Mixing via controllability for randomly forced nonlinear dissipative PDEs

Sergei Kuksin[*] Vahagn Nersesyan[†] Armen Shirikyan[‡]

Abstract

In the paper [KNS20], we studied the problem of mixing for a class of PDEs with a very degenerate bounded noise and established the uniqueness of stationary measure and its exponential stability in the dual-Lipschitz metric. One of the hypotheses imposed on the problem in question required that the unperturbed equation should have exactly one globally stable equilibrium point. In this paper, we relax that condition, assuming only global controllability to a given point. It is proved that the uniqueness of a stationary measure and convergence to it are still valid, whereas the rate of convergence is not necessarily exponential. The result is applicable to randomly forced parabolic-type PDEs, provided that the deterministic part of the external force is in general position, ensuring a regular structure for the attractor of the unperturbed problem.

AMS subject classifications: 35K58, 35R60, 37A25, 37L55, 60G50, 60H15, 76M35, 93B18, 93C20

Keywords: Markov process, stationary measure, mixing, nonlinear parabolic PDEs, Lyapunov function, Haar series

Contents

[*]Institut de Mathémathiques de Jussieu–Paris Rive Gauche, CNRS, Université Paris Diderot, UMR 7586, Sorbonne Paris Cité, F-75013, Paris, France & School of Mathematics, Shandong University, Jinan, PRC & Saint Petersburg State University, Universitetskaya nab., St. Petersburg, Russia; e-mail: Sergei.Kuksin@imj-prg.fr

[†]Laboratoire de Mathématiques, UMR CNRS 8100, UVSQ, Université Paris-Saclay, 45, av. des Etats-Unis, F-78035 Versailles, France & Centre de Recherches Mathématiques, CNRS UMI 3457, Université de Montréal, Montréal, QC, H3C 3J7, Canada; e-mail: Vahagn.Nersesyan@math.uvsq.fr

[‡]Department of Mathematics, CY Cergy Paris Université, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France & Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, QC, H3A 2K6, Canada; e-mail: Armen.Shirikyan@cyu.fr

0 Introduction

In the last twenty years, there was a substantial progress in the question of description of the long-time behaviour of solutions for PDEs with random forcing. The problem is particularly well understood when all the determining modes are directly affected by the stochastic perturbation. In this situation, for a large class of PDEs the resulting random flow possesses a unique stationary distribution, which attracts the laws of all the solutions with an exponential rate. We refer the reader to [FM95, KS00, EMS01, BKL02] for the first results in this direction and to the review papers [ES00, Bri02, Deb13] and the book [KS12] for a detailed discussion of the literature. The question of uniqueness of stationary distribution becomes much more delicate when the random forcing is very degenerate and does not act directly on all the determining modes of the evolution. In this case, the propagation of the randomness under the unperturbed dynamics plays a crucial role and may still ensure the uniqueness and stability of a stationary distribution. There are essentially two mechanisms of propagation—transport and diffusion—and they allowed one to get two groups of results. The first one deals with random forces that are localised in the Fourier space. In this situation, it was proved by Hairer and Mattingly [HM06, HM11] that the Navier–Stokes flow is exponentially mixing in the dual-Lipschitz metric, provided that the random perturbation is white in time. Földes, Glatt-Holtz, Richards, and Thomann [FGRT15] established a similar result for the Boussinesq system, assuming that a degenerate random force acts only on the equation for the temperature. The recent paper [KNS20] deals with various parabolic-type PDEs perturbed by *bounded* observable forces, which allowed for treatment of nonlinearities of arbitrary degree. The second group of results concerns random forces localised in the physical space. They were obtained in [Shi15, Shi20] for the Navier–Stokes equations in an arbitrary domain with a random perturbation distributed either in a subdomain or on the boundary.

The goal of the present paper is to relax a hypothesis in [KNS20] that required the existence of an equilibrium point which is globally asymptotically stable under the unperturbed dynamics. To illustrate our general result, let us consider the following example of a randomly forced parabolic PDE which defines (under suitable hypotheses) a random dynamical system in the Sobolev space $H^s(\mathbb{T}^d)$ with an arbitrary integer $s \geq 0$:

$$\partial_t u - \nu\Delta u + f(u) = h(x) + \eta(t,x), \quad x \in \mathbb{T}^d, \quad d \leq 4. \tag{0.1}$$

Here $\nu > 0$ is a parameter, $f : \mathbb{R} \to \mathbb{R}$ is a polynomial satisfying some natural growth and dissipativity hypotheses (see (4.2) and (4.3)), $h : \mathbb{T}^d \to \mathbb{R}$ is a smooth deterministic function, and η is a finite-dimensional *Haar coloured noise*. More precisely, let us denote by $\{h_0, h_{jl}\}$ the Haar basis in $L^2(0,1)$ defined by the relations[1]

$$h_0(t) = \begin{cases} 1 & \text{for } 0 \leq t < 1, \\ 0 & \text{for } t < 0 \text{ or } t \geq 1, \end{cases}$$

$$h_{jl}(t) = \begin{cases} 0 & \text{for } t < l2^{-j} \text{ or } t \geq (l+1)2^{-j}, \\ 1 & \text{for } l2^{-j} \leq t < \left(l + \frac{1}{2}\right)2^{-j}, \\ -1 & \text{for } \left(l + \frac{1}{2}\right)2^{-j} \leq t < (l+1)2^{-j}, \end{cases}$$

where $j \geq 0$ and $0 \leq l \leq 2^j - 1$. Note that, for any integer $k \geq 1$, the functions

$$\left\{h_0(\cdot - k + 1), h_{jl}, j \geq 0, (k-1)2^j \leq l \leq k2^j - 1\right\}$$

are supported on $[k-1, k]$ and form an orthonormal basis in $L^2([k-1,k])$. We assume that η is a random process that takes values in a sufficiently large[2] finite-dimensional subspace $\mathcal{H}$ of $L^2(\mathbb{T}^d)$ and has the form

$$\eta(t,x) = \sum_{i \in \mathcal{I}} b_i \eta^i(t) \varphi_i(x), \tag{0.2}$$

where $\{\varphi_i\}_{i \in \mathcal{I}}$ is an orthonormal basis in $\mathcal{H}$, $\{b_i\}$ are non-zero numbers, and $\{\eta^i\}$ are independent copies of a random process defined by

$$\tilde{\eta}(t) = \sum_{k=0}^{\infty} \xi_k h_0(t-k) + \sum_{j=0}^{\infty} c_j \sum_{l=0}^{\infty} \xi_{jl} h_{jl}(t). \tag{0.3}$$

In this sum, $\{c_j\}$ is a sequence given by

$$c_j = C\,(j+1)^{-q} \quad \text{for some } C > 0,\ q > 1, \tag{0.4}$$

and $\{\xi_k, \xi_{jl}\}$ are independent identically distributed (i.i.d.) scalar random variables with Lipschitz-continuous density ρ such that $\operatorname{supp}\rho \subset [-1,1]$ and $\rho(0) > 0$. Note that the restrictions of $\tilde{\eta}$ to $[k-1,k]$ (whose translations can be considered as random functions of $t \in [0,1]$) form a sequence of i.i.d. random variables in $L^2(0,1)$. Thus, the random process (0.2) is a concatenation of i.i.d. random variables $\eta_k = \eta|_{[k-1,k]}$ in $L^2([0,1], \mathcal{H})$, and the value u_k of a solution $u(t,x)$ for (0.1) at time $t = k$ depends only on u_{k-1} and η_k. Let us supplement Eq. (0.1) with the initial condition

$$u(0,x) = u_0(x), \tag{0.5}$$

where $u_0 \in H^s(\mathbb{T}^d)$. Under the above hypotheses, the restrictions to integer times of solutions for problem (0.1), (0.5) form a discrete-time Markov process, which is denoted by $(u_k, \mathbb{P}_u)$, and this Markov process is the main subject of our study.

We assume that the space $\mathcal{H}$ and the functions f and h are in general position in the sense that the following two conditions are satisfied.

[1] Note that the Haar basis used in this work differs from that of [Lam96, Section 22] by normalisation.

[2] More precisely, we require $\mathcal{H}$ to be *saturating* in the sense of Definition 4.1. Note that the saturation property does not depend on ν, and therefore we can choose the same finite-dimensional subspace $\mathcal{H}$ for all values of the viscosity.

(S) Stationary states. *The nonlinear elliptic equation*

$$-\nu\Delta w + f(w) = h(x), \quad x \in \mathbb{T}^d \tag{0.6}$$

has finitely many solutions $w_1, \dots, w_N \in H^2(\mathbb{T}^d)$.

Genericity of this condition is proved in Section 5.3, and examples are provided by the criterion established in [CI74, Section 5]; e.g., in our context with $d = 1$, one can take $f(u) = u^3 - u$ and $h = 0$.

The existence of a Lyapunov function (see (4.11)) implies that at least one of the stationary states, say w_N, is *locally asymptotically stable*.[3] This means that, for some number $\delta > 0$, the solutions of the unperturbed equation

$$\partial_t u - \nu\Delta u + f(u) = h(x) \tag{0.7}$$

that are issued from an initial condition u_0 with $\|u_0 - w_N\|_{L^2(\mathbb{T}^d)} \le \delta$ converge uniformly to w_N:

$$\lim_{t\to+\infty} \sup_{u_0\in B(w_N,\delta)} \|u(t) - w_N\|_{H^1(\mathbb{T}^d)} = 0, \tag{0.8}$$

where $B(w, \delta)$ is the ball in L^2 of radius δ centred at w. To formulate the second condition, let us denote by $\mathcal{K}$ the support of the law for the restriction to the interval $[0, 1]$ of the process (0.2) and by $S_n(u_0; \zeta_1, \dots, \zeta_n)$ the value at time $t = n$ of the solution for problem (0.1), (0.5) in which the external force η coincides with ζ_k on the time interval $[k - 1, k]$.

(C) Controllability to the neighbourhood of w_N. *For any* $1 \le i \le N - 1$, *there is an integer* n_i *and functions* $\zeta_{i1}, \dots, \zeta_{in_i} \in \mathcal{K}$ *such that*

$$\|S_{n_i}(w_i; \zeta_{i1}, \dots, \zeta_{in_i}) - w_N\|_{H^1(\mathbb{T}^d)} < \delta. \tag{0.9}$$

The validity of this condition can be derived from Agrachev–Sarychev type approximate controllability results,[4] provided that the support $\mathcal{K}$ is sufficiently large. We also note that if the δ-neighbourhood of w_N is attainable from the other stationary states, then due to the asymptotic stability any neighbourhood of w_N can be reached in a sufficiently large time. The following theorem is a consequence of the main result of this paper on the uniqueness and mixing of a stationary measure for $(u_k, \mathbb{P}_u)$. Its exact formulation and further discussions are presented in Section 4.

Main Theorem. *Under the above conditions, the Markov process* $(u_k, \mathbb{P}_u)$ *has a unique stationary measure* μ *on the space* $H^1(\mathbb{T}^d)$, *and for any other solution* $u(t)$ *of* (0.1), *we have*

$$\|\mathcal{D}(u(k)) - \mu\|_L^* \to 0 \quad \text{as } k \to \infty,$$

where $\|\cdot\|_L^*$ *stands for the dual-Lipschitz metric over the space* $H^1(\mathbb{T}^d)$, *and* $\mathcal{D}(\cdot)$ *denotes the law of a random variable.*

[3] To see this, it suffices to note that the Lyapunov function admits at least one local minimum, and any local minimum is a locally asymptotically stable stationary state.

[4] Theorem 5.5 of the Appendix establishes an approximate controllability property for Eq. (0.1). Namely, it shows that, for any $i \in [\![1, N-1]\!]$, there is an $\mathcal{H}$-valued function ζ_i such that the trajectory of Eq. (0.1) issued from w_i is in the open δ-neighbourhood of w_N at time $t = 1$. Replacing the process η in (0.2) with $a\eta$ and choosing $a \ge 1$ sufficiently large, we can ensure that $\mathcal{K}^a := \operatorname{supp} \mathcal{D}(a\eta)$ contains a function arbitrarily close to ζ_i, so that inequality (0.9) holds with $n_i = 1$.

Among the assumptions of this theorem, a key property is Hypothesis (C) which connects the underlying (deterministic) dynamics with the random perturbation given by η. A natural question is to determine what can be said if (C) is not satisfied. While in the general setting this question is not likely to have a satisfactory answer, when the noise is sufficiently small, one can prove that, for any stable stationary state, there is a unique stationary measure supported in its neighbourhood. In this context, a challenging open problem is the description of the behaviour of stationary measures in the vanishing noise limit, in the spirit of the Freidlin–Wentzell theory developed for finite-dimensional diffusion processes; see Chapter 6 in [FW84]. Note that, in the case of the nonlinear wave equation, with finitely many equilibria and a white-noise perturbation of full range, that limit is studied in [Mar17].

The paper is organised as follows. In Section 1, we formulate and discuss our main theorem on the uniqueness of a stationary measure and mixing for a discrete-time Markov process. In Section 2, we derive some preliminary results needed in the proof of the main theorem, which is established in Section 3. Application to a class of nonlinear parabolic PDEs is presented in Section 4. Finally, the Appendix gathers some auxiliary results.

Acknowledgement

This research was supported by the *Agence Nationale de la Recherche* through the grants ANR-10-BLAN 0102 and ANR-17-CE40-0006-02. SK thanks the *Russian Science Foundation* for support through the grant 18-11-00032. VN and AS were supported by the CNRS PICS *Fluctuation theorems in stochastic systems*. The research of AS was carried out within the MME-DII Center of Excellence (ANR-11-LABX-0023-01) and supported by *Initiative d'excellence Paris-Seine*. The authors thank R. Joly for the proof of the genericity of Hypothesis (S) (see Proposition 5.3) and the anonymous referees for their pertinent remarks which helped to improve the text and to remove a number of inaccuracies.

Notation

For a Polish space X with a metric $d_X(u,v)$, we denote by $B_X(a,R)$ the closed ball of radius $R>0$ centred at $a \in X$ and by $\dot{B}_X(a,R)$ the corresponding open ball. The Borel σ-algebra on X and the set of probability measures are denoted by $\mathcal{B}(X)$ and $\mathcal{P}(X)$, respectively. We shall use the following spaces, norms, and metrics.

$C_b(X)$ denotes the space of bounded continuous functions $f: X \to \mathbb{R}$ endowed with the norm $\|f\|_\infty = \sup_X |f|$, and $L_b(X)$ stands for the space of functions $f \in C_b(X)$ such that

$$\|f\|_L := \|f\|_\infty + \sup_{0<d_X(u,v)\le 1} \frac{|f(u)-f(v)|}{d_X(u,v)} < \infty.$$

In the case of a compact space X, we write $C(X)$ and $L(X)$.

The space $\mathcal{P}(X)$ is endowed with either the *total variation metric* or the *dual-Lipschitz* metric. They are defined by

$$\|\mu_1-\mu_2\|_{\mathrm{var}} := \sup_{\Gamma\in\mathcal{B}(X)} |\mu_1(\Gamma)-\mu_2(\Gamma)| = \frac{1}{2} \sup_{\|f\|_\infty\le 1} |\langle f,\mu_1\rangle - \langle f,\mu_2\rangle|, \qquad (0.10)$$

$$\|\mu_1-\mu_2\|_L^* := \sup_{\|f\|_L\le 1} |\langle f,\mu_1\rangle - \langle f,\mu_2\rangle|, \qquad (0.11)$$

where $\mu_1, \mu_2 \in \mathcal{P}(X)$, and $\langle f, \mu\rangle = \int_X f(u)\,\mu(\mathrm{d}u)$ for $f \in C_b(X)$ and $\mu \in \mathcal{P}(X)$.

$L^p(J, E)$ is the space Borel-measurable functions f on an interval $J \subset \mathbb{R}$ with range in a separable Banach space E such that

$$\|f\|_{L^p(J,E)} = \left(\int_J \|f(t)\|_E^p \mathrm{d}t\right)^{1/p} < \infty;$$

in the case $p = \infty$, this norm should be modified accordingly.

$H^s(D)$ denote the Sobolev space of order $s \geq 0$ with the usual norm $\|\cdot\|_s$.

1 Main result

Let us denote by H and E separable Hilbert spaces and by $S : H \times E \to H$ a continuous mapping. Given a sequence $\{\eta_k\}$ of i.i.d. random variables in E, we consider the random dynamical system (RDS)

$$u_k = S(u_{k-1}, \eta_k), \quad k \geq 1. \tag{1.1}$$

In what follows, we always assume that the law ℓ of the random variables η_k has a compact support $\mathcal{K} \subset E$ and that there is a compact set $X \subset H$ such that $S(X \times \mathcal{K}) \subset X$. Our aim is to study the long-time behaviour of the restriction of the RDS (1.1) to the invariant set X.

For a vector $u \in H$ and a sequence $\{\zeta_k\} \subset E$, we set $S_m(u; \zeta_1, \ldots, \zeta_m) := u_m$, where $\{u_k\}$ is defined recursively by Eq. (1.1) in which $u_0 = u$ and $\eta_k = \zeta_k$. We assume that the hypotheses below hold for the RDS (1.1) and some Hilbert space V compactly embedded into H.

(H$_1$) Regularity. *The mapping S is twice continuously differentiable from $H \times E$ to V, and its derivatives are bounded on bounded subsets. Moreover, for any fixed $u \in H$, the mapping $\eta \mapsto S(u, \eta)$ is analytic from E to H, and all its derivatives $(D_\eta^j S)(u, \eta)$ are continuous functions of (u, η) that are bounded on bounded subsets of $H \times E$.*

(H$_2$) Approximate controllability to a point. *There is $\hat{u} \in X$ such that, for any $\varepsilon > 0$, one can find an integer $m \geq 1$ with the following property: for any $u \in X$ there are $\zeta_1, \ldots, \zeta_m \in \mathcal{K}$ such that*

$$\|S_m(u; \zeta_1, \ldots, \zeta_m) - \hat{u}\| < \varepsilon. \tag{1.2}$$

Given $u \in X$, let us denote by $\mathcal{K}^u$ the set of those $\eta \in E$ for which the image of the derivative $(D_\eta S)(u, \eta)$ is dense in H. It is easy to see that $\mathcal{K}^u$ is a Borel subset in E; see Section 1.1 in [KNS20].

(H$_3$) Approximate controllability of the linearisation. *For any $u \in X$, the set $\mathcal{K}^u$ has full ℓ-measure.*

(H$_4$) Structure of the noise. *There exists an orthonormal basis $\{e_j\}$ in E, independent random variables ξ_{jk}, and real numbers b_j such that*

$$\eta_k = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j, \quad B := \sum_{j=1}^{\infty} b_j^2 < \infty. \tag{1.3}$$

Moreover, the laws of ξ_{jk} have Lipschitz-continuous densities ρ_j with respect to the Lebesgue measure on $\mathbb{R}$.

We refer the reader to Section 1.1 in [KNS20] for a discussion of these conditions and of their relevance in the study of large-time asymptotics of trajectories for PDEs with random forcing. Here we only mention that the approximate controllability hypothesis (H_2) imposed in this paper is weaker than the dissipativity condition of [KNS20] and allows one to treat a much larger class of PDEs that possess several steady states. A drawback is that the main result of this paper does not give any estimate for the rate of convergence (to the unique stationary measure), which remains an interesting open problem.

To formulate our main abstract result, we introduce some notation. Since $\{\eta_k\}$ are i.i.d. random variables, the trajectories of (1.1) issued from X form a discrete-time Markov process, which is denoted by $(u_k, \mathbb{P}_u)$. We shall write $P_k(u, \Gamma)$ for its transition function and $\mathfrak{P}_k : C_b(X) \to C_b(X)$ and $\mathfrak{P}_k^* : \mathcal{P}(X) \to \mathcal{P}(X)$ for the corresponding Markov operators.

Let us recall that a measure $\mu \in \mathcal{P}(H)$ is said to be *stationary* for $(u_k, \mathbb{P}_u)$ if $\mathfrak{P}_1^*\mu = \mu$. The continuity of S implies that $(u_k, \mathbb{P}_u)$ possesses the Feller property, and by the Bogolyubov–Krylov argument and the compactness of X, there is at least one stationary measure. We wish to investigate its uniqueness and stability.

Let $\|\cdot\|_L^*$ be the dual-Lipschitz metric on the space of probability measures on X (see Notation). The following theorem, which is the main result of this paper, describes the behaviour of $\mathfrak{P}_k^*$ as the time goes to infinity.

Theorem 1.1. *Suppose that Hypotheses (H_1)–(H_4) are satisfied. Then the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu \in \mathcal{P}(X)$, and there is a sequence of positive numbers $\{\gamma_k\}$ going to zero as $k \to \infty$ such that*

$$\|\mathfrak{P}_k^*\lambda - \mu\|_L^* \le \gamma_k \quad \text{for all } k \ge 0 \text{ and } \lambda \in \mathcal{P}(X). \tag{1.4}$$

A proof of this result is given in Section 3. Here we discuss very briefly the main idea, postponing the details to Section 3.1.

A sufficient condition for the validity of the conclusions is given by Theorem 5.1 in the Appendix. According to that result, it suffices to check the recurrence and stability properties. The recurrence is a simple consequence of the approximate controllability to the point $\hat{u}$; see Hypothesis (H_2). The proof of stability is much more involved and will follow from two properties, (A) and (B), of Theorem 3.1. Their verification is based on a key new idea of this work, which reduces the required properties to some estimates for sums of independent Bernoulli variables. The latter is discussed in Section 2, together with an auxiliary result on the transformation of the noise space E (which was established in [KNS20]) and a property of continuous probability measures.

2 Preliminary results

2.1 Transformation in the control space

Given a number $\delta > 0$, we set $D_\delta := \{(u, u') \in X \times H : \|u - u'\| \le \delta\}$. Recall that ℓ stands for the law of the random variables η_k; see (1.1). The following proposition is established in Section 3.2 of [KNS20] (see there Proposition 3.3 with $\sigma = 1/4$).

Proposition 2.1. *Suppose that Hypotheses* (H_1), (H_3), *and* (H_4) *are satisfied. Then, for any* $\theta \in (0,1)$, *there are positive numbers* C, β, *and* δ, *a family of Borel subsets* $\{\mathcal{K}^{u,\theta} \subset \mathcal{K}^u\}_{u\in X}$, *and a measurable mapping* $\Phi : X \times H \times E \to E$ *such that* $\Phi^{u,u'}(\eta) = 0$ *if* $\eta \notin \mathcal{K}^{u,\theta}$ *or* $u' = u$, *and*

$$\ell(\mathcal{K}^{u,\theta}) \geq 3/4, \tag{2.1}$$

$$\|\ell - \Psi_*^{u,u'}(\ell)\|_{\mathrm{var}} \leq C\,\|u - u'\|^{\beta}, \tag{2.2}$$

$$\|S(u,\eta) - S(u', \Psi^{u,u'}(\eta))\| \leq \theta\,\|u - u'\|, \tag{2.3}$$

where $\Psi^{u,u'}(\eta) := \eta + \Phi^{u,u'}(\eta)$, $\Psi_*^{u,u'}(\ell)$ *is the image of the measure* ℓ *under* $\Psi^{u,u'}$, *and* $(u,u') \in D_\delta$ *and* $\eta \in \mathcal{K}^{u,\theta}$ *are arbitrary points.*

2.2 Estimates for sums of i.i.d. Bernoulli variables

Let $\{w_k\}_{k\geq 1}$ be a sequence of i.i.d. random variables such that

$$\mathbb{P}\{w_k = 1\} = p, \quad \mathbb{P}\{w_k = -1\} = 1 - p,$$

where $p \in (\frac{1}{2}, 1)$. We define

$$\zeta_k = \sum_{j=1}^{k} w_j, \quad M_k = \zeta_k - (2p-1)k. \tag{2.4}$$

A proof of the following result can be found in [Lam96, Section 12] and [Fel68, Section XIV.2].

Proposition 2.2. **(a)** *For any* $\varepsilon > 0$ *there is a random time* $\tau = \tau(\varepsilon, p) \geq 1$ *and a number* $\alpha = \alpha(\varepsilon, p) > 0$ *such that*

$$M_k \geq -\varepsilon k \quad \text{for } k \geq \tau, \tag{2.5}$$

$$\mathbb{E}\, e^{\alpha\tau} < \infty. \tag{2.6}$$

(b) *For any integer* $l \geq 0$, *we have*

$$\mathbb{P}\big\{\zeta_k > -l \text{ for all } k \geq 0\big\} = 1 - \big(\tfrac{1-p}{p}\big)^l. \tag{2.7}$$

A consequence of this proposition is the following result, which will be used in Section 3.3.

Corollary 2.3. *For any* $c \in (0, 2p-1)$, *there is a sequence* $\{p_l\} \subset \mathbb{R}$ *depending only on* c *and* p *such that*

$$\mathbb{P}\{\zeta_k \geq -l + ck \text{ for all } k \geq 0\} \geq p_l \quad \text{for all } l \geq 1, \tag{2.8}$$

$$p_l \to 1 \quad \text{as } l \to \infty. \tag{2.9}$$

Proof. Applying (2.5) with $\varepsilon = 2p - 1 - c$, we see that $\zeta_k \geq ck$ for $k \geq \tau$. By the Markov inequality and (2.6), we have

$$\mathbb{P}\{\tau > l\} \leq Ce^{-\alpha l} \quad \text{for } l \geq 1.$$

It follows that

$$\mathbb{P}\{\zeta_k \geq ck \text{ for } k \geq l\} \geq 1 - Ce^{-\alpha l} \quad \text{for } l \geq 1. \tag{2.10}$$

On the other hand, it follows from (2.7) that

$$\mathbb{P}\{\zeta_k \geq -l + ck \text{ for } 0 \leq k \leq l\} \geq 1 - \left(\tfrac{1-p}{p}\right)^{[(1-c)l]} \quad \text{for } l \geq 1,$$

where $[a]$ stands for the integer part of a. Combining this with (2.10), we obtain (2.8) with

$$p_l := 1 - \left(\tfrac{1-p}{p}\right)^{[(1-c)l]} - Ce^{-\alpha l}.$$

Since $c < 1$, we have limit (2.9). □

2.3 Continuous probability measures

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We shall say that $\mathbb{P}$ is *continuous* if for any $\Gamma \in \mathcal{F}$ and $p \in [0, \mathbb{P}(\Gamma)]$ there is $\Gamma_p \in \mathcal{F}$ such that $\Gamma_p \subset \Gamma$ and $\mathbb{P}(\Gamma_p) = p$. Given a measurable space $(X, \mathcal{B})$ and measurable mapping $F : \Omega \to X$, we say that $\mathbb{P}$ *admits a disintegration with respect to* $\mathbb{Q} = F_*(\mathbb{P})$ if there is a random probability measure $\{P(x, \cdot)\}_{x\in X}$ on $(\Omega, \mathcal{F})$ such that

$$\mathbb{P}\big(A \cap F^{-1}(B)\big) = \int_B P(x, A)\mathbb{Q}(\mathrm{d}x) \quad \text{for any } A \in \mathcal{F},\ B \in \mathcal{B}. \tag{2.11}$$

The following result provides a simple sufficient condition for continuity of a probability measure.

Lemma 2.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $F : \Omega \to \mathbb{R}$ be a measurable mapping such that $\mathbb{Q} = F_*(\mathbb{P})$ has a density ρ with respect to the Lebesgue measure and $\mathbb{P}$ admits a disintegration $P(s, A)$ with respect to $\mathbb{Q}$. Then $\mathbb{P}$ is continuous.*

Proof. Given $\Gamma \in \mathcal{F}$, we define $\Gamma(r) = \Gamma \cap F^{-1}((-\infty, r]) \in \mathcal{F}$, where $r \in \mathbb{R}$. Then $\mathbb{P}(\Gamma(r))$ converges to 0 as $r \to -\infty$ and to $\mathbb{P}(\Gamma)$ as $r \to +\infty$. Moreover, by (2.11), we have

$$\mathbb{P}\big(\Gamma(r)\big) = \int_{-\infty}^{r} P(s, \Gamma)\rho(s)\,\mathrm{d}s,$$

whence we see that the function $r \mapsto \mathbb{P}(\Gamma(r))$ is continuous. The required result follows from the intermediate value theorem. □

We now apply the above idea to deal with a construction that will be used in Section 3. Namely, let $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$, $i = 1, 2$ be two probability spaces and let $(\Omega, \mathcal{F}, \mathbb{P})$ be their direct product. With a slight abuse of notation, we write $\mathcal{F}_i$ for the sub-σ-algebra on Ω generated by the natural projection $\Omega \to \Omega_i$.

Lemma 2.5. *In addition to the above hypotheses, suppose that the probability space $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ and a function $F : \Omega_2 \to \mathbb{R}$ satisfy the conditions of Lemma 2.4, and let $\Gamma \in \mathcal{F}$ be such that, for some $p \in (0, 1)$,*

$$\mathbb{E}(I_\Gamma \,|\, \mathcal{F}_1) \geq p \quad \mathbb{P}\text{-almost surely.} \tag{2.12}$$

Then there is $\Gamma' \in \mathcal{F}$ such that $\Gamma' \subset \Gamma$ and

$$\mathbb{E}(I_{\Gamma'} \,|\, \mathcal{F}_1) = p \quad \mathbb{P}\text{-almost surely.} \tag{2.13}$$

Proof. We first reformulate the lemma in somewhat different terms. Given $\Gamma \in \mathcal{F}$ and $\omega_1 \in \Omega_1$, we denote

$$\Gamma(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in \Gamma\}.$$

It is straightforward to check that

$$\mathbb{E}(I_\Gamma \,|\, \mathcal{F}_1) = \mathbb{P}_2\bigl(\Gamma(\omega_1)\bigr).$$

Furthermore, the inclusion $\Gamma' \subset \Gamma$ holds if and only if $\Gamma'(\omega_1) \subset \Gamma(\omega_1)$ for any $\omega_1 \in \Omega_1$. Thus, the lemma is equivalent to the following assertion: if $\Gamma \in \mathcal{F}$ is such that $\mathbb{P}_2(\Gamma(\omega_1)) \geq p$ for $\mathbb{P}_1$-a.e. $\omega_1 \in \Omega_1$, then there is $\Gamma' \in \mathcal{F}$ such that $\Gamma'(\omega_1) \subset \Gamma(\omega_1)$ for any $\omega_1 \in \Omega_1$ and $\mathbb{P}_2(\Gamma'(\omega_1)) = p$ for $\mathbb{P}_1$-a.e. $\omega_1 \in \Omega_1$.

Given a real-valued measurable function $r(\omega_1)$, we define

$$\Gamma' = \{(\omega_1, \omega_2) \in \Gamma : F(\omega_2) \leq r(\omega_1)\} \subset \Gamma.$$

Then $\Gamma'(\omega_1) = \Gamma(\omega_1) \cap F^{-1}((-\infty, r(\omega_1)])$, so that

$$\mathbb{E}(I_{\Gamma'} \,|\, \mathcal{F}_1) = \mathbb{P}_2(\Gamma'(\omega_1)) = \int_{-\infty}^{r(\omega_1)} P\bigl(s, \Gamma(\omega_1)\bigr)\rho(s)\,\mathrm{d}s, \tag{2.14}$$

where $P(s, \cdot)$ stands for the disintegration of $\mathbb{P}_2$ with respect to $F_*(\mathbb{P}_2)$, and ρ is the density of $F_*(\mathbb{P}_2)$ with respect to the Lebesgue measure. Consider the measurable function

$$G(\omega_1, t) = \int_{-\infty}^{t} P(s, \Gamma(\omega_1))\rho(s) ds, \quad t \in \mathbb{R}.$$

It is continuous in t, vanishes when $t = -\infty$ and is $\geq p$ when $t = +\infty$. Consider the set $\{t \in \mathbb{R} : G(\omega_1, t) \leq p\}$. This is a measurable set which is the sub-graph of certain measurable function $t = \lambda(\omega_1)$. Choosing $r(\omega_1) = \lambda(\omega_1)$, we see that the right-hand side of (2.14) is identically equal to p, and Γ' is a pre-image of the above-mentioned sub-graph under the measurable mapping $(\omega_1, \omega_2) \mapsto (\omega_1, F(\omega_2))$. We conclude that Γ' is measurable, which completes the proof. ☐

3 Proof of the main theorem

Let us explain the main idea of the proof informally. As was mentioned after the formulation of Theorem 1.1, the crucial point is the proof of the stability inequality (5.2). Theorem 3.1 reduces that inequality to the construction of an auxiliary Markov process (u_k, v_k) in the extended phase space $X \times X$ such that

$$\mathbb{P}\{\|u_1 - v_1\| \leq \theta \|u - u'\|\} \geq p, \quad \mathbb{P}\{\|u_1 - v_1\| > \theta^{-1} \|u - u'\|\} = 0, \tag{3.1}$$

where $\theta \in (0, 1)$ and $p \in (\frac{1}{2}, 1)$ are some parameters, and (u, u') is an initial state. We shall prove that the required stability inequality is a consequence of the almost sure convergence $\|u_k - v_k\| \to 0$ as $k \to \infty$. Relations (3.1) show that, with probability at least $p > 1/2$, the distance between the components contracts by a factor of $\theta < 1$, and if this does not happen, then the distance can expand by a factor of at most θ^{-1}. Stratifying the space $X \times X$ according to the distance between the components of a point (u, u') (see (3.18)), we obtain a $\mathbb{Z}$-valued random walk $\{\xi_k\}$ in a dynamic random environment, and the required property translates as the almost sure convergence of ξ_k to $+\infty$. We then prove that ξ_k can be minorised by a sum of independent Bernoulli variables and use the results of Section 2.2.

3.1 General scheme

We wish to apply a sufficient condition for mixing from [KS12, Section 3.1.2], stated below as Theorem 5.1. To this end, we need to check the recurrence and stability conditions. The recurrence is a consequence of Hypothesis (H_2). Indeed, inequality (1.2) implies that $P_m(u, B_X(\hat{u}, r)) > 0$ for any $u \in X$ and some integer $m = m_r \geq 1$. Since the function $u \mapsto P_m(u, \dot{B}_X(\hat{u}, r))$ is lower semicontinuous and positive, it is separated from zero on the compact set X, so that (5.1) holds. We thus need to prove the stability. We shall always assume that the hypotheses of Theorem 1.1 are satisfied. Recall that, given $\delta > 0$, we write $D_\delta = \{(u, u') \in X \times H : \|u - u'\| \leq \delta\}$. The following result provides a sufficient condition for the validity of (5.2).

Theorem 3.1. *Suppose there is a measurable mapping $\Psi : X \times H \times E \to E$, taking (u, u', η) to $\Psi^{u,u'}(\eta)$, and positive numbers α, β, and $q \in (0, 1)$ such that $\Psi^{u,u}(\eta) = \eta$ for any $u \in X$ and $\eta \in E$, and the following properties hold.*

(A) Stabilisation. *For any $u, u' \in H$, let (u_k, v_k) be defined by*

$$(u_0, v_0) = (u, u'), \tag{3.2}$$
$$(u_k, v_k) = (S(u_{k-1}, \eta_k), S(v_{k-1}, \Psi^{u_{k-1}, v_{k-1}}(\eta_k))). \tag{3.3}$$

Let us introduce the stopping time

$$\tau = \min\big\{k \geq 1 : \|u_k - v_k\| > q^k \|u - u'\|^\alpha\big\} \tag{3.4}$$

and, for any $\delta > 0$, define the quantity

$$p(\delta) = \inf_{(u,u') \in D_\delta} \mathbb{P}\{\tau = +\infty\}.$$

Then

$$\lim_{\delta \to 0} p(\delta) = 1. \tag{3.5}$$

(B) Transformation of measure. *For any $(u, u') \in X \times H$, we have*

$$\|\ell - \Psi_*^{u,u'}(\ell)\|_{\mathrm{var}} \leq C\|u - u'\|^\beta. \tag{3.6}$$

Then condition (5.2) is valid:

$$\lim_{\delta \to 0^+} \sup_{(u,u') \in D_\delta} \sup_{k \geq 0} \|P_k(u, \cdot) - P_k(u', \cdot)\|_L^* = 0. \tag{3.7}$$

Remark 3.2. In terms of the PDE (0.1), the hypotheses of this theorem can be reformulated informally as follows. If we are given two sufficiently close initial conditions, u_0 and u_0', then we can modify the right-hand side η of the equation for the second one so that the difference between the corresponding solutions $u(t)$ and $u'(t)$ goes to zero exponentially fast, unless the realisation of the noise belongs to an exceptional set whose measure goes to zero with the distance $\|u_0 - u_0'\|$. The modification of the right-hand side for u' is obtained by iterations of a map $\Psi^{u,u'}(\eta)$ that does not change much the law of η. As will be established in the proof of the Main Theorem (see the Introduction), the hypotheses of Theorem 3.1 are satisfied for (0.1). Thus, as a by-product, we obtain a stabilisation result for the parabolic PDE (0.1) with a finite-dimensional control η.

Theorem 3.1 is established in Section 3.2. Note that if the constant C in the right-hand side of (3.6) vanishes, then the random variables η_k and $\eta'_k := \Psi^{u_{k-1},v_{k-1}}(\eta_k)$ form a coupling for the pair of measures (ℓ,ℓ), so that $\mathcal{D}(v_k)=\mathcal{D}(u'_k)$, where $\{u'_k, k\geq 0\}$ solves (1.1) with $u'_0=u'$. In this case, we deal with the classical coupling approach to compare $P_k(u,\cdot)$ and $P_k(u',\cdot)$. Our proof of Theorem 1.1 crucially uses the above result with $\|u-u'\|\ll 1$. The right-hand side of (3.6) is not zero in this situation, but it is small, so we deal with a kind of approximate coupling.

To prove Theorem 1.1 given Theorem 3.1, it suffices to construct a measurable mapping Ψ satisfying Conditions (A) and (B). This will be done with the help of Proposition 2.1.

3.2 Proof of Theorem 3.1

Let us define a probability space $(\Omega,\mathcal{F},\mathbb{P})$ by the relations

$$\boldsymbol{\Omega}=\{\boldsymbol{\omega}=(\omega_k)_{k\geq 1}:\omega_k\in E\},\quad \mathcal{F}=\mathcal{B}(\boldsymbol{\Omega}),\quad \mathbb{P}=\bigotimes_{k=1}^{\infty}\ell,$$

where $\boldsymbol{\Omega}$ is endowed with the Tikhonov topology. Let $(u_k(\boldsymbol{\omega}),v_k(\boldsymbol{\omega}))$ be the trajectory of (3.2), (3.3) with $\eta_k\equiv\omega_k$ and let $u'_k(\boldsymbol{\omega})$ be the trajectory of (1.1) with $u_0=u'$ and $\eta_k\equiv\omega_k$. Given $u,u',y,z\in H$, we define mappings $\theta_k:E\to E$, $k\geq 1$ by

$$\theta_k(y,z,\omega)=\begin{cases}\Psi^{y,z}(\omega) & \text{if}\quad \|y-z\|\leq q^{k-1}\|u-u'\|^{\alpha},\\ \omega & \text{if}\quad \|y-z\|> q^{k-1}\|u-u'\|^{\alpha},\end{cases}\tag{3.8}$$

where Ψ is constructed in Proposition 2.1, and consider the mapping

$$\Theta:\boldsymbol{\Omega}\to\boldsymbol{\Omega},\quad \Theta(\boldsymbol{\omega})=\big(\theta_k(u_{k-1}(\boldsymbol{\omega}),v_{k-1}(\boldsymbol{\omega}),\omega_k)\big)_{k\geq 1}.$$

Clearly, $\{u_k(\boldsymbol{\omega})\}_{k\geq 0}$ is a trajectory of (1.1) with $u_0=u$, and

$$v_k(\boldsymbol{\omega})=u'_k(\Theta(\boldsymbol{\omega}))\quad\text{for } k\geq 1,\ \boldsymbol{\omega}\in\{\tau=+\infty\}.\tag{3.9}$$

We now write

$$\|P_k(u,\cdot)-P_k(u',\cdot)\|_L^*\leq\|P_k(u,\cdot)-\mathcal{D}(v_k)\|_L^*+\|\mathcal{D}(v_k)-P_k(u',\cdot)\|_L^*\tag{3.10}$$

and estimate the two terms on the right-hand side. For $(u,u')\in D_\delta$, we have

$$\begin{aligned}\|P_k(u,\cdot)-\mathcal{D}(v_k)\|_L^* &= \sup_{\|F\|_L\leq 1}\big|\mathbb{E}(F(u_k)-F(v_k))\big|\\ &\leq 2\mathbb{P}\{\tau<\infty\}+\mathbb{E}\big(I_{\{\tau=\infty\}}\|u_k-v_k\|\big)\\ &\leq 2(1-p(\delta))+\delta^{\alpha}q^k.\end{aligned}\tag{3.11}$$

To estimate the second term on the right-hand side of (3.10), we use the following simple result, in which $G=\{\tau=+\infty\}$. Its proof is given at the end of this section.

Lemma 3.3. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let X be a Polish space, and let $U,V:\Omega\to X$ two random variables. Suppose there is a measurable mapping $\Theta:\Omega\to\Omega$ such that*

$$U(\Theta(\omega))=V(\omega)\quad\textit{for }\omega\in G,\tag{3.12}$$

where $G\in\mathcal{F}$. Then

$$\|\mathcal{D}(U)-\mathcal{D}(V)\|_{\mathrm{var}}\leq 2\,\mathbb{P}(G^c)+\|\mathbb{P}-\Theta_*(\mathbb{P})\|_{\mathrm{var}}.\tag{3.13}$$

In view of (3.9) and (3.13), we have

$$\begin{aligned}\|\mathcal{D}(v_k) - P_k(u',\cdot)\|_L^* &\le 2\,\|\mathcal{D}(v_k) - P_k(u',\cdot)\|_{\mathrm{var}}\\ &\le 4\,\mathbb{P}\{\tau<\infty\} + 2\|\mathbb{P} - \Theta_*(\mathbb{P})\|_{\mathrm{var}}.\end{aligned} \tag{3.14}$$

The first term on the right-hand side does not exceed $4(1-p(\delta))$. Substituting (3.14) and (3.11) in (3.10), we see that (3.7) will be established if we show that

$$\sup_{(u,u')\in D_\delta} \|\mathbb{P} - \Theta_*(\mathbb{P})\|_{\mathrm{var}} \to 0 \quad \text{as } \delta\to 0. \tag{3.15}$$

To prove this, we use the second relation in (0.10) to calculate the total variation distance between two measures $\mu_1,\mu_2\in\mathcal{P}(\boldsymbol{\Omega})$. Obviously, it suffices to consider the functions F belonging to a dense subset of $C(\boldsymbol{\Omega})$ and satisfying the inequality $\|F\|_\infty\le 1$. Hence, the supremum can be taken over all functions depending on finitely many coordinates.

We thus fix any integer $m\ge 1$ and consider an arbitrary continuous function $F:\boldsymbol{\Omega}\to\mathbb{R}$ of the form $F(\boldsymbol{\omega}) = F(\omega_1,\dots,\omega_m)$ with $\|F\|_\infty\le 1$. Then

$$\begin{aligned}\langle F,\mathbb{P}-\Theta_*(\mathbb{P})\rangle &= \mathbb{E}\,\{F(\omega_1,\dots,\omega_m) - F(\theta_1(u,u',\omega_1),\dots,\theta_m(u_{m-1},v_{m-1},\omega_m)\}\\ &= \sum_{k=1}^m \mathbb{E}F_k(u,u',\omega_1,\dots,\omega_m),\end{aligned} \tag{3.16}$$

where we set

$$\begin{aligned}F_k(u,u',\omega_1,\dots,\omega_m) &= F(\theta_1(u,u',\omega_1),\dots,\theta_{k-1}(u_{k-2},v_{k-2},\omega_{k-1}),\omega_k,\dots,\omega_m)\\ &\quad - F(\theta_1(u,u',\omega_1),\dots,\theta_k(u_{k-1},v_{k-1},\omega_k),\omega_{k+1},\dots,\omega_m).\end{aligned}$$

Let $\mathcal{F}_k\subset\mathcal{F}$ be the σ-algebra generated by the first k coordinates. Setting

$$\Delta_k = F(x_1,\dots,x_{k-1},\omega_k,\dots,\omega_m) - F(x_1,\dots,x_{k-1},\theta_k(y,z,\omega_k),\omega_{k+1},\dots,\omega_m),$$

we note that

$$|\mathbb{E}\,\Delta_k| \le \|\ell - \theta_{k*}(y,z,\ell)\|_{\mathrm{var}} \le I_{[0,q^{k-1}\|u-u'\|^\alpha]}(\|y-z\|)\|\ell - \Psi_*^{y,z}(\ell)\|_{\mathrm{var}},$$

where we used (3.8). Combining this with (3.6), we derive

$$\big|\mathbb{E}\big(F_k(u,u')\,|\,\mathcal{F}_{k-1}\big)\big| = |\mathbb{E}\,\Delta_k| \le Cq^{\beta(k-1)}\|u-u'\|^{\alpha\beta},$$

where one takes $x_j = \theta_j(u_{j-1},v_{j-1},\omega_j)$, $y=u_{k-1}$, and $z=v_{k-1}$ in the middle term after calculating the mean value. Substituting this into (3.16), we obtain

$$|\langle F,\mathbb{P}\rangle - \langle F,\Theta_*(\mathbb{P})\rangle| \le \mathbb{E}\sum_{k=1}^m \big|\mathbb{E}\big(F_k(u,u')\,|\,\mathcal{F}_{k-1}\big)\big| \le C_1\|u-u'\|^{\alpha\beta}.$$

Taking the supremum over F with $\|F\|_\infty\le 1$, we see that (3.15) holds.

Proof of Lemma 3.3. Let $\mu=\mathcal{D}(U)$ and $\nu=\mathcal{D}(V)$. Then, for any $\Gamma\in\mathcal{F}$, we have

$$\begin{aligned}\mu(\Gamma)-\nu(\Gamma) &= \mathbb{E}\big(I_\Gamma(U) - I_\Gamma(V)\big)\\ &\le \mathbb{P}(G^c) + \mathbb{E}\big\{(I_\Gamma(U) - I_\Gamma(U\circ\Theta))I_G\big\}\\ &\le 2\,\mathbb{P}(G^c) + \mathbb{E}\big(I_\Gamma(U) - I_\Gamma(U\circ\Theta)\big) \le 2\,\mathbb{P}(G^c) + \|\mathbb{P}-\Theta_*(\mathbb{P})\|_{\mathrm{var}},\end{aligned}$$

where we used (3.12) for the first inequality. A similar argument enables one to bound $\nu(\Gamma)-\mu(\Gamma)$ by the same expression. Since $\Gamma\in\mathcal{F}$ was arbitrary, these two estimates imply the required inequality (3.13). □

3.3 Completion of the proof

We need to prove that Property (A) of Theorem 3.1 is satisfied for the Markov process (1.1) and the mapping Ψ constructed in Proposition 2.1 with an appropriate choice of θ. To this end, we fix $R>0$ so large that $X\subset B_H(R-1)$ and $\mathcal{K}\subset B_E(R)$, and choose $\theta<1$ such that

$$\|S(u,\eta)-S(u',\eta)\|\le\theta^{-1}\|u-u'\|\quad\text{for } u,u'\in B_H(R),\ \eta\in B_E(R).\tag{3.17}$$

Let us denote by $\delta>0$ and $\Psi: X\times H\times E\to E$ the number and mapping constructed in Proposition 2.1. Given $(u,u')\in X\times H$, let (u_k,v_k) be the random sequence given by (3.2), (3.3). Without loss of generality, we assume that the underlying probability space $(\Omega,\mathcal{F},\mathbb{P})$ coincides with the tensor product of countably many copies of $(E,\mathcal{B}(E),\ell)$ and denote by $\{\mathcal{F}_k\}_{k\ge1}$ the corresponding filtration. For any $(u,u')\in D_\delta$, let $N=N(u,u')\ge0$ be the smallest integer such that

$$\theta^{-N}\|u-u'\|\ge\delta.$$

We define the sets $\boldsymbol{X}_n$, $n\ge -N$ by the relation

$$\boldsymbol{X}_n=\{(v,v')\in D_\delta:\theta^{n+1}\|u-u'\|<\|v-v'\|\le\theta^n\|u-u'\|\}.\tag{3.18}$$

It is clear that the union of the sets $\cup_{n\ge-N}\boldsymbol{X}_n$ and the diagonal $\{(v,v):v\in X\}$ coincides with D_δ. Given $(u,u')\in D_\delta$, let us consider a random sequence $\{\xi_k\}_{k\ge0}$ given by [5]

$$\xi_k=\begin{cases}+\infty & \text{if } u_k=v_k,\\ n & \text{if } (u_k,v_k)\in\boldsymbol{X}_n,\\ -N-1 & \text{if } (u_k,v_k)\notin D_\delta.\end{cases}$$

In particular, we have $\xi_0=0$, and if $\xi_m=+\infty$ for some integer $m\ge1$, then $\xi_k=+\infty$ for $k\ge m$ (since $\Psi^{u,u}(\eta)=\eta$ for any $u\in X$ and $\eta\in E$). Suppose we have proved that

$$\mathbb{P}\{\xi_k\ge -l+ck \text{ for all } k\ge1\}\ge p_l\quad\text{for } \|u-u'\|\le\delta\theta^{2l},\tag{3.19}$$

where the sequence $\{p_l\}$ and the number $c>0$ do not depend on (u,u'), and $p_l\to1$ and $l\to\infty$. Then, in view of (3.18), on the set $\{\xi_k\ge-l+ck\}$, we have

$$\|u_k-v_k\|\le\theta^{-l+ck}\|u-u'\|\le\delta^{1/2}\theta^{ck}\|u-u'\|^{1/2}\le\theta^{ck}\|u-u'\|^{1/2},$$

since we can assume that $\delta<1$. It follows that if we take $q=\theta^c$ and $\alpha=\frac12$, then the random time τ defined by (3.4) will satisfy the inequality $\mathbb{P}\{\tau=+\infty\}\ge p_l$. We thus obtain (3.5). Hence, it remains to prove (3.19). To this end, we shall use Corollary 2.3.

If $\|u-u'\|\le\delta\theta^{2l}$ and $(u_{k-1},v_{k-1})\in X_n$ for some integer $n\ge-2l$, then $\|u_{k-1}-v_{k-1}\|\le\delta$. So inequality (2.3) applies, and combining it with (2.1) and (3.17), we see that

$$\mathbb{P}\{\xi_k-\xi_{k-1}\ge1\,|\,\mathcal{F}_{k-1}\}\ge\frac34\quad\text{on the set } \{\xi_{k-1}\ge-2l\},\tag{3.20}$$

$$\mathbb{P}\{\xi_k-\xi_{k-1}\ge-1\,|\,\mathcal{F}_{k-1}\}=1\quad\text{almost surely},\tag{3.21}$$

where $k\ge1$ is an arbitrary integer. Let us consider the event

$$\Gamma_k:=\{\xi_k-\xi_{k-1}\ge1\}.$$

[5] To simplify the notation, we do not indicate the dependence on (u,u') for $\boldsymbol{X}_n$ and ξ_k (as well as for the events Γ_k,Γ_k' and random variables w_k,ζ_k defined below).

It follows from (3.20) and (3.21) that, with probability 1,

$$
\begin{aligned}
\mathbb{E}\{I_{\Gamma_k} \,|\, \mathcal{F}_{k-1}\} &= \mathbb{E}\left\{I_{\Gamma_k}\left(I_{\{\xi_{k-1}\geq -2l\}} + I_{\{\xi_{k-1}<-2l\}}\right) \,|\, \mathcal{F}_{k-1}\right\} \\
&\geq I_{\{\xi_{k-1}\geq -2l\}}\mathbb{P}\{\xi_k - \xi_{k-1} \geq 1 \,|\, \mathcal{F}_{k-1}\} \\
&\quad + I_{\{\xi_{k-1}<-2l\}}\mathbb{P}\{\xi_k - \xi_{k-1} \geq -1 \,|\, \mathcal{F}_{k-1}\} \geq \frac{3}{4}.
\end{aligned}
$$

It is easy to see that the conditions of Lemma 2.5 are satisfied with the following choice of the probability spaces and the function F: the space $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ is the tensor product of $k-1$ copies of $(E, \mathcal{B}(E), \ell)$, $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ coincides with $(E, \mathcal{B}(E), \ell)$, and $F : E \to \mathbb{R}$ is the orthogonal projection to the vector space of e_1; see Hypothesis (H_4). Hence, there is a subset $\Gamma_k' \subset \Gamma_k$ such that $\mathbb{E}\{I_{\Gamma_k'} \,|\, \mathcal{F}_{k-1}\} = \frac{3}{4}$ almost surely. Define a random variable w_k by

$$
w_k = \begin{cases} 1 & \text{for } \omega \in \Gamma_k', \\ -1 & \text{for } \omega \in \Omega \setminus \Gamma_k'. \end{cases}
$$

The construction implies that the conditional law of w_k given $\mathcal{F}_{k-1}$ satisfies the relations

$$
\mathbb{P}\{w_k = 1 \,|\, \mathcal{F}_{k-1}\} = p, \quad \mathbb{P}\{w_k = -1 \,|\, \mathcal{F}_{k-1}\} = 1 - p, \tag{3.22}
$$

where $p = \frac{3}{4}$. In particular, w_k is an $\mathcal{F}_k$-measurable Bernoulli variable with parameter p. Moreover, since the conditional law of w_k given $\mathcal{F}_{k-1}$ is almost surely constant, for $s = -1, 1$ and $\Gamma \in \mathcal{F}_{k-1}$, we can write

$$
\mathbb{P}\big(\{w_k = s\} \cap \Gamma\big) = \mathbb{E}\big(I_\Gamma \, \mathbb{P}\{w_k = s \,|\, \mathcal{F}_{k-1}\}\big) = \mathbb{P}\{w_k = s\}\, \mathbb{P}(\Gamma),
$$

whence we see that w_k is independent of $\mathcal{F}_{k-1}$. Thus, the sequence $\{w_k\}_{k\geq 1}$ satisfies the hypotheses of Section 2.2.

Let us set $\zeta_k = w_1 + \cdots + w_k$. Applying Corollary 2.3, we find a number $c > 0$ and a sequence $\{p_l\}$ converging to 1 as $l \to \infty$ such that

$$
\mathbb{P}\{\zeta_k \geq -l + ck \text{ for all } k \geq 1\} \geq p_l. \tag{3.23}
$$

Now note that, on the event in (3.23), we have $\xi_k \geq \zeta_k \geq -l + ck$, whence we conclude that (3.19) is valid. This completes the proof of Theorem 1.1.

4 Application

In this section, we apply Theorem 1.1 to a parabolic PDE with a degenerate random perturbation. Namely, we consider Eq. (0.1) in which $f : \mathbb{R} \to \mathbb{R}$ is a polynomial of an odd degree $p \geq 3$ with positive leading coefficient:

$$
f(u) = \sum_{n=0}^{p} c_n u^n, \tag{4.1}
$$

where $c_p > 0$, and $c_0, c_1, \ldots, c_{p-1} \in \mathbb{R}$ are arbitrary. In this case, it is easy to see that f satisfies the inequalities

$$
-C \leq f'(u) \leq C(1 + |u|)^{p-1}, \tag{4.2}
$$

$$
f(u)u \geq c\,|u|^{p+1} - C, \tag{4.3}
$$

where $u \in \mathbb{R}$ is arbitrary, and $C, c > 0$ are some constants. We shall confine ourselves to the case $p = 5$ and $d = 3$, although all the results below remain true (with simple adaptations) in the case

$$\begin{cases} p \geq 3 & \text{for } d = 1, 2, \\ 3 \leq p \leq \frac{d+2}{d-2} & \text{for } d = 3, 4. \end{cases} \tag{4.4}$$

We assume that $h \in H^1(\mathbb{T}^3)$ is a fixed function and η is a random process of the form

$$\eta(t, x) = \sum_{k=1}^{\infty} \mathbb{I}_{[k-1,k)}(t)\eta_k(t - k + 1, x), \tag{4.5}$$

where $\mathbb{I}_{[k-1,k)}$ is the indicator function of the interval $[k-1, k)$, and η_k are i.i.d. random variables in $L^2(J, H)$ with $J := [0, 1]$ and $H := H^1(\mathbb{T}^3)$.

Under the above hypotheses, problem (0.1), (0.5) is well posed[6] in H; see Section 5 in [BV92, Chapter I]. Denoting by $S(u_0, \eta)$ the map taking $(u_0, \eta|_J)$ to the solution u at time $t = 1$, we see that the functions $u_k = u(k, \cdot)$ satisfy relation (1.1). Hence, the framework discussed in Section 1 applies to the randomly forced PDE (0.1). To be able to apply Theorem 1.1 to our setting, we shall have to restrict the RDS (1.1) to an invariant set $X \subset H^2(\mathbb{T}^3)$ considered as a subset in the phase space H, while the noise space E will be a closed subspace in $L^2(J, H)$.

Let us recall some definitions that were used in [KNS20] in the context of the Navier–Stokes system and complex Ginzburg–Landau equations. Given a finite-dimensional subspace $\mathcal{H} \subset H^2 := H^2(\mathbb{T}^3)$, we define by recurrence a non-decreasing sequence subspaces $\mathcal{H}_k \subset H^2$ as follows:

$$\mathcal{H}_0 := \mathcal{H}, \quad \mathcal{H}_{k+1} := \operatorname{span}\{\eta, \zeta\xi : \eta, \zeta \in \mathcal{H}_k, \xi \in \mathcal{H}\}, \quad k \geq 0. \tag{4.6}$$

Definition 4.1. A subspace $\mathcal{H} \subset H^2$ is said to be *saturating* if the union of $\{\mathcal{H}_k\}_{k\geq 0}$ is dense in $L^2(\mathbb{T}^3)$.

Examples of saturating spaces are provided by Proposition 5.2. Note that the saturation property does not depend on the number $\nu > 0$ or on the polynomial f. Let us denote by $(\cdot, \cdot)$ the scalar product in $L^2(\mathbb{T}^3)$ and fix an interval $J_T = [0, T]$.

Definition 4.2. A function $\zeta \in L^2(J_T, \mathcal{H})$ is said to be *observable* if for any Lipschitz-continuous functions $a_i : J_T \to \mathbb{R}$, $i \in \mathcal{I}$ and any continuous function $b : J_T \to \mathbb{R}$ the equality [7]

$$\sum_{i\in\mathcal{I}} a_i(t)(\zeta(t), \varphi_i) - b(t) = 0 \quad \text{in } L^2(J_T)$$

implies that a_i, $i \in \mathcal{I}$ and b vanish identically. A probability measure ℓ on $L^2(J_T, \mathcal{H})$ is said to be *observable* if ℓ-almost every trajectory in $L^2(J_T, \mathcal{H})$ is observable.

We now formulate the hypotheses imposed on the random process (4.5). We assume that it takes values in a finite-dimensional saturating subspace $\mathcal{H} \subset H^2$. Let us fix an orthonormal basis $\{\varphi_i\}_{i\in\mathcal{I}}$ in $\mathcal{H}$, and denote by E_i the space of square-integrable functions on J with range in $\operatorname{span}(\varphi_i)$, so that $E := L^2(J, \mathcal{H})$ is representable as the orthogonal sum of $\{E_i\}_{i\in\mathcal{I}}$. We assume that $\ell = \mathcal{D}(\eta_k)$ has a compact support $\mathcal{K} \subset E$ containing the origin and satisfies the two hypotheses below.

[6] The problem is well posed also in $L^2(\mathbb{T}^3)$. However, we need a higher Sobolev space to ensure the required regularity properties of the resolving operator; see Hypothesis (H_1).

[7] It is easy to see that the observability of a function does not depend on the particular choice of the basis $\{\varphi_i\}$ in $\mathcal{H}$; see Remark 1.4 in [KNS20].

Decomposability. *The measure ℓ is representable as the tensor product of its projections ℓ_i to E_i. Moreover, the measures ℓ_i are decomposable in the following sense: there is an orthonormal basis in E_i such that the measure ℓ_i is representable as the tensor product of its projections to the one-dimensional subspaces spanned by the basis vectors. Finally, for any $i \in \mathcal{I}$ the corresponding one-dimensional projections of ℓ_i possess Lipschitz-continuous densities with respect to the Lebesgue measure.*

Observability. *There is $T \in (0,1)$ such that the projection ℓ' of the measure ℓ to the interval[8] $J_T = [0,T]$ is observable.*

We refer the reader to Section 5 in [KNS20] for a discussion of decomposability and observability properties and examples. In particular, it is shown there that both properties are satisfied for the Haar coloured noise given by (0.2), (0.3).

Let $(u_k, \mathbb{P}_u)$ be the Markov process in $H = H^1(\mathbb{T}^3)$ obtained by restricting the solutions of Eq. (0.1) to integer times, and let $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ be the associated Markov semigroups. The following theorem is the main result of this section.

Theorem 4.3. *In addition to the above assumptions, suppose that the saturating subspace $\mathcal{H}$ contains the function identically equal to 1, and the dynamics of Eq. (0.1) satisfies Hypotheses (S) and (C) of the Introduction. Then, for any $\nu > 0$, the process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu_\nu \in \mathcal{P}(H)$, and there is a sequence of positive numbers $\{\gamma_k\}$ going to zero as $k \to \infty$ such that*

$$\|\mathfrak{P}_k^*\lambda - \mu\|_L^* \leq \gamma_k \quad \textit{for all } k \geq 1 \textit{ and } \lambda \in \mathcal{P}(H), \tag{4.7}$$

where $\|\cdot\|_L^*$ denotes the dual-Lipschitz norm over the space H.

Remark 4.4. The convergence to the stationary distribution μ remains valid for measures λ on $L^2(\mathbb{T}^3)$. Indeed, as was mentioned above, Eq. (0.1) is well posed in $L^2(\mathbb{T}^3)$, and in view of the parabolic regularisation property (e.g., see Section 5 in [BV92, Chapter 1]) and the strong nonlinear dissipation (see (4.10)), the space $H = H^1(\mathbb{T}^3)$ is absorbing in the sense that

$$\mathbb{P}_u\{u_k \in H \text{ for } k \geq 1\} = 1 \quad \text{for any } u \in L^2(\mathbb{T}^3). \tag{4.8}$$

It follows that $\mathfrak{P}_1^*\lambda$ is a probability measure on H. Hence, applying inequality (4.7) to the measure $\mathfrak{P}_1^*\lambda$, we see that it is valid for $\lambda \in \mathcal{P}(L^2(\mathbb{T}^3))$, with γ_k replaced by γ_{k-1}.

Before proving Theorem 4.3, let us consider a concrete example of a stochastic force for which the conclusion holds. To this end, we shall use some results described in the Appendix (see Sections 5.2–5.4).

Example 4.5. Let us denote by $\mathcal{I} \subset \mathbb{Z}^3$ the symmetric set defined in Proposition 5.2 and by $\mathcal{H}$ the corresponding 7-dimensional subspace of trigonometric functions. We consider the process

$$\eta^a(t,x) = a \sum_{l \in \mathcal{I}} b_l \eta^l(t) e_l(x),$$

where $a > 0$ is a (large) parameter, $b_l \in \mathbb{R}$ are non-zero numbers, $\{e_l\}_{l \in \mathcal{I}}$ is the basis of $\mathcal{H} = \mathcal{H}(\mathcal{I})$ defined in Section 5.2, and $\{\eta^l\}_{l \in \mathcal{I}}$ are independent Haar processes,

[8] In other words, ℓ' is the image of ℓ under the map restricting a function in $L^2(J, \mathcal{H})$ to the interval J_T.

see (0.3). Let us fix any $\nu > 0$ and use Proposition 5.3 to find a subset $\mathcal{G}_\nu \subset H^1(\mathbb{T}^3)$ of Baire's second category such that Hypothesis (S) is satisfied for any $h \in \mathcal{G}_\nu$. We fix any $h \in H^1(\mathbb{T}^3)$ with that property and denote by $w_1, \dots, w_N$ the corresponding set of solutions for (0.6). As was explained in the Introduction, one of these solutions is locally asymptotically stable under the dynamics of the unperturbed equation (0.7), and there is no loss of generality in assuming that w_N possesses that property. Let $\delta > 0$ be a number such that the solutions of (0.7) issued from the δ-neighbourhood of w_N satisfy (0.8). In view of Theorem 5.5, for any $i \in [\![1, N-1]\!]$, there is a smooth $\mathcal{H}$-valued function ζ_i such that

$$\|u(1; w_i, \zeta_i) - w_N\|_{H^1} < \delta, \tag{4.9}$$

where $u(t; v, \eta)$ stands for the solution of (0.1) corresponding to the initial state $v \in H^1(\mathbb{T}^3)$ and the external force η. Let $\mathcal{K}^a \subset L^2(J, \mathcal{H})$ be the support of the law ℓ^a for the restriction of η^a to the interval $J = [0, 1]$. Since the Haar functions $\{h_0, h_{jl}\}$ entering (0.3) form a basis in $L^2(J)$, and the density ρ of the random variables ξ_k, ξ_{jl} is positive at zero, choosing $a > 0$ sufficiently large, we can approximate the functions ζ_i, within any accuracy in $L^2(J, \mathcal{H})$, by elements of $\mathcal{K}^a$. It follows that inequalities (4.9) remain valid for some suitable functions $\zeta_i \in \mathcal{K}^a$, provided that $a \gg 1$. Thus, Hypothesis (C) is also fulfilled. Finally, as is explained in Section 5 of [KNS20], the measure ℓ^a possesses the decomposability and observability properties. Hence, for any $\nu > 0$ and $h \in \mathcal{G}_\nu$, we can find $a_0(\nu, h) > 0$ such that the conclusion of Theorem 4.3 is valid for $a \geq a_0(\nu, h)$.

Proof of Theorem 4.3. Recall that $S : H \times E \to H$, $u_0 \mapsto u(1)$ stands for the time-1 resolving operator for problem (0.1), (0.5). Due to the superlinear growth of f and parabolic regularisation property, there is a number $K > 0$ such that

$$\|S(u, \eta)\|_2 \leq K \quad \text{for any } u \in H, \eta \in \mathcal{K}; \tag{4.10}$$

see [JNPS15, Lemma 2.10]. The theorem will be established if we check Hypotheses (H_1)–(H_4) of Theorem 1.1 for $H = H^1(\mathbb{T}^3)$, $V = H^2(\mathbb{T}^3)$, $E = L^2(J, \mathcal{H})$, and $X = B_{H^2}(K)$. By construction, X is compact in H, and the inclusion $S(X \times \mathcal{K}) \subset X$ follows from (4.10). Hypothesis (H_1) on the regularity of S is well known to hold for Eq. (0.1) (e.g., see Section 7.4 in [KNS20] and [Kuk82]), and Hypothesis (H_4) is satisfied in view of the decomposability assumption. The remaining hypotheses are checked in the following two steps.

Step 1. Checking Hypothesis (H_2). By Hypothesis (S), Eq. (0.6) has finitely many stationary states $w_1, \dots, w_N$. As in the Introduction, w_N is locally asymptotically stable and $\delta > 0$ is its stability radius in $L^2(\mathbb{T}^3)$. We claim that Hypothesis (H_2) is valid with $\hat{u} = w_N$. To see this, we first establish (1.2) for $u \in W := \{w_1, \dots, w_{N-1}\}$ and an arbitrary $\varepsilon > 0$. Let us fix any $i \in [\![1, N-1]\!]$ and use Hypothesis (C) to find an integer $n_i \geq 1$ and vectors $\zeta_{i1}, \dots, \zeta_{in_i} \in \mathcal{K}$ such that (0.9) holds. Since the solutions of (0.7) that are issued from the δ-neighbourhood of w_N converge uniformly to w_N, we can find an integer $m \gg 1$ such that inequality (1.2) (in which $\|\cdot\|$ is the H^1-norm) holds for $u = w_i$ and $\hat{u} = w_N$, provided that $\zeta_j = \zeta_{ij}$ for $1 \leq j \leq n_i$ and $\zeta_j = 0$ for $n_i + 1 \leq j \leq m$.

To check (H_2) for arbitrary initial condition $u \in X$, we use the existence of a global Lyapunov function for the unperturbed equation (0.7). Namely, let us set

$$\Phi(u) = \int_{\mathbb{T}^3} \Big(\frac{\nu}{2}|\nabla u|^2 + F(u) - hu\Big)\mathrm{d}x, \tag{4.11}$$

where $F(u) = \int_0^u f(s)\mathrm{d}s$. Then, for any solution $u(t)$ of Eq. (0.7), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(u(t)) = \int_{\mathbb{T}^3} \partial_t u\,(\nu\Delta u - f(u) + h)\,\mathrm{d}x = -\int_{\mathbb{T}^3} (\partial_t u)^2\,\mathrm{d}x \le 0.$$

Thus, the function $t \mapsto \Phi(u(t))$ is non-increasing, and it is constant on a non-degenerate interval if and only if $u \equiv w_i$ for some $1 \le i \le N$. Thus, Φ is a global Lyapunov function for (0.7).

We now use a standard approach to prove that the ω-limit set of any solution $u(t)$ of Eq. (0.7) coincides with one of the stationary states (e.g., see Section 2 in [BV92, Chapter 3]). A simple compactness argument will then show that, for any $r > 0$ and a sufficiently large integer $T > 0$, the union U_r of the balls $B_{L^2}(w_i, r)$, $1 \le i \le N$, contains the function $u(T)$, where $u(t)$ is the trajectory issued from any initial point $u_0 \in X \setminus U_r$. Since $0 \in \mathcal{K}$, this will imply the validity of Hypothesis (H_2).

To prove the required property, we first note that, for any $u_0 \in X$, the trajectory $\{u(t), t \ge 0\}$ is contained in the compact set X, so that the corresponding ω-limit set $\omega(u_0)$ is a non-empty compact subset in H. Hence, for any $w \in \omega(u_0)$ we can find a sequence $t_n \to \infty$ such that $u(t_n) \to w$ in H as $n \to \infty$. By the continuity of $\Phi : H \to \mathbb{R}$ and the monotonicity of $\Phi(u(t))$, we have

$$\Phi(w) = \lim_{n\to\infty} \Phi(u(t_n)) = \inf_{t\ge 0} \Phi(u(t)).$$

On the other hand, the continuity of $S(\cdot, 0) : H \to H$ implies that

$$\Phi(S(w,0)) = \lim_{n\to\infty} \Phi(S(u(t_n),0)) = \lim_{n\to\infty} \Phi(u(t_n+1)) = \inf_{t\ge 0} \Phi(u(t)).$$

This shows that $\Phi(w) = \Phi(S(w,0))$, so that w is a stationary solution for (0.7). Since $\omega(u_0)$ is a connected subset, it must coincide with one of the stationary solutions.

Step 2. Checking Hypothesis (H_3). The verification of this hypothesis is similar to the cases of the Navier–Stokes system and complex Ginzburg–Landau equations considered in [KNS20, Section 4]. Let us recall that the nonlinear term $f : \mathbb{R} \to \mathbb{R}$ has the form (4.1), in which $p = 5$, $c_5 > 0$, and $c_0, \dots, c_4 \in \mathbb{R}$. It defines a smooth mapping in H^2, whose derivative is a multiplication operator given by

$$f'(u; v) = f'(u)v = \left(\sum_{n=1}^{5} n c_n u^{n-1}\right) v.$$

We need to show that, for any $u \in X$ and ℓ-a.e. $\eta \in E$, the image of the derivative $(D_\eta S)(u,\eta) : E \to H$ is a dense subspace. Let us fix $u \in X$ and $\eta \in E$, denote by $\tilde{u} \in L^2(J, H^3) \cap W^{1,2}(J, H^1)$ the solution of (0.1), (0.5), and consider the linearised problem

$$\dot{v} - \nu\Delta v + f'(\tilde{u}(t))v = g, \quad v(s) = v_0, \tag{4.12}$$

where $v_0 \in H$ and $g \in L^2(J, \mathcal{H})$. We denote by $v(t; u_0, g)$ the solution of problem (4.12) with $s = 0$. As is explained in Step 2 of the proof of Theorem 4.7 in [KNS20], it suffices to prove that the vector space $\{v(T; 0, g) : g \in L^2(J_T, \mathcal{H})\}$ is dense in $L^2(\mathbb{T}^3)$ for some $T \in (0, 1)$.

Let $R^{\tilde{u}}(t,s) : H \to H$ with $0 \le s \le t \le 1$ be the resolving operator for problem (4.12) with $g \equiv 0$ and let $T \in (0,1)$ be the number entering the observability hypothesis. We define the Gramian $G^{\tilde{u}} : H \to H$ by

$$G^{\tilde{u}} := \int_0^T R^{\tilde{u}}(T,t)\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* \mathrm{d}t, \tag{4.13}$$

where $R^{\tilde{u}}(T,t)^* : H \to H$ is the adjoint of $R^{\tilde{u}}(T,t)$, and $\mathsf{P}_{\mathcal{H}} : H \to H$ is the projection to $\mathcal{H}$. Together with Eq. (4.12), let us consider its dual problem, which is a backward parabolic equation:

$$\dot{w} + \nu\Delta w - f'(\tilde{u}(t))w = 0, \quad w(T) = w_0. \tag{4.14}$$

This problem has a unique solution $w \in L^1(J, H^1) \cap W^{1,2}(J, H^{-1})$ given by

$$w(t) = R^{\tilde{u}}(T,t)^* w_0. \tag{4.15}$$

In view of Theorem 2.5 in [Zab08, Part IV], the required density property in $L^2(\mathbb{T}^3)$ is valid if and only if

$$\mathrm{Ker}(G^{\tilde{u}}) = \{0\}. \tag{4.16}$$

We claim that this equality holds for any $u \in X$ and ℓ-a.e. $\eta \in E$. To prove this, we shall show that all the elements of $\mathrm{Ker}(G^{\tilde{u}})$ are orthogonal to $\mathcal{H}_k$ for any $k \geq 0$. Since $\cup_{k\geq 0}\mathcal{H}_k$ is dense in $L^2(\mathbb{T}^3)$, this will imply (4.16).

We argue by induction on $k \geq 0$. Let us take any $w_0 \in \mathrm{Ker}(G^{\tilde{u}})$. By (4.13),

$$(G^{\tilde{u}} w_0, w_0) = \int_0^T \|\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* w_0\|^2 \mathrm{d}t = 0.$$

This implies that $\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* w_0 \equiv 0$, and hence, for any $\zeta \in \mathcal{H}_0$, we have

$$(\zeta, R^{\tilde{u}}(T,t)^* w_0) = 0 \quad \text{for } t \in J_T. \tag{4.17}$$

Taking $t = T$, we see that w_0 is orthogonal to $\mathcal{H}_0$. Assuming that the function w_0 is orthogonal to $\mathcal{H}_k$, let us prove its orthogonality to $\mathcal{H}_{k+1}$. We differentiate (4.17) in time and use (4.14) and (4.15) to derive

$$\big(-\nu\Delta\zeta + f'(\tilde{u}(t))\zeta, w(t)\big) = 0 \quad \text{for } t \in J_T.$$

Differentiating this equality in time and using (4.14), we obtain

$$\big(-\nu\Delta\zeta + f'(\tilde{u})\zeta, \dot{w}\big) - \big(f^{(2)}(\tilde{u}; \zeta, -\nu\Delta\tilde{u} + f(\tilde{u}) - h), w\big) + \sum_{i\in\mathcal{I}} \big(f^{(2)}(\tilde{u}; \zeta, \varphi_i), w\big)\eta^i(t) = 0,$$

where $\eta^i(t) = (\eta(t), \varphi_i)$ and $f^{(k)}(u; \cdot)$ is the k^{th} derivative of $f(u)$ (so that $f^{(k)} = 0$ for $k \geq 6$). Setting

$$a_i(t) = \big(f^{(2)}(\tilde{u}; \zeta, \varphi_i), w\big),$$
$$b(t) = \big(-\nu\Delta\zeta + f'(\tilde{u})\zeta, \dot{w}\big) - \big(f^{(2)}(\tilde{u}; \zeta, -\nu\Delta\tilde{u} + f(\tilde{u}) - h), w(t)\big),$$

we get the equality

$$b(t) + \sum_{i\in\mathcal{I}} a_i(t)\eta^i(t) = 0 \quad \text{for } t \in J_T,$$

where a_i are Lipschitz-continuous functions and b is continuous. The observability of ℓ implies that

$$\big(f^{(2)}(\tilde{u}(t); \zeta, \varphi_i), w(t)\big) = 0 \quad \text{for } i \in \mathcal{I},\ t \in J_T.$$

Applying exactly the same argument three more times, we derive

$$\big(f^{(5)}(\zeta, \varphi_i, \varphi_j, \varphi_m, \varphi_n), w(t)\big) = 0 \quad \text{for } i, j, m, n \in \mathcal{I},\ t \in J_T.$$

Taking $t = T$, we see that $w(T) = w_0$ is orthogonal to the space $\mathcal{V}$ spanned by $\{(f^{(5)}(\zeta, \varphi_i, \varphi_j, \varphi_m, \varphi_n)\}$. As the space $\mathcal{H}$ contains the function identically equal to 1, we can take $\varphi_j = \varphi_m = \varphi_n = 1$, in which case

$$f^{(5)}(\zeta, \varphi, 1, 1, 1) = 120\, c_5 \zeta\varphi.$$

The latter implies that $\mathcal{V}$ contains all the products $\zeta\xi$ with $\zeta \in \mathcal{H}_k$ and $\xi \in \mathcal{H}$. Combining this with the induction hypothesis, we conclude that w_0 is orthogonal to $\mathcal{H}_{k+1}$. This completes the proof of Theorem 4.3. $\square$

5 Appendix

5.1 Sufficient conditions for mixing

Let X be a compact metric space and let $(u_k, \mathbb{P}_u)$ be a discrete-time Markov process in X possessing the Feller property. We denote by $P_k(u, \Gamma)$ the corresponding transition function, and by $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ the Markov operators. The following theorem is a straightforward consequence of Theorem 3.1.3 in [KS12].

Theorem 5.1. *Suppose that the following two conditions are satisfied for some point $\hat{u} \in X$.*

Recurrence. *For any $r > 0$, there is an integer $m \geq 1$ and a number $p > 0$ such that*

$$P_m\big(u, B_X(\hat{u}, r)\big) \geq p \quad \text{for any } u \in X. \tag{5.1}$$

Stability. *There is a positive function $\delta(\varepsilon)$ going to zero as $\varepsilon \to 0^+$ such that*

$$\sup_{k \geq 0} \|P_k(u, \cdot) - P_k(u', \cdot)\|_L^* \leq \delta(\varepsilon) \quad \text{for any } u, u' \in B_X(\hat{u}, \varepsilon). \tag{5.2}$$

Then the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu \in \mathcal{P}(X)$, and convergence (1.4) holds.

To establish this theorem, it suffices to take two independent copies of the Markov process $(u_k, \mathbb{P}_u)$ and use standard techniques (based on the Borel–Cantelli lemma) to show that the first hitting time of any ball around $(\hat{u}, \hat{u})$ is almost surely finite and has a finite exponential moment; combining this with the stability property, we obtain the required result. Since the corresponding argument is well known (e.g., see Section 3.3 in [KS12]), we do note give more details.

5.2 Saturating subspaces

As in Section 4, we consider only the case $d = 3$; the other dimensions can be treated by similar arguments. For any non-zero vector $l = (l_1, l_2, l_3) \in \mathbb{Z}^3$, we set

$$c_l(x) = \cos\langle l, x\rangle, \quad s_l(x) = \sin\langle l, x\rangle, \quad x \in \mathbb{T}^3,$$

where $\langle l, x\rangle = l_1x_1 + l_2x_2 + l_3x_3$. Let us define an orthogonal basis $\{e_l\}$ in $L^2(\mathbb{T}^3)$ by the relation

$$e_l(x) = \begin{cases} c_l(x) & \text{if } l_1 > 0 \text{ or } l_1 = 0,\ l_2 > 0 \text{ or } l_1 = l_2 = 0,\ l_3 \geq 0, \\ s_l(x) & \text{if } l_1 < 0 \text{ or } l_1 = 0,\ l_2 < 0 \text{ or } l_1 = l_2 = 0,\ l_3 < 0. \end{cases}$$

Let $\mathcal{I} \subset \mathbb{Z}^3$ be a finite symmetric set (i.e., $-\mathcal{I} = \mathcal{I}$) containing the origin. We define

$$\mathcal{H}(\mathcal{I}) := \operatorname{span}\{e_l : l \in \mathcal{I}\} \tag{5.3}$$

and denote by $\mathcal{H}_k(\mathcal{I})$ the sets $\mathcal{H}_k$ given by (4.6) with $\mathcal{H} = \mathcal{H}(\mathcal{I})$. Recall that $\mathcal{I}$ is called a *generator* if all the vectors in $\mathbb{Z}^3$ are finite linear combinations of elements of $\mathcal{I}$ with integer coefficients.

Proposition 5.2. *The subspace $\mathcal{H}(\mathcal{I})$ is saturating if and only if $\mathcal{I}$ is a generator. In particular, the set $\mathcal{I} = \{(0,0,0), (\pm 1,0,0), (0,\pm 1,0), (0,0,\pm 1)\}$ gives rise to the 7-dimensional saturating subspace $\mathcal{H}(\mathcal{I})$.*

Proof. To prove the sufficiency of the condition, we note that

$$c_l(x)c_r(x) = \frac{1}{2}\big(c_{l-r}(x) + c_{l+r}(x)\big), \quad s_l(x)s_r(x) = \frac{1}{2}\big(c_{l-r}(x) - c_{l+r}(x)\big). \tag{5.4}$$

If $c_r, s_r \in \mathcal{H}(\mathcal{I})$ and $c_l, s_l \in \mathcal{H}_k(\mathcal{I})$, then (5.4) implies that $c_{l+r}, c_{l-r} \in \mathcal{H}_{k+1}(\mathcal{I})$. A similar argument shows that $s_{l+r}, s_{l-r} \in \mathcal{H}_{k+1}(\mathcal{I})$. Since $\mathcal{I}$ is a generator, we see that all the vectors of the basis $\{e_l\}$ can be obtained from the elements of $\mathcal{H}(\mathcal{I})$ after finitely many iterations.

To prove the necessity, assume that $\mathcal{I}$ is not a generator. Then there is a vector $m \in \mathbb{Z}^3$ that is not a finite linear combination of elements of $\mathcal{I}$ with integer coefficients. It is easy to see that the functions c_m and s_m are orthogonal to $\cup_{k\geq 0}\mathcal{H}_k(\mathcal{I})$. This shows that $\mathcal{H}(\mathcal{I})$ is not saturating and completes the proof of the proposition. ☐

5.3 Genericity of Hypothesis (S)

Proposition 5.3. *Let $\nu > 0$ be any number and let f be a real polynomial satisfying conditions (4.2)–(4.4) with $d = 3$. Then there is a subset $\mathcal{G}_\nu \subset H^1(\mathbb{T}^3)$ of Baire's second category such that, for any $h \in \mathcal{G}_\nu$, the nonlinear equation*

$$-\nu\Delta w + f(w) = h(x), \quad x \in \mathbb{T}^3 \tag{5.5}$$

has finitely many solutions.

Before proceeding with the proof, let us recall the formulation of an infinite-dimensional version of Sard's theorem and some related definitions (see [Sma65]). Let X and Y be Banach spaces. A linear operator $L : X \to Y$ is said to be *Fredholm* if its image is closed, and the dimension of its kernel and the co-dimension of its image are finite. The *index* of L is defined by

$$\operatorname{Ind} L := \dim(\operatorname{Ker} L) - \operatorname{codim}(\operatorname{Im} L).$$

It is well known that if $L : X \to Y$ is a Fredholm operator and $K : X \to Y$ is a compact linear operator, then $L + K$ is also Fredholm, and $\operatorname{Ind} L = \operatorname{Ind}(L + K)$. A C^1-smooth map $F : X \to Y$ is said to be *Fredholm* if for any $w \in X$ the derivative $DF(w) : X \to Y$ is a Fredholm operator. The *index* of F is the index of the operator $DF(w)$ at some $w \in X$ (it is independent of the choice of w). A point $y \in Y$ is called a *regular value* for F if $F^{-1}(y) = \varnothing$ or $DF(w) : X \to Y$ is surjective for any $w \in F^{-1}(y)$. The following result is due to Smale [Sma65, Corollary 1.5].

Theorem 5.4. *Let $F : X \to Y$ be a C^k-smooth Fredholm map such that $k > \max\{\operatorname{Ind} F, 0\}$. Then its set of regular values is of Baire's second category.*

Proof of Proposition 5.3. Let us consider the map

$$F : H^3(\mathbb{T}^3) \to H^1(\mathbb{T}^3), \quad w \mapsto -\nu\Delta w + f(w).$$

We have $\mathrm{Ind}(-\nu\Delta) = 0$, so $\mathrm{Ind}(-\nu\Delta + Df(w)) = 0$ for any $w \in H^3(\mathbb{T}^3)$, since the derivative $Df(w) : H^3(\mathbb{T}^3) \to H^1(\mathbb{T}^3)$ (acting as the operator of multiplication by $f'(w)$) is compact. Smale's theorem implies the existence of a set $\mathcal{G}_\nu \subset H^1(\mathbb{T}^3)$ of Baire's second category such that $DF(w) : H^3(\mathbb{T}^3) \to H^1(\mathbb{T}^3)$ is surjective for any solution w of Eq. (5.5) with $h \in \mathcal{G}_\nu$. Since the index is zero, it follows that the derivative $DF(w)$ is an isomorphism between the spaces $H^3(\mathbb{T}^3)$ and $H^1(\mathbb{T}^3)$ for any solution $w \in H^3(\mathbb{T}^3)$ of (5.5). Applying the inverse function theorem, we conclude that the solutions are isolated points in $H^3(\mathbb{T}^3)$. On the other hand, the elliptic regularity implies that the family of all solutions for Eq. (5.5) is a compact set in $H^3(\mathbb{T}^3)$, so there can be only finitely many of them. ☐

5.4 Approximate controllability of parabolic PDEs

In this section, we discuss briefly the approximate controllability for Eq. (0.1) established in [Ner20]. This type of results were obtained by Agrachev and Sarychev [AS05, AS06] for the 2D Navier–Stokes and Euler equations on the torus and later extended to the 3D case in [Shi06, Ner10]. We assume that the nonlinearity $f : \mathbb{R} \to \mathbb{R}$ is a polynomial satisfying the hypotheses of Section 4. The space $\mathcal{H}(\mathcal{I})$ is defined by (5.3) for some finite symmetric set $\mathcal{I} \subset \mathbb{Z}^d$ containing the origin, with an obvious modification of the functions e_l for $d = 1, 2, 4$.

Theorem 5.5. *In addition to the above hypotheses, assume that $\mathcal{I}$ is a generator for $\mathbb{Z}^d$ and $h \in H^1(\mathbb{T}^d)$ is a given function. Then Eq. (0.1) is approximately controllable in $H^1(\mathbb{T}^d)$, i.e., for any $\nu > 0$, $\varepsilon > 0$, and $u_0 \in L^2(\mathbb{T}^d)$, $u_1 \in H^1(\mathbb{T}^d)$, there is a function $\zeta \in L^2([0,1], \mathcal{H}(\mathcal{I}))$ such that the solution of Eq. (0.1) with initial condition $u(0) = u_0$ satisfies the inequality*

$$\|u(1) - u_1\|_{H^1(\mathbb{T}^d)} < \varepsilon.$$

This result is essentially Theorem 2.5 of [Ner20], dealing with the case when the problem in question is not necessarily well posed and assuming that $u_0, u_1 \in H^2(\mathbb{T}^3)$. Under our hypotheses, Eq. (0.1) is well posed, and using a simple approximation argument, we can prove the validity of Theorem 5.5.

References

[AS05] A. A. Agrachev and A. V. Sarychev, *Navier–Stokes equations: controllability by means of low modes forcing*, J. Math. Fluid Mech. **7** (2005), no. 1, 108–152.

[AS06] ———, *Controllability of 2D Euler and Navier–Stokes equations by degenerate forcing*, Comm. Math. Phys. **265** (2006), no. 3, 673–697.

[BKL02] J. Bricmont, A. Kupiainen, and R. Lefevere, *Exponential mixing of the 2D stochastic Navier–Stokes dynamics*, Comm. Math. Phys. **230** (2002), no. 1, 87–132.

[Bri02] J. Bricmont, *Ergodicity and mixing for stochastic partial differential equations*, Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002) (Beijing), Higher Ed. Press, 2002, pp. 567–585.

[BV92] A. V. Babin and M. I. Vishik, *Attractors of Evolution Equations*, North-Holland Publishing, Amsterdam, 1992.

[CI74] N. Chafee and E. F. Infante, *A bifurcation problem for a nonlinear partial differential equation of parabolic type*, Applicable Anal. **4** (1974), 17–37.

[Deb13] A. Debussche, *Ergodicity results for the stochastic Navier–Stokes equations: an introduction*, Topics in mathematical fluid mechanics, Lecture Notes in Math., vol. 2073, Springer, Heidelberg, 2013, pp. 23–108.

[EMS01] W. E, J. C. Mattingly, and Ya. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation*, Comm. Math. Phys. **224** (2001), no. 1, 83–106.

[ES00] W. E and Ya. G. Sinaĭ, *New results in mathematical and statistical hydrodynamics*, Russian Math. Surveys **55** (2000), no. 4(334), 635–666.

[Fel68] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. I.*, John Wiley & Sons, New York, 1968.

[FGRT15] J. Földes, N. Glatt-Holtz, G. Richards, and E. Thomann, *Ergodic and mixing properties of the Boussinesq equations with a degenerate random forcing*, J. Funct. Anal. **269** (2015), no. 8, 2427–2504.

[FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2D Navier–Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), no. 1, 119–141.

[FW84] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York–Berlin, 1984.

[HM06] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), no. 3, 993–1032.

[HM11] ______, *A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs*, Electron. J. Probab. **16** (2011), no. 23, 658–738.

[JNPS15] V. Jakšić, V. Nersesyan, C.-A. Pillet, and A. Shirikyan, *Large deviations and Gallavotti–Cohen principle for dissipative PDE's with rough noise*, Comm. Math. Phys. **336** (2015), no. 1, 131–170.

[KNS20] S. Kuksin, V. Nersesyan, and A. Shirikyan, *Exponential mixing for a class of dissipative PDEs with bounded degenerate noise*, Geom. Funct. Anal. **30** (2020), no. 1, 126–187.

[KS00] S. Kuksin and A. Shirikyan, *Stochastic dissipative PDEs and Gibbs measures*, Comm. Math. Phys. **213** (2000), no. 2, 291–330.

[KS12] ______, *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, Cambridge, 2012.

[Kuk82] S. B. Kuksin, *Diffeomorphisms of function spaces that correspond to quasi-linear parabolic equations*, Mat. Sb. (N.S.) **117(159)** (1982), no. 3, 359–378, 431.

[Lam96] J. W. Lamperti, *Probability*, John Wiley & Sons, New York, 1996.

[Mar17] D. Martirosyan, *Large deviations for stationary measures of stochastic nonlinear wave equations with smooth white noise*, Comm. Pure Appl. Math. **70** (2017), no. 9, 1754–1797.

[Ner10] H. Nersisyan, *Controllability of 3D incompressible Euler equations by a finite-dimensional external force*, ESAIM Control Optim. Calc. Var. **16** (2010), no. 3, 677–694.

[Ner20] V. Nersesyan, *Approximate controllability of nonlinear parabolic PDEs in arbitrary space dimension*, Math. Control Relat. Fields (2020), under revision.

[Shi06] A. Shirikyan, *Approximate controllability of three-dimensional Navier–Stokes equations*, Comm. Math. Phys. **266** (2006), no. 1, 123–151.

[Shi15] ———, *Control and mixing for 2D Navier–Stokes equations with space-time localised noise*, Ann. Sci. Éc. Norm. Supér. (4) **48** (2015), no. 2, 253–280.

[Shi20] ———, *Controllability implies mixing II. Convergence in the dual-Lipschitz metric*, J. Eur. Math. Soc. (2020), accepted for publication.

[Sma65] S. Smale, *An infinite dimensional version of Sard's theorem*, Amer. J. Math. **87** (1965), 861–866.

[Zab08] J. Zabczyk, *Mathematical Control Theory*, Modern Birkhäuser Classics, Birkhäuser, Boston, MA, 2008.