Exponential mixing for a class of dissipative PDEs with bounded degenerate noise

Sergei Kuksin, Vahagn Nersesyan, Armen Shirikyan

Exponential mixing for a class of dissipative PDEs with bounded degenerate noise

Sergei Kuksin* Vahagn Nersesyan† Armen Shirikyan‡

February 9, 2018

Abstract

We study a class of discrete-time random dynamical systems with compact phase space. Assuming that the deterministic counterpart of the system in question possesses a dissipation property, its linearisation is approximately controllable, and the driving noise has a decomposable structure, we prove that the corresponding family of Markov processes has a unique stationary measure, which is exponentially mixing in the dual-Lipschitz metric. The abstract result is applicable to nonlinear dissipative PDEs perturbed by a random force which affects only a few Fourier modes and belongs to a certain class of random processes. We assume that the nonlinear PDE in question is well posed, its nonlinearity is non-degenerate in the sense of the control theory, and the random force is a regular and bounded function of time which satisfies some decomposability and observability hypotheses. This class of forces includes random Haar series, where coefficients for high Haar modes decay sufficiently fast. In particular, the result applies to the 2D Navier–Stokes system and the nonlinear complex Ginzburg–Landau equations. The proof of the abstract theorem uses the coupling method, enhanced by the Newton–Kantorovich–Kolmogorov fast convergence.

AMS subject classifications: 35Q10, 35Q56, 35R60, 37A25, 37L55, 60H15, 76M35

Keywords: Markov process, stationary measure, mixing, Navier–Stokes system, Ginzburg–Landau equation, Newton–Kantorovich–Kolmogorov fast convergence, approximate controllability, Haar series

*CNRS, Institut de Mathémathiques de Jussieu–Paris Rive Gauche, UMR 7586, Université Paris Diderot, Sorbonne Paris Cité, F-75013, Paris, France & School of Mathematics, Shandong University, Shanda Nanlu, 27, 250100, PRC; 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; e-mail: Vahagn.Nersesyan@math.uvsq.fr

‡Department of Mathematics, University of Cergy–Pontoise, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France & Centre de Recherches Mathématiques, CNRS UMI 3457, Université de Montréal, Montréal, QC, H3C 3J7, Canada; e-mail: Armen.Shirikyan@u-cergy.fr

Contents

0 Introduction

The problem of uniqueness of a stationary measure for randomly forced dissipative PDEs attracted a lot of attention in the last twenty years. It is by now well understood that when all determining modes of the unforced PDEs are directly affected by the noise, the problem has a unique stationary distribution, which is exponentially stable as $t \to \infty$ in an appropriate metric. We refer the reader to

the papers [FM95, KS00, EMS01, BKL02] for the first achievements and to the book [KS12] for a detailed description of the results in this setting. The case in which the random perturbation does not act directly on the determining modes of the flow is much less understood (see the literature review below), and it is the subject of the present article.

To describe our results, in this introduction we confine ourselves to the 2D Navier–Stokes system perturbed by a Haar coloured noise. Namely, we consider the equations

$$\partial_t u + \langle u, \nabla\rangle u - \nu\Delta u + \nabla p = \eta(t,x), \quad \operatorname{div} u = 0, \tag{0.1}$$

where $u = (u_1, u_2)$ and p are unknown functions, $\nu > 0$ is a parameter, Δ is the Laplace operator, $\langle u, \nabla\rangle = u_1\partial_1 + u_2\partial_2$, and η is an external (random) force. We assume that all the functions are 2π-periodic with respect to the space variables $x = (x_1, x_2)$ and have zero mean value. Let us introduce the trigonometric basis $\{\varphi_i\}$ in the space of divergence-free vector fields with zero mean value on the 2D torus $\mathbb{T}^2$ and write η in the form

$$\eta(t,x) = \sum_{i=1}^{N} b_i \eta^i(t)\varphi_i(x), \tag{0.2}$$

where b_i are non-zero numbers, and η^i are independent bounded real-valued random processes that are distributed as a fixed process $\tilde\eta(t)$ constructed as follows. Let $\{h_0, h_{jl}\}$ be the Haar system defined by relations (5.5) and (5.6); cf. Section 22 in [Lam96]. We set

$$\tilde\eta(t) = \sum_{k=0}^{\infty} \xi_k h_0(t-k) + \sum_{j=1}^{\infty} c_j \sum_{l=0}^{\infty} \xi_{jl} h_{jl}(t), \tag{0.3}$$

where $\{c_j\}$ is a sequence of non-zero numbers going to zero[1] sufficiently fast, and $\{\xi_k, \xi_{jl}\}$ is a family of independent identically distributed (i.i.d.) real-valued random variables. Processes of the form (0.2) are called *coloured noises* and are widely used in engineering sciences; see [Van06]. We thus consider the dynamics of the Navier–Stokes system subject to the coloured noise (0.2). Because of time correlations of finite depth, the trajectories of (0.1) do not form a Markov process. However, their restrictions to integer times do, and our aim is to describe the large-time behaviour of the corresponding discrete-time processes. The following theorem is one of the main results of this paper.

Main Theorem. *In addition to the above hypotheses, assume that $c_j = Cj^{-q}$ for all $j \geq 1$ and some $C > 0$ and $q > 1$, and the law of the random variables $\{\xi_k, \xi_{jl}\}$ has a Lipschitz-continuous density ρ such that $0 \in \operatorname{supp}\rho \subset [-1, 1]$. Then there is an integer $N_0 \geq 1$ such that, for any $N \geq N_0$ and $\nu > 0$, the*

[1] In the case when $c_j = 2^{j/2}$ for $j \geq 1$ and $\{\xi_k, \xi_{jl}\}$ are independent random variables with centred normal law of unit dispersion, the series (0.3) converges to the white noise; see Theorem 1 in [Lam96, Section 22].

Markov process obtained by restricting the trajectories of (0.1), (0.2) *to integer times has a unique stationary measure* μ_ν. *Moreover, for any solution* $u(t)$, *we have*

$$\mathcal{D}(u(k)) \rightharpoonup \mu_\nu \quad \text{as } k \to \infty, \tag{0.4}$$

where $\mathcal{D}(\zeta)$ *stands for the law of a random variable* ζ, *and the weak convergence in* (0.4) *is understood in the sense of measures on the space of square-integrable vector fields on* $\mathbb{T}^2$. *Finally, convergence* (0.4) *holds exponentially fast in the dual-Lipschitz metric (defined below in* Notation and conventions*).*

To prove this theorem, we shall establish a general result on exponential mixing for discrete-time Markov processes and show that it applies to the problem (0.1), (0.2). Moreover, our general result is applicable to other dissipative PDEs, such as various complex Ginzburg–Landau equations and multidimensional Burgers system in non-potential[2] setting (the latter will not be treated in this paper for reasons of space). As was mentioned above, there are only a few works dealing with highly degenerate noise not acting directly on the determining modes of the unperturbed dynamics. Namely, the existence of densities for finite-dimensional projections of stationary measures for solutions of the Navier–Stokes system was studied in [MP06, AKSS07]. Hairer and Mattingly [HM06, HM11] investigated the Navier–Stokes system perturbed by a finite-dimensional white noise and established the uniqueness of stationary measure and its exponential stability in the dual-Lipschitz metric. Földes at al. [FGRT15] proved a similar result for the Boussinesq system, assuming that the random noise acts only on the equation for temperature. Finally, the case in which the random perturbation is localised in the physical space and time was studied in [Shi15] (see also [Shi18] for the boundary-driven Navier–Stokes system).

Bounded random forces versus white in time forces

In our work we study nonlinear PDEs perturbed by random forces that are, as a function of time, bounded processes of the type of *random Haar series* (see (0.3)), while it is somewhat more traditional in mathematical physics to use the forces that are random processes white in time. What are the advantages and disadvantages of the former class of forces compared to the latter? A first disadvantage is the tradition: one hundred years ago, in Langevin's era, the white in time forces were successfully used to model systems from statistical physics, and since then they were exploited in other problems, usually without serious discussion of their adequacy. Secondly, white-forced equations have useful algebraical features coming from Ito's formula. On the other hand, bounded random forces, exactly due to their boundedness, serve better to build models for some specific physical problems (e.g., they are being used in modern meteorology). Secondly, they have a number of serious analytical advantages. Namely, the corresponding stochastic equations always are well posed if so are the deterministic equations, and—what is more important—for

[2]The one-dimensional Burgers equation and multi-dimensional Burgers system in the potential setting can be treated by softer tools; see [Sin91, Bor16, Shi17].

systems with the random forces which we advocate, the mixing property can be established for significantly broader class of PDEs. Indeed, if the nonlinearity of the unperturbed deterministic equation is Hamiltonian and polynomial, and the random perturbation is white in time, then the existing techniques apply to establish the mixing only if the nonlinearity is at most cubic.[3] At the same time, the main theorem of our work can be used to prove the mixing property for equations with nonlinearity of any degree, see in Section 4 below. For the 2D stochastic Navier–Stokes equation (0.1) (where the nonlinearity is quadratic) with degenerate white in time force η, the exponential mixing is proved in the papers [HM06, HM11] based on an infinite-dimensional version of the Malliavin calculus developed in [MP06]. If the random force η is bounded and degenerate, then the proof of the exponential mixing, presented in our work, is significantly shorter and, we believe, conceptually clearer. A subclass of the random forces which we consider—the random Haar series (0.3)—has a number of similarities with the white forces (and the latter may be obtained as a limiting case of the former, see footnote 1). In particular, the forces (0.3) have independent components with arbitrarily short time scales, which simplifies the verification for them of various non-degeneracy properties (e.g., see Section 5.2, where we show that these forces are Lipschitz-observable).

The paper is organised as follows. In Section 1, we formulate our main result on the uniqueness and exponential mixing of a stationary measure for discrete-time Markov processes possessing some controllability properties and describe briefly its applications. To simplify the reading of the rest of the paper, we also describe the general philosophy of the proof of the main result and discuss some analogies between the coupling schemes for PDEs with regular noise that are used in our proof and the Newton–Kantorovich–Kolomogorov fast convergence. Section 2 is devoted to establishing some auxiliary tools used in the proof of the main theorem. The latter is given in Section 3. In Section 4, we apply our result to the 2D Navier–Stokes system and the complex Ginzburg–Landau equation perturbed by a random noise. Sections 5 and 6 describe some classes of random noises that are allowed in our approach. Finally, the Appendix gathers some auxiliary results used in the main text.

Acknowledgements

We thank V. I. Bogachev for drawing our attention to the papers [Mal16, AY17], and the *Institut Henri Poincaré* in Paris for hosting our working group *Control Theory and Stochastic Analysis*, at which this article was initiated. This research was supported by the *Agence Nationale de la Recherche* through the grants ANR-10-BLAN 0102 and ANR-17-CE40-0006-02. 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*.

[3] See the paper [KN13], where the mixing is proved for the white-forced cubic CGL equations, and the difficulty coming from nonlinearities of higher degree is explained.

Notation and conventions

Let X be a Polish (i.e., complete separable metric) space with a distance $d_X(u,v)$ and the Borel σ-algebra $\mathcal{B}(X)$. We denote by $B_X(a,R)$ the closed ball of radius $R>0$ centred at $a\in X$ and write $\dot B_X(a,R)$ for the corresponding open ball. If X is a Banach space and $a=0$, we write $B_X(R)$ instead of $B_X(0,R)$. We use the following notation in which all abstract Banach and metric spaces are assumed to be separable:

$\mathcal{L}(E,F)$ is the space of bounded linear operators between two Banach spaces E and F. It is endowed with the operator norm.

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

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

with an obvious modification for $p=\infty$.

$C_b(X)$ is the space of bounded continuous functions $f:X\to\mathbb{R}$ endowed with the norm $\|f\|_\infty=\sup_X|f|$.

$L_b(X)$ is the space of functions $f\in C_b(X)$ such that

$$\|f\|_L:=\|f\|_\infty+\sup_{0<d_X(u,v)\le1}\frac{|f(u)-f(v)|}{d_X(u,v)}<\infty.$$

$\mathcal{P}(X)$ denotes the set of Borel probability measures on X. For any $\mu\in\mathcal{P}(X)$ and μ-integrable function $f:X\to\mathbb{R}$, we set

$$\langle f,\mu\rangle=\int_X f(u)\,\mu(\mathrm{d}u).$$

The *total variation metric* on $\mathcal{P}(X)$ is defined by

$$\|\mu_1-\mu_2\|_{\mathrm{var}}:=\sup_{\Gamma\in\mathcal{B}(X)}|\mu_1(\Gamma)-\mu_2(\Gamma)|=\frac12\sup_{\|f\|_\infty\le1}|\langle f,\mu_1\rangle-\langle f,\mu_2\rangle|\,. \quad (0.5)$$

We shall also use the *dual-Lipschitz* metric

$$\|\mu_1-\mu_2\|_L^*:=\sup_{\|f\|_L\le1}|\langle f,\mu_1\rangle-\langle f,\mu_2\rangle|\,.$$

Note that $\|\mu_1-\mu_2\|_L^*\le2\|\mu_1-\mu_2\|_{\mathrm{var}}$.

We denote by C, C_1, etc. unessential positive constants.

If B_1 and B_2 are *real* Banach spaces and $\mathcal{O}\subset B_1$ is an open domain, then analyticity of a map $F:\mathcal{O}\to B_2$ is understood in the sense of Fréchet. In addition, for an analytic map F, we always assume that

the norms of all the derivatives D^kF are bounded on bounded subsets of $\mathcal{O}$. (0.6)

Moreover, if F depends on a parameter u varying in a compact metric space X, then we assume that *all the derivatives $D^k_\eta F(u,\eta)$ are bounded on bounded subsets, uniformly in $u\in X$, and are continuous functions of (u,η).*

1 Mixing for Markovian random dynamical systems

1.1 Setting of the problem

Let H and E be separable Hilbert spaces and let $S : H \times E \to H$ be a continuous mapping. We consider the random dynamical system (RDS) given by

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

where $\{\eta_k\}$ is a sequence of i.i.d. random variables in E. Let us denote by ℓ the law of η_k and assume that it has a compact support $\mathcal{K} \subset E$. Suppose there is a compact set $X \subset H$ such that $S(X \times \mathcal{K}) \subset X$, so that one can consider the restriction of the RDS (1.1) to X. The hypotheses imposed on η_k imply that the trajectories of (1.1) form a discrete-time Markov process in X; we shall denote it by $(u_k, \mathbb{P}_u)$, where $\mathbb{P}_u$ is the probability measure corresponding to the trajectories issued from u (e.g., see Section 2.5.B in [KS91] or Section 1.3.1 in [KS12]). We write $P_k(u, \Gamma)$ for its transition function and denote by $\mathfrak{P}_k : C_b(X) \to C_b(X)$ and $\mathfrak{P}_k^* : \mathcal{P}(X) \to \mathcal{P}(X)$ the corresponding Markov operators. Recall that a measure $\mu \in \mathcal{P}(H)$ is said to be *stationary* for $(u_k, \mathbb{P}_u)$ if $\mathfrak{P}_1^* \mu = \mu$. Since X is compact, the Bogolyubov–Krylov argument implies that there is at least one stationary measure. Our goal is to study its uniqueness and long-time stability under the dynamics. In what follows, we assume that the four hypotheses below are satisfied.

($\mathbf{H}_1$) Regularity. *The mapping $S : H \times E \to H$ is twice continuously differentiable, and its second order derivative is bounded on bounded subsets. Moreover, the mapping $\eta \mapsto S(u, \eta)$ is analytic for any $u \in H$, and the derivatives $(D_\eta^j S)(u, \eta)$ are bounded on bounded subsets of $H \times E$; cf. (0.6). Finally, there is a Banach space V compactly embedded into H such that the derivative $(D_u S)(u, \eta)$ acts continuously from H to V, and the corresponding operator norm is bounded on bounded subsets of $H \times E$.*

($\mathbf{H}_2$) Dissipativity. *There is a number $a \in (0, 1)$ and vectors $\hat{\eta} \in \mathcal{K}$ and $\hat{u} \in X$ such that*

$$\|S(u, \hat{\eta}) - \hat{u}\| \leq a\|u - \hat{u}\| \quad \text{for any } u \in X. \tag{1.2}$$

To formulate the third hypothesis, for any point $u \in X$, we denote by $\mathcal{K}^u$ the set of those $\eta \in \mathcal{K}$ for which the image of $(D_\eta S)(u, \eta) : E \to H$ is dense in H. It is easy to see that $\mathcal{K}^u$ is a section of a Borel subset in the product space $X \times E$ and, hence, is a Borel subset of E.[4]

[4] To see this, consider a measurable space $(Y, \mathcal{Y})$ and a measurable mapping $A : Y \to \mathcal{L}(E, H)$ and denote by $\mathcal{G}_A$ the set of points $y \in Y$ for which the image of $A(y)$ is dense in H is measurable. Choosing countable dense subsets $\{f_i\} \subset E$ and $\{h_j\} \subset H$, it is easy to see that $\mathcal{G}_A = \bigl\{y \in Y : \inf_{i \geq 1} \|A(y) f_i - h_j\|_H = 0 \text{ for any } j \geq 1\bigr\}$. In the case under study, we have $Y = H \times E$, $A = (D_\eta S)(u, \eta)$, and $\mathcal{K}^u = \mathcal{K} \cap \mathcal{G}^u$, where $\mathcal{G}^u = \{\eta \in E : (u, \eta) \in \mathcal{G}_A\}$.

($\mathbf{H}_3$) Approximate controllability of linearisation. *For any $u \in X$, we have $\ell(\mathcal{K}^u) = 1$.*

($\mathbf{H}_4$) Decomposability of the noise. *There is an orthonormal basis $\{e_j\}$ in E such that*

$$\eta_k = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j, \tag{1.3}$$

where ξ_{jk} are independent random variables and b_j are positive numbers with the property such that $|\xi_{jk}| \leq 1$ a.s., and

$$\sum_{j=1}^{\infty} b_j^2 < \infty. \tag{1.4}$$

Moreover, there are Lipschitz-continuous functions $\rho_j : \mathbb{R} \to \mathbb{R}$ such that

$$\mathcal{D}(\xi_{jk}) = \rho_j(r)\mathrm{d}r \quad \text{for all } j \geq 1. \tag{1.5}$$

Before turning to the formulation of our main result, let us make some comments about the above hypotheses. The dissipation property ensures that, for a particular choice of the driving force, the dynamics has a drift towards a fixed point $\hat{u} \in X$. In the context of randomly forced PDEs, this condition is certainly satisfied if the unperturbed deterministic equation has a stationary solution which is globally exponentially stable. The approximate controllability implies that, for a typical realisation of the driving force η, the linearised system $D_\eta S$ exhibits the energy transfer to high frequencies, without specifying the rate of the transfer. As we show in the proof, this condition, together with some analyticity argument, implies that, with high probability, the trajectories of (1.1) can be locally stabilised by a finite-dimensional modification of the driving force. Finally, the hypothesis on the structure of the noise implies that the random forcing possesses a weak non-degeneracy property.

Let us also mention that the controllability of the linearised operator (or, equivalently, the existence of its right inverse) is well known to be important when studying mixing properties for random dynamical systems. In particular, it arises in the Malliavin calculus and plays an important role when proving the absolute continuity of the laws of solution of SDE with respect to the Lebesgue measure; see Chapter 2 in [Nua06]. The approximate controllability of the linearised equation was used by Hairer and Mattingly [HM06, HM11] in their proof of exponential mixing of the 2D Navier–Stokes system perturbed by a degenerate noise, white in time and finite-dimensional in x.

A nonlinear PDE of parabolic type, perturbed by a random force, leads to a system (1.1) via the following construction. Let us formally write the PDE in question as a differential equation in a function space H:

$$\dot{u}(t) = V(u) + \zeta(t), \tag{1.6}$$

where $V(u)$ is an unbounded nonlinear vector field in H (which is not necessarily defined everywhere), and $\zeta(t)$ is a bounded random process in a subspace $\mathcal{H} \subset H$.

Assume that a.e. trajectory of ζ is a locally square-integrable function of time such that the random variables $\{\zeta|_{[k-1,k)}\}_{k\geq 1}$ are independent. Next, for $0\leq\tau\leq 1$ and $k\geq 1$, define $\eta_k(\tau)=\zeta(k-1+\tau)$, so that $\eta_k\in E:=L^2(0,1;\mathcal{H})$ almost surely, and denote by $S:H\times E\to H$ the mapping which takes a pair $(v,\zeta|_{[0,1]})$ to $u(1)$, where $u(t)$ is a solution of (1.6) satisfying $u(0)=v$. With this notation, a solution of (1.6) evaluated at integer times $t\geq 0$ satisfies (1.1). Assumption (H_1) means that Eq. (1.6) is well posed in H and possesses a regularising property usual for parabolic equations, (H_2) holds for $\hat{u}=0$ if the trajectories of (1.6) with $\zeta\equiv 0$ converge to zero as $t\to+\infty$ (which is often the case), and (H_4) specifies the noise. Finally, (H_3) holds if the nonlinearity V satisfies a Hörmander-type condition and may be checked for many nonlinear PDEs; see Section 4 for two examples.

1.2 Main result and examples

Let us denote by $\|\cdot\|_L^*$ the dual-Lipschitz metric in the space of probability measures on X (cf. Notation and conventions). The following theorem is the main result of this paper.

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 are positive numbers C and γ such that*

$$\|\mathfrak{P}_k^*\lambda-\mu\|_L^*\leq Ce^{-\gamma k}\quad \textit{for all } k\geq 0 \textit{ and } \lambda\in\mathcal{P}(X). \tag{1.7}$$

A proof of this result is given in Section 3. It is based on an application of the Kantorovich functional method, described in Section 3.1.1 of [KS12] and repeated here as Theorem 7.1 in the appendix. By that result, to prove the theorem it suffices to check the contraction inequality (7.4) for some Kantorovich functional $\mathcal{K}_F$. This requires subtle analysis based on ideas from the optimal control and theory of analytic functions. It is carried out in Section 3, after developing some auxiliary tools (of independent interest) in Section 2. In Section 4, we discuss in detail two examples which are briefly sketched below.

Example 1.2 (Navier–Stokes system). Let $\mathbb{T}^2_a=\mathbb{R}^2/(2\pi a_1)\mathbb{Z}\oplus(2\pi a_2)\mathbb{Z}$ be a rectangular torus, where $a=(a_1,a_2)$ is a vector with positive coordinates, and let

$$H=\biggl\{u\in L^2(\mathbb{T}^2_a,\mathbb{R}^2):\operatorname{div}u=0\text{ in }\mathbb{T}^2_a,\int_{\mathbb{T}^2_a}u(x)\,\mathrm{d}x=0\biggr\}. \tag{1.8}$$

We consider the Navier–Stokes system in $\mathbb{T}^2_a$ perturbed by a random process. Applying the Leray projection $\Pi:L^2(\mathbb{T}^2_a,\mathbb{R}^2)\to H$ to the equations for the velocity field, we reduce the system to the following nonlocal PDE:

$$\partial_t u+\nu Lu+B(u)=\eta(t,x). \tag{1.9}$$

Here $\nu>0$ is the viscosity coefficient, $L=-\Pi\Delta$ is the Stokes operator, $B(u)=\Pi(\langle u,\nabla\rangle u)$, and η is a random process which is assumed to be of

the form

$$\eta(t,x)=\sum_{k=1}^{\infty}\mathbb{I}_{[k-1,k)}(t)\eta_k(t-k,x), \tag{1.10}$$

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

To formulate some further hypotheses on η_k, let us take a finite or infinite dimensional Hilbert space $\mathcal{H}$ that is continuously embedded into $H^3(\mathbb{T}^2_a,\mathbb{R}^2)$ and denote $E:=L^2(J,\mathcal{H})$. We shall assume that $\{\eta_k\}$ are i.i.d. random variables in E whose law $\ell=\mathcal{D}(\eta_k)$ has a compact support $\mathcal{K}\subset E$, containing the origin and satisfying the two conditions below.

Decomposability. *The random variables η_k are decomposable in the sense that they satisfy Hypothesis* (H_4).

We emphasise that the decomposability can be regarded as a condition on the measure ℓ, since any other random variable with law ℓ has the same structure as η_k. To describe the second condition, we need the concepts of observable[5] function and observable measure.

Definition 1.3. Let $(\cdot,\cdot)_{\mathcal{H}}$ be an inner product and $\{\varphi_i\}_{i\in\mathcal{I}}$ an orthonormal basis in $\mathcal{H}$. We say that a function $\zeta\in L^2(J,\mathcal{H})$ is *Lipschitz-observable* with respect to $\{\varphi_i\}_{i\in\mathcal{I}}$ if for any Lipschitz-continuous functions $a_i: J\to\mathbb{R}$, $i\in\mathcal{I}$ and continuous function $b: J\to\mathbb{R}$ such that

$$\sum_{i\in\mathcal{I}}\|a_i\|^2_{C(J)}<\infty \tag{1.11}$$

the equality[6]

$$\sum_{i\in\mathcal{I}}a_i(t)(\zeta(t),\varphi_i)_{\mathcal{H}}-b(t)=0\quad\text{in } L^1(J) \tag{1.12}$$

implies that a_i, $i\in\mathcal{I}$ and b vanish identically.

A probability measure ℓ on $L^2(J,\mathcal{H})$ is said to be *Lipcshitz-observable* with respect to $\{\varphi_i\}$ if ℓ-almost every trajectory $\eta\in L^2(J,\mathcal{H})$ is Lipschitz-observable with respect to $\{\varphi_i\}$.

Observability. *The measure ℓ is Lipschitz-observable with respect to an orthonormal basis $\{\varphi_i\}$ of the space $\mathcal{H}$.*

The properties of decomposability and observability are not very restrictive, and some examples are given in Section 5. In particular, as we show in Section 5.2, the coloured noise (0.2), (0.3) satisfies both.

[5]The concept of observability is widely used in the control theory and means, roughly speaking, that if a functional of a non-zero solution of a homogeneous linear differential equation vanishes identically in time, then it must be zero. In Definition 1.3 below, we have a similar property for functions: the left-hand side of (1.12) defines an affine function, and if it vanishes on ζ, then it must be zero.

[6]The convergence in $L^1(J)$ of the series in (1.12) (when $\mathcal{I}$ is infinite) follows easily from (1.11).

Remark 1.4. If $\dim \mathcal{H} < \infty$, then property of Lipschitz-obvervability does not depend on the basis. Indeed, if we change $\{\varphi_i\}_{i\in\mathcal{I}}$ to another orthogonal basis in $\mathcal{H}$, then the vector-function $a(t) = (a_1(t), \dots a_N(t))$ will be replaced by $a'(t) = Ua(t)$, where $U : \mathbb{R}^N \to \mathbb{R}^N$ is an orthogonal transformation. The function $a'(t)$ vanishes identically in t if and only if $a(t)$ does, and the components of a' are Lipschitz-continuous if and only if those of a are. This implies the required assertion.

We also note that, in the case of a finite $\mathcal{I}$, the Lipschitz-observability means that the elements $\{(\zeta(t), \varphi_i)\}$ of the space $L^2(J)/C(J)$ are linearly independent over the ring of Lipschitz-continuous functions on J.

We shall prove that Theorem 1.1 applied to the Navier–Stokes system (1.9) implies the following result: if the law $\ell = \mathcal{D}(\eta_k)$ is decomposable and observable, and if the space $\mathcal{H}$ possesses a *saturation property* (see Section 4.1), then the Markov process associated with (1.9) has a unique stationary measure, which is exponentially mixing in the dual-Lipschitz metric.

Example 1.5 (Ginzburg–Landau equation). Let $\mathbb{T}^3 \subset \mathbb{R}^3$ be a rectangular torus. We consider the equation

$$\partial_t u - (\nu + i)\Delta u + \gamma u + ic|u|^{p-1}u = \eta(t, x), \tag{1.13}$$

where $u = u(t, x)$ is an unknown complex-valued function, ν, γ, and c are positive parameters, and $p \in \{3, 5\}$. Equation (1.13) is well posed in the Sobolev space $H^1(\mathbb{T}^3, \mathbb{C})$. Assuming again that the random force η has the form (1.10) and satisfies the decomposability and observability hypotheses, we shall prove uniqueness and exponential mixing of stationary measure for (1.13).

1.3 Coupling in infinite dimension and Newton–Kantorovich–Kolmogorov fast convergence

In this subsection, we describe the general scheme of the proof of Theorem 1.1, a coupling construction, and discuss its relation to the technique of fast convergence due to Newton–Kantorovich–Kolmogorov. To prove convergence (1.7), it suffices to verify that any two trajectories $\{u_k\}$ and $\{u'_k\}$ with random initial data u and u' converge in distribution when $k \to \infty$. Very often this property is proven with the help of a coupling argument, originated in 1930's in the work of Doeblin [Doe40] (see also [Doe00]), and we recall now the main idea.

Let us consider the following dynamics in the space $X \times X$:

$$\begin{aligned}(u_0, v_0) &= (u, u'),\\ (u_k, v_k) &= (S(u_{k-1}, \eta_k), S(v_{k-1}, \eta'_k)), \quad k \geq 1,\end{aligned}$$

where (η_k, η'_k) is a sequence of independent random variables such that

$$\mathcal{D}(\eta'_k) = \mathcal{D}(\eta_k) = \ell \quad \text{for all } k \geq 1. \tag{1.14}$$

It is clear that $\{u_k\}$ is a trajectory of (1.1) starting from u, while $\{v_k\}$ coincides with that issued from u' in the sense of law: $\mathcal{D}(v_k) = \mathfrak{P}_k^* \mathcal{D}(u') = \mathcal{D}(u'_k)$. Our

goal is to choose a sequence $\{\eta'_k\}$ satisfying (1.14) such that, with probability 1, $\{v_k\}$ is asymtotically close to $\{u_k\}$ as $k \to \infty$. To construct η'_k, we fix a small parameter $\delta_0 > 0$ and distinguish between the following two cases:

(a) If $\|u - u'\| > \delta_0$, we choose for η'_k an independent copy of η_1 as long as $\|u_{k-1} - v_{k-1}\| > \delta_0$. Due to the dissipativity, at some random time τ, we shall have $\|v_\tau - u_\tau\| \le \delta_0$. The Markov time τ is no bigger than the first instance when both trajectories are in the $\frac{\delta_0}{2}$-neighbourhood of $\hat{u}$, and the latter can be controlled due to Hypotheses (H_2).

(b) The case $\|u - u'\| \le \delta_0$ contains the main difficulty. We seek η'_1 in the form $\eta'_1 = \Psi^{u,u'}(\eta_1)$, where the transformation $\Psi^{u,u'}$ satisfying

$$\Psi^{u,u'}_*(\ell) = \ell \tag{1.15}$$

has to be constructed. Relation (1.15) implies (1.14) (with $k = 1$), and the goal is to find $\Psi^{u,u'}$ such that the inequality $\|u_1 - v_1\| \ll \|u - u'\|$ holds with high probability.

In the best case, we may have $v_1 = u_1$ almost surely, that is

$$S(u, \eta_1) - S(u', \eta'_1) = 0, \quad \text{a.s.} \tag{1.16}$$

This is not likely to be possible because, for deterministic initial states u and u', it would imply that $\mathcal{D}(u_1) = \mathcal{D}(v_1) = \mathcal{D}(u'_1)$, which is not necessarily the case.

The situation is reminiscent of that treated by Kolmogorov's celebrated theorem on nearly-integrable Hamiltonians

$$H_\delta(p, q) = h_1(p) + \delta f_1(p, q), \qquad (p, q) \in B^n \times \mathbb{T}^n, \tag{1.17}$$

where $B^n \subset \mathbb{R}^n$ is a ball and $\delta > 0$ is a small parameter. Since the hamiltonian dynamics for a Hamiltonian depending only on p is integrable, a naive idea to study that corresponding to H_δ is to find a symplectic transformation S_δ such that

$$H_\delta \circ S_\delta(p, q) = h_\delta(p). \tag{1.18}$$

However, it was shown by Poincaré that, in general, such a transformation S_δ does not exist. Kolmogorov's well-known idea to bypass this obstruction is to achieve (1.18) only up to a higher order term:

$$H_\delta \circ S_1(p, q) = h_2(p) + \delta^2 f_2(p, q). \tag{1.19}$$

This relation holds if the symplectic transformation S_1 is a time-1 flow of a Hamiltonian $\delta g_1(p, q)$, with some function g_1 satisfying the linear homological equation

$$\{h_1(p), g_1(p, g)\} = f_1(p, q) - \langle f_1 \rangle(p), \tag{1.20}$$

where $\langle \cdot \rangle$ stands for the averaging in $q \in \mathbb{T}^n$. If h_1 meets a mild non-degeneracy condition, then (1.20) can be solved with a disparity of order δ, for all q and for p

outside a small-measure set $B_1 \subset B^n$. The corresponding transformation S_1 reduces H_δ to a Hamiltonian (1.19) which is much closer to be integrable than H_δ. Iterating this argument, Kolmogorov constructed symplectomorphisms $\{S_1 \circ \cdots \circ S_j\}_{j\geq 1}$ transforming H_δ to some Hamiltonians $H_{\delta,j}$ that are δ^{2^j}-close to be integrable, for all q and for p outside a set $B_1 \cup \cdots \cup B_j \subset B$. When $j \to \infty$, the transformations $S_1 \circ \cdots \circ S_j$ converge, super-exponentially fast, to a limiting transformation S_δ which satisfies (1.18) for all p outside the set $\cup_j B_j$ which turns out to be of a small measure. This implies the assertions made by Kolmogorov in his seminal paper[7] [Kol54].

Going back to our problem and denoting $\delta = \|u - u'\| \leq \delta_0$, we write $u' = u + \delta v$, where $\|v\| = 1$, and seek a transformation of the form $\Psi^{u,u'} = \mathrm{Id} + \delta\Phi$. Similar to Kolmogorov's approach, let us rewrite relation (1.16) as

$$\delta\big(D_\eta S(u,\eta_1)\Phi(\eta_1) - S'(u,u',\eta_1)\big) + O(\delta^2) = 0,$$

where $S' = \delta^{-1}(S(u+\delta v,\eta_1) - S(u,\eta_1))$ is of order 1. Neglecting the term $O(\delta^2)$, consider the equation

$$D_\eta S(u,\eta_1)\Phi(\eta_1) = S'(u,u',\eta_1). \tag{1.21}$$

This is the homological equation of our proof, analogous to Eq. (1.20) from Kolmogorov's theorem. Equation (1.21), as well as (1.20), cannot be solved exactly or approximatively for a.a. η_1. However, in view of Hypothesis (H_3), it can be solved approximately for $\eta_1 \notin \mathcal{N}^u$, where $\mathcal{N}^u$ is a suitable "bad" set of small ℓ-measure in the support $\mathcal{K}$ of ℓ. It is proved in Section 2.2 that, for any $\varepsilon > 0$, an approximate solution $\Phi(\eta_1)$ solving (1.21) up to a term of order ε can be found in the form

$$\Phi(\eta_1) = R_\varepsilon(u,\eta_1)S',$$

where $R_\varepsilon(u,\eta_1)$ is a finite-dimensional linear operator satisfying the inequality

$$\|R_\varepsilon(u,\eta_1)\| \leq C(\varepsilon, \ell(\mathcal{N}^u)) < \infty \tag{1.22}$$

with some function $C(\varepsilon, r)$ going to $+\infty$ as $\varepsilon \to 0$ or $r \to 0$. Setting

$$\Psi^{u,u'}(\eta_1) = \eta_1 + \delta\Phi(\eta_1) = \eta_1 + \delta R_\varepsilon(u,\eta_1)S'(u,u',\eta_1),$$

we make the left-hand side of (1.16) of order

$$\delta_1 = O(\delta\varepsilon) + \delta^2 C(\varepsilon, \ell(\mathcal{N}^u))^2 \quad \text{for} \quad \eta_1 \notin \mathcal{N}^u. \tag{1.23}$$

In the KAM theory, the usual strategy is to choose $\varepsilon = \delta^{\gamma_1}$ with some $\gamma_1 > 0$. Then, if we knew that $\delta^2 C(\delta^{\gamma_1}, \ell(\mathcal{N}^u))^2 \leq \delta^{1+\gamma_2}$ with some $\gamma_2 > 0$, this would lead to a super-exponential convergence $\|v_k - u_k\| \to 0$, typical for the theory.

[7] The paper contained a scheme of the proof, and a complete proof appeared only 10 years later in the works by Arnold and Moser, resulting in creation of the KAM theory; see [Arn63, Mos66].

However, such a choice is now impossible, since the constant $C(\varepsilon, \ell(\mathcal{N}^u))$ is practically out of control. Instead, we derive from (1.23) that

$$\delta_1 \le C_1 \delta \big(\varepsilon + \delta_0 C(\varepsilon, \ell(\mathcal{N}^u))^2\big)$$

and choose $\varepsilon = (4C_1)^{-1}$ and $\delta_0 = \big(4C_1 C(\varepsilon, \ell(\mathcal{N}^u))^2\big)^{-1}$. This implies that $\delta_1 \le \frac{1}{2}\delta$ and leads to an exponential convergence $\|v_k - u_k\| \to 0$, which is sufficient for our purposes.

Construction of vectors $\{\eta'_k\}_{k\ge 1}$ encounters two[8] difficulties, which we describe for the first step:

(1) The transformation $\eta_1 \mapsto \Psi^{u,u'}(\eta_1)$ does not preserve the measure ℓ, so (1.15) does not necessarily hold.

(2) The transformation $\Psi^{u,u'}(\eta_1)$ is not defined for $\eta_1 \in \mathcal{N}^u$.

The first difficulty is overcome due to the observation that, in our situation, the distance between the measures ℓ and $\Psi^{u,u'}_* \ell$ is of order $\big(\delta \|R_\varepsilon(u, \eta_1)\|\big)^{\varkappa}$ with some $\varkappa > 0$, which is small by (1.22). Thus, even though the laws of v_1 and u'_1 are not the same, the two are close, which allows to bound the distance between the laws of u_1 and u'_1 by the triangle inequality, provided that v_1 is close to u_1.

To handle the second difficulty, we extend the definition of $\Psi^{u,u'}$ to $\mathcal{N}^u$ as follows:

(c) If $\eta_1 \in \mathcal{N}^u$, then $\Psi^{u,u'}(\eta_1) = \eta_1$.

Since the mapping S is Lipschitz on the compact set $X \times \mathcal{K}$, for $\eta_1 \in \mathcal{N}^u$ we have $\|u_1 - v_1\| \le C\delta$. We then iterate Steps (a)–(c) depending on the value of the difference $\|u_k - v_k\|$.

The growth of the constant $C(\varepsilon, r)$ as $r \to 0$ is difficult to control, and we cannot make the ℓ-measure set $\mathcal{N}^u$ very small. As a consequence, case (c) happens rather often. This slows down the fast convergence, usually associated with the quadratic scheme. However, combining this construction with some techniques based on the study of the behaviour of Kantorovich functionals on a pair of trajectories enables one to prove that the Markov operator defines a contraction on the space of measures. This proves Theorem 1.1.

To summarise, the proof of Theorem 1.1 is based on the classical coupling scheme, enhanced with the quadratic convergence à la Kolmogorov to cope with difficulties (1) and (2) described above. The realisation of this scheme meets serious analytic difficulties. In Section 3, we implement the scheme, using some technical lemmas proved in Section 2. Moreover, these ideas apply also to the case in which the dissipativity hypothesis (H_2) is replaced by a weaker condition of global approximate controllability to a point; that possibility will be analysed in a subsequent publication.

[8]The first difficulty does not have an analogue in the KAM theory, whereas the second is almost always present and manifests itself in the fact that the homological equation (1.20) cannot be solved for all actions $p \in B$.

2 Auxiliary tools

In this section, we establish some results that will be used in the proof of Theorem 1.1. Their demonstration is based on well known methods. However, since the results are not available in the literature in the form we need, we provide rather detailed proofs.

2.1 Transformation of probability measures under piecewise Lipschitz mappings

In what follows, we denote by E a separable Hilbert space, by $\mathcal{K} \subset E$ a compact subset, and by $\Psi : \mathcal{K} \to E$ a mapping of the form $\Psi(\eta) = \eta + \Phi(\eta)$, where $\Phi : \mathcal{K} \to E$ is a Borel-measurable "small" mapping vanishing outside $\mathcal{K}$. Our goal is to study the transformation of measures on $\mathcal{K}$ under Ψ. We shall consider the situation in which the mappings Φ and Ψ are piecewise Lipschitz in the following sense: their restrictions to a "large" closed subset $\mathcal{K}_1 \subset \mathcal{K}$ is Lipschitz, while outside $\mathcal{K}_1$ they are, respectively, the zero-map and the identity-map.

Let E be the direct sum of closed subspaces $\mathcal{E}$ and $\mathcal{E}'$, where $\dim \mathcal{E} < +\infty$, and let $\mathsf{P}_{\mathcal{E}}$ and $\mathsf{P}_{\mathcal{E}'}$ be the associated projections. We assume that the image of Φ is contained in $\mathcal{E}$. Given any subset $A \subset E$, we denote by A' its projection to $\mathcal{E}'$, and for any $w \in A'$, we write $A(w) = \{v \in \mathcal{E} : v + w \in A\}$. Let $\ell \in \mathcal{P}(E)$ be a measure that is supported by $\mathcal{K}$ and is representable as the tensor product of its projections $\ell_{\mathcal{E}}$ and $\ell_{\mathcal{E}'}$ under $\mathsf{P}_{\mathcal{E}}$ and $\mathsf{P}_{\mathcal{E}'}$, respectively. We assume that $\ell_{\mathcal{E}}$ has a Lipschiz-continuous density ρ with respect to the Lebesgue measure on $\mathcal{E}$.

Theorem 2.1. *Let $\Phi : \mathcal{K} \to \mathcal{E} \subset E$ be a Borel-measurable mapping that possesses the following properties:*

(a) *There is a positive number $\varkappa$ and a closed subset $\mathcal{K}_1 \subset \mathcal{K}$ such that $\Phi\big|_{\mathcal{K}\setminus\mathcal{K}_1} = 0$ and*

$$\|\Phi(\eta)\| \le \varkappa, \quad \|\Phi(\eta) - \Phi(\eta')\| \le \varkappa\|\eta - \eta'\| \quad \text{for } \eta, \eta' \in \mathcal{K}_1. \tag{2.1}$$

(b) *There are positive numbers c and γ such that, for any $w \in \mathcal{K}_1'$ and $r \in [0,1]$, we have*

$$\mathrm{Leb}\big\{v \in \mathcal{E} : \mathrm{dist}\big(v, \partial_w \mathcal{K}_1\big) \le r\big\} \le c\, r^{\gamma}, \tag{2.2}$$

where $\partial_w \mathcal{K}_1 = \mathcal{K}_1(w) \cap \overline{\mathcal{K}_1(w)^c}$ and $\mathcal{K}_1(w)^c = \mathcal{K}(w) \setminus \mathcal{K}_1(w)$.

Let $\ell \in \mathcal{P}(E)$ be a measure satisfying the above hypotheses such that $\operatorname{supp} \ell \subset \mathcal{K}$. *Then there are positive numbers C and β depending only on ℓ, c, and γ such that*

$$\|\ell - \Psi_*(\ell)\|_{\mathrm{var}} \le C\varkappa^{\beta}. \tag{2.3}$$

Proof. Note that the assertion of the theorem is trivial if $\varkappa$ is separated from zero, so that we shall consider the case $\varkappa \le \frac{1}{2}$. We wish to estimate the supremum of

the absolute value of the expression $\langle f\circ\Psi,\ell\rangle-\langle f,\ell\rangle$ over all indicator functions f of Borel sets in $\mathcal{K}$. To this end, we use the Fubini theorem to write

$$\langle f\circ\Psi,\ell\rangle=\int_{\mathcal{K}'}\ell_{\mathcal{E}'}(\mathrm{d}w)\int_{\mathcal{K}(w)}f(v+w+\Phi(v+w))\rho(v)\mathrm{d}v, \tag{2.4}$$

$$\langle f,\ell\rangle=\int_{\mathcal{K}'}\ell_{\mathcal{E}'}(\mathrm{d}w)\int_{\mathcal{K}(w)}f(v+w)\rho(v)\mathrm{d}v. \tag{2.5}$$

Suppose we have shown that

$$\left|\int_{\mathcal{K}_1(w)}\big(f(v+w+\Phi(v+w))-f(v+w)\big)\rho(v)\mathrm{d}v\right|\le C_1\varkappa^\beta, \tag{2.6}$$

where C_1 and β do not depend on w, f, and $\varkappa$. In this case, taking the absolute value of the difference between (2.4) and (2.5), using that Φ is zero outside $\mathcal{K}_1$, and estimating the interior integral with the help of (2.6), for any indicator function f we derive

$$|\langle f\circ\Psi,\ell\rangle-\langle f,\ell\rangle|\le C\varkappa^\beta.$$

Since f is arbitrary, we arrive at the required estimate (2.3). Thus, we need to establish (2.6).

We first outline the main idea. Suppose that $\mathcal{K}_1(w)$ coincides with the whole space $\mathcal{E}$. In this case, we can make a change of variable $v\mapsto v+\Phi(v+w)=v'$ and rewrite the integral $\int_{\mathcal{E}}f(v+w+\Phi(v+w))\rho(v)\mathrm{d}v$ in the form

$$\int_{\mathcal{E}}f(v'+w)\frac{\rho(\Theta_w(v'))}{\det\big(I+(D\Phi)(\Theta_w(v')+w)\big)}\mathrm{d}v',$$

where $\Theta_w(v')$ is the inverse of $v+\Phi(v+w)$ with respect to v. This expression is easy to compare with $\int_{\mathcal{E}}f(v+w)\rho(v)\mathrm{d}v$ due to inequalities (2.1). However, the set $\mathcal{K}_1(w)$ may have a complicated structure, and to carry out the above mentioned change of variables, we need to extend Φ to the whole space and to introduce some truncations not to change much the values of the integrals.

Let us turn to the accurate proof of (2.6). We first extend the mapping Φ from $\mathcal{K}_1$ to E. To this end, we use the following result, whose proof can be found in [Val45] (see also Section 2.10.43 of [Fed69] for the finite-dimensional case, which is not sufficient for our purposes).

Proposition 2.2 (Kirszbraun theorem)**.** *Let E_1 and E_2 be two Hilbert spaces, let $A\subset E_1$ be a set and $\Phi:A\to E_2$ be a Lipschitz-continuous function, with a Lipschitz constant $\varkappa$. Then there is a function $\tilde\Phi:E_1\to E_2$ that coincides with Φ on A and is Lipschitz continuous with the same constant $\varkappa$.*

Let us denote by $\tilde\Phi:E\to\mathcal{E}$ a Lipschitz-continuous function, with Lipschitz constant $\le\varkappa$, that coincides with Φ on $\mathcal{K}_1$. Since $\mathcal{K}_1$ is compact, we can multiply $\tilde\Phi$ by a cut-off function, so that there is no loss of generality in assuming that $\tilde\Phi$ has a bounded support and Lipschitz constant $\le C\varkappa$, where $C>0$ does not depend on $\varkappa$.

To introduce truncations, we define the sets

$$\mathcal{K}_r^+(w) = \{v \in \mathcal{K}(w) : \operatorname{dist}(v, \mathcal{K}_1(w)) \le r\},$$
$$\mathcal{K}_r^-(w) = \{v \in \mathcal{K}(w) : \operatorname{dist}(v, \mathcal{K}_1(w)^c) \le r\},$$

where $r > 0$ is a small parameter chosen below. Let us consider the functions

$$\chi_r^+(v,w) = \frac{\operatorname{dist}(v, \mathcal{K}_r^+(w)^c)}{\operatorname{dist}(v, \mathcal{K}_r^+(w)^c) + \operatorname{dist}(v, \mathcal{K}_1(w))},$$
$$\chi_r^-(v,w) = \frac{\operatorname{dist}(v, \mathcal{K}_1(w)^c)}{\operatorname{dist}(v, \mathcal{K}_1(w)^c) + \operatorname{dist}(v, \mathcal{K}_r^-(w)^c)}.$$

These are non-negative functions bounded by 1 and Lipschitz continuous with constant r^{-1}. Since f is a non-negative function and the support of ρ is equal to $\mathcal{K}(w)$, we can write

$$\begin{aligned}\int_{\mathcal{K}_1(w)} f(v+w+\Phi(v+w))\rho(v)\mathrm{d}v &\le \int_{\mathcal{E}} f(v+w+\tilde{\Phi}(v+w))\chi_r^+(v,w)\rho(v)\mathrm{d}v,\\ \int_{\mathcal{K}_1(w)} f(v+w)\rho(v)\mathrm{d}v &\ge \int_{\mathcal{E}} f(v+w)\chi_r^-(v,w)\rho(v)\mathrm{d}v\\ &\ge \int_{\mathcal{E}} f(v+w)\chi_r^+(v,w)\rho(v)\mathrm{d}v - c\,\|\rho\|_\infty r^\gamma,\end{aligned}$$

where we used the fact that $|\chi_r^-(v,w) - \chi_r^+(v,w)|$ is a function bounded by 1 and supported in the set entering the left-hand side of (2.2). Now note that, for $\varkappa \le \frac{1}{2}$, the mapping $v \mapsto v + \tilde{\Phi}(v+w)$ is a bi-Lipschitz homeomorphism of $\mathcal{E}$. Therefore, denoting by $\delta(f)$ the expression under the absolute value in (2.6) and using, for instance, Theorem 3.2.5 in [Fed69] to make a change of variable, we obtain

$$\begin{aligned}\delta(f) &\le \int_{\mathcal{E}} f(v+w)\frac{\chi_r^+(\Theta_w(v),w)\rho(\Theta_w(v))}{\det\big(I + (D\tilde{\Phi})(\Theta_w(v)+w)\big)}\,\mathrm{d}v\\ &\quad - \int_{\mathcal{E}} f(v+w)\chi_r^+(v,w)\rho(v)\mathrm{d}v + C_1 r^\gamma\\ &\le \int_B f(v+w)|\Delta(v,w)|\mathrm{d}v + C_1 r^\gamma, \qquad (2.7)\end{aligned}$$

where $B \subset \mathcal{E}$ is a large ball containing the supports of ρ and $\rho \circ \Theta_w$, and we set

$$\Delta(v,w) = \frac{\chi_r^+(\Theta_w(v),w)\rho(\Theta_w(v))}{\det\big(I + (D\tilde{\Phi})(\Theta_w(v)+w)\big)} - \chi_r^+(v,w)\rho(v).$$

Since $v \mapsto \Theta_w(v)$ is a 2-Lipschitz function satisfying the inequality $|\Theta_w(v) - v| \le \varkappa$ for all $v \in \mathcal{E}$, we have $|\Delta(v,w)| \le C_2\varkappa(1 + r^{-1})$. Substituting this into (2.7), we derive

$$\delta(f) \le C_3\varkappa(1 + r^{-1}) + C_1 r^\gamma.$$

Choosing $r = \varkappa^{1/(1+\gamma)}$, we get

$$\delta(f) \le C_4 \varkappa^{\gamma/(1+\gamma)}. \tag{2.8}$$

A similar argument shows that $\delta(f) \ge -C_5 \varkappa^{\gamma/(1+\gamma)}$. Combining this with (2.8), we arrive at inequality (2.6) with $\beta = \frac{\gamma}{1+\gamma}$. This completes the proof of the theorem. □

2.2 Approximate right inverse of linear operators with dense image

Let F and H be separable Hilbert spaces and let $A : F \to H$ be a continuous linear operator. Consider the equation $A\zeta = f$. In general, it does not have a solution, and when it does, the solution may not be unique. The following result shows that, under some additional conditions, one may construct an approximate solution that linearly depends on f.

Proposition 2.3. *In addition to the above hypotheses, let the image of A be dense in H and let V be a Banach space compactly embedded into H. Then for any $\varepsilon > 0$ there is a continuous linear operator $R_\varepsilon : H \to F$ with a finite-dimensional range such that*

$$\|AR_\varepsilon f - f\|_H \le \varepsilon \|f\|_V \quad \text{for any } f \in V. \tag{2.9}$$

Proof. Let us define the operator $G = AA^* : H \to H$. Since the image $\mathrm{Im}(A)$ is dense in H, the kernel of the self-adjoint operator G is trivial, and therefore the image $\mathrm{Im}(G)$ is dense in H. Let us recall that the operator A^*G^{-1} defined on $\mathrm{Im}(G)$ is called the *Moore–Penrose pseudo-inverse* and singles out the solution of the least norm for the equation $A\zeta = f$ (when it exists). We shall need the following lemma giving a natural approximation of the right inverse of A.

Lemma 2.4. *Let $G : H \to H$ be a non-negative self-adjoint operator. Then the mapping $(0, +\infty) \ni \gamma \mapsto (G + \gamma)^{-1}$ is a well-defined smooth operator function such that $\Delta_f(\gamma) := \|G(G + \gamma)^{-1} f - f\|^2$ decreases with γ for every $f \in H$. Moreover, the norms of the operators $G(G + \gamma)^{-1}$ and $(G + \gamma)^{-1}$ are bounded, respectively, by 1 and γ^{-1}, and if G has dense image*

$$\lim_{\gamma \to 0} \|G(G + \gamma)^{-1} f - f\| = 0 \quad \text{for any } f \in H. \tag{2.10}$$

We now construct R_ε by truncating $A^*(G + \gamma)^{-1}$ to ensure that the image is finite-dimensional. Namely, choosing an orthonormal basis $\{f_j\}$ in F and denoting by P_M the orthogonal projection to the vector space spanned by the first M vectors, we define $R_{\gamma,M} = \mathsf{P}_M A^*(G + \gamma)^{-1}$. We now fix any $\varepsilon > 0$. By Lemma 2.4, for any $f \in H$ there is $\gamma_\varepsilon(f) > 0$ such that

$$\|AA^*(G + \gamma)^{-1} f - f\| \le \frac{\varepsilon}{3} \quad \text{for } 0 < \gamma \le \gamma_\varepsilon(f).$$

Since the norm of the operator $AA^*(G+\gamma)^{-1}$ is bounded by 1, for any $f \in H$ there is $\delta_\varepsilon(f) > 0$ such that

$$\|AA^*(G+\gamma)^{-1}g - g\| \le \frac{2\varepsilon}{3} \quad \text{for } 0 < \gamma \le \gamma_\varepsilon(f),\ \|g - f\| \le \delta_\varepsilon(f). \tag{2.11}$$

The open balls $\{O_f := \dot{B}_H(f, \delta_\varepsilon(f))\}_{f\in H}$ form a covering of the compact set $B_V(1) \subset H$. Choosing a finite sub-covering $\{O_{f_j}, 1 \le j \le m\}$ and setting $\gamma_\varepsilon := \min\{\gamma_\varepsilon(f_j), 1 \le j \le m\}$, we derive from (2.11) that

$$\|AA^*(G+\gamma_\varepsilon)^{-1}f - f\| \le \frac{2\varepsilon}{3} \quad \text{for } f \in B_V(1).$$

Since the sequence $\{AR_{\gamma_\varepsilon,M}\}_{M\ge1}$ converges to $AA^*(G+\gamma_\varepsilon)^{-1}$ as $M \to \infty$ in the strong operator topology and the convergence is uniform on compact subsets, we can find $M_\varepsilon \ge 1$ such that

$$\|AR_{\gamma_\varepsilon,M_\varepsilon}f - f\|_H \le \varepsilon \quad \text{for } f \in B_V(1).$$

By homogeneity, this implies (2.9) with $R_\varepsilon = R_{\gamma_\varepsilon,M_\varepsilon}$. □

Proof of Lemma 2.4. Since $G \ge 0$, it follows that $(G+\gamma I)^{-1}$ is well defined and smooth in $\gamma > 0$. When proving that Δ_f decreases with γ, we can assume, by the spectral theorem, that G acts in a Lebesgue space $L^2(\mathfrak{X}, \lambda)$ as the multiplication by a bounded non-negative function $a(x)$. In this case,

$$\Delta_f(\gamma) = \|G(G+\gamma I)^{-1}f - f\|^2 = \int_{\mathfrak{X}} \frac{\gamma^2|f(x)|^2}{(a(x)+\gamma)^2}\,\lambda(\mathrm{d}x).$$

It remains to note that the integrand is an increasing function of $\gamma > 0$.

The above representation of the operator G readily implies the assertions concerning the norms of $G(G+\gamma)^{-1}$ and $(G+\gamma)^{-1}$. To prove (2.10) for operators with a dense image, we first note that

$$G(G+\gamma)^{-1} - I = -\gamma(G+\gamma)^{-1} \tag{2.12}$$

Since the norm of the operator $\gamma(G+\gamma)^{-1}$ is bounded by 1, it suffices to prove that the right-hand side of (2.12) goes to zero for a dense subset of vectors $f \in H$. Let us take any f in the (dense) image of G. Then there is $h \in H$ such that $f = Gh$, so that

$$\|\gamma(G+\gamma)^{-1}f\| \le \gamma\|h\| + \gamma^2\|(G+\gamma)^{-1}h\| \le 2\gamma\|h\|.$$

This implies (2.10) and completes the proof. □

In what follows, we shall need a version of Proposition 2.3 for the case when the operator A depends on a parameter and degenerates for some of its values. Namely, let X be a compact metric space, let E be a separable Hilbert space, and let $\ell \in \mathcal{P}(E)$ be a Borel measure with a compact support $\mathcal{K}$. Consider a continuous mapping $A : X \times E \to \mathcal{L}(F, H)$ such that $A(u, \cdot) : E \to \mathcal{L}(F, H)$

is analytic for any $u \in X$ (we recall (0.6) and the convention for analytic mappings with parameter). As in Proposition 2.3, we denote by V a Banach space compactly embedded into H. Finally, we fix an orthonormal basis $\{f_j\}$ in F and denote by F_N the vector span of $f_1, \dots, f_N$.

Proposition 2.5. *In addition to the above hypotheses, let us assume that, for any $u \in X$, there is a set of full measure $\mathcal{K}^u \subset \mathcal{K}$ such that the image of the linear operator $A(u,\eta)$ is dense in H for any $\eta \in \mathcal{K}^u$. Then, for any $\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2) \in (0,1)^2$, there is an integer $M_{\vec{\varepsilon}} \geq 1$, positive numbers $\nu_{\varepsilon_2}, C_{\vec{\varepsilon}}$, and a non-negative continuous function $\mathfrak{F}_{\vec{\varepsilon}}(u,\eta)$, defined on $X \times E$ and analytic in η, such that the following properties hold.*

Bound on the measure. *The ℓ-measure of the sets*

$$\mathcal{K}^u_{\vec{\varepsilon}} := \{\eta \in \mathcal{K} : \mathfrak{F}_{\vec{\varepsilon}}(u,\eta) \leq \nu_{\varepsilon_2}\} \tag{2.13}$$

satisfies the inequality

$$\ell(\mathcal{K}^u_{\vec{\varepsilon}}) \geq 1 - \varepsilon_1 \quad \text{for } u \in X. \tag{2.14}$$

Right inverse. *Let us define the compact set*

$$\mathcal{D}_{\vec{\varepsilon}} = \{(u,\eta) \in X \times \mathcal{K} : \mathfrak{F}_{\vec{\varepsilon}}(u,\eta) \leq 2\nu_{\varepsilon_2}\}. \tag{2.15}$$

Then there is a continuous mapping $R_{\vec{\varepsilon}} : \mathcal{D}_{\vec{\varepsilon}} \to \mathcal{L}(H,F)$ such that

$$\operatorname{Im}\big(R_{\vec{\varepsilon}}(u,\eta)\big) \subset F_{M_{\vec{\varepsilon}}}, \quad \|R_{\vec{\varepsilon}}(u,\eta)\|_{\mathcal{L}(H,F)} \leq C_{\vec{\varepsilon}} \quad \text{for } (u,\eta) \in \mathcal{D}_{\vec{\varepsilon}}, \tag{2.16}$$

$$\|A(u,\eta)R_{\vec{\varepsilon}}(u,\eta)f - f\|_H \leq \varepsilon_2 \|f\|_V \quad \text{for } (u,\eta) \in \mathcal{D}_{\vec{\varepsilon}},\ f \in V. \tag{2.17}$$

Proof. We essentially repeat the proof of Proposition 2.3, following the dependence on the parameters. Namely, we set $G(u,\eta) = A(u,\eta)A(u,\eta)^*$, and given an integer $M \geq 1$ and a number $\gamma > 0$, define

$$R_\gamma(u,\eta) = A(u,\eta)^*\big(G(u,\eta) + \gamma I\big)^{-1}, \quad R_{M,\gamma}(u,\eta) = \mathsf{P}_M R_\gamma(u,\eta), \tag{2.18}$$

where $\mathsf{P}_M : F \to F$ denotes the orthogonal projection to F_M. We shall prove that, for any given $\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2) \in (0,1)^2$ and an appropriate choice of M and γ, the operator $R_{M,\gamma}$ possesses all the required properties.

The fact that the image of $R_{M,\gamma}$ is contained in the subspace F_M follows immediately from the definition. Furthermore, since the norm of the inverse $(G+\gamma I)^{-1}$ is bounded by γ^{-1}, we have

$$\|R_{M,\gamma}(u,\eta)\|_{\mathcal{L}(H,F)} \leq C_1\gamma^{-1}, \quad C_1 = \sup_{(u,\eta)\in X\times\mathcal{K}} \|A(u,\eta)\|_{\mathcal{L}(F,H)},$$

so that the inequality in (2.16) holds for any fixed M and γ. Thus, we need to construct a function $\mathfrak{F}_{\vec{\varepsilon}}$ and to prove (2.14) and (2.17).

Let us show that for any $\gamma > 0$ there is an integer $M_{\gamma,\vec{\varepsilon}} \geq 1$ with the following property: if for some closed subset $\mathcal{D}_{\vec{\varepsilon}} \subset X \times \mathcal{K}$ we have

$$\sup_{(u,\eta)\in\mathcal{D}_{\vec{\varepsilon}}} \|A(u,\eta)R_\gamma(u,\eta)f - f\| < \varepsilon_2 \quad \text{for } f \in B_V(1), \tag{2.19}$$

then inequality (2.17) with $R_{\vec{\varepsilon}} = R_{M,\gamma}$ is true for $M \geq M_{\gamma,\vec{\varepsilon}}$. Indeed, let us note that

$$\Delta_{M,\gamma}(u,\eta)f := AR_\gamma f - AR_{M,\gamma}f = A(u,\eta)(I - \mathsf{P}_M)R_\gamma(u,\eta).$$

Since the mapping $(u,\eta,f) \mapsto R_\gamma(u,\eta)f$ acting from $X \times \mathcal{K} \times H$ to H is continuous, the image of the compact set $\mathcal{D}_{\vec{\varepsilon}} \times B_V(1)$ is compact. The convergence of P_M to I in the strong operator topology implies that

$$\sup_{(u,\eta,f)\in\mathcal{D}_{\vec{\varepsilon}}\times B_V(1)} \|\Delta_{M,\gamma}(u,\eta)f\| \to 0 \quad \text{as } M \to \infty.$$

Combining this with (2.19), we see that (2.17) is true for $M \geq M_\gamma$. We thus need to construct a function $\mathfrak{F}_{\vec{\varepsilon}}$ and numbers $\gamma, \nu_{\varepsilon_2} > 0$ for which (2.14) and (2.19) hold.

To this end, note that, by Lemma 2.4, for any $u \in X$, $\eta \in \mathcal{K}^u$, and $f \in H$,

$$\lim_{\gamma\to 0} \|A(u,\eta)R_\gamma(u,\eta)f - f\| = 0. \tag{2.20}$$

Let $\{f_j, 1 \leq j \leq N\} \subset H$ be an $(\varepsilon_2/4)$-net for the compact set $B_V(1) \subset H$. For $\gamma > 0$, we define the continuous function

$$\mathfrak{F}_\gamma(u,\eta) = \sum_{j=1}^{N} \|A(u,\eta)R_\gamma(u,\eta)f_j - f_j\|^2 \tag{2.21}$$

and notice that it is analytic in η and decreases with $\gamma > 0$ for each (u,η) in view of Lemma 2.4. Furthermore, setting $\nu_{\varepsilon_2} = \varepsilon_2^2/32$, we see that if $\mathfrak{F}_\gamma(u,\eta) \leq 2\nu_{\varepsilon_2}$, then for any $f \in B_V(1)$ we have

$$\begin{aligned} \|A(u,\eta)R_\gamma(u,\eta)f - f\| &\leq \min_{1\leq j\leq N} \big(\|AR_\gamma(f - f_j)\| + \|f - f_j\|\big) + \mathfrak{F}_\gamma(u,\eta)^{1/2} \\ &\leq 3\varepsilon_2/4, \end{aligned} \tag{2.22}$$

where we used the fact that the norm of the operator $G(G + \gamma I)^{-1}$ is bounded by 1. Suppose we have established the following lemma.

Lemma 2.6. *Under the hypotheses of Proposition 2.5, for any $\nu > 0$ we have*

$$\ell\big(\{\eta \in \mathcal{K} : \mathfrak{F}_\gamma(u,\eta) < \nu\}\big) \to 1 \quad \textit{uniformly in } u \in X \textit{ as } \gamma \to 0^+. \tag{2.23}$$

Applying this result with $\nu = \nu_{\varepsilon_2}$, we can find $\gamma = \gamma(\vec{\varepsilon}) > 0$ such that (2.14) holds for the set $\mathcal{K}^u_{\vec{\varepsilon}}$ defined by relation (2.13) with $\mathfrak{F}_{\vec{\varepsilon}} = \mathfrak{F}_{\gamma(\vec{\varepsilon})}$. It remains to note that inequality (2.19) follows immediately from (2.22). This completes the proof of the proposition. □

Proof of Lemma 2.6. By assumption, the image of $A(u,\eta)$ is dense for any $u \in X$ and $\eta \in \mathcal{K}^u$. It follows from (2.20) that

$$\mathfrak{F}_\gamma(u,\eta) \to 0 \quad \text{as } \gamma \to 0 \text{ for each } \eta \in \mathcal{K}^u.$$

Since $\ell(\mathcal{K}^u) = 1$ and the almost sure convergence implies convergence in probability, we see that the family of functions $\{m_\gamma : X \to [0,1]\}_{\gamma\in(0,1)}$ defined by

$$m_\gamma(u) = \ell\big(\{\eta \in \mathcal{K} : \mathfrak{F}_\gamma(u,\eta) < \nu\}\big) \tag{2.24}$$

converges to 1 as $\gamma \to 0^+$ for any $u \in X$ and $\nu > 0$. We need to prove that this convergence is uniform. Suppose we have established the following two properties:

Semicontinuity. For any $\gamma \in (0,1)$, the function $m_\gamma : X \to [0,1]$ is lower semicontinuous.

Monotonicity. For any $\gamma_1 \le \gamma_2$ and $u \in X$, we have $m_{\gamma_1}(u) \ge m_{\gamma_2}(u)$.

In this case, the required uniform convergence follows from Dini's theorem for a sequence of increasing functions, which remains true when the functions are lower semicontinuous; see [Dud02, Theorem 2.4.10].

Let us prove the semicontinuity. We denote by $\mu_u \in \mathcal{P}(\mathbb{R})$ the image of ℓ under the mapping $\mathfrak{F}_\gamma(u,\cdot)$. Thus, for any bounded continuous function $g : \mathbb{R} \to \mathbb{R}$, we have

$$\int_{\mathbb{R}} g(r)\mu_u(\mathrm{d}r) = \int_E g\big(\mathfrak{F}_\gamma(u,\eta)\big)\ell(\mathrm{d}\eta).$$

It follows from the Lebesgue theorem on dominated convergence that the function $u \mapsto \mu_u$ acting from X to the space $\mathcal{P}(\mathbb{R})$ endowed with the weak topology is continuous. By the portemanteau theorem (see [Dud02, Theorem 11.1.1]), for any open set $O \subset \mathbb{R}$ the function $u \to \mu_u(O)$ is lower semicontinuous. It remains to note that $m_\gamma(u) = \mu_u(O)$ with $O = (-\infty, \nu)$.

To prove the monotonicity, it suffices to note that, by Lemma 2.4, the function $\Delta_f(\gamma) := \|G(G+\gamma)^{-1}f - f\|^2$ decreases with γ, so that the same is true for $\mathfrak{F}_\gamma(u,\eta)$. The proof of the lemma is complete. □

2.3 Measure of a tubular neighbourhood of a nodal set

As before, let X be a compact metric space, let E be a separable Hilbert space, and let $\ell \in \mathcal{P}(E)$ be a probability measure with a compact support $\mathcal{K}$. We assume that ℓ is *decomposable* in the sense that there is an orthonormal basis $\{e_j\}$ in E such that ℓ can be represented as the direct of product its projections to the one-dimensional subspaces spanned by e_j. Let us fix a bounded convex open set $\mathcal{O} \supset \mathcal{K}$ and consider a continuous function $\mathfrak{F} : X \times \mathcal{O} \to \mathbb{R}$ such that $\mathfrak{F}(u,\cdot) : \mathcal{O} \to \mathbb{R}$ is analytic and not identically zero for any $u \in X$.

Lemma 2.7. *In addition to the above hypotheses, let us assume that the one-dimensional projections of ℓ to the spaces* $\operatorname{span}(e_j)$ *possess continuous densities*

with respect to the Lebesgue measure. Then there are positive numbers C and c such that, for any $u \in X$ and $r \in [0,1]$,

$$\ell\big(\{\eta \in \mathcal{O} : |\mathfrak{F}(u,\eta)| \le r\}\big) \le C\, r^c. \tag{2.25}$$

As an immediate consequence of this lemma, we obtain the following estimate for the measure of a tubular neighbourhood of the nodal set for an analytic function. Namely, for any $u \in X$, we denote

$$\mathcal{N}(u) = \{\eta \in \mathcal{O} : \mathfrak{F}(u,\eta) = 0\}.$$

Corollary 2.8. *Under the hypotheses of Lemma 2.7, there are numbers $C, c > 0$ such that, for any $u \in X$ and $r \in [0,1]$,*

$$\ell\big(\{\eta \in \mathcal{O} : \operatorname{dist}(\eta, \mathcal{N}(u)) \le r\}\big) \le C\, r^c. \tag{2.26}$$

Proof. We first recall that $\|(D_\eta \mathfrak{F})(u,\eta)\|_{\mathcal{L}(E)} \le C < \infty$ for $(u,\eta) \in X \times \mathcal{O}$, in view of our convention concerning analytic functions. By the convexity of $\mathcal{O}$, it follows that

$$\{\eta \in \mathcal{O} : \operatorname{dist}(\eta, \mathcal{N}(u)) \le r\} \subset \{\eta \in \mathcal{O} : |\mathfrak{F}(u,\eta)| \le C' r\},$$

where $C' > 0$ does not depend on $u \in X$ and $r \in [0,1]$. The required result is now implied by (2.25). □

Proof of Lemma 2.7. It suffices to establish (2.25) for $r \le r_0$, for a suitable $r_0 > 0$. Let us denote by $\mathcal{Z}(u,r)$ the set in the left-hand side of (2.25) and take any $u^0 \in X$ and $\eta^0 \in \mathcal{K}$. If $\mathfrak{F}(u^0,\eta^0) \ne 0$, then for sufficiently small balls $X \supset O_{u^0} \ni u^0$ and $E \supset O_{\eta^0} \ni \eta^0$, we have

$$|\mathfrak{F}(u,\eta)| > \sigma(u^0,\eta^0) > 0 \quad \text{for } u \in O_{u^0},\ \eta \in O_{\eta^0}. \tag{2.27}$$

Now assume that $\mathfrak{F}(u^0,\eta^0) = 0$. Since $\mathfrak{F}(u^0,\cdot)$ is analytic and not identically zero, we can find a vector $e \in E$ belonging to the span of finitely many first vectors of the basis $\{e_j\}$ such that the function $t \mapsto f_{u^0,\eta^0}(t) := \mathfrak{F}(u^0,\eta^0 + te)$ does not vanish identically. It follows that $f^{(m)}_{u^0,\eta^0}(0) \ne 0$ for some $m = m(u^0,\eta^0)$, so that we can find a number $\delta > 0$ and open balls $O_{u^0} \ni u^0$ and $O_{\eta^0} \ni \eta^0$ such that

$$|f^{(m)}_{u,\eta}(t)| \ge \gamma(u^0,\eta^0) > 0 \quad \text{for } u \in O_{u^0},\ \eta \in O_{\eta^0},\ |t| \le \delta.$$

This implies that (see Lemma 2 in [Bak86] or Lemma B.1 in [Eli02])

$$\operatorname{Leb}\big(\{t \in [-\delta,\delta] : |f(t)| \le r\}\big) \le C(u^0,\eta^0)\, r^{1/m} \tag{2.28}$$

for $u \in O_{u^0}$, $\eta \in O_{\eta^0}$, and $r \in [0,1]$. Since the projection of ℓ to any subspace spanned by finitely many vectors of the basis $\{e_j\}$ possesses a bounded density with respect to the Lebesgue measure, applying the Fubuni theorem, we conclude from (2.28) that

$$\ell\big(\mathcal{Z}(u,r) \cap O_{\eta^0}\big) \le C(u^0,\eta^0)\, r^{c(u^0,\eta^0)}. \tag{2.29}$$

Now let us choose a finite system of domains $O_{u^j} \times O_{\eta^j}$ which covers the compact set $X \times \mathcal{K}$ such that (2.27) or (2.29) hold with $u^0 = u^j$ and $\eta^0 = \eta^j$. Denoting by r_0 the minimum of all involved constants $\sigma(u^j, \eta^j)$, by C the maximum of the constants $C(u^j, \eta^j)$, and by c the minimum of all exponents $c(u^j, \eta^j)$, we see that (2.25) follows from (2.27) and (2.29). □

In Section 5.2, we shall need the following particular case of Lemma 2.7 when E is a finite-dimensional space and X is a singleton.

Corollary 2.9. *Let $\mathfrak{F} : \mathbb{R}^n \to \mathbb{R}$ be a non-zero analytic function and let $\xi_1, \dots, \xi_n$ be independent random variables whose joint law possesses a density (with respect to the Lebesgue measure) that is bounded by $\lambda\rho$, where $\lambda > 0$ is a number, and ρ is a continuous function with compact support. Then*

$$\mathbb{P}\{|\mathfrak{F}(\xi_1, \dots, \xi_n)| \le r\} \le C(\rho)\, r^c \lambda \quad \text{for any } r \in [0, 1], \tag{2.30}$$

where $C(\rho)$ and c are positive numbers depending on $\mathfrak{F}$ (but not on λ and the random variables $\xi_1, \dots, \xi_n$).

3 Proof of the main result

3.1 General scheme and reduction to a coupling

We wish to apply a sufficient condition for exponential mixing from [KS12, Section 3.1.1], stated below as Theorem 7.1. To this end, we shall construct a symmetric measurable function $F : X \times X \to \mathbb{R}_+$ satisfying (7.1) and prove that (7.4) holds for it. Replacing S by the mapping $\widetilde{S}(u, \eta) = S(u + \hat{u}, \eta + \hat{\eta}) - \hat{u}$, we may assume without loss of generality that $\hat{u} = 0$ and $\hat{\eta} = 0$.

Our construction will depend on four parameters $q, b \in (0, 1)$ and $R, d > 0$, three of which are fixed now. Namely, let $b = \frac{a+1}{2}$, where $a \in (0, 1)$ is the number in (1.2), let $R > 0$ be such that $X \subset B_H(R)$, and let $q \in (0, 1)$ satisfy the inequality

$$\|S(u_1, \zeta) - S(u_2, \zeta)\| \le q^{-1} \|u_1 - u_2\| \quad \text{for any } u_1, u_2 \in X,\ \zeta \in \mathcal{K}; \tag{3.1}$$

the existence of such a number q is implied by Hypothesis (H_1). We denote by $N = N(d) \ge 1$ the least integer satisfying the inequality $b^N R \le d/2$, define $\boldsymbol{X} = X \times X$, and introduce the sets

$$\boldsymbol{X}_\infty = \big\{(u, u') \in \boldsymbol{X} : u = u'\big\}, \tag{3.2}$$

$$\boldsymbol{X}_n = \big\{(u, u') \in \boldsymbol{X} : q^{n+1} d < \|u - u'\| \le q^n d\big\}, \tag{3.3}$$

$$\boldsymbol{X}_k = \big\{(u, u') \in \boldsymbol{X} : \|u - u'\| > d, Rb^{N+k+1} < \|u\| \vee \|u'\| \le Rb^{N+k}\big\}, \tag{3.4}$$

where $n \ge 0$ and $-N \le k \le -1$. It is straightforward to check that $\boldsymbol{X}$ is the union of the sets $\{\boldsymbol{X}_n, -N \le n \le \infty\}$. Recall that $P_k(u, \Gamma)$ stands for the transition function of the Markov process defined by (1.1). A key observation when proving (7.4) is the following result.

Proposition 3.1. *Under the hypotheses of Theorem 1.1, there are $\gamma \in (0,1]$ and $C > 0$ such that, for any $\nu \in (0,1)$, there is $d_0 \in (0,1)$ possessing the following property: for any $d \in (0,d_0)$ we can construct a number $p \in (0,1)$, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and measurable functions $V, V' : X \times X \times \Omega \to X$ (a coupling) such that the following assertions hold.*

(a) *For any $(u,u') \in \boldsymbol{X}$, the laws of $V(u,u';\cdot)$ and $V'(u,u';\cdot)$ coincide with $P_1(u,\cdot)$ and $P_1(u',\cdot)$, respectively. Moreover, $V(u,u;\cdot) = V'(u,u;\cdot)$ almost surely for any $(u,u) \in \boldsymbol{X}_\infty$.*

(b) *For any $(u,u') \in \boldsymbol{X}_n$, we have*

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \geq n+1\big\} \geq 1-\nu, \tag{3.5}$$

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \leq n-2\big\} \leq C\,\|u-u'\|^\gamma, \tag{3.6}$$

where $n \geq 0$ in (3.5) and $n \geq 1$ in (3.6).

(c) *For $-N \leq k \leq -1$ and $(u,u') \in \boldsymbol{X}_k$, we have*

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \geq k+1\big\} \geq p. \tag{3.7}$$

Taking this result for granted, let us complete the proof of the theorem. Let $\gamma \in (0,1]$ and $C > 0$ be the numbers constructed in Proposition 3.1. We fix $\nu > 0$ so small that[9]

$$q^{\gamma/2} + q^{-\gamma/2}\nu < 1, \tag{3.8}$$

$$q^{\gamma/2} + 3\nu < 1. \tag{3.9}$$

Let $d_0 > 0$ be the constant constructed in Proposition 3.1 for the above choice of ν and let $d \in (0,d_0)$ be a number that will be chosen below; once it is fixed, the integer N and the sets $\boldsymbol{X}_n$ with $-N \leq n \leq +\infty$ are uniquely determined. We define

$$F(u,u') = \begin{cases} 0 & \text{for } (u,u') \in \boldsymbol{X}_\infty, \\ (q^n d)^{\gamma/2} & \text{for } (u,u') \in \boldsymbol{X}_n, \\ M_k & \text{for } (u,u') \in \boldsymbol{X}_k, \end{cases} \tag{3.10}$$

where $n \geq 0$, $-N \leq k \leq -1$, and $M_k \geq 2d^{\gamma/2}$ is a decreasing sequence to be chosen below. It is straightforward to see that F satisfies (7.1). We shall prove that inequality (7.4) holds with some $\varkappa \in (0,1)$.

To this end, we first reduce the proof to the particular case in which μ_1, μ_2 are Dirac masses. Namely, suppose we have proved that

$$\mathcal{K}_F(\mathfrak{P}_1^*\delta_u, \mathfrak{P}_1^*\delta_{u'}) \leq \varkappa \mathcal{K}_F(\delta_u,\delta_{u'}) = \varkappa F(u,u') \quad \text{for any } (u,u') \in \boldsymbol{X}. \tag{3.11}$$

Let us take any measures $\mu_1, \mu_2 \in \mathcal{P}(X)$. For any $\theta > 0$, there are X-valued random variables ξ_1, ξ_2 such that

$$\mathcal{K}_F(\mu_1,\mu_2) \geq \mathbb{E}\,F(\xi_1,\xi_2) - \theta. \tag{3.12}$$

[9] The role of inequalities (3.8) and (3.9) will be clarified below.

Now let $(V(u,u'),V'(u,u'))$ be the random variables constructed in Proposition 3.1. They can be assumed to be independent of (ξ_1,ξ_2). In this case, the pair $(V(\xi_1,\xi_2),V'(\xi_1,\xi_2))$ is a coupling for $(\mathfrak{P}_1^*\mu_1,\mathfrak{P}_1^*\mu_2)$. Using again the independence and relations (3.11) and (3.12), we derive

$$\begin{aligned}\mathcal{K}_F(\mathfrak{P}_1^*\mu_1,\mathfrak{P}_1^*\mu_2) &\le \mathbb{E}F\big(V(\xi_1,\xi_2),V'(\xi_1,\xi_2)\big)\\ &\le \varkappa\,\mathbb{E}F(\xi_1,\xi_2)=\varkappa\,(\mathcal{K}_F(\mu_1,\mu_2)+\theta).\end{aligned}$$

Since $\theta>0$ was arbitrary, this proves (7.4).

To establish (3.11), notice that there is nothing to prove when $(u,u')\in\boldsymbol{X}_\infty$. For $(u,u')\notin\boldsymbol{X}_\infty$, we abbreviate $F(u,u')=:F$ and distinguish between four cases. For the reader's convenience, we outlined in Figure 1 the mutual dependence of parameters and the way they are chosen, assuming that b, R, q, C, and γ are already fixed (see (3.1) and Proposition 3.1).

Case 1: $(u,u')\in\boldsymbol{X}_n$ *with* $n\ge1$, *so* $F=(q^nd)^{\gamma/2}$. Since (V,V') is a coupling for $(\mathfrak{P}_1^*\delta_u,\mathfrak{P}_1^*\delta_{u'})$, it follows from of (3.5), (3.6), and (3.10) that

$$\begin{aligned}\mathcal{K}_F(\mathfrak{P}_1^*\delta_u,\mathfrak{P}_1^*\delta_{u'}) &\le \mathbb{E}F\big(V(u,u'),V'(u,u')\big)\\ &\le q^{\gamma/2}F\,\mathbb{P}(G_n^1)+M_{-N}\,\mathbb{P}(G_n^2)+q^{-\gamma/2}F\,\mathbb{P}(G_n^3)\\ &\le F\big(q^{\gamma/2}\,\mathbb{P}(G_n^1)+q^{-\gamma/2}\mathbb{P}(G_n^3)+M_{-N}(q^nd)^{-\gamma/2}\mathbb{P}(G_n^2)\big)\\ &=:F\varkappa_1,\end{aligned}$$

where G_n^1 and G_n^2 denote the events on the left-hand sides of (3.5) and (3.6), respectively, G_n^3 is the complement of $G_n^1\cup G_n^2$ corresponding to the event $\{(V,V')\in\boldsymbol{X}_n\cup\boldsymbol{X}_{n-1}\}$, and we used the fact that $\{M_k\}$ is a decreasing sequence. The required inequality (3.11) will be established if we prove that $\varkappa_1<1$, uniformly in $n\ge1$. To this end, notice that $\mathbb{P}(G_n^3)\le\nu$ in view of (3.5), so that

$$q^{\gamma/2}\,\mathbb{P}(G_n^1)+q^{-\gamma/2}\mathbb{P}(G_n^3)\le q^{\gamma/2}+q^{-\gamma/2}\nu. \tag{3.13}$$

Furthermore, it follows from (3.6) that

$$M_{-N}(q^nd)^{-\gamma/2}\mathbb{P}(G_n^2)\le CM_{-N}(q^nd)^{\gamma/2}\le CM_{-N}d^{\gamma/2}.$$

Combining this with (3.13) and (3.8), we see that $\varkappa_1<1$, provided that

$$CM_{-N}d^{\gamma/2}<1-q^{\gamma/2}-q^{-\gamma/2}\nu. \tag{3.14}$$

Let us note that N depends on the choice of d, so that the parameters M_{-N} and d are not independent. Our choice of M_k will ensure that $M_{-N}\le3d^{\gamma/2}$, so that (3.14) will be satisfied if

$$3Cd^{\gamma}<1-q^{\gamma/2}-q^{-\gamma/2}\nu. \tag{3.15}$$

In what follows, we fix $d\in(0,d_0)$ satisfying (3.15). Together with b, they determine $N\ge1$ as the least positive integer satisfying $b^NR\le d/2$.

Case 2: $(u,u')\in \boldsymbol{X}_0$, *so* $F=d^{\gamma/2}$. Let us set

$$M_k = 2d^{\gamma/2} + \varepsilon(B^{N-1} - B^{N+k}), \tag{3.16}$$

where $\varepsilon>0$ and $B>1$ will be chosen below. Arguing as above, using (3.5), and assuming that $\varepsilon \le B^{1-N}d^{\gamma/2}$ (so that $M_{-N}\le 3d^{\gamma/2}$), we derive

$$\begin{aligned}\mathcal{K}_F(\mathfrak{P}_1^*\delta_u,\mathfrak{P}_1^*\delta_{u'}) &\le q^{\gamma/2}F\,\mathbb{P}(G_0^1) + M_{-N}\big(1-\mathbb{P}(G_0^1)\big)\\ &\le F\big(q^{\gamma/2}+d^{-\gamma/2}M_{-N}\nu\big) \le F\big(q^{\gamma/2}+3\nu\big) =: F\varkappa_2.\end{aligned}$$

In view of (3.9), we have $\varkappa_2<1$.

Case 3: $(u,u')\in \boldsymbol{X}_{-1}$, *so* $F=M_{-1}=2d^{\gamma/2}$. It follows from (3.7) with $m=-1$ that

$$\begin{aligned}\mathcal{K}_F(\mathfrak{P}_1^*\delta_u,\mathfrak{P}_1^*\delta_{u'}) &\le pd^{\gamma/2} + (1-p)\big(2d^{\gamma/2}+\varepsilon(B^{N-1}-1)\big)\\ &\le F\big(1-\tfrac{p}{2}+\varepsilon(2d^{\gamma/2})^{-1}(1-p)B^{N-1}\big) =: F\varkappa_3.\end{aligned}$$

It is straightforward to check that $\varkappa_3\le 1-\frac{p}{4}<1$, provided that

$$\varepsilon \le p(1-p)^{-1}B^{-N}d^{\gamma/2}, \quad B\ge 2.$$

Case 4: $(u,u')\in \boldsymbol{X}_k$ *with* $-N\le k\le -2$, *so* $F=M_k$. Using (3.7) and (3.16), we derive

$$\begin{aligned}\mathcal{K}_F(\mathfrak{P}_1^*\delta_u,\mathfrak{P}_1^*\delta_{u'}) &\le p\big(2d^{\gamma/2}+\varepsilon(B^{N-1}-B^{N+k+1})\big)\\ &\quad + (1-p)\big(2d^{\gamma/2}+\varepsilon(B^{N-1}-1)\big)\\ &= 2d^{\gamma/2}+\varepsilon(B^{N-1}-pB^{N+k+1}-1+p).\end{aligned}$$

Let us set $B=2/p\ge 2$. Then the right-most term in the above inequality does not exceed $\varkappa_4 F$ with

$$\varkappa_4 = 1-\frac{1}{3}\varepsilon d^{-\gamma/2}.$$

Comparing the restrictions imposed on the parameters, we see that (3.11) holds with $\varkappa=\max\{\varkappa_i, 1\le i\le 4\}$, provided that

$$\varepsilon = B^{-N}d^{\gamma/2}\min\big\{p(1-p)^{-1}, B\big\}. \tag{3.17}$$

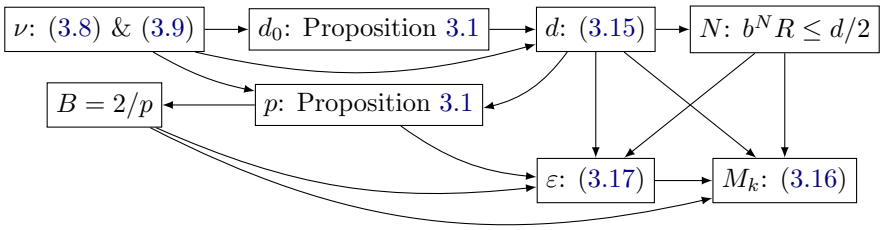

Figure 1: The parameters are chosen in the following order: $\nu, d_0, d, N, p, B, \varepsilon$

Relation (3.11) implies inequality (7.4), and the exponential mixing (1.7) follows. Thus, to complete the proof of Theorem 1.1, it remains to establish Proposition 3.1. To this end, we first prove two auxiliary results. The construction of (V, V') is given in Section 3.3.

3.2 Two auxiliary results

Recall that the number $a \in (0,1)$ is defined in Hypothesis (H_2), which is assumed to hold with $\hat{u} = 0$ and $\hat{\eta} = 0$, and that $b = \frac{1+a}{2}$.

Proposition 3.2. *Let ζ be a random variable with law ℓ. Then, for any $r > 0$ and any $u, u' \in X$ with $\|u\| \vee \|u'\| \geq r$, we have*

$$\mathbb{P}\big\{\|S(u,\zeta)\| \vee \|S(u',\zeta)\| \leq b(\|u\| \vee \|u'\|)\big\} > 0. \tag{3.18}$$

Proof. By Hypothesis (H_2) and the Lipschitz property of $S(u,\cdot) : \mathcal{K} \to H$, for any $u, u' \in X$ with $\|u\| \vee \|u'\| \geq r$ and $\eta \in \mathcal{K}$ we have

$$\begin{aligned} \|S(u,\zeta)\| &\leq \|S(u,0)\| + C\|\zeta\|_E \leq a\|u\| + C(r^{-1}\|u\| \vee \|u'\|)\|\zeta\| \\ &\leq \big(a + Cr^{-1}\|\zeta\|\big)\|u\| \vee \|u'\|. \end{aligned} \tag{3.19}$$

The right-most term of this inequality does not exceed $b(\|u\| \vee \|u'\|)$ provided that $\|\zeta\| \leq C^{-1}r(b-a)$, and a similar estimate holds for $\|S(u',\zeta)\|$. It follows that the probability on the left-hand side of (3.18) is minorised by $\mathbb{P}\{\|\zeta\| \leq C^{-1}r(b-a)\}$. This quantity is positive because $\hat{\eta} = 0$ is in the support of the law ℓ. □

Given a number $\delta > 0$, we set $D_\delta = \{(u,u') \in X \times H : \|u - u'\| \leq \delta\}$.

Proposition 3.3. *For any $\sigma, \theta \in (0,1)$ there are positive numbers C, β, and δ, a Borel-measurable mapping $\Phi : X \times H \times E \to E$, and a family of subsets $\{\mathcal{K}^{u,\sigma,\theta} \subset \mathcal{K}^u\}_{u \in X}$ such that $\Phi^{u,u'}(\eta) = 0$ if $\eta \notin \mathcal{K}^{u,\sigma,\theta}$ or $u' = u$, and we have the following inequalities, in which $\Psi^{u,u'}(\eta) = \eta + \Phi^{u,u'}(\eta)$:*

$$\ell(\mathcal{K}^{u,\sigma,\theta}) \geq 1 - \sigma, \tag{3.20}$$

$$\|\ell - \Psi_*^{u,u'}(\ell)\|_{\mathrm{var}} \leq C\,\|u - u'\|^\beta, \tag{3.21}$$

$$\|S(u,\eta) - S(u',\Psi^{u,u'}(\eta))\| \leq \theta\,\|u - u'\|, \tag{3.22}$$

where $(u,u') \in D_\delta$ and $\eta \in \mathcal{K}^{u,\sigma,\theta}$.

Proof. We first outline the main idea. We seek a vector $\zeta \in E$ depending on u, u', and η such that

$$\|S(u,\eta) - S(u',\eta+\zeta)\| \leq \theta\|u - u'\|, \tag{3.23}$$

provided that $\|u - u'\|$ is sufficiently small. Since S is a C^2-function of its arguments, we can write

$$S(u',\eta+\zeta) = S(u,\eta) + (D_u S)(u,\eta)(u'-u) + (D_\eta S)(u,\eta)\zeta + r(u,u',\eta,\zeta), \tag{3.24}$$

where r is a remainder term of order $\|u-u'\|^2+\|\zeta\|^2$. We see that a good choice of ζ would be defining it as a solution of the equation

$$(D_\eta S)(u,\eta)\zeta = -(D_u S)(u,\eta)(u'-u).$$

This equation is not necessarily soluble. However, by Hypothesis (H_3), the image of $(D_\eta S)(u,\eta)$ is dense in H. Therefore, by Proposition 2.5, for any $\varepsilon>0$, we can construct an approximate right inverse $R_\varepsilon(u,\eta): H\to E$ such that

$$\|(D_\eta S)(u,\eta)R_\varepsilon(u,\eta)f - f\| \le \varepsilon\|f\|_V \quad \text{for any } f\in V, \tag{3.25}$$

where V is the Banach space in (H_1). The mapping $(D_u S)(u,\eta)$ is continuous from H to V, so that in (3.25) we can take $f=-(D_u S)(u,\eta)(u'-u)$. We shall show that, for a sufficiently small $\varepsilon=\varepsilon(\theta)>0$, the mapping

$$\Phi^{u,u'}(\eta) = -R_\varepsilon(u,\eta)(D_u S)(u,\eta)(u'-u) \tag{3.26}$$

satisfies all required properties. Let us turn to an accurate proof, which is divided into four steps.

Step 1. Construction of Φ. Let us fix a small parameter $\varepsilon>0$ that will be chosen later. Let ν_ε, $\mathcal{K}^u_\varepsilon$, $\mathcal{D}_\varepsilon$, $\mathfrak{F}_\varepsilon$, and R_ε be the objects described in Proposition 2.5 with $\vec{\varepsilon}=(\varepsilon,\varepsilon)$, $E=F$, and $A(u,\eta)=(D_\eta S)(u,\eta)$. We now construct the sets $\mathcal{K}^{u,\sigma,\theta}\subset\mathcal{K}^u$ on which $\Phi^{u,u'}(\cdot)$ will be defined by (3.26). The main point is to choose them in such a way that inequality (2.2) is true with $\mathcal{K}_1=\mathcal{K}^{u,\sigma,\theta}$ (so that we can apply Theorem 2.1 to prove (3.21)).

Let us recall that $\mathcal{K}^u_\varepsilon$, $\mathcal{D}_\varepsilon$, and R_ε have the form

$$\mathcal{K}^u_\varepsilon = \{\eta\in\mathcal{K}: \mathfrak{F}_\varepsilon(u,\eta)\le\nu_\varepsilon\} \quad \text{for } u\in X, \tag{3.27}$$

$$\mathcal{D}_\varepsilon = \{(u,\eta)\in X\times\mathcal{K}: \mathfrak{F}_\varepsilon(u,\eta)\le 2\nu_\varepsilon\}, \tag{3.28}$$

$$R_\varepsilon(u,\eta) = \mathsf{P}_M R_\gamma(u,\eta) \quad \text{for } (u,\eta)\in\mathcal{D}_\varepsilon, \tag{3.29}$$

where the operator R_γ is defined in (2.18), and the number $\gamma=\gamma(\varepsilon)>0$ and the integer $M=M(\varepsilon)\ge 1$ are chosen appropriately. Let us denote by $\mathcal{D}'_\varepsilon$ the projection of $\mathcal{D}_\varepsilon$ to the space $X\times E^\perp_M$. In other words,

$$\mathcal{D}'_\varepsilon = \{(u,w)\in X\times E^\perp_M : \text{there is } v\in E_M \text{ such that } w+v\in\mathcal{K}^u_\varepsilon\}.$$

Since $\mathcal{D}_\varepsilon$ is compact, so is its projection $\mathcal{D}'_\varepsilon$. We shall need the following lemma established at the end of this subsection.

Lemma 3.4. *The set $\mathcal{D}'_\varepsilon$ can be represented as the disjoint union of finitely many sets $\mathcal{D}'_1,\dots,\mathcal{D}'_m$ such that the following property holds: for any integer $l\in[1,m]$ there is $\nu_l\in(\nu_\varepsilon,3\nu_\varepsilon/2)$ such that, for any $(u,w)\in\overline{\mathcal{D}'_l}$, the function $v\mapsto\mathfrak{F}_\varepsilon(u,w+v)-\nu_l$ is not identically zero.*

We now set

$$\mathcal{K}^{u,\sigma,\theta} = \bigcup_{l=1}^m \big\{\eta\in\mathcal{K}: \eta=v+w, (u,w)\in\mathcal{D}'_l, \mathfrak{F}_\varepsilon(u,w+v)\le\nu_l\big\} \tag{3.30}$$

and introduce a function $\Phi : X \times H \times E \to E$ as follows: $\Phi^{u,u'}(\eta)$ is defined by (3.26) if $\eta \in \mathcal{K}^{u,\sigma,\theta}$, and $\Phi^{u,u'}(\eta) = 0$ otherwise. This is a measurable mapping with range in E_M. We claim that (3.20)–(3.22) hold for an appropriate choice of ε.

Step 2. Proof of (3.20). Up to now, the parameter $\varepsilon > 0$ was arbitrary. Let us choose it so small that $\varepsilon \le \sigma$. In view of (2.14), the required inequality will be established if we prove that $\mathcal{K}^{u,\sigma,\theta} \supset \mathcal{K}^u_\varepsilon$. But this relation immediately follows from (3.27) and (3.30) since $\nu_l > \nu_\varepsilon$.

Step 3. Proof of (3.22). We first note that (3.28) and (3.30) imply the inclusion $\mathcal{K}^{u,\sigma,\theta} \subset \{\eta \in \mathcal{K} : (u,\eta) \in \mathcal{D}_\varepsilon\}$. In view of relations (3.24) and (3.25), in which $\zeta = \Phi^{u,u'}(\eta)$ and $f = -(D_u S)(u,\eta)(u'-u)$, for $u \in X$, $u' \in H$, and $\eta \in \mathcal{K}^{u,\sigma,\theta}$, we have

$$\begin{aligned}\|S(u,\eta) - S(u', \eta + \Phi^{u,u'}(\eta))\| \\ \le \varepsilon \, \|(D_u S)(u,\eta)(u'-u)\|_V + \|r(u,u',\eta,\Phi^{u,u'}(\eta))\|. \quad (3.31)\end{aligned}$$

Hypothesis (H_1) implies that

$$\|(D_u S)(u,\eta)(u'-u)\|_V \le C_1 \|u'-u\|, \quad (3.32)$$

where we denote by C_i positive numbers not depending on ε, u, u', and η. Furthermore, since S is a C^2 function whose second derivative is bounded on bounded subsets, for any $\rho > 0$ and $\|u\| + \|u'\| + \|\eta\|_E + \|\zeta\|_E \le \rho$, we have

$$\|r(u,u',\eta,\zeta)\| \le C_2\big(\|u-u'\|^2 + \|\zeta\|_E^2\big). \quad (3.33)$$

Recalling the definition of $\Phi^{u,u'}$ (see (3.26)) and using the inequality in (2.16), we derive

$$\|\Phi^{u,u'}(\eta)\| \le C_3(\varepsilon)\|u-u'\| \quad \text{for } u \in X,\ \eta \in \mathcal{K}^{u,\sigma,\theta},\ u' \in H. \quad (3.34)$$

Substituting (3.32)–(3.34) into (3.31), we obtain

$$\|S(u,\eta) - S(u', \eta + \Phi^{u,u'}(\eta))\| \le \big(C_4\varepsilon + C_5(\varepsilon)\|u-u'\|\big)\|u-u'\|.$$

For a given $\theta \in (0,1)$, we can choose first ε and then δ so that $(C_4\varepsilon + C_5(\varepsilon)\delta) \le \theta$. We thus obtain (3.22) for $(u,u') \in D_\delta$ and $\eta \in \mathcal{K}^{u,\sigma,\theta}$.

Step 4. Proof of (3.21). We shall use Theorem 2.1 with $\mathcal{K}_1 = \mathcal{K}^{u,\sigma,\theta}$. Let us fix $(u,u') \in D_\delta$. The mapping $\Phi^{u,u'}(\cdot) : \mathcal{K} \to E$ is measurable, and its image is contained in E_M. By Hypothesis (H_4), the measure ℓ can be written as the direct product of its projections to E_M and $E_M^\perp$. Inequality (3.21) will be established if we show that Properties (a) and (b) of Theorem 2.1 are true for $\Phi^{u,u'}(\cdot)$ with $\varkappa = C\|u-u'\|$, where C is an absolute constant.

By construction, $\Phi^{u,u'}$ vanishes outside $\mathcal{K}^{u,\sigma,\theta}$. In view of (3.34), the first inequality in (2.1) is satisfied. To prove[10] the second, let us fix any smooth

[10] Notice that we cannot apply the mean value theorem, since we do not know if $\mathcal{K}^{u,\sigma,\theta}$ is convex.

function $h(t)$ that is equal to 1 for $|t| \le 3\nu_\varepsilon/2$ and vanishes for $|t| \ge 2\nu_\varepsilon$. Then we can write

$$\Phi^{u,u'}(\eta) = -h\big(\mathfrak{F}_\varepsilon(u,\eta)\big) R_\varepsilon(u,\eta)(D_u S)(u,\eta)(u'-u) \quad \text{for } \eta \in \mathcal{K}^{u,\sigma,\theta}. \tag{3.35}$$

The right-hand side of (3.35) is obviously locally Lipschitz in $\eta \in E$, with a Lipschitz constant proportional to $\|u-u'\|$. It follows that $\Phi^{u,u'}$ satisfies the Lipschitz condition on $\mathcal{K}^{u,\sigma,\theta}$ with a constant of the form $C\|u-u'\|$. We have thus established (a).

To prove (b), we denote by $\mathcal{K}^l$, $l = 1, \dots, m$, the sets entering the right-hand side of (3.30). Let us note

$$\partial_w \mathcal{K}^{u,\sigma,\theta} = \bigcup_{l=1}^m \partial_w \mathcal{K}^l,$$

so that it suffices to establish (2.2) for each of the sets $\mathcal{K}^l$. Denoting by O_M a sufficiently large ball in E_M, it is easy to see that

$$\begin{aligned} \partial_w \mathcal{K}^l &= \{v \in \mathcal{K}^{u,\sigma,\theta}(w) : \mathfrak{F}_\varepsilon(u, w+v) = \nu_l\} \\ &\subset \{v \in O_M : \mathfrak{F}_\varepsilon(u, w+v) - \nu_l = 0\}. \end{aligned}$$

Since the analytic function $O_M \ni v \mapsto \mathfrak{F}_\varepsilon(u, w+v) - \nu_l$ does not vanish identically for $(u,w) \in \overline{\mathcal{D}}_l$, applying Corollary 2.8 to it and to the Lebesgue measure on E_M, we conclude that (2.2) is valid for $\partial_w \mathcal{K}^l$. This completes the proof of Proposition 3.3. □

Proof of Lemma 3.4. For any $(u,w) \in \mathcal{D}'_\varepsilon$, the function $v \mapsto \mathfrak{F}_\varepsilon(u, w+v)$ is analytic on the finite-dimensional space E_M. Therefore, by Sard's theorem, almost every real number is a regular value for it. Hence, we can find a number $\nu_{u,w} \in (\nu_\varepsilon, 3\nu_\varepsilon/2)$ such that each of the functions $v \mapsto \mathfrak{F}_\varepsilon(u, w+v) - \nu_{u,w}$ is not identically equal to zero. By continuity, $v \mapsto \mathfrak{F}_\varepsilon(u', w'+v) - \nu_{u,w}$ will not vanish identically, provided that (u', w') belongs to a non-degenerate closed ball $B_{u,w}$ centred at (u,w). Since the corresponding open balls form a covering of the compact set $\mathcal{D}'_\varepsilon$, we can find a finite sub-covering B_{u_j,w_j}, $j = 1, \dots, m$. We can now set $\mathcal{D}'_l = (\mathcal{D}'_\varepsilon \cap B_{u_l,w_l}) \setminus \big(\cup_{j=1}^{l-1} \mathcal{D}'_j\big)$, where the union is empty for $l = 1$. □

3.3 Proof of Proposition 3.1

Let us fix an arbitrary $\nu \in (0,1)$. To define the mappings V and V', we first consider the case in which either $u = u'$ or $\|u-u'\| > d$, where $d > 0$ is arbitrary for the moment. Let us denote by ζ a random variable such that $\mathcal{D}(\zeta) = \ell$. We set

$$V(u,u',\cdot) = S(u,\zeta), \quad V'(u,u',\cdot) = S(u',\zeta).$$

Then (V, V') satisfies property (a). Moreover, it follows from Proposition 3.2 that

$$\mathbb{P}\big\{\|V\| \vee \|V\| \le b(\|u\| \vee \|u'\|)\big\} \ge p, \tag{3.36}$$

where $p > 0$ is a number depending only on d (but not on the vectors $u, u' \in X$). Recalling (3.4), we see that (3.7) holds.

Let us turn to the case $\|u - u'\| \le d$, assuming that $d \le \delta$, where $\delta > 0$ is the number constructed in Proposition 3.3 with some parameters σ and θ to be chosen below. We denote by η an E-valued random variable with law ℓ and by $\mathfrak{M}_{u,u'} \in \mathcal{P}(E \times E)$ the law of the pair $(\eta, \Psi^{u,u'}(\eta))$, where $\Psi^{u,u'} : E \to E$ is the mapping constructed in Proposition 3.3. Let $\mathfrak{N}_{u,u'} \in \mathcal{P}(E \times E)$ be a maximal coupling for the pair $(\Psi_*^{u,u'}(\ell), \ell)$. That is, $\mathfrak{N}_{u,u'}$ is a probability measure on $E \times E$ with marginals $\Psi_*^{u,u'}(\ell)$ and ℓ such that

$$\mathfrak{N}_{u,u'}\big(\{(\zeta, \zeta') \in E \times E : \zeta \ne \zeta'\}\big) = \|\Psi_*^{u,u'}(\ell) - \ell\|_{\mathrm{var}}. \tag{3.37}$$

In view of Theorem 1.2.28 in [KS12], we can choose $\mathfrak{N}_{u,u'}$ to be a random probability measure on $E \times E$ with the underlying space D_δ. By construction, the projections of $\mathfrak{M}_{u,u'}$ and $\mathfrak{N}_{u,u'}$ to, respectively, the second and the first components coincide for any $(u, u') \in D_\delta$. Therefore, by Theorem 7.3 and the remark following it, there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and measurable functions $\zeta, \hat\zeta, \zeta' : D_\delta \times \Omega \to E$ such that, for any $(u, u') \in D_\delta$,

$$\mathcal{D}\big(\zeta(u, u', \cdot), \hat\zeta(u, u', \cdot)\big) = \mathfrak{M}_{u,u'}, \quad \mathcal{D}\big(\hat\zeta(u, u', \cdot), \zeta'(u, u', \cdot)\big) = \mathfrak{N}_{u,u'}. \tag{3.38}$$

The very definition of ζ and ζ' implies that their laws coincide with ℓ for any $(u, u') \in D_\delta$. Therefore, defining the functions

$$V(u, u', \omega) = S(u, \zeta(u, u', \omega)), \quad V'(u, u', \omega) = S(u, \zeta'(u, u', \omega)),$$

we see that their laws coincide with $P_1(u, \cdot)$ and $P_1(u', \cdot)$, respectively. Let us prove (3.5) and (3.6).

To this end, for any $(u, u') \in D_\delta$ we introduce the events[11]

$$\begin{aligned}
\Omega_1^{u,u'} &:= \{\zeta \in \mathcal{K}^{u,\sigma,\theta}\} \cap \{\hat\zeta(u, u') = \zeta'(u, u')\},\\
\Omega_2^{u,u'} &:= \{\zeta \in \mathcal{K}^{u,\sigma,\theta}\} \cap \{\hat\zeta(u, u') \ne \zeta'(u, u')\},\\
\Omega_3^{u,u'} &:= \{\zeta \notin \mathcal{K}^{u,\sigma,\theta}\}.
\end{aligned}$$

These events form a partition of the probability space Ω, and it follows from (3.20), (3.21), (3.37), and (3.38) that

$$\mathbb{P}(\Omega_1^{u,u'}) \ge 1 - \sigma - C_1\delta^\beta, \tag{3.39}$$

$$\mathbb{P}(\Omega_2^{u,u'}) \le C_1\|u - u'\|^\beta, \tag{3.40}$$

where C_1, β, and δ depend on σ and θ. Moreover, it follows from (3.22) and (3.1) that, on the event $\Omega_1^{u,u'}$, we have

$$\|V(u, u') - V'(u, u')\| = \|S(u, \zeta(u, u')) - S(u', \hat\zeta(u, u'))\| \le \theta\|u - u'\|, \tag{3.41}$$

[11] We shall often omit the argument ω to simplify formulas.

whereas on $\Omega_3^{u,u'}$,

$$\|V(u,u') - V'(u,u')\| = \|S(u,\zeta(u,u')) - S(u',\zeta(u,u'))\| \le q^{-1}\|u-u'\|. \quad (3.42)$$

We now specify our choice of the parameters. Let

$$\sigma = \frac{\nu}{2}, \quad \theta = q, \quad (3.43)$$

let $C =: C_1$, β and δ be the constants constructed in Proposition 3.3, and let $d_0 \in (0,\delta)$ be so small that $C_1 d_0^\delta \le \frac{\nu}{2}$. It follows from (3.41), (3.42), and (3.43) that, for $(u,u') \in \boldsymbol{X}_n$, we have

$$\{(V,V') \in \boldsymbol{X}_m \text{ for some } m \ge n+1\} \supset \Omega_1^{u,u'},$$
$$\{(V,V') \in \boldsymbol{X}_m \text{ for some } m \le n-2\} \subset \Omega_2^{u,u'},$$

where $n \ge 0$ for the first inclusion and $n \ge 1$ for the second. The required inequalities (3.5) and (3.6) (with $\gamma = \beta$ and $C = C_1$) follow now from (3.39) and (3.40), respectively. This completes the proof of Proposition 3.1.

4 Applications

4.1 Two-dimensional Navier–Stokes system

We consider the Navier–Stokes system on the torus $\mathbb{T}^2_a = \mathbb{R}^2/(2\pi a_1)\mathbb{Z} \oplus (2\pi a_2)\mathbb{Z}$, written in the form (1.9) after projecting to the space H of divergence-free square-integrable vector fields on $\mathbb{T}^2_a$ with zero mean value. The random forcing is assumed to be of the form (1.10), where $\{\eta_k\}$ is a sequence of i.i.d. random variables in $E := L^2(J,\mathcal{H})$ for some Hilbert space $\mathcal{H}$ embedded into H. Let $S(u_0,\zeta)$ be the time-1 resolving operator for (1.9). That is, S maps $H \times E$ to H and takes (u_0,η) to $u(1)$, where $u(t)$ stands for the solution of (1.9) satisfying the initial condition $u(0) = u_0$. Let us denote by $H^s(\mathbb{T}^2_a,\mathbb{R}^2)$ the usual Sobolev space of order $s \in \mathbb{Z}$ and by H^s its subspace of divergence-free vector functions. Standard dissipativity and regularisation results enable one to prove that, for any compact subset $\mathcal{K} \subset E$, there is a compact absorbing set $X \subset H$, bounded and closed in H^2, such that $S(X \times \mathcal{K}) \subset X$; see [KS01, Section 2] or [Shi15, Section 4.1]. Thus, if $\{\eta_k\}$ is a sequence of i.i.d. random variables in E whose law ℓ has a compact support $\mathcal{K} \subset E$, then (1.1) defines a homogeneous Markov process $(u_k,\mathbb{P}_u)$ in H for which X is a closed invariant subset. Our aim is to study the large-time asymptotics of $(u_k,\mathbb{P}_u)$.

To formulate the main result, we need the concept of a *saturating subspace*. Given a subspace $\mathcal{H} \subset H^3$, we define a non-decreasing sequence of closed subspaces $\mathcal{H}_k \subset H^3$, $k \ge 0$, by the following rule:

- $\mathcal{H}_0$ coincides with the closure of $\mathcal{H}$ in H^3;

- if $\mathcal{H}_k$ is already defined, then $\mathcal{H}_{k+1}$ is the closure in H^3 of the subspace $\mathcal{H}' \subset H^3$ such that, for any $\eta_1 \in \mathcal{H}'$, there is an integer $n \geq 1$ and vectors $\eta, \zeta_l \in \mathcal{H}_k$ and $\xi_l \in \mathcal{H}$ with $1 \leq l \leq n$ satisfying

$$\eta_1 = \eta + \sum_{l=1}^{n} Q(\zeta_l, \xi_l), \tag{4.1}$$

where $Q(\zeta, \xi) = \Pi(\langle \zeta, \nabla \rangle \xi + \langle \xi, \nabla \rangle \zeta)$.

Definition 4.1. A subspace $\mathcal{H} \subset H^3$ is said to be *saturating* if the union of $\{\mathcal{H}_k\}_{k \geq 0}$ is dense in H.[12]

Theorem 4.2. *Let $\mathcal{H} \subset H^3$ be a saturating subspace that is a Hilbert space with an orthonormal basis $\{\varphi_i\}_{i \in \mathcal{I}}$ satisfying*

$$\sum_{i \in \mathcal{I}} \|\varphi_i\|_1^2 < \infty, \tag{4.2}$$

and let η be a random process of the form (1.10) that satisfies the decomposability and observability hypotheses introduced in Example 1.2. Assume, in addition, that the support of the law ℓ of η_k contains the origin. Then, for any $\nu > 0$, the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu_\nu \in \mathcal{P}(X)$, which satisfies inequality (1.7).[13]

Proof. In view of Theorem 1.1, it suffices to check Hypotheses (H_1)–(H_4), in which H is defined by (1.8), and $E = L^2(J, \mathcal{H})$. The regularity condition is a consequence of some well-known properties of the resolving semigroup for the 2D Navier–Stokes system (see Section 6 in [BV92, Chapter 1] and [Kuk82]). Condition (H_2) with $\hat{\eta} = 0$ and $\hat{u} = 0$ follows immediately from the dissipativity of the homogeneous Navier–Stokes system. Hypothesis (H_4) is implied by the decomposability condition on the noise (see Example 1.2). Thus, we only need to check (H_3). The proof is divided into three steps.

Step 1. Reduction. We recall that $X \subset H^2$ stands for a closed bounded set that is invariant and absorbing under the dynamics of Eq. (1.9) restricted to the integer lattice. Given $u \in X$ and $\eta \in E$, let $\tilde{u} \in L^2(J, H^3) \cap W^{1,2}(J, H^1)$ be the solution of Eq. (1.9) with the initial condition u at time $t = 0$ and let $R^{\tilde{u}}(t, s) : H \to H$ (with $0 \leq s \leq t \leq 1$) be the two-parameter process solving the linearised problem

$$\dot{v} + \nu L v + Q(\tilde{u}(t))v = 0, \quad v(s) = v_0, \tag{4.3}$$

[12] We emphasise that condition of saturation depends on the parameters a_1 and a_2 of the torus and the quadratic term Q, but not on the viscosity ν.

[13] Even though the convergence to the stationary measure is formulated in Theorem 4.2 for initial functions supported by X, the exponential convergence to μ_ν holds for any initial measure with a finite first moment. Our choice of presentation is dictated only by the simplicity of formulation and the fact that generalisation to arbitrary measures can be carried out by a well-known simple argument.

where $Q(\tilde{u})v = Q(\tilde{u}, v)$ and $v_0 \in H$. In other words, $R^{\tilde{u}}(t,s)$ takes v_0 to $v(t)$, where $v : [s,1] \to H$ is the solution of (4.3). Denote by $A(u,\eta)$ the operator $D_\eta S(u,\eta)$. This is the resolving operator for the non-homogeneous equation (4.3) (that is, zero in the right-hand side should be replaced by ζ) with $v_0 = 0$. It can be written as

$$A(u,\eta) : E \to H, \quad \zeta \mapsto \int_0^1 R^{\tilde{u}}(1,s)\zeta(s)\,\mathrm{d}s.$$

Denoting by $\mathfrak{L}(u,\eta)$ its image, we need to prove that $\mathfrak{L}(u,\eta)$ is dense in H for any fixed $u \in X$ and ℓ-a.e. $\eta \in \mathcal{K}$. In view of Theorem 2.5 in [Zab08, Part IV], the density of $\mathfrak{L}(u,\eta)$ is equivalent to the triviality of the kernel for the operator [14]

$$G^{\tilde{u}} := \int_0^1 R^{\tilde{u}}(1,t)\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(1,t)^* \mathrm{d}t, \tag{4.4}$$

where $R^{\tilde{u}}(1,t)^* : H \to H$ is the adjoint of $R^{\tilde{u}}(1,t)$, and $\mathsf{P}_{\mathcal{H}} : H \to H$ stands for the orthogonal projection to the closure of $\mathcal{H}$ in H.

Step 2. Description of $R^{\tilde{u}}(1,t)^*$. Together with (4.3), let us consider the dual problem

$$\dot{w} - \nu L w - Q^*(\tilde{u}(t))w = 0, \quad w(1) = w_0, \tag{4.5}$$

where $Q^*(\tilde{u})$ is the (formal) L^2-adjoint of $Q(\tilde{u})$ given by

$$Q^*(\tilde{u})w = -\Pi\big(\langle \tilde{u}, \nabla\rangle w + (\nabla \otimes w)\tilde{u}\big),$$

where $\nabla \otimes w$ is a 2×2 matrix with the elements $\partial_i w_j$. Problem (4.5) is a linear backward parabolic equation, and for any $w_0 \in H$, it possesses a unique solution $w \in L^1(J, H^1) \cap W^{1,2}(J, H^{-1})$. It is well known, and can be proved easily, that w can be written as

$$w(t) = R^{\tilde{u}}(1,t)^* w_0. \tag{4.6}$$

Step 3. Inductive argument. We wish to prove that $\mathrm{Ker}(G^{\tilde{u}}) = \{0\}$ for any $u \in X$ and ℓ-a.e. $\eta \in E$. Since the union of $\mathcal{H}_k$ is dense in H, it suffices to show that any element of $\mathrm{Ker}(G^{\tilde{u}})$ is orthogonal to $\mathcal{H}_k$ for all $k \geq 0$.

Let us fix any realisation $\eta \in \mathcal{K}$ that is observable with respect to $\{\varphi_i\}_{i\in\mathcal{I}}$, and suppose that $w_0 \in \mathrm{Ker}(G^{\tilde{u}})$. In view of (4.4), we have

$$(G^{\tilde{u}} w_0, w_0) = \int_0^1 \|\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(1,t)^* w_0\|^2 \mathrm{d}t = 0,$$

whence we see that $\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(1,t)^* w_0 = 0$ for any $t \in J$. Thus,

$$(\zeta, R^{\tilde{u}}(1,t)^* w_0) = 0 \quad \text{for } t \in J, \tag{4.7}$$

where $\zeta \in \mathcal{H}_0$ is arbitrary. Taking $t = 1$, we see that w_0 must be orthogonal to $\mathcal{H}_0$. Suppose we have proved that relation (4.7) holds for all $\zeta \in \mathcal{H}_k$ with

[14] The operator $G^{\tilde{u}}$ is called the *Gramian* of the linear control problem associated with (4.3). It is a central object for studying controllability properties and enables one to formulate them in terms of observability of solutions for the dual problem; see Chapter 1 in [Cor07].

some integer $k \geq 0$ (so that taking again $t = 1$ we see that w_0 must be orthogonal to $\mathcal{H}_k$). By induction, the verification of (H_3) will be completed once we show that (4.7) holds for all $\zeta \in \mathcal{H}_{k+1}$.

For any given $\zeta \in \mathcal{H}_k$, differentiating (4.7) in time and using (4.5), we derive[15]

$$\big(\nu L\zeta + Q(\tilde{u}(t))\zeta, w(t)\big) = 0 \quad \text{for } t \in J, \tag{4.8}$$

where w is given by (4.6). If we set

$$y(t) = \tilde{u}(t) - \int_0^t \eta(s)\mathrm{d}s = \tilde{u}(t) - \sum_{i\in\mathcal{I}} \varphi_i(x) \int_0^t \eta^i(s)\,\mathrm{d}s, \tag{4.9}$$

where $\eta^i(t) = (\eta(t), \varphi_i)_{\mathcal{H}}$, then (4.8) can be rewritten as

$$\big(\nu L\zeta + Q(y(t))\zeta, w(t)\big) + \sum_{i\in\mathcal{I}} \big(Q(\varphi_i)\zeta, w(t)\big) \int_0^t \eta^i(s)\mathrm{d}s = 0.$$

Differentiating this relation in time and setting

$$\begin{aligned} a_i(t) &= \big(Q(\varphi_i)\zeta, w(t)\big), \\ b(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\big(\nu L\zeta + Q(y(t))\zeta, w(t)\big) + \sum_{i\in\mathcal{I}} \big(Q(\varphi_i)\zeta, \dot{w}(t)\big) \int_0^t \eta^i(s)\mathrm{d}s, \end{aligned}$$

we derive

$$b(t) + \sum_{i\in\mathcal{I}} a_i(t)\eta^i(t) = 0 \quad \text{for all } t \in J. \tag{4.10}$$

The function b is continuous, and $\{a_i, i \in \mathcal{I}\}$ are differentiable. Moreover, condition (4.2) implies that (1.11) holds. By observability of ℓ, it follows that $a_i \equiv 0$ for all $i \in \mathcal{I}$, whence, recalling the definition of $\mathcal{H}_{k+1}$, we conclude that (4.7) holds for any $\zeta \in \mathcal{H}_{k+1}$. This completes the proof of Theorem 4.2. □

4.2 Complex Ginzburg–Landau equation

We now turn to Eq. (1.13). To simplify the presentation, we confine ourselves to the most interesting case $p = 5$. In addition, we shall consider only finite-dimensional random forces; the extension to infinite-dimensional forces is simpler than in the case of the Navier–Stokes system.

We thus fix a vector $a = (a_1, a_2, a_3)$ with positive coordinates and denote by $\mathbb{Z}^3_a$ the integer lattice generated by the vectors $a_k\iota_k$, where $(\iota_1, \iota_2, \iota_3)$ is the standard basis in $\mathbb{R}^3$. Let us set $\mathbb{T}^3_a = \mathbb{R}^3/2\pi\mathbb{Z}^3_a$ and consider the equation

$$\partial_t u - (\nu + i)\Delta u + \gamma u + ic|u|^4 u = \eta(t, x), \quad x \in \mathbb{T}^3_a. \tag{4.11}$$

[15] Here and henceforth, the time derivatives of the functions $\tilde{u}$ and w should be understood in appropriate functional spaces. We leave it to the reader to check that the regularity of $\tilde{u}$ and w is sufficient to differentiate all the inner products.

Here $u = u(t,x)$ is an unknown complex-valued function, ν, γ, and c are given positive numbers, and η is a random process of the form (1.10), where $\{\eta_k\}$ is a sequence of i.i.d. random variables in $L^2(J,H^2)$, and we set $H^s := H^s(\mathbb{T}^3_a,\mathbb{C})$ for $s\in\mathbb{Z}$. Equation (4.11) is supplemented with the initial condition

$$u(0) = u_0 \in H^1. \tag{4.12}$$

We shall consider $L^2 := L^2(\mathbb{T}^3_a,\mathbb{C})$ as a real Hilbert space with the inner product

$$(u,v) = \operatorname{Re}\int_{\mathbb{T}^3_a} u\bar v\,\mathrm{d}x$$

and endow the Sobolev spaces H^s with the associated norms and inner products. As in the case of the Navier–Stokes system, if the law of η_k has a compact support $\mathcal{K}$ in $L^2(J,H^2)$, then discrete-time Markov process $(u_k,\mathbb{P}_u)$ associated with (1.1) (in which $S : H^1\times L^2(J,H^2)\to H^1$ is the time-1 resolving operator for (4.11)) possesses a compact absorbing set $X\subset H^1$, that is closed and bounded in H^2.

Let us introduce a concept of a *saturating subspace* for (4.11). Given a finite-dimensional subspace $\mathcal{H}\subset H^2$ invariant under complex conjugation (that is, $\bar\zeta\in\mathcal{H}$ for all $\zeta\in\mathcal{H}$), we define a sequence of closed subspaces $\mathcal{H}_k\subset H^2$, $k\ge 0$, by the following rule:

- $\mathcal{H}_0$ coincides with $\mathcal{H}$;
- if $\mathcal{H}_k$ is already defined, then $\mathcal{H}_{k+1}$ is the vector span of $\mathcal{H}_k$ and the products $\zeta\xi$ with $\zeta\in\mathcal{H}_k$ and $\xi\in\mathcal{H}$.

It is straightforward to check that $\{\mathcal{H}_k\}$ is a non-decreasing sequence of finite-dimensional subspaces in H^2 that are invariant under complex conjugation.

Definition 4.3. The subspace $\mathcal{H}\subset H^2$ is said to be *saturating* if the union of $\{\mathcal{H}_k\}_{k\ge0}$ is dense in L^2.

Let us write $B(u) = ic|u|^4u$ for the nonlinear term in (4.11) and note that it is a real-analytic function in the space H^2, and its derivative $Q(u) : H^2\to H^2$ acts essentially as a multiplication operator:

$$Q(u)v = ic\bigl(3|u|^4v + 2|u|^2u^2\bar v\bigr).$$

Theorem 4.4. *Let $\mathcal{H}\subset H^2$ be a finite-dimensional subspace that is invariant under complex conjugation, contains the function identically equal to* 1*, and is saturating. Let $\{\varphi_i\}_{i\in\mathcal{I}}$ be an orthonormal basis in $\mathcal{H}$ and let η be a random process of the form (1.10) in which $\{\eta_k\}$ is an i.i.d. sequence in $E = L^2(J,\mathcal{H})$ such that the following conditions are fulfilled.*

Decomposability. *The law $\ell\in\mathcal{P}(E)$ of η_k has a compact support $\mathcal{K}\subset E$ containing the origin and satisfies the decomposability hypotheses of Example 1.2.*

Observability. *There is $T \in (0,1)$ such that the law ℓ' of the restriction of the random variables η_k to the interval $J' = [0,T]$ is Lipschitz-observabile with respect to $\{\varphi_i\}$.*

Then, for any $\nu > 0$, the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu_\nu \in \mathcal{P}(X)$, which satisfies inequality (1.7).

Remark 4.5. Theorem 4.4 remains valid in a slightly more general setting: it suffices to require that $\mathcal{H}$ should contain a subspace $\mathcal{H}'$ satisfying the hypotheses imposed on $\mathcal{H}$ in the theorem. The proof of this observation is straightforward, and we skip the details.

Proof of Theorem 4.4. As in the case of the Navier–Stokes system, we only need to check that, for any $u \in X$ and ℓ-a.e. realisation of $\eta \in E$, the image of the derivative $(DS)(u,\eta) : H^1 \to H^1$ is dense. To this end, we 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 (4.11) issued from u, and consider the linearised problem

$$\dot{v} + Lv + Q(\tilde{u})v = g, \quad v(0) = v_0, \tag{4.13}$$

where $L = -(\nu + i)\Delta + \gamma$, $g \in E$, and $v_0 \in H^1$. Let us denote by $v(t; v_0, g)$ the solution of (4.13). We need to prove that, for ℓ-a.e. $\eta \in E$, the vector space $\{v(1; 0, g), g \in E\}$ is dense in H^1. In view of parabolic regularisation, the resolving operator for (4.13) with $g \equiv 0$ is continuous from L^2 to H^1 on an arbitrary interval $J' = [T, 1]$ with $0 < T < 1$. Since the restriction at $t = 1$ of the space of solutions $v(t)$ for the homogeneous equation on J' with $v(T) \in L^2$ is dense in H^1 (see Proposition 7.4), the required property will be established if we prove that the vector space $\{v(T; 0, g), g \in E\}$ is dense in L^2. To this end, we repeat the scheme used in the case of the Navier–Stokes system. The difference is that the operator Q is no longer linear in $\tilde{u}$, which makes the argument slightly more involved.

Let us denote by $R^{\tilde{u}}(t,s) : L^2 \to L^2$ (with $0 \le s \le t \le T$) the two-parameter process solving the linearised problem (4.13) with $g \equiv 0$ and let

$$G^{\tilde{u}} := \int_0^T R^{\tilde{u}}(T,t) \mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* \mathrm{d}t, \tag{4.14}$$

where $R^{\tilde{u}}(T,t)^* : L^2 \to L^2$ is the adjoint of $R^{\tilde{u}}(T,t)$, and $\mathsf{P}_{\mathcal{H}} : L^2 \to L^2$ stands for the orthogonal projection to $\mathcal{H}$. We wish to prove that $\mathrm{Ker}(G^{\tilde{u}}) = \{0\}$ for any $u \in X$ and ℓ-a.e. $\eta \in E$. Since the union of $\mathcal{H}_k$ is dense in H, it suffices to show that any element of $\mathrm{Ker}(G^{\tilde{u}})$ is orthogonal to $\mathcal{H}_k$ for all $k \ge 0$.

To this end, we fix $u \in X$ and take any $w_0 \in \mathrm{Ker}(G^{\tilde{u}})$. In view of (4.14), we have $\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* w_0 = 0$ for any $t \in J'$, so that

$$(\zeta, R^{\tilde{u}}(T,t)^* w_0) = 0 \quad \text{for } t \in J', \tag{4.15}$$

where $\zeta \in \mathcal{H}_0$ is an arbitrary vector. Assuming that relation (4.15) is true for any $\zeta \in \mathcal{H}_k$, we now prove its validity for $\zeta \in \mathcal{H}_{k+1}$. Once this is established, we can complete the proof by taking $t = T$.

Differentiating (4.15) in time and using an analogue of relation (4.6) for (4.13), we derive

$$\big(L\zeta + Q(\tilde{u}(t))\zeta, w(t)\big) = 0 \quad \text{for } t \in J', \tag{4.16}$$

where $w(t)$ is given by (4.6). Let us write

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

where $\eta^i(t) = (\eta(t), \varphi_i)_{\mathcal{H}}$. Differentiating (4.16) in time and using (4.11), we obtain

$$\big(L\zeta + Q(\tilde{u})\zeta, \dot{w}\big) - \big(B_2(\tilde{u}; \zeta, L\tilde{u} + B(\tilde{u})), w\big) + \sum_{i\in\mathcal{I}} \big(B_2(\tilde{u}; \zeta, \varphi_i), w\big)\eta^i(t) = 0, \tag{4.17}$$

where $B_k(u;\cdot)$ stands for the k^{th} derivative of $B(u)$ (so that $Q = B_1$, and $B_k = 0$ for $k \geq 6$). We thus obtain relation (4.10), in which $J = J'$,

$$\begin{aligned} a_i(t) &= \big(B_2(\tilde{u}(t); \zeta, \varphi_i), w(t)\big), \\ b(t) &= \big(L\zeta + Q(\tilde{u}(t))\zeta, \dot{w}(t)\big) - \big(B_2(\tilde{u}(t); \zeta, L\tilde{u}(t) + B(\tilde{u}(t))), w(t)\big). \end{aligned}$$

These are Lipschitz-continuous and continuous functions, respectively, and the observability of ℓ' implies that (cf. (4.16))

$$\big(B_2(\tilde{u}(t); \zeta, \varphi_i), w(t)\big) = 0 \quad \text{for } i \in \mathcal{I},\ t \in J'.$$

Applying exactly the same argument three more times, we see that

$$\big(B_5(\zeta, \varphi_i, \varphi_j, \varphi_m, \varphi_n), w(t)\big) = 0 \quad \text{for } i, j, m, n \in \mathcal{I},\ t \in J', \tag{4.18}$$

where we used the fact that the fifth derivative of $B(u)$ does not depend on u. We see that $w(t)$ must be orthogonal to the vector space $\mathcal{V}$ spanned by $\{(B_5(\zeta, \varphi_i, \varphi_j, \varphi_m, \varphi_n)\}$. Now note that

$$B_5(\zeta, \varphi, 1, 1, 1) = 12ic\,(3\zeta\varphi + \bar{\zeta}\bar{\varphi} + 3\bar{\zeta}\varphi + 3\zeta\bar{\varphi}). \tag{4.19}$$

Since both $\mathcal{H}$ and $\mathcal{H}_k$ are invariant under complex conjugation, it follows from relation (4.19) that $\mathcal{V}$ must contain all the products $\zeta\xi$ with $\zeta \in \mathcal{H}_k$ and $\xi \in \mathcal{H}$. Thus, we have $\mathcal{H}_{k+1} \subset \mathcal{V}$, which completes the proof. □

5 Observable processes with decomposable laws

5.1 Observable functions

Let $\mathcal{H}$ be a (finite or infinite-dimensional) Hilbert space with an inner product $(\cdot,\cdot)$, let $\{\varphi_i\}_{i\in\mathcal{I}}$ be an orthonormal basis in $\mathcal{H}$, and let $\eta : J \to \mathcal{H}$ be a function in $L^2(J, \mathcal{H})$. Recall that the concept of an observable function was introduced in Definition 1.3. The following two results provide sufficient conditions for a function to be observable.

Proposition 5.1. *Let $\eta : J \to \mathcal{H}$ be a bounded measurable function such that, for any $i \in \mathcal{I}$, the projection (η, φ_i) has left and right limits at any point of J and is discontinuous on a countable dense set $\mathbb{D}_i$ and continuous on $J \setminus \mathbb{D}_i$. Suppose, in addition, that $\mathbb{D}_i \cap \mathbb{D}_j = \varnothing$ for $i \neq j$. Then η is observable with respect to $\{\varphi_i\}$.*

Proof. We confine ourselves to the infinite-dimensional case. Let a_i be Lipschitz-continuous functions and b a continuous function satisfying (1.11) and (1.12). The boundedness of η implies that the series $\sum_i a_i(t)(\eta(t), \varphi_i)$ converges uniformly in $t \in J$. It follows that its sum $\zeta(t)$ has left and right limits at any point of J. Let us fix any $i \in \mathcal{I}$ and $s \in \mathbb{D}_i$. Since $s \notin \mathbb{D}_j$ for $j \neq i$ and $(\eta(t), \varphi_j)$ is continuous on $J \setminus \mathbb{D}_j$, we see that

$$\Delta\zeta(s) := \zeta(s^+) - \zeta(s^-) = a_i(s)\big(\eta(s^+) - \eta(s^-), \varphi_i\big). \tag{5.1}$$

Since b is continuous on J, it follows from (1.12) that the expression in (5.1) must vanish. By the hypotheses of the proposition, the function (η, φ_i) has a jump at s, whence we conclude that $a_i(s) = 0$. Since $s \in \mathbb{D}_i$ is arbitrary and $\mathbb{D}_i$ is dense, the function a_i is identically zero for any $i \in \mathcal{I}$, so that b also vanishes. □

We now consider the case of a finite set $\mathcal{I}$ and write $\mathcal{I} = \{1, \dots, N\}$. Let us fix an orthonormal basis $\{\varphi_i\}$ and set $\eta = (\eta^1, \dots, \eta^N)$.

Proposition 5.2. *Let the functions η^i have left and right limits at any point of J, and denote by $\Delta\eta^i(t)$ the jump of η^i at t. Suppose that, for any $s \in J$ and $\varepsilon \in (0, 1]$, there are $t_1^\varepsilon, \dots, t_N^\varepsilon \in [s - \varepsilon, s + \varepsilon]$ such that the $N \times N$ matrix $R_\varepsilon(s)$ with the entries $r_l^i := \Delta\eta^i(t_l^\varepsilon)$ is invertible and satisfies the inequality*

$$\big\|R_\varepsilon^{-1}(s)\big\| \leq C\varepsilon^{-\theta}, \tag{5.2}$$

where $C > 0$ and $\theta \in (0, 1)$ do not depend on ε. Then η is observable.

Proof. Let $a_1, \dots, a_N$ be real-valued Lipschitz-continuous functions and let b a continuous function such that (1.12) holds. We fix an arbitrary $s \in J$ and, given $\varepsilon > 0$, find points $t_1^\varepsilon, \dots, t_N^\varepsilon$ satisfying the properties mentioned in the statement. It follows from (1.12) that

$$\sum_{i=1}^N r_l^i a_i(t_l^\varepsilon) = 0, \quad l = 1, \dots, N.$$

Setting $a = (a_1, \dots, a_N)$ and using the Lipschitz-continuity of a, we derive the relation

$$R_\varepsilon(s)a(s) = d_\varepsilon(s),$$

where $|d_\varepsilon(s)| \leq C_1\varepsilon$. Applying the matrix $R_\varepsilon^{-1}(s)$ and recalling (5.2), we obtain $|a(s)| \leq C_2\varepsilon^{1-\theta}$, where C_2 does not depend on ε. Since s and ε were arbitrary, we see that $a \equiv 0$ and, hence, $b \equiv 0$. □

5.2 Processes satisfying the hypotheses of Section 1.2

We now construct some examples of stochastic processes on $[0,1]$ that possess the decomposability and observability properties. The decomposability will be a simple consequence of the explicit form of the processes, whereas observability will follow from a finer analysis based on Propositions 5.1 and 5.2. In all the examples, we deal with the observability on the interval $J=[0,1]$; however, exactly the same argument shows that the result remains true on the interval $[0,T]$ with any $T\in(0,1)$.

Jump process

Let $\mathcal{H}$ be a Hilbert space with an orthonormal basis $\{\varphi_i\}_{i\in\mathcal{I}}$. Suppose that, for any $i\in\mathcal{I}$, we are given an orthonormal basis $\{\psi_l^i\}_{l\geq 1}$ in $L^2(J)$ such that the following properties hold:

- **(a)** the function ψ_l^i is continuous outside a finite set $\mathbb{D}_l^i$ and has different left and right limits at the points of $\mathbb{D}_l^i$;
- **(b)** the union $\cup_l \mathbb{D}_l^i := \mathbb{D}_i$ is dense in J for each $i\in\mathcal{I}$;
- **(c)** the intersection $\mathbb{D}_{i_1}\cap\mathbb{D}_{i_2}$ is empty for $i_1\neq i_2$.

Let us define a stochastic process by the relation

$$\eta(t)=\sum_{i\in\mathcal{I}}\eta^i(t)\varphi_i,\quad \eta^i(t)=\sum_{l=1}^{\infty}b_{il}\xi_l^i\psi_l^i(t), \tag{5.3}$$

where ξ_l^i are independent random variables whose laws possess Lipschitz-continuous densities supported by $[-1,1]$ and $b_{il}\in\mathbb{R}$ are non-zero numbers such that

$$\sum_{i\in\mathcal{I}}\sum_{l=1}^{\infty}\bigl(b_{il}^2+|b_{il}|\,\|\psi_l^i\|_\infty\bigr)<\infty. \tag{5.4}$$

It is obvious that the law ℓ of η is a decomposable measure [16] on $L^2(J,\mathcal{H})$. We claim that ℓ is observable with respect to $\{\varphi_i\}$. Indeed, it follows from (5.4) that the series in (5.3) converges uniformly in t and ω, so that the trajectories of η^i have left and right limits at any point and are continuous outside $\mathbb{D}_i$. If we prove that they are a.s. discontinuous at any point of $\mathbb{D}_i$, then Proposition 5.1 will immediately imply the observability of η with respect to $\{\varphi_i\}$.

Let us fix any point $s\in\mathbb{D}_i$ and calculate the jump of η^i at s. It follows from (5.3) that

$$\Delta\eta^i(s)=\sum_{l=1}^{\infty}b_{il}\xi_l^i\bigl(\psi_l^i(s^+)-\psi_l^i(s^-)\bigr).$$

[16] Note that the index j in (1.3) is now replaced by the pair (l,i), and vectors e_j are the products $\psi_l^i(t)\varphi_i$.

This series converges absolutely, and at least one of the terms $\psi_l^i(s^+) - \psi_l^i(s^-)$ is non-zero. Since the random variables ξ_l^i are independent, and laws have densities, we conclude that so does the law of $\Delta\eta^i(s)$. Therefore, $\Delta\eta^i(s) \neq 0$ with probability 1.

Haar process with exponentially decaying coefficients

Let us recall the definition of the Haar system $\{h_0, h_{jl}\}$. We set

$$h_0(t) = \begin{cases} 1 & \text{for } 0 \le t < 1, \\ 0 & \text{for } t < 0 \text{ or } t \ge 1, \end{cases} \tag{5.5}$$

$$h_{jl}(t) = \begin{cases} 0 & \text{for } t < l2^{-j} \text{ or } t \ge (l+1)2^{-j}, \\ 1 & \text{for } l2^{-j} \le t < \big(l + \frac{1}{2}\big)2^{-j}, \\ -1 & \text{for } \big(l + \frac{1}{2}\big)2^{-j} \le t < (l+1)2^{-j}, \end{cases} \tag{5.6}$$

where $j \ge 1$ and $l \ge 0$ are integers. It is well known that the restrictions of the functions $\{h_0, h_{jl}, j \ge 0, 0 \le l \le 2^j - 1\}$ to the interval $J = [0,1]$ form an orthogonal basis in $L^2(J)$; see Section 22 in [Lam96]. This implies that the system of functions $\{h_0(\cdot + k), h_{jl}(\cdot)\}$, where $j \ge 1$ and $k, l \ge 0$, form an orthogonal basis of $L^2(\mathbb{R}_+)$.

Let us assume that $\mathcal{H}$ is an N-dimensional Euclidean space. We fix an orthonormal basis $\{\varphi_i, 1 \le i \le N\}$ and consider the following process in $\mathcal{H}$:

$$\eta(t) = \sum_{i=1}^{N} b_i \Big(\xi_0^i h_0(t) + \sum_{j=1}^{\infty} \sum_{l=0}^{2^j - 1} c_j \xi_{jl}^i h_{jl}(t) \Big) \varphi_i, \quad t \in J, \tag{5.7}$$

where $b_i \in \mathbb{R}$ are non-zero numbers, $c_j = A^{-j}$ with some $A > 1$, and ξ_{jl}^i are i.i.d. scalar random variables such that $|\xi_{jl}^i| \le 1$ almost surely, and their law possesses a continous density with respect to the Lebesgue measure. For any $j \ge 1$, any point of J belongs to the support of at most two functions h_{jl}, so that the series in (5.7) absolutely converges, uniformly with respect to $t \in J$. Since the functions $\{h_0\varphi_i, h_{jl}\varphi_i\}$ form an orthonormal basis in $L^2(J, \mathcal{H})$, the law $\ell \in \mathcal{P}(L^2(J, \mathcal{H}))$ of η is decomposable, and any random variable with law ℓ has the form (5.7). Let us prove that, if $A > 1$ is sufficiently close to 1, then a.e. trajectory of η is observable.

We shall apply Proposition 5.2. Let us fix a point $s \in J$ and a number $\varepsilon > 0$ and consider the points

$$\tau_l^j := \big(l + \tfrac{1}{2}\big)2^{-j}, \quad 0 \le l \le 2^j - 1. \tag{5.8}$$

It is clear that if an integer j of order $C_1 \log \frac{1}{\varepsilon}$ is fixed, then there are N points

$$\tau_{l_1}^j =: \tau_1, \quad \dots, \quad \tau_{l_N}^j =: \tau_N, \quad l_r = l_0 + r,$$

that belong to the ε-neighbourhood of s. Consider the matrix $R_\varepsilon(s)$ corresponding to the points $\tau_1, \dots, \tau_N$. Its entries are $r_m^i = \Delta\eta^i(\tau_m) = \Delta\eta^i(\tau_{l_m}^j)$,

where

$$\Delta\eta^i(\tau_l^j) = b_i\biggl(-2c_j\xi_{jl}^i + \sum_{r=1}^{\infty} c_{j+r}\bigl(\xi^i_{j+r,(2l+1)2^{r-1}} - \xi^i_{j+r,(2l+1)2^{r-1}-1}\bigr)\biggr)$$
$$= b_i A^{-j}\zeta_l^i(j). \tag{5.9}$$

Here $\{\zeta_l^i(j)\}$ are i.i.d. random variables whose law has a continuous density with respect to the Lebesgue measure. It follows that the determinant of the matrix $R_\varepsilon(s)$ can be written as

$$\det R_\varepsilon(s) = b_1\cdots b_N A^{-Nj}\Sigma(j), \tag{5.10}$$

where $\Sigma(j)$ is the determinant of the matrix $(\zeta^i_{l_m}(j), 1\le i,m\le N)$. Notice that the entries of this matrix are i.i.d. random variables whose law does not depend on j. Since the determinant is a non-zero polynomial of the matrix entries, Corollary 2.9 applies and gives the inequality

$$\mathbb{P}\{|\Sigma(j)|\le r\}\le Cr^c \quad \text{for any } r\in[0,1], \tag{5.11}$$

where C and c are positive numbers not depending on j. Taking $r = e^{-\delta j}$, where $\delta>0$ is sufficiently small and will be chosen below, we derive

$$\mathbb{P}\{|\Sigma(j)|\le e^{-\delta j}\}\le Ce^{-c\delta j} \quad \text{for any } j\ge 0.$$

By the Borel–Cantelli lemma, there is an almost surely finite random integer $j_0\ge 1$ such that

$$|\Sigma(j)| > e^{-\delta j} \quad \text{for } j\ge j_0.$$

Combining this with (5.10), we derive

$$|\det R_\varepsilon(s)|\ge C_2 e^{-j(N\log A+\delta)}, \tag{5.12}$$

where $j\ge j_0$ is arbitrary. For $j\sim C_1\log\frac{1}{\varepsilon}$ (such a choice is possible for a.e. ω and sufficiently small ε), we obtain

$$|\det R_\varepsilon(s)|\ge C_2\varepsilon^\theta,$$

where $\theta = C_1(N\log A+\delta)$. Since the entries of $R_\varepsilon(s)$ are a.s. bounded by a universal number, we conclude that (5.2) holds. Taking $\delta = \log A$ and $A>1$ sufficiently close to 1, we see that $\theta<1$, and Proposition 5.2 implies the required result.

Haar process with polynomially decaying coefficients

We consider again process (5.7), in which the numbers c_j go to zero at a polynomial rate:

$$c_j = Cj^{-q} \quad \text{for all } j\ge 1, \tag{5.13}$$

where $C>0$ and $q>1$ are some numbers. For the same reasons as in the previous case, the series in (5.7) absolutely converges, uniformly in $t\in J$, and

the law of η is a decomposable measure on $L^2(J,\mathcal{H})$. We claim that, for any $T \in (0,1]$, a.e. trajectory of η is observable on $J' = [0,T]$. To prove it, we apply the same argument as before. The first line of (5.9) remains true, and the jumps take the form $\Delta\eta^i(\tau_l^j) = Cb_i j^{1-q}\zeta_l^i(j)$, where

$$\zeta_l^i(j) = -2j^{-1}\xi_{jl}^i + j^{-1}\sum_{r=1}^{\infty}(1+rj^{-1})^{-q}\bigl(\xi^i_{j+r,(2l+1)2^{r-1}} - \xi^i_{j+r,(2l+1)2^{r-1}-1}\bigr).$$

It follows that the determinant of $R_\varepsilon(s) = (r_m^i)$ can be written as (cf. (5.10))

$$\det R_\varepsilon(s) = C^N b_1 \cdots b_N j^{-N(q-1)}\Sigma(j),$$

where $\Sigma(j)$ is the determinant of the matrix $(\zeta^i_{l_m}(j))$. A simple calculation shows that $\{\zeta^i_{l_m}(j)\}$ are independent random variables that are a.s. bounded by a universal number, and their laws do not depend on i and m. Suppose we have shown that the laws of $(\zeta_l^i(j))$ possess densities ρ_j (with respect to the Lebesgue measure) satisfying the inequality

$$\rho_j(r) \le j\rho(r) \quad \text{for all } j \ge 1,\ r \in \mathbb{R}, \tag{5.14}$$

where $\rho : \mathbb{R} \to \mathbb{R}$ is a continuous function with compact support. Applying Corollary 2.9, we conclude that inequality (5.11) remains valid with C replaced by Cj^N. Hence, by the same argument as before, we obtain (cf. (5.12))

$$|\det R_\varepsilon(s)| \ge C_2 e^{-\delta j - N(q-1)\log j}. \tag{5.15}$$

The proof can now be completed as in the case of exponentially decaying coefficients.

It remains to prove (5.14). To this end, we write $\zeta_l^i(j) = -2j^{-1}\xi_{jl}^i + \eta_l^i(j)$, where $\eta_l^i(j)$ is a bounded random variable independent of ξ_{jl}^i. By the hypotheses, the law of ξ_{jl}^i has a bounded continuous density not depending on i, j, and l. Let us denote by M its maximum. Then the density $\tilde\rho_j$ of the law for $-2j^{-1}\xi_{jl}^i$ is bounded by $Mj/2$. Since ρ_j is a convolution with $\tilde\rho_j$, it is bounded by the same constant. On the other hand, we know that $\zeta_l^i(j)$ is almost surely bounded by a universal number. It follows that (5.14) holds for some continuous function ρ with compact support.

6 Examples of saturating subspaces

In this section, we discuss some algebraic conditions that ensure the saturating property of a given subspace. This type of conditions are rather well known in the control theory and Malliavin calculus for PDEs; see [EM01, AS06, HM06, MP06].

Navier–Stokes system

Let us endow the space $\mathbb{R}^2$ with the scalar product

$$\langle x, y\rangle_a = \sum_{i=1}^{2} a_i^{-1} x_i\, y_i$$

and the corresponding norm $|x|_a = \sqrt{\langle x, x\rangle_a}$. In the case $a = (1,1)$, we write $\langle\cdot,\cdot\rangle$ and $|\cdot|$. Let $\mathbb{Z}^2_*$ be the set of non-zero integer vectors $l = (l_1, l_2) \in \mathbb{Z}^2$. For $l \in \mathbb{Z}^2_*$, we define the functions

$$e^a_l(x) = \begin{cases} c^a_l(x) & \text{if } l_1 > 0 \text{ or } l_1 = 0,\ l_2 > 0, \\ s^a_l(x) & \text{if } l_1 < 0 \text{ or } l_1 = 0,\ l_2 < 0 \end{cases}$$

on $\mathbb{T}^2_a$, where

$$c^a_l(x) = l^{\perp_a} \cos\langle l, x\rangle_a, \quad s^a_l(x) = l^{\perp_a} \sin\langle l, x\rangle_a, \quad l^{\perp_a} = (-a_2^{-1} l_2, a_1^{-1} l_1).$$

For any subset $\mathcal{I} \subset \mathbb{Z}^2$, let $\mathbb{Z}^2_{\mathcal{I}}$ be the set of all vectors in $\mathbb{Z}^2$ that can be represented as finite linear combinations of elements of $\mathcal{I}$ with integer coefficients. We shall say that $\mathcal{I}$ is a *generator* if $\mathbb{Z}^2_{\mathcal{I}} = \mathbb{Z}^2$.

Let $\mathcal{I} \subset \mathbb{Z}^2_*$ be a finite symmetric set (i.e., $-\mathcal{I} = \mathcal{I}$) and let

$$\mathcal{H}(\mathcal{I}) = \operatorname{span}\{e^a_l : l \in \mathcal{I}\}. \tag{6.1}$$

We denote by $\mathcal{H}_k(\mathcal{I}) \subset H^3$ the spaces defined in Section 4.1 with $\mathcal{H} = \mathcal{H}(\mathcal{I})$.

Proposition 6.1. *Let $\mathcal{I} \subset \mathbb{Z}^2_*$ be a finite symmetric set. Then the space $\mathcal{H}(\mathcal{I})$ is saturating if and only if $\mathcal{I}$ is a generator and contains at least two non-parallel elements m and n such that $|m^{\perp_a}| \neq |n^{\perp_a}|$.*

This result implies that the space $\mathcal{H}(\mathcal{I})$ is saturating when

$$\begin{aligned} &\mathcal{I} = \{(1,0), (-1,0), (1,1), (-1,-1)\} && \text{for } a = (1,1), \\ &\mathcal{I} = \{(1,0), (-1,0), (0,1), (0,-1)\} && \text{for } a = (\lambda, 1) \text{ with } \lambda \neq 1. \end{aligned}$$

Proof of Proposition 6.1. Step 1. We first establish some relations for Leray's projection Π and the quadratic function Q defined in Section 4.1. For any $l \in \mathbb{R}^2_*$, let us denote by $P(l) : \mathbb{R}^2 \to \mathbb{R}^2$ the orthogonal projection onto $\operatorname{span}(l)$ with respect to the scalar product $\langle\cdot,\cdot\rangle$, so that $P(l)b = |l|^{-2}\langle b, l\rangle l$. We claim that

$$\Pi\big(b \cos\langle l, x\rangle_a\big) = P(l^{\perp_a}) b \cos\langle l, x\rangle_a, \tag{6.2}$$

$$\Pi\big(b \sin\langle l, x\rangle_a\big) = P(l^{\perp_a}) b \sin\langle l, x\rangle_a, \tag{6.3}$$

$$\begin{aligned} 2Q(c^a_l, c^a_r) = {} & \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} - \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \sin\langle l - r, x\rangle_a \\ & - \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} + \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \sin\langle l + r, x\rangle_a, \end{aligned} \tag{6.4}$$

$$\begin{aligned} 2Q(c^a_l, s^a_r) = {} & \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} + \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \cos\langle l + r, x\rangle_a \\ & + \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} - \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \cos\langle l - r, x\rangle_a, \end{aligned} \tag{6.5}$$

$$\begin{aligned} 2Q(s^a_l, s^a_r) = {} & \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} - \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \sin\langle l - r, x\rangle_a \\ & + \Pi\big(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} + \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\big) \sin\langle l + r, x\rangle_a, \end{aligned} \tag{6.6}$$

We confine ourselves to the proof of (6.2) and (6.4), since the other relations can be established in a similar way. Let us recall that Leray's projection of a

function $v \in L^2(\mathbb{T}^2_a, \mathbb{R}^2)$ can be written as $\Pi v = v - \nabla(\Delta^{-1} \operatorname{div} v)$, where Δ^{-1} is the inverse of the Laplacian with range in the space of functions with zero mean value. Relation (6.2) is now implied by the following simple formulas:

$$\operatorname{div}\bigl(b \cos\langle l, x\rangle_a\bigr) = -\langle b, l\rangle_a \sin\langle l, x\rangle_a\,, \quad \Delta^{-1} \sin\langle l, x\rangle_a = -\bigl|l^{\perp_a}\bigr|^{-2} \sin\langle l, x\rangle_a\,.$$

To prove (6.4), we write

$$\langle c^a_l, \nabla\rangle c^a_r = \langle l^{\perp_a}, r\rangle_a r^{\perp_a} \cos\langle l, x\rangle_a \sin\langle r, x\rangle_a.$$

Interchanging the roles of l and r, using elementary formulas for the product of trigonometric functions, and recalling the definition of Q, we arrive at (6.4).

Step 2. Let us prove that if $l, r \in \mathbb{Z}^2_*$ are non-parallel vectors such that $|l^{\perp_a}| \neq |r^{\perp_a}|$, $c^a_r, s^a_r \in \mathcal{H}(\mathcal{I})$, and $c^a_l, s^a_l \in \mathcal{H}_k(\mathcal{I})$ for some $k \geq 0$, then the functions $c^a_{l+r}, c^a_{l-r}, s^a_{l+r}, s^a_{l-r}$ belong to $\mathcal{H}_{k+1}(\mathcal{I})$. Indeed, a direct verification shows that

$$P((l-r)^{\perp_a})\bigl(\langle l^{\perp_a}, r\rangle_a r^{\perp_a} - \langle r^{\perp_a}, l\rangle_a l^{\perp_a}\bigr) = |(l-r)^{\perp_a}|^{-2} \langle l^{\perp_a}, r\rangle_a \bigl(|l^{\perp_a}|^2 - |r^{\perp_a}|^2\bigr)(l-r)^{\perp_a}. \quad (6.7)$$

Let $C_a(l, r)$ be the coefficient in front of $(l-r)^{\perp_a}$. Combining (6.4), (6.6), (6.3), and (6.7), for any $\alpha, \beta \in \mathbb{R}$, we derive

$$Q(c^a_l, 2\alpha c^a_r) + Q(s^a_l, 2\beta s^a_r) = C_a(l, r)(\alpha+\beta)(l-r)^{\perp_a} \sin\langle l-r, x\rangle_a - C_a(l, -r)(\alpha-\beta)(l+r)^{\perp_a} \sin\langle l+r, x\rangle_a. \quad (6.8)$$

Since l and r are non-parallel vectors such that $|l^{\perp_a}| \neq |r^{\perp_a}|$, the numbers $C_a(l, r)$ and $C_a(l, -r)$ are non-zero. This and relation (6.8) readily imply that $s^a_{l+r}, s^a_{l-r} \in \mathcal{H}_{k+1}(\mathcal{I})$. A similar argument using (6.2), (6.5), and (6.7) shows that $c^a_{l+r}, c^a_{l-r} \in \mathcal{H}_{k+1}(\mathcal{I})$.

Step 3. We now take any symmetric set $\mathcal{I} \subset \mathbb{Z}^2_*$ and prove the necessity of the hypotheses of Proposition 6.1. If $\mathcal{I}$ is not a generator, then there is a vector $l \in \mathbb{Z}^2 \setminus \mathbb{Z}^2_{\mathcal{I}}$. It follows from (6.2)–(6.6) that the function c^a_l is orthogonal to $\mathcal{H}_\infty(\mathcal{I})$ in H, so $\mathcal{H}(\mathcal{I})$ is not saturating. Furthermore, if any pair of elements $(m, n) \in \mathcal{I}$ either are parallel or satisfy the relation $|m^{\perp_a}| \neq |n^{\perp_a}|$, then (6.4)–(6.7) imply that $\mathcal{H}_k(\mathcal{I}) = \mathcal{H}(\mathcal{I})$ for all $k \geq 0$, so that $\mathcal{H}(\mathcal{I})$ is not saturating.

To prove the sufficiency, let us assume that $\mathcal{I}$ is a generator containing at least two non-parallel elements m and n such that $|m^{\perp_a}| \neq |n^{\perp_a}|$. We define a sequence of symmetric sets $\mathcal{I}_k \subset \mathbb{Z}^2_*$ by the following rule: $\mathcal{I}_0 = \mathcal{I}$, and for $k \geq 1$,

$$\mathcal{I}_k = \mathcal{I}_{k-1} \cup \{l + r : l \in \mathcal{I}_{k-1}, r \in \mathcal{I}, |l^{\perp_a}| \neq |r^{\perp_a}|, l \text{ and } r \text{ are not parallel}\}.$$

Let $\mathcal{I}_\infty$ be the union of the sets $\mathcal{I}_k$ with $k \geq 0$ and of the point $(0, 0)$. The proof of Proposition 1 in [AS06] (see also Proposition 4.4 in [HM06]) implies that $\mathcal{I}_\infty$ coincides with $\mathbb{Z}^2_{\mathcal{I}} = \mathbb{Z}^2$. It follows that the union of $\mathcal{H}(\mathcal{I}_k)$ with $k \geq 0$ contains all divergence-free trigonometric polynomials, whence we see that $\mathcal{H}(\mathcal{I})$ is saturating. □

Ginzburg–Landau equation

Let us fix a vector $a = (a_1, a_2, a_3)$ with positive components and define the functions

$$c_l^a(x) = \cos\langle l, x\rangle_a, \quad s_l^a(x) = \sin\langle l, x\rangle_a \quad \text{for } l \in \mathbb{Z}^3_*,$$

where $\langle l, x\rangle_a = \sum_i a_i^{-1} l_i x_i$. We now set $e_0^a \equiv 1$ and

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

Given a finite symmetric set $\mathcal{I} \subset \mathbb{Z}^3$ containing the origin, introduce the space (cf. (6.1))

$$\mathcal{H}(\mathcal{I}) = \operatorname{span}\{e_l^a, ie_l^a : l \in \mathcal{I}\}.$$

We denote by $\mathcal{H}_k(\mathcal{I})$ the sets $\mathcal{H}_k$ defined in Section 4.2 for $\mathcal{H} = \mathcal{H}(\mathcal{I})$.

Proposition 6.2. *The subspace $\mathcal{H}(\mathcal{I})$ is saturating if and only if $\mathcal{I}$ is a generator. In particular, the set $\mathcal{I} = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}$ gives rise to the* 14*-dimensional saturating subspace $\mathcal{H}(\mathcal{I})$.*

Proof. The necessity of the condition is straightforward (cf. the case of the Navier–Stokes system), so that we confine ourselves to the proof of sufficiency. It is enough to show that if $l, r \in \mathbb{Z}^3_*$ are two vectors such that $c_r^a, s_r^a \in \mathcal{H}(\mathcal{I})$ and $c_l^a, s_l^a \in \mathcal{H}_k(\mathcal{I})$ for some $k \geq 0$, then the functions $c_{l+r}^a, c_{l-r}^a, s_{l+r}^a, s_{l-r}^a$ belong to $\mathcal{H}_{k+1}(\mathcal{I})$.

To see this, let us note that

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

It follows that $c_{l+r}^a, c_{l-r}^a \in \mathcal{H}_{k+1}(\mathcal{I})$. A similar argument applies to the functions s_{l+r}^a and s_{l-r}^a. This completes the proof of the proposition. □

7 Appendix

7.1 Sufficient condition for exponential 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. Let us consider a bounded Borel-measurable symmetric function $F : X \times X \to \mathbb{R}_+$ such that

$$F(u_1, u_2) \geq c\, d_X(u_1, u_2)^\beta \quad \text{for any } u_1, u_2 \in X, \tag{7.1}$$

where $\beta \leq 1$ and c are positive numbers not depending on u_1 and u_2. Recall that the *Kantorovich functional* associated with F is defined by

$$\mathcal{K}_F(\mu_1, \mu_2) = \inf_{\xi_1, \xi_2} \mathbb{E}\, F(\xi_1, \xi_2), \tag{7.2}$$

where the infimum is taken over all X-valued random variables ξ_1, ξ_2 such that $\mathcal{D}(\xi_1) = \mu_1$ and $\mathcal{D}(\xi_2) = \mu_2$. It follows from inequality (7.1) and the definition of the dual-Lipschitz distance that

$$\|\mu_1 - \mu_2\|_L^* \le C\mathcal{K}_F(\mu_1, \mu_2) \quad \text{for any } \mu_1, \mu_2 \in \mathcal{P}(X), \tag{7.3}$$

where $C = c^{-1}\operatorname{diam}(X)^{1-\beta}$. A proof of the following theorem[17] can be found in Section 3.1.1 of [KS12].

Theorem 7.1. *Suppose there is a number $\varkappa \in (0,1)$ and a bounded Borel-measurable function $F : X \times X \to \mathbb{R}_+$ satisfying (7.1) such that*

$$\mathcal{K}_F(\mathfrak{P}_1^*\mu_1, \mathfrak{P}_1^*\mu_2) \le \varkappa\mathcal{K}_F(\mu_1, \mu_2) \quad \textit{for any } \mu_1, \mu_2 \in \mathcal{P}(X). \tag{7.4}$$

Then the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu \in \mathcal{P}(X)$, which satisfies inequality (1.7) with some positive numbers C and γ.

7.2 Estimate for the total variation distance

Lemma 7.2. *Let U and V be two random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with range in a Polish space X and let $\Theta : \Omega \to \Omega$ be a measurable mapping such that*

$$U(\Theta(\omega)) = V(\omega) \quad \textit{for } \omega \in G, \tag{7.5}$$

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

$$\|\mathcal{D}(U) - \mathcal{D}(V)\|_{\mathrm{var}} \le 2\,\mathbb{P}(G^c) + \|\mathbb{P} - \Theta_*(\mathbb{P})\|_{\mathrm{var}}. \tag{7.6}$$

Proof. 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 (7.5) 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 (7.6). □

7.3 Measurable version of gluing lemma

Let X be a Polish space and let $(W, \mathcal{W})$ be a measurable space. Recall that a family $\{\mu_w, w \in W\} \subset \mathcal{P}(X)$ is called a *random probability measure* (RPM) on X with the underlying space $(W, \mathcal{W})$ if for any $\Gamma \in \mathcal{B}(X)$, the mapping $w \mapsto \mu_w(\Gamma)$ is measurable from W to $\mathbb{R}$. The following result can be obtained by repeating the argument of the proof of Lemma 7.6 in [Vil03], using a measurable version of the disintegration theorem (see [Mal16, AY17]).

[17] In [KS12], the result is proved with $\beta = 1$. However, using (7.3), it is straightforward to check that the proof remains valid for any $\beta \in (0,1)$.

Theorem 7.3. *Let X, Y, and Z be Polish spaces and let $\{\mathfrak{M}_w\}$ and $\{\mathfrak{N}_w\}$ be RPM measures on $X \times Y$ and $Y \times Z$, respectively, with the same underlying space $(W, \mathcal{W})$ such that the marginals of $\mathfrak{M}_w$ and $\mathfrak{N}_w$ on Y are the same for any $w \in W$. Then there is a RPM $\{\mathfrak{L}_w\}$ on $X \times Y \times Z$ such that, for any $w \in W$, the marginals of $\mathfrak{L}_w$ on $X \times Y$ and $Y \times Z$ coincide with $\mathfrak{M}_w$ and $\mathfrak{N}_w$, respectively.*

This result can be reformulated as follows: under the hypotheses of the theorem, there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and measurable functions $\xi_w(\omega)$, $\eta_w(\omega)$, and $\zeta_w(\omega)$ defined on $W \times \Omega$ and taking values in X, Y, and Z, respectively, such that

$$\mathcal{D}(\xi_w, \eta_w) = \mathfrak{M}_w, \quad \mathcal{D}(\eta_w, \zeta_w) = \mathfrak{N}_w \quad \text{for any } w \in W.$$

7.4 Density of solutions for linear parabolic PDEs

Retaining the notation of Section 2.2, let us consider the following linear parabolic PDE on $J = [0, 1]$:

$$\dot{v} - \mu \Delta v + b(t, x)v + c(t, x)\bar{v} = 0, \quad x \in \mathbb{T}^3_a, \tag{7.7}$$

where $b, c \in H^1(J \times \mathbb{T}^3_a, \mathbb{C})$ are bounded functions, and $\mu \in \mathbb{C}$ is a parameter with a positive real part. We denote by $\mathcal{U}$ the space of all $v \in L^2(J, H^2) \cap W^{1,2}(J, L^2)$ satisfying (7.7) and, for $s \in J$, set $\mathcal{U}_s = \{u(s, \cdot) : u \in \mathcal{U}\} \subset H^1$. The following result is a variation of the well-known property of approximate controllability of (7.7) by a starting control.

Proposition 7.4. *For any $s \in J$, the subspace $\mathcal{U}_s$ is dense in H^1.*

Proof. The claim is trivial for $s = 0$, so without loss of generality, let us assume that $s = 1$. We only outline the proof, which is based on a classical argument.

For $\tau \in [0, 1)$, let us denote by $\{R(t, \tau), \tau \le t \le 1\}$ the resolving process for Eq. (7.7) with an initial condition specified at $t = \tau$. This means that, for any $v_0 \in H^1$, the function $v(t) = R(t, \tau)v_0$ is the unique solution of (7.7) in the space $L^2([\tau, 1], H^2) \cap W^{1,2}([\tau, 1], L^2)$ such that $v(\tau) = v_0$. We denote by $R(1, t)^*$ the adjoint of $R(1, t)$ with respect to the inner product in H^1. Setting $L = I - \Delta$, it is straightforward to check that, for any $w_1 \in H^1$, the function $w : t \mapsto R(1, t)^* w_1$ is the unique solution in $L^2(J, H^2) \cap W^{1,2}(J, L^2)$ for the dual problem

$$\dot{w} + \bar{\mu} \Delta w - L^{-1}(\bar{b} L w) - L^{-1}(c L \bar{w}) = 0, \quad w(1) = w_1, \tag{7.8}$$

and the inner product $(v(t), w(t))_1$ does not depend on t.

Suppose now that $\mathcal{U}_1$ is not dense in H^1. Then there is a non-zero $w_1 \in H^1$ such that

$$0 = \big(R(1, 0)v_0, w_1\big)_1 = \big(v_0, R(1, 0)^* w_1\big)_1 \quad \text{for any } v_0 \in H^1, \tag{7.9}$$

whence we conclude that $R(1, 0)^* w_1 = 0$. Thus, the solution $w(t)$ of problem (7.8) vanishes at $t = 0$. By the backward uniqueness for (7.8) (e.g., see Section II.8 in [BV92]), we conclude that $w_1 = 0$. This completes the proof of the proposition. □

References

[AKSS07] A. Agrachev, S. Kuksin, A. Sarychev, and A. Shirikyan, *On finite-dimensional projections of distributions for solutions of randomly forced 2D Navier–Stokes equations*, Ann. Inst. H. Poincaré Probab. Statist. **43** (2007), no. 4, 399–415.

[Arn63] V. I. Arnold, *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian*, Uspehi Mat. Nauk **18** (1963), no. 5 (113), 13–40.

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

[AY17] G. A. Alekseev and E. V. Yurova, *On Gaussian conditional measures depending on a parameter*, Theory Stoch. Processes **22 (38)** (2017), no. 2, to appear.

[Bak86] V. I. Bakhtin, *Averaging in multifrequency systems*, Funct. Anal. Appl. **20** (1986), no. 2, 83–88.

[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.

[Bor16] A. Boritchev, *Multidimensional potential Burgers turbulence*, Comm. Math. Phys. **342** (2016), no. 2, 441–489.

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

[Cor07] J.-M. Coron, *Control and Nonlinearity*, American Mathematical Society, Providence, RI, 2007.

[Doe40] W. Doeblin, *Éléments d'une théorie générale des chaînes simples constantes de Markoff*, Ann. Sci. École Norm. Sup. (3) **57** (1940), 61–111.

[Doe00] W. Doeblin, *Sur l'équation de Kolmogoroff*, C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), no. Special Issue, 1059–1128.

[Dud02] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.

[Eli02] L. H. Eliasson, *Perturbations of linear quasi-periodic system*, Dynamical systems and small divisors (Cetraro, 1998), Lect. Notes Math., vol. 1784, Springer, Berlin, 2002, pp. 1–60.

[EM01] W. E and J. C. Mattingly, *Ergodicity for the Navier–Stokes equation with degenerate random forcing: Finite-dimensional approximation*, Comm. Pure Appl. Math. **54** (2001), 1386–1402.

[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.

[Fed69] H. Federer, *Geometric Measure Theory*, Springer, New York, 1969.

[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.

[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.

[KN13] S. Kuksin and V. Nersesyan, *Stochastic CGL equations without linear dispersion in any space dimension*, Stochastic PDE: Anal. Comp. **1** (2013), no. 3, 389–423.

[Kol54] A. N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk SSSR (N.S.) **98** (1954), 527–530.

[KS91] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991.

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

[KS01] ———, *A coupling approach to randomly forced nonlinear PDE's. I*, Comm. Math. Phys. **221** (2001), no. 2, 351–366.

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

[Kuk82] S. B. Kuksin, *Diffeomorphisms of function spaces that correspond to quasilinear 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.

[Mal16] I. I. Malofeev, *Measurable dependence of conditional measures on a parameter*, Dokl. Akad. Nauk **470** (2016), no. 1, 13–17.

[Mos66] J. Moser, *A rapidly convergent iteration method and non-linear partial differential equations. I, II*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 265–315, 499–535.

[MP06] J. C. Mattingly and É. Pardoux, *Malliavin calculus for the stochastic 2D Navier–Stokes equation*, Comm. Pure Appl. Math. **59** (2006), no. 12, 1742–1790.

[Nua06] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer-Verlag, Berlin, 2006.

[Shi15] A. Shirikyan, *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.

[Shi17] ———, *Mixing for the Burgers equation driven by a localised two-dimensional stochastic forcing*, Evolution Equations: Long Time Behavior and Control (K. Ammari and S. Gerbi, eds.), Cambridge University Press, Cambridge, 2017, pp. 179–194.

[Shi18] ———, *Controllability implies mixing II. Convergence in the dual-Lipschitz metric*, Preprint (2018).

[Sin91] Ya. G. Sinaĭ, *Two results concerning asymptotic behavior of solutions of the Burgers equation with force*, J. Statist. Phys. **64** (1991), no. 1-2, 1–12.

[Val45] F. A. Valentine, *A Lipschitz condition preserving extension for a vector function*, Amer. J. Math. **67** (1945), 83–93.

[Van06] W. C. Van Etten, *Introduction to Random Signals and Noise*, John Wiley & Sons, 2006.

[Vil03] C. Villani, *Topics in Optimal Transportation*, AMS, Providence, RI, 2003.

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