space-time white noise.

It has been established in [46] that under certain mild conditions, for any initial condition $\xi \in K_\alpha$, the equation (3.3) reflected at $-\alpha \le 0$ admits a unique solution. This solution is a pair (X^α, η^α) wherein X^α is a K_α-valued process adapted to W starting at ξ, and η^α is an adapted nonnegative random measure satisfying

$$\int_{\mathbb{R}_+}\int_0^1 \left(\alpha + X_t^\alpha(x)\right) \eta^\alpha(\mathrm{d}t, \mathrm{d}x) = 0.$$

In terms of ergodic results for (3.3), it is natural to wonder how the singularity of the reflection is affecting invariant measures. Let us thus say a few words about invariant measures for a system analogous to (3.3), yet with no reflection. In other words, let us consider an equation such as (2.7) with $B = \mathrm{Id}$ and F of the form (3.4). For any $\xi \in H_0$ (recall that H_0 is defined by (3.1)), we set

$$G(\xi) := \int_0^1 \mathrm{d}x \int_0^{\xi(x)} f(x, y)\mathrm{d}y. \tag{3.5}$$

Now we define an underlying measure μ on H_0 as the distribution of the Brownian bridge. It is a well-known fact (see the proof of Theorem 3.1 below for more details about this assertion) that an invariant measure μ^G for (2.7) with $B = \mathrm{Id}$ is given by

$$\mu^G(\mathrm{d}\xi) := e^{2G(x)}\mu(\mathrm{d}x). \tag{3.6}$$

In his paper [53], Zambotti shows how the above expansion for the invariant measure is perturbed by the reflection. The result can be summarized as follows:

Theorem 3.1 *(Theorem 5 of [53]) Let B_t be the standard Brownian motion in $\mathbb{R}^3$. Let ν be the probability measure of the 3-d Bessel bridge, namely, the law of $\{|B_t|\}_{t\in[0,1]}$ conditioned to $B_1 = 0$. The set K_0 is introduced in (3.2) with $\alpha = 0$. We consider a measure ν^G on K_0 defined as follows: for all $\xi \in K_0$,*

$$\nu^G(\mathrm{d}\xi) := e^{2G(\xi)}\nu(\mathrm{d}\xi), \tag{3.7}$$

where G is given by (3.5). Then, ν^G is an invariant measure for (3.3) reflected at 0.

Proof Just as in Theorem 2.1, we only provide a brief overview of the proof strategy, directing interested readers to [53] for a more detailed exposition.
Step 1: Approximating the equation. Let $\alpha > 0$. For any $\epsilon > 0$, consider the following equation

$$\mathrm{d}X_t^{\alpha,\epsilon} = \left[AX_t^{\alpha,\epsilon} + F(X_t^{\alpha,\epsilon}) + \frac{(\alpha + X_t^{\alpha,\epsilon})^-}{\epsilon}\right]\mathrm{d}t + \mathrm{d}W_t, \tag{3.8}$$

with $r^- := \max\{-r, 0\}$ for all $r \in \mathbb{R}$. Then, it is known that for any initial condition $X_0 = \xi \in H_0$, the equation (3.8) has a unique solution in H_0. Furthermore, assuming the initial condition $\xi \in K_\alpha$, the same monotonicity arguments as in [46] assert that as $\epsilon \downarrow 0$, the following statements hold:

(i) For any $(t, x) \in \mathbb{R}_+ \times [0, 1]$, $X_t^{\alpha,\epsilon}(x)$ increases to a random variable $X_t^\alpha(x)$ such that $X_t^\alpha = \left\{X_t^\alpha(x); x \in [0, 1]\right\} \in K_\alpha$ for all $t > 0$;

(ii) $\left\{\eta^{\alpha,\epsilon}(\mathrm{d}t, \mathrm{d}x) := \epsilon^{-1}(\alpha + X_t^{\alpha,\epsilon}(x))^- \mathrm{d}t\mathrm{d}x\right\}_{\epsilon>0}$ converge weakly to a nonnegative measure η^α on $\mathbb{R}_+ \times [0, 1]$;

(iii) The pair (X^α, η^α) is a solution to (3.3) reflected at $-\alpha$.

Step 2: Invariant measure for the OU process. Let A be any infinitesimal generator of a strongly continuous semigroup on H_0 as in Section 2.1. Then the *Ornstein-Uhlenbeck process* (OU process), which is the solution to

$$\mathrm{d}Z_t = AZ_t\mathrm{d}t + \mathrm{d}W_t\,,$$

has an invariant measure μ, which is a centered Gaussian measure on H_0 with covariance operator $\int_0^\infty e^{2At}\mathrm{d}t = (-2A)^{-1}$. This fact had been proved in Da Prato-Zabczyk's monograph [21, Theorem 6.2.1]. As alluded to in (3.6), in case $A = \frac{1}{2}\Delta$, the Gaussian measure $\mu = \mathcal{N}\left(0, (-\Delta)^{-1}\right)$ on H_0, concentrated on $H_0 \cap C([0, 1])$, coincides with the distribution of the 1-d Brownian bridge $\mathcal{B}(t) \overset{(d)}{=} B_t - tB_1$ for all $t \in [0, 1]$ (where B stands for a 1-d Brownian motion). This can be seen by the following equality, valid for every pair $h, k \in H_0$:

$$\begin{aligned}\mathbb{E}\left[\langle\mathcal{B}, h\rangle\langle\mathcal{B}, k\rangle\right] &= \int_0^1 \mathrm{d}t \int_0^1 \mathrm{d}s\, \mathbb{E}\left[(B_s - sB_1)(B_t - tB_1)\right] h(s)k(t) \\ &= \int_0^1 \mathrm{d}t \int_0^1 \mathrm{d}s\,(s \wedge t - st)\, h(s)k(t) = \int_0^1 (-\Delta)^{-1}h(t)k(t)\mathrm{d}t.\end{aligned} \tag{3.9}$$

Notice that in (3.9), the last inequality can be briefly justified as follows: for every $h \in H_0$, the solution to the Dirichlet problem

$$\Delta g(t) = h(t), \quad \text{with boundary condition } g(0) = g(1) = 0,$$

can be written as

$$g(t) = -\int_0^1 \mathrm{d}s\,(s \wedge t - st)\, h(s).$$

Step 3: Invariant measure for(3.8). We now justify the claim in (3.6) in our context. Let G be given as in (3.5). It is easy to see that for any $\xi, \xi' \in H_0$, we have

$$\langle\nabla G(\xi), \xi'\rangle_{H_0} = \int_0^1 f(x, \xi(x))\,\xi'(x)\,\mathrm{d}x = \langle f(\cdot, \xi(\cdot)), \xi'\rangle_{H_0}.$$

In the same way, if V_α is a mapping on H_0 defined by

$$V_\alpha(\xi) := \frac{1}{2}\int_0^1 \left[(\alpha + \xi(x))^-\right]^2 \mathrm{d}x, \quad \text{for all} \quad \xi \in H_0,$$

then it is readily checked that for all $\xi, \xi' \in H_0$, the following holds

$$\langle \nabla V_\alpha(\xi), \xi' \rangle_{H_0} = -\left\langle (\alpha + \xi(\cdot))^- , \xi' \right\rangle_{H_0} .$$

As a consequence, (3.8) can be written as the following gradient system:

$$\mathrm{d}X_t^{\alpha,\epsilon} = \left[AX_t^{\alpha,\epsilon} + \nabla G(X_t^{\alpha,\epsilon}) - \frac{1}{\epsilon} \nabla V_\alpha(X_t^{\alpha,\epsilon}) \right] \mathrm{d}t + \mathrm{d}W_t.$$

Therefore, it follows from [21, Section 8.6] that equation (3.8) has an invariant measure of the form

$$\nu_{\alpha,\epsilon}^G(\mathrm{d}\xi) := \frac{1}{Z_{\alpha,\epsilon}} e^{2G(\xi)-(2V_\alpha(\xi)/\epsilon)} \mu(\mathrm{d}\xi), \tag{3.10}$$

where, under some integrability conditions (not specified here for the sake of conciseness), the renormalization constant $Z_{\alpha,\epsilon}$ defined below is a finite positive number:

$$Z_{\alpha,\epsilon} := \int_{H_0} e^{2G(\xi)-(2V_\alpha(\xi)/\epsilon)} \mu(\mathrm{d}\xi) \in (0, \infty).$$

Step 4: Convergence of $\nu_{\alpha,\epsilon}^G$. Our claim (3.7) will be obtained by taking limits in both α and ϵ in (3.10). Let us start by taking $\epsilon \downarrow 0$. It follows that

$$\lim_{\epsilon \downarrow 0} 2V_\alpha(\xi)/\epsilon = \begin{cases} 0, & \text{if } \xi \in K_\alpha, \\ \infty, & \text{if } \xi \in (H_0 \setminus K_\alpha) \cap C([0,1]). \end{cases} \tag{3.11}$$

Furthermore, recall that μ is a Gaussian measure supported on $H_0 \cap C([0, 1])$. Plugging relation (3.11) into (3.10), it is thus not hard to see that the following limit holds true:

$$\nu_{\alpha,\epsilon}^G(\mathrm{d}\xi) \to \nu_\alpha^G(\mathrm{d}\xi) := \frac{1}{Z_\alpha} e^{2G(\xi)} \nu_\alpha(\mathrm{d}\xi), \quad \text{as} \quad \epsilon \downarrow 0,$$

where the measure ν_α and the normalization constant Z_α are defined by

$$\nu_\alpha(\mathrm{d}\xi) := \frac{1}{\mu(K_\alpha)} \mathbf{1}_{K_\alpha}(\xi) \mu(\mathrm{d}\xi) \quad \text{and} \quad Z_\alpha := \int_{K_\alpha} e^{2G(\xi)} \nu_\alpha(\mathrm{d}\xi). \tag{3.12}$$

The limit in α in relation (3.12) is more involved. It is treated separately, and leads to Proposition 3.3 below. This proposition asserts that ν_α converges weakly to the law of a Bessel bridge. Together with relation (3.12), this concludes the proof of Theorem 3.1.

Remark 3.2 Recall that μ coincides with the distribution of the 1-d Brownian bridge. Thus by [48, Chapter III, Exercise (3.14)], $\mu(K_\alpha)$ for $\alpha \geq 0$ can be explicitly computed as follows:

$$\mu(K_\alpha) = 1 - e^{-2\alpha^2}. \tag{3.13}$$

Hence $\mu(K_\alpha) \sim 2\alpha^2$ as $\alpha \to 0$, which dictates the singularity of ν_α in (3.12). Sorting out this singularity is the content of Proposition 3.3.

Proposition 3.3 *The measures* $\{\nu_\alpha(\mathrm{d}\xi);\ \alpha > 0\}$ *converge weakly to* $\nu(\mathrm{d}\xi)$*, the law of the 3-d Bessel bridge, as* $\alpha \downarrow 0$.

Proof Proposition 3.3 had already been proved in [29]. Here, we provide an alternative proof as presented in [53]. The proof is based on the next theorem quoted from [5].

Theorem 3.4 *([5, Theorem 1]) Let* $e = \{e_\tau;\ \tau \in [0,1]\}$ *be a 3-d Bessel bridge, and let* ζ *be a random variable uniformly distributed on* $[0,1]$ *and independent of e. Consider a process* $\beta = \{\beta_\tau;\ \tau \in [0,1]\}$ *defined by*

$$\beta_\tau := e_{\tau \oplus \zeta} - e_\zeta, \quad \textit{with } \tau \oplus \zeta := \tau + \zeta \pmod 1.$$

Then β *is a 1-d Brownian bridge.*

Let us now go back to the analysis for ν_α. Thanks to (3.12), one can write, that for a generic $\varphi \in C_b(H_0)$,

$$\int_{H_0} \varphi(\xi)\nu_\alpha(\mathrm{d}\xi) = \frac{1}{\mu(K_\alpha)} \int_{K_\alpha} \varphi(\xi)\mu(\mathrm{d}\xi),$$

where we recall that μ is the distribution of the 1-*d* Brownian bridge. Now we apply Theorem 3.4 in order to express μ in terms of the distribution ν of the 3-*d* Bessel bridge. We get

$$\int_{H_0} \varphi(\xi)\nu_\alpha(\mathrm{d}\xi) = \frac{1}{\mu(K_\alpha)} \int_0^1 \mathrm{d}r \int_{K_0} \varphi(\xi_{\cdot\oplus r} - \xi_r)\mathbf{1}_{K_\alpha}(\xi_{\cdot+r} - \xi_r)\, \nu(\mathrm{d}\xi). \tag{3.14}$$

In addition, if $\xi \in K_0$, where K_0 is defined as in (3.2) with $\alpha = 0$, we have $\xi \geq 0$ and $\xi(0) = \xi(1) = 0$. Therefore, for any $r \in [0,1]$ we have

$$\xi_{\cdot\oplus r} - \xi_r \geq -\alpha \iff \xi_\cdot - \xi_r \geq -\alpha \iff \xi_r \leq \alpha + \xi_\cdot \iff \xi_r \leq \alpha.$$

Plugging this elementary relation into (3.14), we obtain

$$\int_{H_0} \varphi(\xi)\nu_\alpha(\mathrm{d}\xi) = \frac{1}{\mu(K_\alpha)} \int_0^1 \mathrm{d}r \int_{K_0} \varphi(\xi_{\cdot\oplus r} - \xi_r)\mathbf{1}_{[0,\alpha]}(\xi_r)\, \nu(\mathrm{d}\xi). \tag{3.15}$$

We now proceed by splitting the right hand side of (3.15) in two terms:

$$\int_{H_0} \varphi(\xi)\nu_\alpha(\mathrm{d}\xi) = \frac{1}{\mu(K_\alpha)}\big(I_1(\alpha) + I_2(\alpha)\big), \tag{3.16}$$

where we recall from (3.13) that $\mu(K_\alpha) = 1 - e^{-2\alpha^2}$, and where

$$I_1(\alpha) := \int_0^{1/2} \mathrm{d}r \int_{K_0} \varphi(\xi_{\cdot\oplus r} - \xi_r)\mathbf{1}_{[0,\alpha]}(\xi_r)\nu(\mathrm{d}\xi), \tag{3.17}$$

$$I_2(\alpha) := \int_{1/2}^1 \mathrm{d}r \int_{K_0} \varphi(\xi_{\cdot\oplus r} - \xi_r)\mathbf{1}_{[0,\alpha]}(\xi_r)\nu(\mathrm{d}\xi). \tag{3.18}$$

In the sequel, we will only handle the term $I_1(\alpha)$, while the other term $I_2(\alpha)$ can be treated exactly along the same lines. The approach to addressing $I_1(\alpha)$ involves conditioning the integration in $I_1(\alpha)$ to the values of ξ_r. To this aim, we recall that classical considerations on Bessel bridge show that the density of ξ_r is of the form

$$\nu\left(\xi_r \in \mathrm{d}y\right) = g_{r(1-r)}(y)\mathrm{d}y, \quad \text{with } g_\tau(y) := \sqrt{\frac{2}{\pi\tau^3}}\, y^2 \exp\left(-\frac{y^2}{2\tau}\right), \ \tau > 0, \ y \geq 0. \tag{3.19}$$

Denoting by $\delta_\varphi(r, y)$ the integration of φ with respect to ν conditional on $\xi_r = y$, we end up with

$$I_1(\alpha) = \int_0^{1/2} \mathrm{d}r \int_0^\alpha \mathrm{d}y\, g_{r(1-r)}(y)\delta_\varphi(r, y).$$

Recalling (3.19), by preforming a change of variable $y/\sqrt{r(1-r)} \mapsto y$, we can deduce that

$$I_1(\alpha) = \int_0^{1/2} \mathrm{d}r \int_0^{\frac{\alpha}{r(1-r)}} \mathrm{d}y\, g_1(y)\, \delta_\varphi\left(r, [r(1-r)]^{1/2} y\right).$$

Moreover, $r(1-r) \leq 1/4$ for every $r \in [0, 1/2]$. Some elementary computations thus imply a new decomposition for $I_1(\alpha)$ of the form

$$I_1(\alpha) = I_{1,1}(\alpha) + I_{1,2}(\alpha) \tag{3.20}$$

where

$$I_{1,1}(\alpha) := \int_0^{1/2} \mathrm{d}r \int_0^{2\alpha} \mathrm{d}y\ g_1(y)\delta_\varphi\left(r, [r(1-r)]^{1/2} y\right), \tag{3.21}$$

$$I_{1,2}(\alpha) := \int_{2\alpha}^\infty \mathrm{d}y\ g_1(y) \int_0^{\rho(\alpha,y)} \mathrm{d}r\ \delta_\varphi\left(r, [r(1-r)]^{1/2} y\right), \tag{3.22}$$

and the function ρ being defined as

$$\rho(\alpha, y) := \frac{1}{2}\left(1 - \sqrt{1 - (2\alpha/y)^2}\right). \tag{3.23}$$

In particular in the term $I_{1,2}(\alpha)$ of the previous decomposition, we have exchanged the integration order in r and y.

Let us analyze the term $I_{1,1}(\alpha)$ in (3.21). In view of (3.13) and $g_1(y)$ as defined in (3.19) with $\tau = 1$, some easy considerations on definite integrals yield

$$\lim_{\alpha\downarrow 0}\frac{I_{1,1}(\alpha)}{\mu(K_\alpha)} = \lim_{\alpha\downarrow 0}\frac{\sqrt{2/\pi}}{1-e^{-2\alpha^2}}\int_0^{2\alpha} \mathrm{d}y\, y^2 e^{-y^2/2}\int_0^{1/2} \mathrm{d}r\delta_\varphi\left(r, [r(1-r)]^{1/2}y\right) = 0. \tag{3.24}$$

As far as the term $I_{1,2}(\alpha)$ in (3.22) is concerned, it still deserves a detailed analysis for which we can refer to [53]. Let us just mention that, since $\rho(\alpha, y)$ as spelled out in (3.23) is approximately $(\alpha/y)^2$ whenever $\alpha \to 0$, then

$$\begin{aligned}\lim_{\alpha\downarrow 0}\frac{I_{1,2}(\alpha)}{\mu(K_\alpha)} &= \lim_{\alpha\downarrow 0}\frac{\sqrt{2/\pi}}{1-e^{-2\alpha^2}}\int_{2\alpha}^{\infty} \mathrm{d}y\; y^2 e^{-y^2/2}\int_0^{(\alpha/y)^2} \mathrm{d}r\;\delta_\varphi\left(r, [r(1-r)]^{1/2}y\right)\\ &= \lim_{\alpha\downarrow 0}\left(\frac{\sqrt{2/\pi}}{2\alpha^2}\alpha^2\int_{2\alpha}^{\infty} \mathrm{d}y\; e^{-y^2/2}(y/\alpha)^2\int_0^{(\alpha/y)^2} \mathrm{d}r\right)\\ &\quad\times\lim_{r\downarrow 0}\delta_\varphi\big(r, [r(1-r)]^{1/2}y\big)\\ &= \frac{1}{2}\int_{K_0}\varphi(\xi)\nu(\mathrm{d}\xi),\end{aligned} \tag{3.25}$$

where we recall that ν is the law of the 3-d Bessel bridge. The last identity in relation (3.25) stems from the fact that

$$\lim_{r\downarrow 0}\delta_\varphi\left(r, [r(1-r)]^{1/2}y\right) = \lim_{r\uparrow 1}\delta_\varphi\left(r, [r(1-r)]^{1/2}y\right) = \int_{K_0}\varphi(\xi)\nu(\mathrm{d}\xi).$$

The detailed proof the aforementioned result can be found in [53, Lemma 6]. Intuitively, this result arises because as $r \downarrow 0$ or $\uparrow 1$, the conditioning event $\{\xi_r = [r(1-r)]^{1/2}\, y\}$ approaches the entire sample space.

Summarizing our considerations, plugging (3.24) and (3.25) into (3.20), we find that

$$\lim_{\alpha\downarrow 0}\frac{I_1(\alpha)}{\mu(K_\alpha)} = \frac{1}{2}\int_{K_0}\varphi(\xi)\nu(\mathrm{d}\xi). \tag{3.26}$$

As mentioned above, $I_2(\alpha)$ in (3.18) can be handled in the same way. Reporting (3.26) into the decomposition (3.16), this concludes the proof of Theorem 3.1.

Remark 3.5 As seen from our sketch of the proof for Theorem 3.1, the identification of an invariant measure for the stochastic heat equation reflected at 0 hinges upon several key components. First, it relies on identifying an invariant measure for the Ornstein-Uhlenbeck process and associated gradient systems (this step is borrowed again from Da Prato-Zabczyk's series of books, see [21, Theorem 6.2.1 and Section

8.6]). Additionally, the relationship between the 1-d Brownian bridge conditioning on nonnegative sample paths and the 3-d Bessel bridge (see Proposition 3.3) plays a crucial role, which provides a nice representation of the invariant measure as in Theorem 3.1. The invariant measure presented in Theorem 3.1 is unique, which can be deduced by comparing the asymptotic behaviors of $X_t^{\xi_1}$ and $X_t^{\xi_2}$, which are solutions to (3.3) reflected at 0 with initial datum ξ_1 and ξ_2 respectively. While Zambotti's original work [53] does not explicitly state this uniqueness, it was later affirmed in a broader context by Yang and Zhang [52, Theorem 2.2] by adopting a coupling method from Mueller [45]. Regarding the equation (3.3), Zambotti and his collaborators have conducted additional studies, such as deriving an integration by parts formula with respect to the 3-d Bessel bridge in [54], and exploring fine properties of the contact set $\{(t, x); u(t, x) = 0\}$ in [26, 55]. In addition to the aforementioned works, further investigations into ergodicity and invariant measures for stochastic PDEs with reflection have been pursued in various scholarly articles, see, e.g., [39, 47, 58].

4 Ergodicity with degenerate noise

We will now examine a second departure from the Da Prato-Zabczyk standard setting, which concerns degenerate situations for the noise. Namely we have mentioned that uniqueness and ergodicity of the invariant measure occurs under the non-degeneracy assumption (2.16). In particular, this implies that W should be truly infinite-dimensional (or otherwise stated $\mathrm{Rank}(BB^*) = \infty$).

In this section we will describe a celebrated result by Hairer and Mattingly [37] in which ergodicity is achieved in spite of a finite-dimensional noise. The result in [37] is proved for stochastic Navier-Stokes equations, which are well-known systems describing the time evolution of an incompressible fluid. We consider here a version of the equation for $x \in \mathbb{R}^2$, written with the notation of Section 2 as

$$\mathrm{d}u_t + (u_t \cdot \nabla)\, u_t = \nu \triangle u_t - \nabla p_t + \xi_t, \quad \mathrm{div}\, u = 0, \tag{4.1}$$

where $u_t(x) \in \mathbb{R}^2$ denotes the value of the velocity field at time t and position x, $p_t(x)$ denotes the pressure, and $\xi_t(x)$ is an external force field acting on the fluid. In the stochastic setting the driving force ξ is a Gaussian field which is white in time and colored in space.

The existence of an invariant measure for the stochastic PDE (4.1) can be proved by some "soft" techniques using the regularization and dissipativity properties of the flow [20, 32]. However, showing the uniqueness of the invariant measure is a challenging problem for at least two reasons:

(i) The proof of uniqueness requires a detailed analysis of the nonlinearity.
(ii) In fact the nonlinearity in (4.1) can balance the degeneracy of W and produces the unique invariant measure. It is thus really at the heart of the analysis.

This section is devoted to a brief explanation of those mechanisms. It is structured as follows: we will first fix some notation and state the main result of [37] in Section 4.1. Section 4.2 will then be devoted to a brief explanation of the main idea in the proof.

4.1 Setup and ergodicity of Navier-Stokes equations

the framework used in [37] is the so-called vorticity formulation of the Navier-Stokes equation, which will yield a formulation close to (2.7). Consider (4.1) on the torus $\mathbb{T}^2 := [-\pi, \pi]^2$ driven by a Gaussian noise ξ. For a divergence-free velocity field, the vorticity X is defined by $X := \nabla \wedge u = \partial_2 u_1 - \partial_1 u_2$. Note that the velocity and vorticity formulations are equivalent since u can be recovered from X and the condition $\nabla \cdot u = 0$. With this notation it can be proved that the vorticity formulation for the stochastic Navier-Stokes equation is

$$\mathrm{d}X_t = \nu \Delta X_t \, \mathrm{d}t + B(X, X) \, \mathrm{d}t + Q \, \mathrm{d}W_t, \qquad X_0 = x_0 \,, \tag{4.2}$$

where Δ is the Laplacian with periodic boundary conditions and $B(u, w) = -(u \cdot \nabla) w$ is the usual Navier-Stokes nonlinearity. In what follows, we will focus on the vorticity formulation of the problem given by equation (4.2). Our state space for X will be $H = L_0^2(\mathbb{T}^2)$, the space of real-valued square-integrable functions on the torus with vanishing mean.

We now turn to a description of the finite dimensional noise $Q \mathrm{d}W_t$ in (4.2). To this aim, we start by introducing a convenient way to index the Fourier basis of H. Namely we write $\mathbb{Z}^2 \backslash \{(0,0)\} = \mathbb{Z}_+^2 \cup \mathbb{Z}_-^2$, where

$$\mathbb{Z}_-^2 := \left\{ (k_1, k_2) \in \mathbb{Z}^2 | -k \in \mathbb{Z}_+^2 \right\} \quad \text{and}$$
$$\mathbb{Z}_+^2 := \left\{ (k_1, k_2) \in \mathbb{Z}^2 | k_2 > 0 \right\} \cup \left\{ (k_1, 0) \in \mathbb{Z}^2 | k_1 > 0 \right\}.$$

Our set of oscillatory functions is defined in the following way for all $k \in \mathbb{Z}^2 \backslash \{(0,0)\}$: for all $x \in \mathbb{T}^2$ we set

$$f_k(x) = \begin{cases} \sin(k \cdot x), & \text{if } k \in \mathbb{Z}_+^2; \\ \cos(k \cdot x), & \text{if } k \in \mathbb{Z}_-^2. \end{cases} \tag{4.3}$$

For the remainder of the section, we also fix a set

$$\mathcal{Z}_0 = \{k_n; n = 1, \ldots, m\} \subset \mathbb{Z}^2 \backslash \{(0,0)\}, \tag{4.4}$$

which corresponds to the set of driving modes of equation (4.2). We are now ready to introduce our finite dimensional noise.

Definition 4.1 In equation (4.2), the finite dimensional noise QW_t is defined as follows: we start from a $\mathbb{R}^m$-valued Wiener process W defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that f_k is given in (4.3) and denote by $\{e_n; n = 1, \ldots, m\}$ the standard basis of $\mathbb{R}^m$. Define $Q : \mathbb{R}^m \to H$ by $Qe_n = q_n f_{k_n}$, where the q_n are some strictly positive numbers, and the wave numbers k_n are given by the set $\mathcal{Z}_0$. Then for

$t \geq 0$ and $x \in \mathbb{T}^2$ we have

$$QW_t(x) = \sum_{n=1}^{m} q_n W_t^n f_{k_n}(x). \tag{4.5}$$

Notice that QW_t is an element of $H = L_0^2(\mathbb{T}^2)$.

The main result of [37], achieving uniqueness for the invariant measure in spite of having a finite-dimensional noise in (4.5), can now be summarized as follows.

Theorem 4.2 *Let $\mathcal{Z}_0$ satisfy the following assumptions:*

A1. *There exist at least two elements in $\mathcal{Z}_0$ with different Euclidean norms.*
A2. *The set of integer linear combinations of elements of $\mathcal{Z}_0$ generates $\mathbb{Z}^2$.*

Then equation (4.2) has a unique invariant measure on H.

Remark 4.3 Condition **A2** above is equivalent to the easily verifiable condition that the greatest common divisor of the set $\{\det(k, l) : k, l \in \mathcal{Z}_0\}$ is 1, where $\det(k, l)$ is the determinant of the 2×2 matrix with columns k and l.

Example 4.4 Let $\mathcal{Z}_0 = \{(1, 0), (-1, 0), (1, 1), (-1, -1)\}$. It is clear that $\mathcal{Z}_0$ satisfies the assumptions of Theorem 4.2. Therefore, equation (4.2) is ergodic with degenerate driving noise

$$\begin{aligned} QW(t, x) = {} & W_1(t) \sin(x_1) + W_2(t) \cos(x_1) + W_3(t) \sin(x_1 + x_2) \\ & + W_4(t) \cos(x_1 + x_2). \end{aligned}$$

4.2 Main idea in the proof of Theorem 4.2

The general criterion for uniqueness of the stationary measure is taken again from Da Prato-Zabczyk's monograph [21]. Namely it is a well-known and much-used fact that the strong Feller property, combined with some irreducibility of the transition probability, implies the uniqueness of the invariant measure. Denote by P_t the underlying semigroup as featured in (2.10). In order to establish the strong Feller property of P_t, one usually resorts to an integration by parts argument in the Malliavin calculus sense, which we now describe.

4.2.1 Malliavin calculus approach in a non-degenerate case

Having in mind that the strong Feller property is a crucial step towards uniqueness of the invariant measure, let us recall a basic criterion allowing to establish this type of result (see [21, Lemma 7.1.5]).

Proposition 4.5 *A semigroup P_t on a Hilbert space H is strong Feller if, for all $\varphi : H \to \mathbb{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty$ finite and for all $t > 0$ one has*

$$|\nabla P_t \varphi(x)| \leq C(\|x\|) \, \|\varphi\|_\infty, \tag{4.6}$$

where $C : \mathbb{R}_+ \to \mathbb{R}$ *is a fixed nondecreasing function.*

Interestingly enough, strong Feller properties can be deduced from Malliavin calculus considerations. To this aim, let us introduce some additional notation. First denote by $\Phi_t : C([0,t];\mathbb{R}^m) \times H \to H$ the Itô map such that the solution to (4.2) can be written $X_t = \Phi_t(W, X_0)$ for every initial condition $X_0 \in H$ and almost every realization of W. Then the infinitesimal variation of X_t with respect to a perturbation $\xi \in H$ of the initial condition is given by

$$J_t\xi := \lim_{\varepsilon\to 0} \frac{\Phi_t(W, X_0 + \varepsilon\xi) - \Phi_t(W, X_0)}{\varepsilon}. \tag{4.7}$$

In addition, the infinitesimal variation of X_t with respect to a perturbation of the Wiener process in the direction of $V(s) = \int_0^s v(r)\mathrm{d}r$ with $v \in L^2([0,T];H)$ is given by the Malliavin derivative

$$\mathcal{D}^v X_t = \lim_{\varepsilon\to 0} \frac{\Phi_t(W + \varepsilon V, X_0) - \Phi_t(W, X_0)}{\varepsilon}. \tag{4.8}$$

We now state the lemma relating Malliavin derivatives and the strong Feller property. Notice that it is proved here in a rather informal way.

Lemma 4.6 *Let us assume that for every direction* $\xi \in H$ *one can find* $v = v(\xi)$ *lying in the space* $L^2([0,T],H)$*, such that*

$$\mathcal{D}^v X_t = J_t\xi.$$

Then the semigroup P_t *enjoys the strong Feller property as given in Proposition 4.5.*

Proof By definition of the semigroup P_t, we have

$$\langle \nabla P_t\varphi(X_0), \xi\rangle = \mathbb{E}\left[\nabla\varphi(X_t)\, J_t\xi\right] = \mathbb{E}\left[\nabla\varphi(X_t)\, \mathcal{D}^v X_t\right],$$

where the second identity stems from our assumption $\mathcal{D}^v X_t = J_t\xi$. Now it is easily seen that $\nabla\varphi(X_t)\mathcal{D}^v X_t = \mathcal{D}^v(\varphi(X_t))$. Hence a standard integration by parts in the Malliavin calculus sense yields

$$\langle \nabla P_t\varphi(X_0), \xi\rangle = \mathbb{E}\left[\varphi(X_t)\int_0^t v(s)\mathrm{d}W_s\right].$$

Therefore we get

$$|\langle \nabla P_t\varphi(X_0), \xi\rangle| \le \|\varphi\|_\infty \|v\|_{L^2([0,T],H)},$$

from which inequality (4.6) is easily deduced. This concludes the proof of Lemma 4.6.

Thanks to Lemma 4.6, the strong Feller property for X_t is reduced to a relation involving its Malliavin derivative. However, the ability to find a v such that $\mathcal{D}^v X_t = J_t \xi$ relies on some invertibility of the Malliavin matrix on a proper space. This requirement poses some major technical difficulty in an infinite dimensional setting. In addition, a more fundamental question is whether one should expect the strong Feller property for an infinitely dimensional system at all when the noise is degenerate. Indeed, it seems that the only result showing the strong Feller property for an infinite dimensional system where the covariance of the noise does not have a dense range is given in [37]. However, this still requires the forcing to act in a non-degenerate way on a subspace of finite codimension. A different approach is thus necessary for degenerate noises like the one described in Definition 4.1.

4.2.2 Malliavin calculus approach in the degenerate case

A key observation of Hairer and Mattingly [37] is that the strong Feller property is neither essential nor natural for the study of ergodicity in dissipative infinite-dimensional systems. To provide an alternative, they introduced the following weaker *asymptotic strong Feller* property which is satisfied by the system under consideration and is sufficient to give ergodicity.

Let d be a pseudo-metric on $\mathcal{X}$. Given two positive finite Borel measures μ_1, μ_2 on $\mathcal{X}$ with equal mass, denote by $\mathcal{C}(\mu_1, \mu_2)$ the set of positive measure on $\mathcal{X}^2$ with marginals μ_1 and μ_2. Define

$$\|\mu_1 - \mu_2\|_d := \inf_{\mu \in \mathcal{C}(\mu_1,\mu_2)} \int_{\mathcal{X}^2} d(x, y)\mu(\mathrm{d}x, \mathrm{d}y).$$

Definition 4.7 A Markov transition semigroup P_t on a Polish space $\mathcal{X}$ is called asymptotically strong Feller at x if there exists a totally separating system of pseudo-metrics $\{d_n; n \geq 1\}$ for $\mathcal{X}$ and a sequence $\{t_n; n \geq 1\}$ with $t_n > 0$ such that

$$\liminf_{U \in \mathcal{U}_x} \limsup_{n \to \infty} \sup_{y \in U} \left\| P_{t_n}(x, \cdot) - P_{t_n}(y, \cdot) \right\|_{d_n} = 0,$$

where $\mathcal{U}_x$ is the collection of all open sets containing x. It is called asymptotically strong Feller if this property holds at every $x \in \mathcal{X}$.

Remark 4.8 It is proved in [37, Corollary 3.5] that if $\{d_n; n \geq 1\}$ is a total separating system of pseudo-metrics for $\mathcal{X}$, then $\|\mu_1 - \mu_2\|_{\mathrm{TV}} = \lim_{n\to\infty} \|\mu_1 - \mu_2\|_{d_n}$ for any two positive measures μ_1 and μ_2 with equal mass on $\mathcal{X}$.

Remark 4.9 Note that when $t_n = t$ for all n, the transition probabilities $P_t(x, \cdot)$ are continuous in the total variation topology and thus P_s is strong Feller at times $s \geq t$. In order to show (4.2) satisfies the asymptotic strong Feller property, one will take $t_n = n$ as will be seen below.

The following criterion for the asymptotic strong Feller property, generalizing Proposition 4.5, is then proposed in [37].

Proposition 4.10 *Let $\{t_n; n \geq 1\}$ and $\{\delta_n; n \geq 1\}$ be two positive sequences with $\{t_n\}$ nondecreasing and $\{\delta_n\}$ converging to zero. A semigroup P_t on a Hilbert space $\mathcal{H}$ is asymptotic strong Feller if, for all $\varphi : \mathcal{H} \to \mathbb{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty$ finite, we have*

$$\left|\nabla P_{t_n}\varphi(x)\right| \leq C(\|x\|)\left(\|\varphi\|_\infty + \delta_n \|\nabla\varphi\|_\infty\right), \tag{4.9}$$

where $C : \mathbb{R}_+ \to \mathbb{R}$ is a fixed nondecreasing function.

Remark 4.11 The asymptotic strong Feller property can be related to Malliavin calculus notions. The point is that the Malliavin matrix is not invertible in our degenerate context. Therefore we are not able to construct a $v \in L^2([0, T], \mathbb{R}^m)$ for a fixed time T that produces the same infinitesimal shift in the solution as a perturbation ξ in the initial condition. Instead, we can construct a $v \in L^2([0, \infty); \mathbb{R}^m)$ such that an infinitesimal shift of the noise in the direction v provides asymptotically the same effect as an infinitesimal perturbation in the direction ξ. In other words, one has $\|J_t\xi - \mathcal{D}^{v_{0,t}} X_t\| \to 0$ as $t \to \infty$, where $v_{0,t}$ is the restriction of v on the interval $[0, t]$. Set $\rho_t := J_t\xi - \mathcal{D}^{v_{0,t}} X_t$. Then one has the approximate integration by parts formula:

$$\begin{aligned}\langle \nabla P_t\varphi(X_0), \xi\rangle &= \mathbb{E}\left((\nabla\varphi)(X_t) J_t\xi\right)\\ &= \mathbb{E}\left((\nabla\varphi)(X_t)\mathcal{D}^{v_{0,t}} X_t\right) + \mathbb{E}\left((\nabla\varphi)(X_t)\rho_t\right)\\ &= \mathbb{E}\left(\varphi(X_t)\int_0^t v(s)\mathrm{d}W_s\right) + \mathbb{E}\left((\nabla\varphi)(X_t)\rho_t\right),\end{aligned} \tag{4.10}$$

from which it follows that

$$|\langle \nabla P_t\varphi(X_0), \xi\rangle| \leq \|\varphi\|_\infty \times \mathbb{E}\left(\left|\int_0^t v(s)\mathrm{d}W_s\right|\right) + \|\nabla\varphi\|_\infty \times \mathbb{E}\left(\|\rho_t\|\right).$$

If $\mathbb{E}(\|\rho_t\|)$ is small, one easily derives (4.9) from (4.10). Notice that the explicit construction of v in (4.10) is highly non-trivial and rather technical. We refer the interested readers to the original paper for details. Instead, we make the following remark regarding the construction of v.

Remark 4.12 In order to construct a suitable v one needs to better incorporate the pathwise smoothing which the dynamics possesses at small scales. Due to the degenerate nature of the noise, v will be constructed as a non-adapted process. The first term in (4.10) is thus understood as a Skorohod integral. In order to provide a good estimate of this term, one has to have good control of the "low modes" when they are not directly forced by the noise and when Girsanov's theorem cannot be used directly (cf. Theorem 4.12 in [37]). This is the heart of the analysis for the structure of the Malliavin matrix for equation (4.2). It exploits the algebraic structure of the nonlinearity, which transmits the randomness to the non-directly exited unstable directions. This results in an associated diffusion which in the end is hypoelliptic (see [42] for more details).

With Remarks 4.11 and 4.12 in hand, let us conclude by stating a proposition which yields the asymptotic Feller property for the stochastic Navier-Stokes equation.

Proposition 4.13 *Let P_t be the semigroup related to equation (4.2). Then for all $\eta > 0$, there exist constants $C, \delta > 0$ such that for every Fréchet differentiable function φ from H to $\mathbb{R}$ one has the bound*

$$\|\nabla P_n\varphi(X_0)\| \leq C \exp\left(\eta\|X_0\|^2\right) \times \left(\|\varphi\|_\infty + \|\nabla\varphi\|_\infty e^{-\delta n}\right), \tag{4.11}$$

for every $X_0 \in H$ and $n \in \mathbb{N}$. Otherwise stated, the asymptotic strong Feller property of Proposition 4.10 holds true for equation (4.2) by considering $t_n = n$ and $\delta_n = e^{-\delta n}$.

We close this section by recalling that the asymptotic strong Feller property stated in Proposition (4.13) leads to ergodicity. We skip those considerations for sake of conciseness and we refer to [37] for more details. Also notice the reference [41], based on control type arguments.

5 Phase transition in ergodicity in $\mathbb{R}^d$ with $d \geq 3$

This section is devoted to another very important degenerate case for equation (2.8). That is, instead of considering a degenerate noise $\dot{W}$ in (2.8), we will assume that the diffusion coefficient B can vanish. This situation occurs in particular when considering an important system called *parabolic Anderson model* (referred to as PAM in the sequel). Our main message here is that in $\mathbb{R}^d$ when $d \geq 3$, one can observe phase transitions (in terms of the noise intensity) between a situation where an invariant measure exists and a very different case with no invariant measure. We begin with a series of preliminary remarks in Section 5.1. Then we state some phase transition results for the parabolic Anderson model and related equations in Section 5.2. Eventually we will derive a new result about convergence in distribution to the invariant measure in Section 5.3.

5.1 Preliminary remarks

So far we have expressed our stochastic PDE's in the Da Prato-Zabczyk infinite dimensional setting. However, most of the results concerning parabolic Anderson models use the multiparametric setting for SPDEs popularized by Walsh [51] and Dalang [24, 25]. The current section will thus use this framework. The reader is referred to [27] for a correspondence between the infinite-dimensional and the multiparametric settings. With this notational warning in mind, in this section we review existence results for the invariant measure of the following stochastic heat equation in $\mathbb{R}^d$ with $d \geq 3$:

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right) u(t,x) = b(x, u(t,x))\dot{W}(t,x), \qquad \text{for } x \in \mathbb{R}^d \text{ and } t > 0. \tag{5.1}$$

In equation (5.1), $b(x,u)$ is uniformly bounded in the first variable and globally Lipschitz continuous in the second variable, i.e., for some constants $L_b > 0$ and $L_0 \geq 0$,

$$|b(x,u) - b(x,v)| \leq L_b|u - v| \quad \text{and} \quad |b(x,0)| \leq L_0 \quad \text{for all } u, v \in \mathbb{R} \text{ and } x \in \mathbb{R}^d. \tag{5.2}$$

Notice that our hypothesis (5.2) allows the degenerate value $b(x,0) = 0$. In particular, it covers the linear case $b(x,u) = \lambda u$, which is the aforementioned PAM (see [8]). In this note, we will specifically focus on the degenerate case:

$$b(x,0) \equiv 0, \quad \text{for all} \quad x \in \mathbb{R}^d. \tag{5.3}$$

The noise $\dot{W}(t,x)$ featured in equation (5.1) is defined in a standard way within the random field framework for stochastic PDE's. Specifically, $\dot{W}$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$ generated by the noise, it is a centered Gaussian noise that is white in time and homogeneously colored in space. Its covariance structure is given by

$$\mathbb{E}\left[W(\psi)\, W(\phi)\right] = \int_0^\infty \mathrm{d}s \int_{\mathbb{R}^d} \Gamma(\mathrm{d}x)(\psi(s,\cdot) * \widetilde{\phi}(s,\cdot))(x), \tag{5.4}$$

where ψ and ϕ are continuous and rapidly decreasing functions, $\widetilde{\phi}(x) := \phi(-x)$, "$*$" refers to the convolution in the spatial variable, and Γ is a nonnegative and nonnegative definite tempered measure on $\mathbb{R}^d$ that is commonly referred to as the *correlation measure*. The Fourier transform[1] of Γ (in the generalized sense) is also a nonnegative and nonnegative definite tempered measure, which is usually called the *spectral measure* and is denoted by $\widehat{f}(\mathrm{d}\xi)$. Moreover, in the case where Γ has a density f, namely $\Gamma(\mathrm{d}x) = f(x)\mathrm{d}x$, we write $\widehat{f}(\mathrm{d}\xi)$ as $\widehat{f}(\xi)\mathrm{d}\xi$. Existence and uniqueness of (5.1) for bounded (resp. rough) initial conditions were established in [24] (resp. [16]) under *Dalang's condition*

$$\Upsilon(\beta) := (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{\widehat{f}(\mathrm{d}\xi)}{\beta + |\xi|^2} < \infty \quad \text{for some (and hence all) } \beta > 0. \tag{5.5}$$

Moreover, existence and uniqueness results in this setting rely on Itô type estimates which are elaborations of (5.4). Namely, for an adapted process $v = \{v(t,x); t \geq 0, x \in \mathbb{R}^d\}$, we have

$$\mathbb{E}\left[\left(\int_s^t \int_{\mathbb{R}^d} v(r,y) W(\mathrm{d}r,\mathrm{d}y)\right)^2\right] = \int_s^t \mathrm{d}r \iint_{\mathbb{R}^{2d}} f(y-y')\, \mathbb{E}[v(r,y)v(r,y')]\, \mathrm{d}y\mathrm{d}y'. \tag{5.6}$$

Ideally one would like to consider general space-time covariance structures like in [38]. However fractional dependences in time yield a much more complicated picture in terms of ergodic behavior, as assessed e.g in [28].

[1] We use the following convention and notation for Fourier transform: $\mathcal{F}\psi(\xi) = \widehat{\psi}(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\psi(x)\mathrm{d}x$ for any Schwarz test function $\psi \in \mathcal{S}(\mathbb{R}^d)$.

In the literature (starting from Da Prato-Zabczyk's contributions summarized in Section 2.2), the existence of invariant measures for SHE is often studied with a drift term and a general second order differential operator $\mathcal{A}$:

$$\left(\frac{\partial}{\partial t} - \mathcal{A}\right) u(t,x) = g(x, u(t,x)) + b(x, u(t,x))\dot{W}(t,x), \quad \text{for } x \in \mathcal{O}, t > 0. \tag{5.7}$$

In order to control the growth of the solution's moments, both the drift term g and the differential operator $\mathcal{A}$ (paired with its domain $\mathcal{O}$ and boundary conditions) have to be sufficiently dissipative to counterbalance the diffusion part that is governed by the diffusion coefficient b and the noise correlation function/measure f. We claim that the setup as given in (5.1) (or the one in Chen and Eisenberg [12]) is among the most challenging ones. This is because:

(i) The semigroup for a bounded domain $\mathcal{O} \subseteq \mathbb{R}^d$ usually has stronger contraction properties than that for the whole space $\mathbb{R}^d$. This bounded domain assumption is the setup for Cerrai [9, 10], Brzeźniak and Gątarek [7], and Bogachev and Röckner [6]. One can find more related references from the above three papers. For the random field approach, Mueller [45] studied the equation on a torus with the operator $\frac{\partial^2}{\partial x^2} - \alpha$, with $\alpha > 0$. This parameter α provides an exponential dissipative effect; see Theorem 1.2 (*ibid.*).

In the following, we will focus on works where the spatial domain is the entire space $\mathbb{R}^d$. There are much fewer works in this setting; notably, we will comment on the following references [2, 12, 30, 36, 43, 44, 50]. Most of the works in this line have been carried out in some weighted $L^2(\mathbb{R}^d)$ space with a weight function ρ, referred as $L^2_\rho(\mathbb{R}^d)$. This setup was initially introduced by Tessitore and Zabczyk [50]. However, there are two exceptions: Eckmann and Hairer used the $L^\infty(\mathbb{R})$ space and Gu and Li [36] used the space of continuous functions denoted by $C\left(\mathbb{R}^d\right)$.

(ii) It is, in general, much harder to handle the degenerate diffusion coefficient case: $b(u) = u$. Eckmann and Hairer [30] studied the SHE on $\mathbb{R}$, albeit with an additive noise (that is $b(\cdot) \equiv 1$ in (5.7)). In that case, a nonlinear drift term of the form $g(u) = u(1-u^2)$ contributes to a sufficient amount of dissipativeness. A different setting is provided by Misiats, Stanzhytskyi *et al* [44]. Namely, they assume that the diffusion coefficient b satisfies the following condition:

$$|b(x, u_1) - b(x, u_2)| \le L\,\varphi(x)|u_1 - u_2|,$$

for some function $\varphi(x)$ that decays fast enough. In a slightly earlier paper of Misiats, Stanzhytskyi *et al* [43], they also require b to be bounded.

Now we will further concentrate on works that allow the degenerate case $b(u) = u$ in order to be able to cover the parabolic Anderson model.

(iii) Lacking a dissipative drift term makes the problem more challenging. Indeed, Assing and Manthey's work [2] allows $b(u) = u$. Within their framework, they can afford a space-time white noise in $d = 1$. However, a trace-class noise (i.e., Q in (2.2) being a trace-class operator) has to be considered in $d \geq 2$. Such general setup is achieved by a strong dissipative condition on the drift term g in the following form: for some constants C and $\kappa > 0$, it holds that

$$ug(u) \leq C - \kappa u^2, \quad \text{for all } u \in \mathbb{R}.$$

In particular, g cannot vanish.

Now let us further narrow down to works without a drift term. In this case, only the following papers are left: [12, 36, 50]. In all these works, without the help of the drift term, one has to use the weak dissipative property of the heat kernel on the whole space coming from the factor $t^{-d/2}$ of the heat kernel in $\mathbb{R}^d$. If we want this dissipative effect to be strong enough, this imposes $d \geq 3$.

(iv) The analysis of a model like PAM is simplified when one considers a trace-class noise in (5.1). This is what is assumed in Gu and Li [36]: the spatial covariance measure Γ in (5.4) is of the form $\Gamma(\mathrm{d}x) = f(x)\mathrm{d}x$, with

$$f(x) = \int_{\mathbb{R}^d} \phi(x+y)\phi(y)\mathrm{d}y, \quad \text{for some } \phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}_+)\,. \tag{5.8}$$

This ensures that the noise is trace-class, since $\int_{\mathbb{R}^d} \widehat{f}(\xi)\mathrm{d}\xi = f(0) = \|\phi\|^2_{L^2(\mathbb{R}^d)} < \infty$. It is much more challenging to include more singular noises which are not trace-class. This is achieved in Tessitore and Zabczyk [50], where the conditions on the noise are given by Hypothesis 2.1 (i) and especially (3.4) of their paper. However, those conditions are very involved and sometimes difficult to check; see Section 5.2 of [12] for a detailed discussion. In contrast, the work by Chen and Eisenberg [12] provides the following concise condition on the correlation function in order to get a phase transition for the existence of an invariant measure:

$$\int_{\mathbb{R}^d} \frac{\widehat{f}(\mathrm{d}\xi)}{|\xi|^2 \wedge |\xi|^{2(1-\alpha)}} < \infty, \quad \text{for some } \alpha \in \left(4\Upsilon(0)L_b^2, 1\right), \tag{5.9}$$

where L_b is the Lipschitz constant as given in (5.2); see Remark 1.2 or (1.19) (*ibid.*). Note that the above condition (5.9) implicitly requires that

$$\Upsilon(0) := (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{\widehat{f}(\mathrm{d}\xi)}{|\xi|^2} < \infty. \tag{5.10}$$

Note also that the correlation in (5.8) satisfies condition (5.9), by noticing that $\widehat{\phi}$ is a Schwarz function. It is worth pointing out that the trace-class noise condition $f(0) < \infty$ neither implies nor is implied by condition (5.9). In particular, one easily finds some noises that are not trace-class but satisfy condition (5.9). We briefly discuss Bessel and Riesz kernels below.

(a) Let $f_s(\cdot)$ denote the family of *Bessel kernels*[2] *with parameter* $s > 0$ on $\mathbb{R}^d$, i.e.,

$$f_s(x) = \mathcal{F}^{-1}\left[\frac{1}{\left(1+|\xi|^2\right)^{s/2}}\right](x), \quad \text{for } x \in \mathbb{R}^d.$$

Since f_s is both nonnegative and nonnegative definite, it can be served either as correlation function or the spectral measure. On the one hand (see Proposition 5.2 of [12]), assuming $d \geq 3$ and using f_s as the correlation function, condition (5.10) (resp. condition (5.9)) is satisfied for all $s > d-2$ (resp. $s > d - 2(1-\alpha)$). On the other hand, f_s is a trace-class operator if and only if $s > d$ since

$$\int_{\mathbb{R}^d} \widehat{f_s}(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^d} \left(1+|\xi|^2\right)^{-s/2} \mathrm{d}\xi < \infty \quad \Longleftrightarrow \quad s > d.$$

Hence, all Bessel kernels $f_s(\cdot)$ with $s \in (d-2(1-\alpha), d]$ provide examples of correlation functions that are not trace-class operators but can ensure existence of a nontrivial invariant measure. On the other hand, when using the Bessel kernel f_s as the spectral measure, as explained in Proposition 5.3 of [12], assuming again $d \geq 3$, conditions (5.10) and (5.9) are satisfied for all $s > 2$ and for all $s > 2(1-\alpha)$, respectively. Hence, correlation functions $\widehat{f_s}(x) = \left(1+|x|^2\right)^{-s/2}$ with $s > 2(1-\alpha)$ provide examples of trace-class operators that have long-range heavy-tail correlations, in contrast to the finite-range correlation in (5.8). Note that f_s, used as correlation functions, also provides long-range correlation but with a light tail, i.e., some exponential tail.

(b) The Riesz kernel $f(x) = |x|^{-\beta}$ plays a prominent role of in the study of the homogeneous Gaussian noise, since it provides heavy-tailed correlations and is singular at zero. In particular, it is not a trace-class operator. However, it is easy to check that it won't satisfy either conditions (5.9) or (5.10). Example 5.10 of [12] provides kernel functions that have similar behaviors to the Riesz kernel with power blow-up near zero at some rate and power decay near infinity at possibly a different rate. To be more precise, for any s_1 and $s_2 \in (0, d)$, the *Riesz-type kernel with parameters* (s_1, s_2) is defined as a combination of the Bessel kernel and its Fourier transform:

$$f_{s_1,s_2}(x) := f_{s_1}(x) + \widehat{f_{s_2}}(x) \quad \text{or equivalently} \quad \widehat{f_{s_1,s_2}}(\xi) := \widehat{f_{s_1}}(\xi) + f_{s_2}(\xi). \tag{5.11}$$

It is clear that f_{s_1,s_2} is both nonnegative and nonnegative definite. From the properties of the Bessel kernel and its Fourier transform, we see that

$$f_{s_1,s_2}(x) \asymp \begin{cases} |x|^{s_1-d} & |x| \to 0, \\ |x|^{-s_2} & |x| \to \infty. \end{cases}$$

[2] *Interested readers can refer Section 1.2.1 of* [35] *for more details of the Bessel kernel.*

Assume $d \geq 3$. As explained in Example 5.10 of [12], condition (5.10) holds provided $s_1 > d-2$ and $s_2 > 2$ and condition (5.9) holds provided that $s_1 > d-2(1-\alpha)$ and $s_2 > 2(1-\alpha)$. Similar to the Bessel kernel case, $f_{s_1,s_2}(x)$ is a trace-class operator if and only if $s_1 > d$. Let us assume $d \geq 3$. Then, when used as the correlation function, the Riesz-type kernel f_{s_1,s_2} with $s_1 \in (d-2(1-\alpha), d)$ and $s_2 > 2(1-\alpha)$ provides non trace-class operators that both have heavy-tail correlations and can ensure existence of a nontrivial invariant measure.

(v) Due to the importance of the Dirac delta initial condition [1] and the stationary initial condition (i.e., the two-sided Brownian motion as the initial data) pointed out in [4], it would be preferable to be able to include more general initial conditions, that are neither bounded at one point nor at infinity. Chen and Dalang [11] introduced the so-called *rough initial conditions* to cover these unbounded initial conditions; see also [14, 16]. To be more precise, a rough initial condition μ refers to a deterministic, locally finite, regular, signed Borel measure that satisfies the following integrability condition at infinity:

$$\int_{\mathbb{R}^d} |\mu|\,(\mathrm{d}x) \exp\left(-a|x|^2\right) < \infty \quad \text{for all } a > 0, \tag{5.12}$$

where $|\mu| = \mu_+ + \mu_-$ and $\mu = \mu_+ - \mu_-$ refer to the *Hahn decomposition* of the measure μ. The work by Chen and Eisenberg [12] allows the rough initial condition.

Remark 5.1 We would like to mention that, in terms of techniques, most of the above works are based on the Krylov-Bogoliubov theorem (see [21, 23, 40] as well as relations (2.10) and (2.11) above). However, the work by Gu and Li [36] uses a different argument by shifting the initial time to negative infinity. As a result, they obtain a stronger result. The convergence there is a strong convergence. The arguments rely critically on the stationarity of the solution $u(t,x)$, which is a consequence of the constant one initial condition. Nevertheless, Gu and Li pointed out that this constant initial condition can be perturbed by an $L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ function; see Remark 3.6 (*ibid.*). In particular, it is not clear in their framework whether the constant initial condition can be perturbed by a Dirac delta measure. We will go back to this approach in Section 5.3.

5.2 Phase transition for the moments and invariant measures

In this part, let us state precisely the phase transition phenomenon under conditions (5.10) and (5.9). Recall that the solution to (5.1) is understood as the mild solution:

$$u(t,x) = J_0(t,x) + \int_0^t \int_{\mathbb{R}^d} p(t-s, x-y)\, b(y, u(s,y)) W(\mathrm{d}s, \mathrm{d}y), \tag{5.13}$$

where the stochastic integral is interpreted as the Walsh integral ([24, 51]) and $J_0(t,x)$ refers to the solution to the homogeneous equation, namely,

$$J_0(t,x) = J_0(t,x;\mu) := \int_{\mathbb{R}^d} p(t,x-y)\mu(\mathrm{d}y) = [p_t * \mu](x). \tag{5.14}$$

Here and throughout the rest of the paper, we use $p(t,x)$, or sometimes $p_t(x)$, to denote the heat kernel: $p(t,x) = p_t(x) := (2\pi t)^{-d/2} \exp\left(-|x|^2/(2t)\right)$.

The first result asserts that under the cone condition (5.17) below, the second moments of the solution exhibit a phase transition depending on the noise intensity, namely, the values of L_b and l_p in (5.2) and (5.17), respectively. Let us first define the *upper and lower (moment) Lyapunov exponents of order* p ($p \geq 2$) by

$$\overline{m}_p(x) := \limsup_{t\to+\infty} \frac{1}{t} \log \mathbb{E}\left(|u(t,x)|^p\right), \quad \underline{m}_p(x) := \liminf_{t\to+\infty} \frac{1}{t} \log \mathbb{E}\left(|u(t,x)|^p\right). \tag{5.15}$$

Then we get the following bounds on the exponents $\sup_{x\in\mathbb{R}^d} \overline{m}(x)$ and $\inf_{x\in\mathbb{R}^d} \underline{m}(x)$ under hypothesis (5.10).

Theorem 5.2 *(Theorem 1.3 of [16]) Let u be the solution to* (5.1) *starting from an initial condition μ, which is a nonnegative Borel measure on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} e^{-\gamma|x|}\mu(\mathrm{d}x) < \infty\,, \quad \textit{for all } \gamma > 0. \tag{5.16}$$

Assume $b(x,u) = b(u)$ in (5.1) *satisfies the cone condition:*

$$l_b := \inf_{x\in\mathbb{R}} \frac{b(x)}{|x|} > 0. \tag{5.17}$$

Recall that the covariance of the noise is given by (5.4) *and is specified by a correlation measure f satisfying* (5.5). *Also recall that $\Upsilon(0)$ is given by* (5.10). *Then the following holds true:*
(i) *If $\Upsilon(0) < \infty$, then for some nonnegative constants $0 < \underline{\lambda}_c \leq \overline{\lambda}_c < \infty$, it holds that*

$$\begin{cases} \sup_{x\in\mathbb{R}^d} \overline{m}_2(x) = 0, & \textit{if } L_b < \underline{\lambda}_c, \\ \inf_{x\in\mathbb{R}^d} \underline{m}_2(x) > 0, & \textit{if } l_b > \overline{\lambda}_c, \end{cases} \tag{5.18}$$

where L_b is the Lipschitz constant as given in (5.2).
(ii) *If $\Upsilon(0) = \infty$, then $u(t,x)$ is fully intermittent, i.e., $\overline{m}_1(x) \equiv 0$ and* $\inf_{x\in\mathbb{R}^d} \underline{m}_2(x) > 0$.
(iii) *The following two conditions are equivalent:*

$$\Upsilon(0) < \infty \quad \Longleftrightarrow \quad d \geq 3 \quad \textit{and} \quad \int_{\mathbb{R}^d} \frac{f(z)}{|z|^{d-2}}\mathrm{d}z < \infty. \tag{5.19}$$

Remark 5.3 Moments estimates like those in Theorem 5.2 are an important step towards ergodicity, as assessed by Theorem 2.1. In the current case, relation (5.18) for $L_b \le \underline{\lambda}_c$ is an indication that a Krylov-Bogoliubov type argument towards ergodicity can be applied. On the other hand, relation (5.18) for $l_b > \overline{\lambda}_c$ rules out any type of ergodic result. This is the announced phase transition phenomenon, in terms of coefficients which represent *noise intensity*.

Remark 5.4 As part of Theorem 5.2, relation (5.19) asserts that the phase transition phenomenon can only occur in dimension 3 or higher.

In Theorem 5.2, we have only considered initial conditions with power growth at infinity, which are tempered/Schwarz measures. The choice of such initial conditions is due to the fact that the corresponding solution to the homogeneous equation $J_0(t, x)$ will only induce trivial contributions to the Lyapunov exponent. However, if we do not focus only on Lyapunov exponents, the paper [16] establishes the following pointwise moment bounds in case of an initial condition with exponential growth:

Theorem 5.5 *Let u be the solution to (5.1) starting from a nonnegative rough initial data μ, namely, μ is is a nonnegative Borel measure on $\mathbb{R}^d$ such that (5.12) holds. Suppose that the diffusion coefficient $b(x, u) = b(u)$ satisfies the following degenerate condition:*

$$b(0) = 0\ , \quad \textit{and} \quad |b(u_2) - b(u_1)| \le L_b|u_2 - u_1|. \tag{5.20}$$

Moreover, suppose that $4L_b^2\Upsilon(0) < 1$, namely, condition (5.10) holds and $L_b < (2\Upsilon(0))^{-1/2}$. Then

$$\mathbb{E}\left[u(t, x)^2\right] \le \frac{J_0^2(t, x)}{L_b^2\left[1 - 4L_b^2\Upsilon(0)\right]}, \quad \textit{for all } t \ge 0 \textit{ and } x \in \mathbb{R}^d, \tag{5.21}$$

where we recall that $J_0(t, x)$ is defined by (5.14).

Remark 5.6 From this point onward, we have chosen to work with a coefficient b which only depends on the variable u. Generalizations to coefficients $b(x, u)$ are possible.

Proof of Theorem 5.5 The proof uses the contraction coefficient L_b in a Gronwall type lemma, very similarly to what will be done in (5.62). We thus spare the details for sake of conciseness. Moreover, this type of estimate has already been carried out in [16]. Namely note that the degenerate condition (5.20) implies that $|b(x, u)| \le L_b|u|$ for all $u \in \mathbb{R}$ uniformly in $x \in \mathbb{R}^d$. Hence, we can invoke part (2) of Theorem 2.4 in [16]. After applying (2.10) of [16] with $x = x'$ followed by (2.19) of [16] with $\nu = 1$ and $\lambda = L_b$, we see that

$$\mathbb{E}\left[u(t, x)^2\right] \le L_b^{-2} J_0^2(t, x) H(t; 2L_b^2),$$

where $H(t; \gamma)$ is defined in (2.14) of [16] with $y = 0$ and $\nu = 1$. Then by the second part of Lemma 2.5 of [16] with $\nu = 1$ and $\gamma = 2L_b^2$, when $\Upsilon(0) < \infty$ and

$4L_b^2\Upsilon(0) < 1$, then

$$H(t; 2L_b^2) \le \frac{1}{1 - 4L_b^2\Upsilon(0)}, \quad \text{for all } t \ge 0.$$

Combining these two bounds proves Theorem 5.5.

Remark 5.7 Here are some additional comments: (1) The most challenging aspect of the paper [16] is the demonstration that the *lower* bound of the Lyapunov exponent is strictly positive, i.e., the second relationship in (5.18). The difficulty comes from the rough initial condition and the lack of Feynman-Kac representations for the moments for the nonlinear SHE. (2) By the Burkholder-Gundy-Davis inequality (see the version in [19, Theorem 1.4]), one can easily extend the boundedness of the second moment to the case of p-th moments with $p \ge 2$ and $L_0 > 0$ in (5.2); see Theorem 1.7 of [14] for more details.

Boundedness of the moments in (5.21) paves the way for the existence of invariant measures. This was already asserted in Da Prato-Zabczyk's setting, as recalled in (2.9) and Remark 5.3. However, one needs to embed the random field solution in some Hilbert space. The moment estimates in this setup are a straightforward application of Theorem 5.2, which is given by Theorem 5.8 below. For any locally integrable and nonnegative function $\rho : \mathbb{R}^d \to \mathbb{R}_+$, denote by $L^2_\rho(\mathbb{R}^d)$ the Hilbert space of ρ-weighted square integrable functions. We will use $\langle \cdot, \cdot \rangle_\rho$ and $\|\cdot\|_\rho$ to denote the corresponding inner product and norm, respectively:

$$\langle f, g\rangle_\rho := \int_{\mathbb{R}^d} f(x)g(x)\rho(x)\mathrm{d}x\,, \quad \text{and} \quad \|f\|_\rho := \int_{\mathbb{R}^d} |f(x)|^2\rho(x)\mathrm{d}x. \tag{5.22}$$

The next theorem states that, under proper conditions on the noise structure and the intensity of the noise (namely, the value of L_b), the time dependence of the moments of the solution in any weighted $L^2(\mathbb{R}^d)$ space comes solely from the contribution of the initial condition.

Theorem 5.8 *(Theorem 1.1 of [12] for the degenerate diffusion coefficient case) Let $u(t, x; \mu)$ be the solution to (5.1) starting from a rough initial condition μ which satisfies (5.12). Assume that*

(i) the diffusion coefficient b satisfies the degenerate condition (5.20);
(ii) $\rho : \mathbb{R}^d \to \mathbb{R}_+$ is a nonnegative $L^1(\mathbb{R}^d)$ function;
(iii) for all $t > 0$, the initial condition μ satisfies the condition that $\mathcal{G}_\rho(t; |\mu|) < \infty$, where

$$\mathcal{G}_\rho(t;\mu) := \int_{\mathbb{R}^d} \left[(p(t,\cdot) * \mu)(x)\right]^2 \rho(x)\,\mathrm{d}x; \tag{5.23}$$

(iv) the spectral measure $\widehat{f}$ satisfies (5.10), i.e., $\Upsilon(0) < \infty$, and the Lipschitz constant L_b is small enough so that $4L_b^2\Upsilon(0) < 1$.

Then there exists a unique $L^2(\Omega)$-continuous solution $u(t,x)$ such that for some constant $C > 0$, which does not depend on t, the following holds:

$$\mathbb{E}\left(\|u(t,\cdot;\mu)\|_\rho^2\right) \le C\mathcal{G}_\rho(t;|\mu|) < \infty, \quad \textit{for any } t > 0, \tag{5.24}$$

where we recall that the norm $\|\cdot\|_\rho$ is given by (5.22).

Recall that for the existence of an invariant measure, we will need moments bounded in t (see (2.9)). However, the moment upper bound in (5.24) may still blow up as $t \to \infty$; see Proposition 5.1 of [12] for one example. Given the continuity of $t \to \mathcal{G}_\rho(t;|\mu|)$, one needs to impose condition (5.28) below.

Note that Theorem 5.8 above imposes no additional restrictions on the weight function, as long as it is nonnegative and integrable. However, due to the non-compactness of the ambient space $\mathbb{R}^d$, in order to extend the heat semigroup to a C_0-semigroup on a weighted $L^2(\mathbb{R}^d)$ space one cannot use arbitrary weight functions. The notion of admissible weight functions plays an important role.

Definition 5.9 ([50]) A function $\rho : \mathbb{R}^d \mapsto \mathbb{R}$ is called an admissible weight function if it is a strictly positive, bounded, continuous, and $L^1(\mathbb{R}^d)$-integrable function such that for all $T > 0$, there exists a constant $C_\rho(T)$ satisfying

$$\big(p(t,\cdot) * \rho(\cdot)\big)(x) \le C_\rho(T)\rho(x) \quad \text{for all } t \in [0,T] \text{ and } x \in \mathbb{R}^d. \tag{5.25}$$

As proved in Proposition 2.1 of [50], the weighted spaces of Definition 5.9 have nice compactness properties. Namely, for any admissible functions ρ and $\tilde{\rho}$, if

$$\int_{\mathbb{R}^d} \frac{\rho(x)}{\tilde{\rho}(x)} \mathrm{d}x < \infty, \tag{5.26}$$

then for all $t > 0$, the heat semigroup is compact from $L_{\tilde{\rho}}(\mathbb{R}^d)$ to $L_\rho(\mathbb{R}^d)$. Canonical choices of the admissible weight functions include the following examples:

$$\rho(x) = \exp(-a|x|) \quad a > 0, \quad \text{and} \quad \widetilde{\rho}(x) = \left(1 + |x|^a\right)^{-1} \quad a > d. \tag{5.27}$$

Finally, the next theorem integrates various components, including initial data, the noise structure (via the correlation function), the intensity of the noise (via the value of L_b), and admissible weight functions, to establish the existence of an invariant measure.

Theorem 5.10 *(Theorem 1.3 of [12] for the degenerate diffusion coefficient case) Let $u(t,x)$ be the solution to (5.1) starting from a rough initial condition μ, namely, μ is a signed Borel measure on $\mathbb{R}^d$ such that (5.12) holds. Assume that*

(i) $b(x,u) = b(u)$ and it verifies condition (5.20).

(ii) there are two admissible weight functions ρ and $\tilde{\rho}$ such that (5.26) holds and

$$\limsup_{t\to\infty} \mathcal{G}_{\tilde{\rho}}(t;|\mu|) < \infty; \tag{5.28}$$

(iii) the spectral measure $\widehat{f}$ satisfies (5.10), i.e., $\Upsilon(0) < \infty$ and the Lipschitz constant L_b is small enough such that $4L_b^2\Upsilon(0) < 1$;

(iv) for some $\alpha \in \left(4L_b^2\Upsilon(0), 1\right)$, the spectral measure $\widehat{f}$ satisfies the following condition

$$\Upsilon_\alpha(\beta) := (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{\widehat{f}(\mathrm{d}\xi)}{\left(\beta + |\xi|^2\right)^{1-\alpha}} < \infty \quad \textit{for some (hence all) } \beta > 0. \tag{5.29}$$

Then we have that

(1) for any $\tau > 0$, the sequence of laws of $\{\mathcal{L}u(t, \cdot; \mu)\}_{t \geq \tau}$ is tight, i.e., for any $\epsilon \in (0, 1)$, there exists a compact set $\mathscr{K} \subset L^2_\rho(\mathbb{R}^d)$ such that

$$\mathscr{L}u(t, \cdot; \mu)(\mathcal{K}) \geq 1 - \epsilon, \qquad \textit{for all } t \geq \tau > 0; \tag{5.30}$$

(2) there exists a nontrivial invariant measure for (5.1).

Remark 5.11 Conditions (ii) and (iii) in Theorem 5.10 can be more compactly written as (5.9). Note that condition (5.9) implicitly implies that $4\Upsilon(0)L_b^2 < 1$, in order to find an α such that (5.29) is satisfied. Once again, let us recall that it has to be interpreted as a small intensity condition on the noisy forcing term.

In the literature, condition (5.29) is often called the *strengthened Dalang's condition*. Since this condition first appeared in the paper by Sanz-Solé and Sarrà [49], we may also call it the *Sanz-Solé-Sarrà condition*. The conditions on the structure and intensity of the noise, as specified in Conditions (iii) and (iv) of the above theorem, can be consolidated into a single condition—condition (5.9).

Remark 5.12 Note that under the degenerate condition (5.20), i.e., $b(0) \equiv 0$, the zero function is an invariant measure for (5.1). This invariant measure is called the *trivial invariant measure*. Once the existence of an invariant measure is obtained via the Bogoliubov theorem (see the proof of Theorem 2.1), it is routine to show that there exists a nontrivial one; see Theorem 4.1 of [50].

Remark 5.13 The expression $4L_b^2\Upsilon(0)$ in both Theorems 5.8 and 5.10 takes the form of $2^7 L_b^2\Upsilon(0)$ in [12] (see (1.10b) and (1.11) *ibid.*). This discrepancy arises from applying the general p-th moment bounds (for $p \geq 2$), as given in (1.14) of [14], to the second moment. These p-th moment bounds were derived using the Burkholder-Gundy-Davis inequality for cases where $b(x, 0)$ may not vanish. However, when focusing specifically on the second moment, its application is unnecessary. Moreover, the degenerate condition (5.20) further simplifies the moment bounds.

Interested readers may refer to Section 5 of [12], where Theorem 5.10 is discussed in detail from several perspectives (initial data, weight functions, correlation functions, and a comparison with the conditions outlined by Tessitore and Zabczyk [50]).

5.3 The asymptotic behavior via Gu-Li's approach

This section is devoted to establishing a new result (to the best of our knowledge). It establishes convergence in law to the invariant distribution under the "*weak disorder*" condition given by (iii) in Theorem 5.10. Specifically, suppose the correlation function f is given by (5.8) and suppose the diffusion coefficient b depends only on the second argument. Gu and Li proved in [36] that, when the spatial dimension $d \geq 3$, the solution $u(t, \cdot)$ to the SHE (5.1) at time t with a flat initial condition converges in law to a stationary random field in space as $t \to \infty$. In this part, we will show that one can extend this result to a broader class of Gaussian noises (namely, those satisfying condition (5.9)), adapting the methods in [36]. A similar question is also considered in [34], where the spatial correlation of the driving noise is uniformly bounded with a Riesz type tail. With respect to the aforementioned references [34, 36] we will also consider a broader class of initial conditions.

Theorem 5.14 *Suppose that conditions (i) and (iii) of Theorem 5.10 hold. Let the following condition (ii') replace condition (ii) of Theorem 5.10:*

(ii') For $t \geq 0$, $x \in \mathbb{R}$ and a Borel measure μ on $\mathbb{R}^d$, let $J_0(t, x; \mu)$ be the solution to the homegenous equation (see (5.14)). We assume that the following limit exists:

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t, x; \mu)| < \infty. \tag{5.31}$$

We also assume that the nonnegative weight function ρ is an element in $L^1(\mathbb{R}^d)$, which does not necessarily need to be admissible (see Definition 5.9). Recall that $u(t, \cdot)$ is the solution to (5.1). Then the following existence and uniqueness statements hold:

(1) There exists a random field $Z = \{Z(x);\ x \in \mathbb{R}^d\}$ such that $Z(\cdot) \in L^2_\rho(\mathbb{R}^d)$ a.s. and

$$u(t, \cdot) \xrightarrow{(d)} Z, \quad \text{as } t \to \infty, \quad \text{in } L^2_\rho(\mathbb{R}^d). \tag{5.32}$$

(2) Suppose u_1 and u_2 are two solutions to (5.1) starting from μ_1 and μ_2, respectively. Assume that both μ_i satisfy (5.31). Let Z_1 and Z_2 be the respective limiting random fields given in part (1). Then we have

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t, x; \mu_1 - \mu_2)| = 0 \quad \Longrightarrow \quad Z_1 \overset{(d)}{=} Z_2\,, \tag{5.33}$$

where the notation $Z_1 \overset{(d)}{=} Z_2$ denotes that the two random fields follow the same law.

Before the proof, we first make some remarks.

Remark 5.15 (Perturbation condition) Recall that in Gu and Li [36], the initial condition needs to be of the form $u(0, x) = \lambda + g(x)$ with $g \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$; see Remark 3.6 (*ibid.*). The above Theorem 5.14 relaxes this perturbation term g. Indeed, the perturbation term g can go well beyond functions in $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. To state the most general perturbation type assumption on the initial condition, suppose that the rough initial measure μ can be decomposed into two parts $\mu = \mu_0 + \mu_1$ such that μ_0 is a rough initial measure that satisfies (5.31), and μ_1, another rough initial measure, is treated as the perturbation. Here is the *perturbation condition*:

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t,x;\mu_1)| = 0. \tag{5.34}$$

Thanks to part (2) of Theorem 5.14, provided the above perturbation condition (5.34) holds, or equivalently, if

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t,x;\mu)| = \lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t,x;\mu_0)|,$$

then the limiting random field Z follows the same law no matter whether the system starts from μ or μ_0. The perturbation condition (5.34) gives a partition of all possible rough initial conditions into equivalent classes with respect to the law of the asymptotic random field Z.

Example 5.16 Following the same setup as Remark 5.15, we provide some examples for which the perturbation condition (5.34) is satisfied:

(a) μ_1 is such that $|\mu_1|$ is a finite Borel measure. One can consider for instance a Dirac delta measure $\mu_1 = \delta_0(x)$, or a linear combination of the form $\mu_1 = \delta_{-1}(x) - \delta_1(x)$. This is clear because $J_0(t,x;|\mu_1|) \le (2\pi t)^{-d/2}\,|\mu_1|(\mathbb{R}^d)$, and this last quantity converges to 0 as $t\to\infty$.

(b) $|\mu_1|$ does not need to be a finite measure. Consider for example $\mu_1(\mathrm{d}x) = |x|^{-\alpha}\mathrm{d}x$, for a given $\alpha\in(0,d)$. In this case, μ_1 is a nonnegative and locally finite measure with total variation being infinity. However, the perturbation condition (5.34) is still satisfied. This fact is proved in Example 5.7 of [12],

$$\sup_{x\in\mathbb{R}^d} J_0(t,x;|\cdot|^{-\alpha}) \le C_\alpha t^{-\alpha/2} \to 0, \quad \text{as } t\to\infty,$$
$$\text{with } C_\alpha := 2^{-\alpha/2}\Gamma((d-\alpha)/2)\Gamma(d/2).$$

(c) Let $\mu_1 = \sum_{k\in\mathbb{Z}^d}\delta_{2\pi k}(x) - (2\pi)^{-d}$. We claim that μ_1 satisfies the perturbation condition (5.34). Indeed, in this case, for all $x\in\mathbb{R}^d$ and $t>0$,

$$J_0(t,x;\mu_1) = G(t,x) - (2\pi)^{-d}, \quad \text{with} \quad G(t,x) := \sum_{k\in\mathbb{Z}^d} p(t,x+2\pi k).$$

Since $J_0(t,x;\mu_1)$ is 2π periodic in each direction of x, we see that for $t\ge 1$,

$$\sup_{x\in\mathbb{R}^d} |J_0(t,x;\mu_1)| = \sup_{x\in[-\pi,\pi]^d} \left|G(t,x) - (2\pi)^{-d}\right| \le Ce^{-t/2} \to 0, \quad \text{as} \quad t\to\infty,$$

where the inequality follows from, e.g., part (4) of Lemma 2.1 in [17]. This proves the claim. As a consequence, the solution $u(t,x)$ to (5.1) starting from the flat initial condition $(2\pi)^{-d}$ and from $\sum_{k\in\mathbb{Z}^d}\delta_{2\pi k}(x)$ both yield the limiting field Z of the same distribution.

Remark 5.17 (Comparison with Theorem 5.10) Assumption (ii') of Theorem 5.14 is stronger than the assumption (ii) of Theorem 5.10. However, Theorem 5.14 does not require the Sanz-Solé-Sarrà condition (5.29), nor does it impose additional requirements on the weight function ρ beyond being nonnegative and integrable. In [15], the authors constructed a correlation function $f(x)$ in $\mathbb{R}^3$ that does not satisfy the Sanz-Solé-Sarrà condition (5.29) by applying a logarithmic correction to the critical case. They established solutions with unbounded oscillations. Nevertheless, the correlation constructed there fulfills the condition $\Upsilon(0) < \infty$; see Theorem 2.1 and its proof *ibid*.

Remark 5.18 Hypothesis (iii) in Theorem 5.10 has been named *weak disorder condition* at the beginning of Section 5.3. This type of condition is related to the random polymer literature and leads to central limit type theorems for the polymer measure; see, e.g., [18]. We plan on pursuing this type of result for general environments W in a subsequent publication.

Before proving Theorem 5.14, let us introduce some additional notation.

Notation 5.19 The function $J_0(t,x;|\mu|)$ defined by (5.14) is smooth on $(0,\infty)\times\mathbb{R}^d$. Hence, condition (5.31) implies that the following two constants are finite:

$$C_\mu := \sup_{x\in\mathbb{R}^d} J_0(1,x;|\mu|) < \infty\,, \quad \text{and} \quad \widehat{C}_\mu := \sup_{(t,x)\in[1,\infty)\times\mathbb{R}^d} J_0(t,x;|\mu|) < \infty. \tag{5.35}$$

It is clear that $C_\mu \le \widehat{C}_\mu$. For all $t\ge 0$, we also define

$$k(t) := \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\ p(t,y)p(t,y')f(y-y') = (2\pi)^{-d}\int_{\mathbb{R}^d}\widehat{f}(\mathrm{d}\xi)\,e^{-t|\xi|^2}, \tag{5.36}$$

where the second equality is an application of the Plancherel theorem. Eventually, we set

$$H(t) := \int_t^\infty k(s)\mathrm{d}s. \tag{5.37}$$

We also label a preliminary result for further use. Its elementary proof is left to the reader for sake of conciseness.

Lemma 5.20 *Assume that the assumption* (5.10) *holds, i.e.,* $\Upsilon(0) < \infty$. *Then the function* $H(t)$ *defined in* (5.37) *is monotone decreasing, such that*

$$H(t)\downarrow 0 \quad \text{as } t\to\infty, \quad \text{and} \quad H(t)\uparrow (2\pi)^{-d}\int_{\mathbb{R}^d}\frac{\widehat{f}(\mathrm{d}\xi)}{|\xi|^2} = \Upsilon(0), \quad \text{as } t\downarrow 0. \tag{5.38}$$

We are now ready to prove the main result of this section.

Proof of item (1) in Theorem 5.14 Recall that we assume the coefficient b to depend on the u variable only. Thanks to the well-posedness results, as established in [16], one can find a unique solution to (5.1) starting from the rough initial condition μ. Also note that in the proof, we use $\|\cdot\|_p$ to denote $L^p(\Omega)$ norms. We now divide our proof in several steps.

Step 1: Dynamical system in negative time. Since we consider the large time behavior, one can restart the system at time 1. Indeed, the Cauchy-Schwarz inequality and the moment bound (5.21) imply that, for all $t > 0$ and $x \in \mathbb{R}^d$,

$$\begin{aligned}\mathbb{E}\left([u(1,\cdot) * p(t,\cdot)(x)]^2\right) &\leq \int_{\mathbb{R}^d} \|u(1,y)\|_2^2\, p(t,x-y)\mathrm{d}y \\ &\leq \frac{1}{L_b^2\left(1-4L_b^2\Upsilon(0)\right)} \int_{\mathbb{R}^d} J_0^2(1,y)p(t,x-y)\mathrm{d}y \leq \frac{C_\mu^2}{L_b^2\left(1-4L_b^2\Upsilon(0)\right)} < \infty,\end{aligned} \tag{5.39}$$

where C_μ is defined by (5.35). Hence,

$$[u(1,\cdot) * p(t,\cdot)](x) < \infty \quad \text{a.s. for all } t > 0 \text{ and } x \in \mathbb{R}^d.$$

This relation implies that $u(1,\cdot)$, considered as an initial condition, satisfies assumption (5.12) a.s. In the following, we will thus restart our system after one unit of time.

For any parameter $K > 0$, we now extend the white noise $\dot{W}$ in time to a two-sided noise indexed by $t \in \mathbb{R}$. In addition, define a process u_K^* by

$$u_K^* = \left\{u_K^*(t,x) :\ (t,x) \in [-K-1,\infty) \times \mathbb{R}^d\right\},$$

as the unique solution to the following equation

$$\begin{cases}\left(\dfrac{\partial}{\partial t} - \dfrac{1}{2}\Delta\right) u_K^*(t,x) = b\left(u_K^*(t,x)\right)\dot{W}(t,x), & (t,x) \in [-K-1,\infty) \times \mathbb{R}^d, \\ u_K^*(-K-1,\cdot) = \mu. \end{cases} \tag{5.40}$$

We will restart the system at time $-K$, namely,

$$u_K^*(-K,x) = J_0(1,x;\mu) + \int_{-K-1}^{-K}\int_{\mathbb{R}^d} p\left(-K-s,x-y\right) b\left(u_K^*(s,y)\right) W(\mathrm{d}s,\mathrm{d}y). \tag{5.41}$$

According to this procedure, the process u_K is now defined as

$$u_K = \left\{u_K(t,x) :\ (t,x) \in [-K,\infty) \times \mathbb{R}^d\right\},$$

given as the unique solution to the following restarted equation:

$$\begin{cases} \left(\dfrac{\partial}{\partial t} - \dfrac{1}{2}\Delta\right) u_K(t,x) = b\left(u_K(t,x)\right) \dot{W}_1(t,x), & (t,x) \in [-K,\infty) \times \mathbb{R}^d, \\ u_K(-K,x) = u_K^*(-K,x), & x \in \mathbb{R}^d, \quad \text{a.s.,} \end{cases} \tag{5.42}$$

where $\dot{W}_1(t,x) = \dot{W}(1+t,x)$ is a time-shifted noise. Note that equivalently, u_K is defined through the following mild formulation:

$$u_K(t,x) = J_0\left(t+K, x; u_K^*(-K,\cdot)\right) + \int_{-K}^{t} \int_{\mathbb{R}^d} p(t-s, x-y) b(u_K(s,y)) W_1(\mathrm{d}s, \mathrm{d}y). \tag{5.43}$$

Due to the time stationarity of noise $\dot{W}$, we see that the random field

$$\widetilde{u} = \left\{\widetilde{u}(t,x) = u_K\left(t-K-1, x\right); (t,x) \in [0,\infty) \times \mathbb{R}^d\right\}$$

shares the same distribution as u—the solution to (5.1). In the following, we are thus reduced to prove the following claim in order to establish the theorem:
Claim: $\left\{u_K(0,x);\, K \geq 0,\ x \in \mathbb{R}^d\right\}$ is a Cauchy sequence in $L^\infty\left(\mathbb{R}^d; L^2(\Omega)\right) \cap L^2\left(\Omega; L_\rho^2\left(\mathbb{R}^d\right)\right)$.
The remainder of the proof is now devoted to prove the above claim.
Step 2: Reduction to an $L^2(\Omega)$-estimate. Let $K > 0$. Denote

$$J(t+K,x) := \sup_{L \geq K} \mathbb{E}\left[\left(u_L(t,x) - u_K(t,x)\right)^2\right], \quad \text{for all } t \geq -K \text{ and } x \in \mathbb{R}^d. \tag{5.44}$$

We also define

$$\mathcal{J}(t+K) := \sup_{x \in \mathbb{R}^d} J(t+K,x). \tag{5.45}$$

Then, the convergence in $L^\infty\left(\mathbb{R}^d; L^2(\Omega)\right)$, as stated in our *Claim* above, holds if and only if, for any finite $t \in \mathbb{R}$,

$$\lim_{K \to \infty} \mathcal{J}(K+t) = 0. \tag{5.46}$$

Moreover, relation (5.46) implies the convergence in $L_\rho^2\left(\mathbb{R}^d; L^2(\Omega)\right)$, as stated in the conclusion (5.32) of Theorem 5.14. This implication is observed by noticing that

$$\int_{\mathbb{R}^d} \mathbb{E}\left[\left(u_L(0,x) - u_K(0,x)\right)^2\right] \rho(x) \mathrm{d}x \leq \mathcal{J}(K) \left\|\rho\right\|_{L^1(\mathbb{R}^d)} \to 0, \quad \text{as } L > K \to \infty.$$

Hence, the Claim is proved once the limit in (5.46) is established.
Step 3: Some additional notation. In the following, we will focus on proving (5.46). Let us first set up some notation. Since $|b(u)| \leq L_b|u|$ thanks to (5.20), we see that the quantity $J(t+K,x)$ defined by (5.44) can be decomposed as

$$J(t+K,x) \leq 4I_0^K(t,x) + 4I_1^K(t,x) + 2I_2^K(t,x), \tag{5.47}$$

where

$$I_0^K(t,x) := \sup_{L \geq K} \mathbb{E}\left[\left|J_0\left(t+L,x;u_L^*(1,\cdot)\right) - J_0\left(t+K,x;u_K^*(1,\cdot)\right)\right|^2\right], \tag{5.48}$$

$$I_1^K(t,x) := L_b^2 \sup_{L \geq K} \mathbb{E}\bigg[\int_{-L}^{-K} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\, p(t-s,x-y)p(t-s,x-y') \\ \times |u_L(s,y)u_K(s,y')| f\left(y-y'\right)\bigg], \tag{5.49}$$

and

$$I_2^K(t,x) := L_b^2 \sup_{L \geq K} \mathbb{E}\bigg[\int_{-K}^{t} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\, p(t-s,x-y)\,|u_L(s,y)-u_K(s,y)| \\ \times f(y-y')p(t-s,x-y')\left|u_L(s,y')-u_K(s,y')\right|\bigg]. \tag{5.50}$$

We now bound the terms I_0^K, I_1^K, and I_2^K separately.
Step 4: Term $I_0^K(t,x)$. Let us first bound the term I_0^K defined by (5.48). To this aim, we first recall from (5.14) that for a given $K>0$, we have

$$J_0\left(t+K,x;u_K^*(-K,\cdot)\right) = \int_{\mathbb{R}^d} p\left(t+K,x-y\right) u_K^*(-K,y)\mathrm{d}y.$$

Then plug in expression (5.41) for u_K^*, and use the semigroup property for the heat kernel. This allows to write

$$J_0\left(t+K,x;u_K^*(-K,\cdot)\right) = J_0(t+K+1,x;\mu) + S_K(t,x), \tag{5.51}$$

where the stochastic integral in (5.51) is defined by

$$S_K(t,x) := \int_{-K-1}^{-K}\int_{\mathbb{R}^d} p\left(t-s,x-y\right) b\left(u_K^*(s,y)\right) W(\mathrm{d}s,\mathrm{d}y). \tag{5.52}$$

Reporting this expression into the definition (5.48) of $I_0^K(t,x)$, some elementary manipulations show that

$$I_0^K(t,x) \le 4\Theta(t+K) + 8 \sup_{L \ge K} \|S_L(t,x)\|_2^2 \tag{5.53}$$

where for all $\tau \ge 0$, the term $\Theta(\tau)$ can be further decomposed as

$$\Theta(\tau) := \sup_{s,r \ge \tau} \sup_{x \in \mathbb{R}^d} |J_0(s+1,x;\mu) - J_0(r+1,x;\mu)|^2 . \tag{5.54}$$

We now proceed to bound the terms in relation (5.53). To begin with, it is clear from condition (5.31) that the term $\Theta(t)$ in (5.54) is such that

$$\sup_{\tau \ge 0} \Theta(\tau) < \infty; \quad \text{and} \quad \Theta(\tau) \downarrow 0, \text{ as } \tau \to \infty. \tag{5.55}$$

Therefore we immediately get $\lim_{K\to\infty} \Theta(t+K) = 0$. Our main task is thus to bound the quantity $S_K(t,x)$ introduced in (5.52). However, this term can be estimated thanks to a direct application of relation (5.6) plus Cauchy-Schwarz's inequality as follows:

$$\begin{aligned}\|S_K(t,x)\|_2^2 \le L_b^2 \int_{-K-1}^{-K} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y' \, p\,(t-s,x-y) \left\|u_K^*(-K,y)\right\|_2 \\ \times f(y-y') p\left(t-s,x-y'\right) \left\|u_K^*(-K,y')\right\|_2 .\end{aligned}$$

Plugging our moment bound (5.21) into the above inequality, we end up with

$$\begin{aligned}\|S_K(t,x)\|_2^2 \le \frac{1}{1-4L_b^2\Upsilon(0)} \int_{-K-1}^{-K} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y' \, p\,(t-s,x-y)\, J_0(1,y;|\mu|) \\ \times f(y-y') p\left(t-s,x-y'\right) J_0(1,y';|\mu|).\end{aligned} \tag{5.56}$$

Therefore, using the constant C_μ defined in (5.35) and the function $H(t)$ in (5.37), we have

$$\begin{aligned}&\sup_{L\ge K} \|S_L(t,x)\|_2^2 \\ &\quad \le \frac{C_\mu^2}{1-4L_b^2\Upsilon(0)} \sup_{L\ge K} \int_{t+L}^{t+L+1} \mathrm{d}s \\ &\quad \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y' \, p\,(s,x-y)\, p\left(s,x-y'\right) f(y-y') \\ &\quad = \frac{C_\mu^2}{1-4L_b^2\Upsilon(0)} \sup_{L\ge K} \left[H(t+L) - H(t+L+1)\right] \\ &\quad \le \frac{C_\mu^2}{1-4L_b^2\Upsilon(0)} H(t+K),\end{aligned} \tag{5.57}$$

where the last equality is due to Lemma 5.20. Summarizing our considerations for the term I_0^K, we gather relations (5.53), (5.55), and (5.57), and get

$$I_0^K(t,x) \le 4\Theta(t+K) + \frac{8C_\mu^2 H(t+K)}{1-4L_b^2\Upsilon(0)}. \tag{5.58}$$

Step 5: Term $I_1^K(t,x)$. For the term I_1^K defined by (5.49), one proceeds similarly to I_0^K. That is we start by using relation (5.6), which yields

$$\begin{aligned} I_1^K(t,x) \le L_b^2 \sup_{L\ge K}\int_{-L}^{-K} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\, p(t-s,x-y)\,\|u_L(s,y)\|_2 \\ \times f(y-y')p(t-s,x-y')\,\|u_L(s,y')\|_2. \end{aligned} \tag{5.59}$$

Next, we notice that $u_K(s,y) := u(s+K+1,y)$, where u is the solution to (5.1). Hence, applying the moment bound (5.21) as for (5.49), we see that

$$\|u_L(s,y)\|_2^2 = \|u(s+L+1,y)\|_2^2 \le \frac{J_0^2(s+L+1,y;|\mu|)}{L_b^2(1-4L_b^2\Upsilon(0))}.$$

Reporting his information into (5.59), we have thus obtained an inequality which mimics relation (5.56):

$$\begin{aligned} I_1^K(t,x) \le \frac{1}{1-4L_b^2\Upsilon(0)} \sup_{L\ge K}\int_{-L}^{-K} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y' \\ p(t-s,x-y)J_0(s+L+1,y;|\mu|) \\ \times f(y-y')p(t-s,x-y')J_0(s+L+1,y';|\mu|). \end{aligned}$$

From there, we basically repeat the calculations leading to (5.57). We let the patient reader check that

$$I_1^K(t,x) \le \frac{\widehat{C}_\mu^2\, H(t+K)}{1-4L_b^2\Upsilon(0)}, \tag{5.60}$$

where we recall that $\widehat{C}_\mu$ is introduced in (5.35) and the function H is given by (5.37). *Step 6: Term* $I_2^K(t,x)$. This term is treated very similarly to $I_0^K(t,x)$ and $I_1^K(t,x)$. Let us just summarize briefly the computations for the sake of completeness. Specifically, starting from the definition of I_2^K, applying relation (5.6) and the Cauchy-Schwarz inequality, we get

$$\begin{aligned} I_2^K(t,x) \le L_b^2 \sup_{L\ge K}\int_{-K}^{t} \mathrm{d}s \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\; p(t-s,y)\,\|u_L(s,x-y)-u_K(s,x-y)\|_2 \\ \times f(y-y')p(t-s,y')\,\|u_L(s,x-y')-u_K(s,x-y')\|_2. \end{aligned}$$

Recalling the notation for $\mathcal{J}(t+K)$ in (5.45) and the definition of $k(s)$ in (5.36), we see that

$$\begin{aligned} I_2^K(t,x) &\le L_b^2 \int_{-K}^{t} \mathrm{d}s\, \mathcal{J}(s+K) \iint_{\mathbb{R}^{2d}} \mathrm{d}y\mathrm{d}y'\, p(t-s,y)f(y-y')p(t-s,y') \\ &= L_b^2 \int_0^{t+K} \mathrm{d}s\, \mathcal{J}(t+K-s)k(s). \end{aligned} \tag{5.61}$$

Step 7: Conclusion In order to conclude the proof of (5.32), combine relations (5.58), (5.60), and (5.61) into (5.47). Also, recall that $\mathcal{J}(t+K)$ is defined by (5.45). We get that for all $t > K$,

$$\mathcal{J}(t+K) \le g_K(t) + 2L_b^2 \int_0^{t+K} \mathcal{J}(t+K-s)k(s)\mathrm{d}s, \tag{5.62}$$

where the function g_K is defined as

$$g_K(t) := 16\Theta(t+K) + \frac{4\left(8C_\mu^2 + \widehat{C}_\mu^2\right) H(t+K)}{1 - 4L_b^2 \Upsilon(0)}. \tag{5.63}$$

Now recall that Θ and H are respectively introduced in (5.54) and (5.37). Moreover, owing to (5.55) and (5.38), we have that g_K in (5.63) decreases to 0 as $t \to \infty$. Hence one can safely apply (A.5) in order to get $\lim_{K\to\infty} \mathcal{J}(t+K) = 0$ for any fixed $t \in \mathbb{R}$ provided that $4L_b^2\Upsilon(0) < 1$, which is guaranteed by condition (iii) of Theorem 5.14. This thereby confirms (5.46). This proves part (1) of Theorem 5.14.

Proof of item (2) in Theorem 5.14 For $i = 1, 2$, let u_i be the solution to (5.1) starting from the rough initial condition μ_i satisfying (5.31). Let Z_i be the respective limiting random fields. Set $u := u_1 - u_2$, $\mu := \mu_1 - \mu_2$, $\widetilde{b}(u) := b(u+u_2) - b(u_2)$, and $Z := Z_1 - Z_2$. Assume that

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}^d} |J_0(t,x;\mu)| = 0. \tag{5.64}$$

Hence, u is the solution to (5.1) with the diffusion coefficient b replaced by $\widetilde{b}$ starting from μ. Note that $\widetilde{b}(0) = 0$ and $\widetilde{b}$ and b share the same Lipschitz constant L_b. Thus, the arguments in *Steps 1–7* (proof of Theorem 5.14 item (1)) still apply for the current u. We thus get a limiting random field called Z. Under this setup, it suffices to prove that $Z \overset{(d)}{=} 0$, i.e., the limiting random field follows the same law as the trivial (vanishing) field. To this aim, we can resort again to a similar seven-step proof as for part (1). Below, we only highlight the modifications. In particular, *Step 1* remains unchanged, and we recall that u_K denotes the solution restarted from negative time $-K$ after the system has run for one unit of time from μ. In *Step 2*, all occurrences of u_L and suprema over $L \ge K$ should be removed. Notably, the quantity $\mathcal{J}(t+K)$ defined in (5.45) now takes the following form

$$\mathcal{J}(t+K)=\sup_{x\in\mathbb{R}^d}\mathbb{E}\left[u_K^2(t,x)\right]. \tag{5.65}$$

Proceed similarly in *Step 3*. Now, in the expression of I_0^K (see(5.48)), only one term remains—the one involving u_K^*. Additionally, $I_1^K \equiv 0$ (see (5.49)). In *Step 4*, the upper bound for $\|S_K\|_2^2$ as given in (5.57) remains valid. The expression $\Theta(\tau)$ given in (5.54) should be replaced by

$$\widetilde{\Theta}(\tau):=\sup_{s\geq\tau}\sup_{x\in\mathbb{R}^d}|J_0(s+1,x;\mu)|^2. \tag{5.66}$$

Condition (5.64) implies that both properties in (5.55) with Θ replaced by $\widetilde{\Theta}$ still hold. One can skip *Step 5* since $I_1^K \equiv 0$. One can also repeat *Step 6* in the proof of item (1) above, with $\mathcal{J}(r)$ replaced by the quantity (5.65). This yields an estimate for I_2^K which is similar to (5.61). Then, in *Step 7* we obtain (5.62) and (5.63), but with Θ replaced by $\widetilde{\Theta}$. Finally, since $\widetilde{\Theta}$ now vanishes as $\tau\to\infty$, we conclude in the same way as in *Step 7* that $\lim_{K\to\infty}\mathcal{J}(t+K)=0$. Thus, for any $t\in\mathbb{R}$ and $x\in\mathbb{R}^d$ fixed, $u_K(t,x)\to 0$ in $L^2(\Omega)$ as $K\to\infty$ and hence the convergence also holds in law. By part (1) of Theorem 5.14, we also know that $u_K(t,x)\to Z(x)$ in law as $K\to\infty$. Therefore, the two limits have to be equal a.s., namely,

$$Z(x)=0,\quad \text{a.s. for all } x\in\mathbb{R}^d,$$

which immediately implies that the finite dimensional distributions of the field Z are the same as those of the zero field. This completes the proof of part (2) of Theorem 5.14.

Remark 5.21 Our result in Theorem 5.14 is stated as a convergence in $L^2_\rho(\mathbb{R}^d)$. Under condition (iv) of Theorem 5.10, and considering some weighted *Garsia-Rodemich-Rumsey* type lemmas, we could have improved the topology in which the convergence holds. Namely one can obtain a convergence in distribution in weighted Hölder spaces of the form $C^\beta_\rho(\mathbb{R}^d)$, for a proper Hölder exponent β. However, since the use of weighted Garsia-Rodemich-Rumsey inequalities is cumbersome, we have postponed this objective to a further publication for sake of conciseness.

A A Gronwall-type lemma

In this appendix, we prove a generic Gronwall-type lemma, which plays a key role in the proof of Theorem 5.14.

Lemma A.1 *Let $g(\cdot)$ be a nonnegative and monotone function on $[0,\infty)$ that decreases to 0 as $t\to\infty$, and such that $g(0)<\infty$. Let $k(\cdot)$ be a nonnegative and integrable* *function on* $[0,\infty)$ *and define*

$$h(t) := \int_t^\infty k(s)\mathrm{d}s. \tag{A.1}$$

Then, we have:

(i) For any nonnegative integer n and $t \geq 0$*, it holds that*

$$\int_0^t g\left(\frac{t-s}{2^n}\right) k(s)\mathrm{d}s \leq g(0)h\left(\frac{t}{2^{n+1}}\right) + h(0)g\left(\frac{t}{2^{n+1}}\right). \tag{A.2}$$

In particular, when $g = h$ *as defined in* (A.1)*, the above inequality* (A.2) *reduces to*

$$\int_0^t h\left(\frac{t-s}{2^n}\right) k(s)\mathrm{d}s \leq 2h(0)h\left(\frac{t}{2^{n+1}}\right). \tag{A.3}$$

(ii) For any nonnegative integer n and $t \geq 0$*, it holds that*

$$\begin{aligned}\int_{[0,t]^n_<} \mathrm{d}u_n \cdots \mathrm{d}u_1\, g\left(t-u_n\right) \prod_{i=1}^n k(u_i - u_{i-1}) \\ \leq \left(2^n - 1\right) g(0)h(0)^{n-1} h\left(\frac{t}{2^n}\right) + h(0)^n g\left(\frac{t}{2^n}\right),\end{aligned} \tag{A.4}$$

where $[0,t]^n_< := \{(s_1,\ldots,s_n) \in [0,t]^n; 0 < s_1 < \cdots < s_n < t\}$ *and we have used the convention that* $u_0 \equiv 0$*.*

(iii) Let f be a function on $[0,\infty)$ *such that*

$$f(t) \leq g(t) + \beta \int_0^t f(t-s)k(s)\mathrm{d}s, \quad \textit{for all } t \geq 0 \textit{ with } \beta > 0.$$

Then, whenever $\beta < [2h(0)]^{-1}$*, the following convergence is true,*

$$\lim_{t\to\infty} f(t) = 0. \tag{A.5}$$

Remark A.2 This result generalizes [36, Lemma 3.5], in which $g = h$, and $k(t) = 1 \wedge t^{-d/2}$. The difficulty of the above Gronwall-type lemma stems from not knowing the exact decay rate of the kernel function $k(s)$ as $s \to \infty$, as well as the lack of knowledge regarding whether and how $k(s)$ diverges as s approaches zero (see, e.g., the Riesz-type kernel given in (5.11)). The proof below is based on the dominated convergence argument. For comparison, when dealing with specific forms of the kernel function $k(s)$, such as $k(s) = s^{-1/\alpha}$ for $\alpha > 1$, explicit computations can be carried out; see Lemma A.2 of [13] for more details.

Proof of Lemma A.1 We will prove the 3 items successively.
Part (i). Decompose the integral into two regions:

$$\int_0^t g\left(\frac{t-s}{2^n}\right)k(s)\mathrm{d}s = \left(\int_0^{\frac{t}{2^{n+1}}} + \int_{\frac{t}{2^{n+1}}}^t\right) g\left(\frac{t-s}{2^n}\right)k(s)\mathrm{d}s =: I_1 + I_2. \quad \text{(A.6)}$$

Noting that both functions g and h are monotonically decreasing on $[0, \infty)$, we see that

$$I_1 \le g\left(\frac{t}{2^{n+1}}\right)\int_0^{\frac{t}{2^{n+1}}} k(s)\mathrm{d}s \le h(0)g\left(\frac{t}{2^{n+1}}\right) \quad \text{and} \quad \text{(A.7)}$$

$$I_2 \le g(0)\int_{\frac{t}{2^{n+1}}}^{\infty} k(s)\mathrm{d}s = g(0)h\left(\frac{t}{2^{n+1}}\right). \quad \text{(A.8)}$$

Plugging (A.7) and (A.8) into (A.6) proves part (i) of Lemma A.1.
Part (ii). We will prove (A.4) by induction in n. When $n = 1$, (A.4) is simply a consequence of part (i). Assume that $n \ge 2$ and that (A.4) holds for $n-1$. For $t > u_1 > 0$, by the induction assumption, we see that

$$\begin{aligned}
&\int_{[u_1,t]_<^{n-1}} \mathrm{d}u_n \cdots \mathrm{d}u_2\, g\,(t-u_n)\prod_{i=2}^n k(u_i - u_{i-1}) \\
&\quad = \int_{[0,t-u_1]_<^{n-1}} \mathrm{d}v_n \cdots \mathrm{d}v_2\, g\,(t-u_1-v_n)\prod_{i=2}^n k(v_i - v_{i-1}) \\
&\quad \le \left(2^{n-1}-1\right)g(0)h(0)^{n-2}h\left(\frac{t-u_1}{2^{n-1}}\right) + h(0)^{n-1}g\left(\frac{t-u_1}{2^{n-1}}\right).
\end{aligned}$$

Denote the left-hand-side (resp. right-hand-side) of the inequality (A.4) by $L_n(t)$ (resp. $R_n(t)$). Integrating both sides of the above inequality from 0 to t with respect to the u_1 variable gives

$$\begin{aligned}
L_n(t) \le &\left(2^{n-1}-1\right)g(0)h(0)^{n-2}\int_0^t h\left(\frac{t-u_1}{2^{n-1}}\right)k(u_1)\mathrm{d}u_1 \\
&+ h(0)^{n-1}\int_0^t g\left(\frac{t-u_1}{2^{n-1}}\right)k(u_1)\mathrm{d}u_1 \le R_n(t),
\end{aligned}$$

where we have applied the bounds (A.2) and (A.3) to the above two integrals. This proves part (ii) of Lemma A.1.
Part (iii). By iteration, one can write

$$f(t) \le g(t) + \sum_{n=1}^{\infty}\beta^n \int_0^t \mathrm{d}s_n \int_0^{t-s_n}\mathrm{d}s_{n-1}\cdots\int_0^{t-\sum_{j=2}^n s_j}\mathrm{d}s_1\, g\left(t-\sum_{j=1}^n s_j\right)\prod_{i=1}^n k(s_i)$$

$$=g(t)+\sum_{n=1}^{\infty}\beta^n\int_{[0,t]_<^n}\mathrm{d}u_n\cdots\mathrm{d}u_1\; g(t-u_n)\prod_{i=1}^{n}k(u_i-u_{i-1}).$$

Thus, part (ii) implies that

$$\begin{aligned} f(t)\leq & g(t)+\sum_{n=1}^{\infty}\beta^n\left[\left(2^n-1\right)g(0)h(0)^{n-1}h\left(\frac{t}{2^n}\right)+h(0)^n g\left(\frac{t}{2^n}\right)\right]\\ &\leq g(t)+\max\{1,g(0),h(0)\}\sum_{n=1}^{\infty}2\beta[2\beta h(0)]^{n-1}\left(h\left(\frac{t}{2^n}\right)+g\left(\frac{t}{2^n}\right)\right). \end{aligned} \tag{A.9}$$

By replacing t by 0 in the last line of (A.9) we see that provided that $2\beta h(0) < 1$, the summation is finite. Hence, one can apply the dominated convergence theorem to pass the limit of $t \to \infty$ inside the summation. This completes the whole proof of Lemma A.1.

Acknowledgements S. Tindel is supported for this work by NSF grant DMS-2153915. L. Chen and P. Xia are partially supported by NSF grant DMS-2246850 and L. Chen is partially supported by a travel grant (MPS-TSM-00959981) from the Simons foundation.

References

1. Amir, Gideon, Corwin, Ivan, Quastel, Jeremy: Probability distribution of the free energy of the continuum directed random polymer in 1 + 1 dimensions. Comm. Pure Appl. Math. **64**(4), 466–537 (2011)
2. Assing, Sigurd, Manthey, Ralf: Invariant measures for stochastic heat equations with unbounded coefficients. Stochastic Process. Appl. **103**(2), 237–256 (2003)
3. A. Bensoussan and J.-L. Lions. Applications des inéquations variationnelles en contrôle stochastique. Méthodes Mathématiques de l'Informatique [Mathematical Methods of Information Science], No. 6. Dunod, Paris, 1978
4. Bertini, Lorenzo, Cancrini, Nicoletta: The stochastic heat equation: Feynman-Kac formula and intermittence. J. Statist. Phys. **78**(5–6), 1377–1401 (1995)
5. Biane, Ph.: Relations entre pont et excursion du mouvement brownien réel. Ann. Inst. H. Poincaré Probab. Statist. **22**(1), 1–7 (1986)
6. Bogachev, Vladimir I., Röckner, Michael: Elliptic equations for measures on infinite-dimensional spaces and applications. Probab. Theory Related Fields **120**(4), 445–496 (2001)
7. Zdzisław Brzeźniak and Dariusz Gątarek. Martingale solutions and invariant measures for stochastic evolution equations in Banach spaces. Stochastic Process. Appl., 84(2):187–225, 1999
8. René A. Carmona and S. A. Molchanov. Parabolic Anderson problem and intermittency. Mem. Amer. Math. Soc., 108(518):viii+125, 1994
9. Sandra Cerrai. Second order PDE's in finite and infinite dimension, volume 1762 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001. A probabilistic approach
10. Cerrai, Sandra: Stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term. Probab. Theory Related Fields **125**(2), 271–304 (2003)
11. Chen, Le., Dalang, Robert C.: Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions. Ann. Probab. **43**(6), 3006–3051 (2015)
12. Chen, Le., Eisenberg, Nicholas: Invariant Measures for the Nonlinear Stochastic Heat Equation with No Drift Term. J. Theoret. Probab. **37**(2), 1357–1396 (2024)

13. Le Chen, Yaozhong Hu, and David Nualart. Regularity and strict positivity of densities for the nonlinear stochastic heat equation. Mem. Amer. Math. Soc., 273(1340):v+102, 2021
14. Chen, Le., Huang, Jingyu: Comparison principle for stochastic heat equation on $\mathbb{R}^d$. Ann. Probab. **47**(2), 989–1035 (2019)
15. Le Chen, Jingyu Huang, Davar Khoshnevisan, and Kunwoo Kim. Dense blowup for parabolic SPDEs. Electron. J. Probab., 24:Paper No. 118, 33, 2019
16. Le Chen and Kunwoo Kim. Nonlinear stochastic heat equation driven by spatially colored noise: moments and intermittency. Acta Math. Sci. Ser. B (Engl. Ed.), 39(3):645–668, 2019
17. Le Chen, Cheng Ouyang, and William Vickery. Parabolic anderson model with colored noise on torus. Preprint arXiv:2308.10802, to appear in Bernoulli, August 2023
18. Francis Comets. Directed polymers in random environments, volume 2175 of Lecture Notes in Mathematics. Springer, Cham, 2017. Lecture notes from the 46th Probability Summer School held in Saint-Flour, 2016
19. Conus, Daniel, Khoshnevisan, Davar: On the existence and position of the farthest peaks of a family of stochastic heat and wave equations. Probab. Theory Related Fields **152**(3–4), 681–701 (2012)
20. Ana Bela Cruzeiro: Solutions et mesures invariantes pour des équations d'évolution stochastiques du type Navier-Stokes. Exposition. Math. **7**(1), 73–82 (1989)
21. Da Prato, G., Zabczyk, J.: Ergodicity for infinite-dimensional systems. London Mathematical Society Lecture Note Series, vol. 229. Cambridge University Press, Cambridge (1996)
22. Da Prato, Giuseppe, Zabczyk, Jerzy: Second order partial differential equations in Hilbert spaces. London Mathematical Society Lecture Note Series, vol. 293. Cambridge University Press, Cambridge (2002)
23. Giuseppe Da Prato and Jerzy Zabczyk. Stochastic equations in infinite dimensions, volume 152 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, second edition, 2014
24. Robert C. Dalang. Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. Electron. J. Probab., 4:no. 6, 29, 1999
25. Robert C. Dalang and N. E. Frangos. The stochastic wave equation in two spatial dimensions. Ann. Probab., 26(1):187–212, 1998
26. Robert C. Dalang, C. Mueller, and L. Zambotti. Hitting properties of parabolic s.p.d.e.'s with reflection. Ann. Probab., 34(4):1423–1450, 2006
27. Dalang, Robert C., Quer-Sardanyons, Lluís: Stochastic integrals for spde's: a comparison. Expo. Math. **29**(1), 67–109 (2011)
28. Deya, Aurélien, Panloup, Fabien, Tindel, Samy: Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise. Ann. Probab. **47**(1), 464–518 (2019)
29. Durrett, Richard T., Iglehart, Donald L., Miller, Douglas R.: Weak convergence to Brownian meander and Brownian excursion. Ann. Probability **5**(1), 117–129 (1977)
30. Eckmann, Jean-Pierre., Hairer, Martin: Invariant measures for stochastic partial differential equations in unbounded domains. Nonlinearity **14**(1), 133–151 (2001)
31. Etheridge, Alison M., Labbé, Cyril: Scaling limits of weakly asymmetric interfaces. Comm. Math. Phys. **336**(1), 287–336 (2015)
32. Flandoli, Franco: Dissipativity and invariant measures for stochastic Navier-Stokes equations. NoDEA Nonlinear Differential Equations Appl. **1**(4), 403–423 (1994)
33. Funaki, Tadahisa, Olla, Stefano: Fluctuations for $\nabla\phi$ interface model on a wall. Stochastic Process. Appl. **94**(1), 1–27 (2001)
34. Luca Gerolla, Martin Hairer, and Xue-Mei Li. Fluctuations of stochastic pdes with long-range correlations. Preprint arXiv:2303.09811, March 2023
35. Loukas Grafakos. Modern Fourier analysis, volume 250 of Graduate Texts in Mathematics. Springer, New York, third edition, 2014
36. Yu, Gu., Li, Jiawei: Fluctuations of a nonlinear stochastic heat equation in dimensions three and higher. SIAM J. Math. Anal. **52**(6), 5422–5440 (2020)
37. Martin Hairer and Jonathan C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. Ann. of Math. (2), 164(3):993–1032, 2006
38. Yaozhong Hu, Jingyu Huang, David Nualart, and Samy Tindel. Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. Electron. J. Probab., 20:no. 55, 50, 2015

39. Kalsi, Jasdeep: Existence of invariant measures for reflected stochastic partial differential equations. J. Theoret. Probab. **33**(3), 1755–1767 (2020)
40. Nicolas Kryloff and Nicolas Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. Ann. of Math. (2), 38(1):65–113, 1937
41. Kuksin, Sergei, Shirikyan, Armen: Mathematics of two-dimensional turbulence. Cambridge Tracts in Mathematics, vol. 194. Cambridge University Press, Cambridge (2012)
42. Mattingly, Jonathan C., Pardoux, Étienne.: Malliavin calculus for the stochastic 2D Navier-Stokes equation. Comm. Pure Appl. Math. **59**(12), 1742–1790 (2006)
43. Oleksandr Misiats, Oleksandr Stanzhytskyi, and Nung Kwan Yip. Existence and uniqueness of invariant measures for stochastic reaction-diffusion equations in unbounded domains. J. Theoret. Probab., 29(3):996–1026, 2016
44. Oleksandr Misiats, Oleksandr Stanzhytskyi, and Nung Kwan Yip. Invariant measures for stochastic reaction-diffusion equations with weakly dissipative nonlinearities. Stochastics, 92(8):1197–1222, 2020
45. Mueller, Carl: Coupling and invariant measures for the heat equation with noise. Ann. Probab. **21**(4), 2189–2199 (1993)
46. Nualart, D., Pardoux, É.: White noise driven quasilinear SPDEs with reflection. Probab. Theory Related Fields **93**(1), 77–89 (1992)
47. Otobe, Yoshiki: Invariant measures for SPDEs with reflection. J. Math. Sci. Univ. Tokyo **11**(4), 425–446 (2004)
48. Revuz, Daniel, Yor, Marc: Continuous martingales and Brownian motion. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293. Springer-Verlag, Berlin (1991)
49. Marta Sanz-Solé and Mònica Sarrà. Hölder continuity for the stochastic heat equation with spatially correlated noise. In Seminar on Stochastic Analysis, Random Fields and Applications, III (Ascona, 1999), volume 52 of Progr. Probab., pages 259–268. Birkhäuser, Basel, 2002
50. Gianmario Tessitore and Jerzy Zabczyk. Invariant measures for stochastic heat equations. Probab. Math. Statist., 18(2, Acta Univ. Wratislav. No. 2111):271–287, 1998
51. John B. Walsh. An introduction to stochastic partial differential equations. In École d'été de probabilités de Saint-Flour, XIV—1984, volume 1180 of Lecture Notes in Math., pages 265–439. Springer, Berlin, 1986
52. Yang, Juan, Zhang, Tusheng: Existence and uniqueness of invariant measures for SPDEs with two reflecting walls. J. Theoret. Probab. **27**(3), 863–877 (2014)
53. Zambotti, Lorenzo: A reflected stochastic heat equation as symmetric dynamics with respect to the 3-d Bessel bridge. J. Funct. Anal. **180**(1), 195–209 (2001)
54. Zambotti, Lorenzo: Integration by parts formulae on convex sets of paths and applications to SPDEs with reflection. Probab. Theory Related Fields **123**(4), 579–600 (2002)
55. Zambotti, Lorenzo: Occupation densities for SPDEs with reflection. Ann. Probab. **32**(1A), 191–215 (2004)
56. Lorenzo Zambotti. Random obstacle problems, volume 2181 of Lecture Notes in Mathematics. Springer, Cham, 2017. Lecture notes from the 45th Probability Summer School held in Saint-Flour, 2015
57. Zambotti, Lorenzo: A brief and personal history of stochastic partial differential equations. Discrete Contin. Dyn. Syst. **41**(1), 471–487 (2021)
58. Zhang, Tusheng: Large deviations for invariant measures of SPDEs with two reflecting walls. Stochastic Process. Appl. **122**(10), 3425–3444 (2012)