HAL
open science

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

Sergei Kuksin, Vahagn Nersesyan, Armen Shirikyan

To cite this version:

Sergei Kuksin, Vahagn Nersesyan, Armen Shirikyan. Exponential mixing for a class of dissipative PDEs with bounded degenerate noise. Geometric And Functional Analysis, 2020, 30, pp.126-187. 10.1007/s00039-020-00525-5 . hal-02386027

HAL Id: hal-02386027

https://hal.science/hal-02386027v1

Submitted on 29 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Sergei Kuksin* Vahagn Nersesyan† Armen Shirikyan‡

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 is bounded and 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 bounded random force which affects only a few Fourier modes. 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 the 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 equations, Newton–Kantorovich–Kolmogorov fast convergence, approximate controllability, Haar series

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

†Laboratoire de Mathématiques, UMR CNRS 8100, UVSQ, Université Paris-Saclay, 45, av. des Etats-Unis, F-78035 Versailles, France; 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 & Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, QC, H3A 2K6, Canada; e-mail: Armen.Shirikyan@u-cergy.fr

Contents

0 Introduction **2**

1 Mixing for Markovian random dynamical systems **8**
1.1 Setting of the problem 8
1.2 Main result and examples 10
1.3 Coupling in infinite dimension and Newton–Kantorovich–Kolmogorov fast convergence 12

2 Proof of the main result **16**
2.1 General scheme 16
2.2 Reduction to a coupling 19
2.3 Proof of Theorem 2.2 22
2.4 Image of measures under piecewise Lipschitz maps 24
2.5 Approximate right inverse of linear operators with dense image . 27
2.6 Proof of Proposition 2.3 29

3 Some auxiliary results **32**
3.1 Proof of Theorem 2.8 33
3.2 Measure of a tubular neighbourhood of a nodal set 35

4 Applications **37**
4.1 Two-dimensional Navier–Stokes system 37
4.2 Complex Ginzburg–Landau equation 42

5 Observable processes with decomposable laws **46**
5.1 Observable functions 46
5.2 Processes satisfying the hypotheses of Section 4.1 47

6 Examples of saturating subspaces **50**

7 Appendix **54**
7.1 Measurable version of gluing lemma 54
7.2 Density of sets of solutions for a linear parabolic PDE 54
7.3 Resolving operator for the Navier–Stokes system 55
7.4 Resolving operator for the complex Ginzburg–Landau equation . 56

Bibliography **60**

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 PDE 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 when 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 discussing two examples when our theory is applicable (other equations will be considered in the main text). Namely, we consider the 2D Navier–Stokes system and the Ginzburg–Landau equations, both perturbed by a Haar coloured noise. Thus, we study the following problems:

2D Navier–Stokes system:

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

where $x \in \mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$, $u = (u_1, u_2)$ and p are unknown velocity field and pressure, $\nu > 0$ is a parameter, Δ is the Laplace operator, and $\langle u, \nabla \rangle = u_1\partial_1 + u_2\partial_2$.

3D Ginzburg–Landau equation:

$$\partial_t u - (\nu + i)\Delta u + \gamma u + i|u|^{2m} u = \eta(t,x), \tag{0.2}$$

where $x \in \mathbb{T}^3 = \mathbb{R}^3/2\pi\mathbb{Z}^3$, $m = 1$ or 2, $u = u(t,x)$ is a complex-valued unknown function, and $\nu, \gamma > 0$ are some parameters.

In both cases, η is an external (random) force. Furthermore, when dealing with the Navier–Stokes system, we assume that all the functions have zero mean value in x.

Let us introduce the trigonometric basis $\{\varphi_i\}$ in the space of 2π-periodic (vector) functions of x.[1] We write the external force η in the form

$$\eta(t,x) = \sum_{i=1}^{N} b_i \eta^i(t) \varphi_i(x), \tag{0.3}$$

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 L^∞-normalised Haar system defined by relations (5.5) and (5.6); cf. Section 22 in [Lam96]. We set

$$\tilde{\eta}(t) = \sum_{k=0}^{\infty} \xi_k h_0(t-k) + \sum_{j=1}^{\infty} c_j \sum_{l=0}^{\infty} \xi_{jl} h_{jl}(t), \tag{0.4}$$

[1] In the case of the Navier–Stokes system, the basis $\{\varphi_i\}$ is composed of divergence-free $\mathbb{R}^2$-valued functions on $\mathbb{T}^2$ with zero mean value, while in the case of the Ginzburg–Landau equation, they are complex-valued functions on $\mathbb{T}^3$.

where $\{c_j\}$ is a sequence of non-zero numbers going to zero[2] 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 (0.3) are called *coloured noises* and are widely used in engineering sciences; e.g., see [Van06, Section 6.1]. We thus consider the dynamics of Equations (0.1) and (0.2) subject to the coloured noise (0.3). Because of the time correlations of finite depth, the corresponding trajectories 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 a consequence of the main result of this paper (see Theorem 1.1) and is valid for both the Navier–Stokes system and the complex Ginzburg–Landau equation.

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

$$\mathcal{D}(u(k)) \rightharpoonup \mu_\nu \quad \text{as } k \to \infty,\ k \in \mathbb{N}, \tag{0.5}$$

where $\mathcal{D}(\zeta)$ stands for the law of a random variable ζ, and the weak convergence in (0.5) is understood in the sense of measures on the space of square-integrable (vector) functions on the torus. Finally, convergence (0.5) holds exponentially fast in the dual-Lipschitz metric (defined below in Notation and conventions*).*

If $u_\nu(t), t \geq 0$, is a solution of (0.1) or (0.2) such that $\mathcal{D}(u_\nu(0)) = \mu_\nu$, then $u_\nu(t)$ is a statistically periodic process, $\mathcal{D}(u_\nu(t)) = \mathcal{D}(u_\nu(t+1))$ for each $t \geq 0$, and by Theorem 0.1, for any solution $u(t)$ the dual-Lipschitz distance between $\mathcal{D}(u(t))$ and $\mathcal{D}(u_\nu(t))$ goes to zero as $t \to \infty$.

To prove Theorem 0.1, we shall establish a general result on exponential mixing for discrete-time Markov processes and show that it applies to the problems (0.1) and (0.2). Moreover, our general result is applicable to other dissipative PDEs, such as reaction-diffusion equations with polynomial nonlinearities, viscous primitive equations of large scale ocean dynamics, and multidimensional Burgers system in non-potential[3] setting (these examples are not 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

[2]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 (0.4) converges to the white noise; see Theorem 1 in [Lam96, Section 22]. Moreover, by Donsker's invariance principle (see [Bil99, Section 8]), the integral of (0.4) converges to the Brownian motion on large time scales.

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

noise and established the uniqueness of stationary measure and its exponential stability in the dual-Lipschitz metric (see also [HMS11]). 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 [Shi19] for the boundary-driven Navier–Stokes system).

This work is a continuation of the research, started in [AKSS07] in order to prove the mixing for systems with degenerate noise by the basic Markov techniques and methods of control theory, instead of the Malliavin calculus used in [HM06, HM11, FGRT15]. Namely, in [AKSS07] we evoked those techniques to establish that the finite-dimensional projections of distributions of solutions for systems with various degenerate noises have densities against the Lebesque measure. The results of [AKSS07] do not provide any information about the regularity of densities, in contrast to the infinite-dimensional version of the Malliavin calculus developed in [MP06]. Its further development was a key ingredient of the proof of ergodicity in [HM06, HM11]. In particular, the question of almost sure invertibility of Malliavin's matrix plays a crucial role, and Tikhonov's regularisation is used to construct its approximate inverse. It is well known that, in the control theory, Malliavin's matrix corresponds to the Gramian, responsible for the controllability of the linearised system. This observation led us to the idea to use the linearised controllability in Doeblin's scheme to prove the mixing. Carrying out this idea in the form of a quadratic convergence procedure, we arrived at a "soft" proof of mixing in systems with various random forces, in difference with the "hard" proof, based on the exact formulas from the Malliavin calculus. The latter requires the forces to be Gaussian. In this paper, we work with systems stirred by bounded random forces (see below), but we are certain that our approach also applies to systems perturbed by Gaussian forces.

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 (0.4)), while it is 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? The 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, where unbounded forces make no sense (e.g., they are being used in modern meteorology). Secondly, they have a number of serious analytical advantages. Namely, the corresponding stochastic equations are always well posed if so are the deterministic equations, while the equations,

perturbed by white noise, may be not (or their solutions may satisfy only weak a priori estimates, cf. Example 1.4 below). What is more important is that for 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.[4] 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, cf. Eq. (0.2). For the stochastic 2D 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 (0.4)—has a number of similarities with the white forces (and the latter may be obtained as a limiting case of the former, see footnote 2). In particular, the forces (0.4) 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 paper, in Section 1.3 we describe the general philosophy of the proof and discuss some analogies between the coupling scheme for PDEs with noise, used in our proof, and the Newton–Kantorovich–Kolomogorov fast convergence. Section 2 is devoted to the proof of the main result. In Section 4, we apply our result to the 2D Navier–Stokes system and the complex Ginzburg–Landau equations 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 auxiliary results used in the main text.

Acknowledgements

We thank V. I. Bogachev for discussion on the measurable version of gluing lemma, which was established in the papers [BM19b, BM19a] on our request. We also thank 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. SK thanks the *Russian Science Foundation* for support through the grant 18-11-00032. VN and AS

[4] 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.

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*.

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 a separable Banach space E such that

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

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

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

$\mathcal{P}(X)$ 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

$$\|\mu_1-\mu_2\|_{\mathrm{var}}:=\sup_{\Gamma\in\mathcal{B}(X)}|\mu_1(\Gamma)-\mu_2(\Gamma)|=\frac{1}{2}\sup_{\|f\|_\infty\le 1}|\langle f,\mu_1\rangle-\langle f,\mu_2\rangle|\,. \tag{0.6}$$

We shall also use the *dual-Lipschitz* metric

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

Note that $\|\mu_1-\mu_2\|_L^*\le 2\|\mu_1-\mu_2\|_{\mathrm{var}}$.

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

$$\textit{the norms of all the derivatives } D^k F \textit{ are bounded on bounded subsets of } \mathcal{O}. \tag{0.7}$$

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 the 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. *There is a Banach space V compactly embedded into H such that the image of S is contained in V, the mapping $S : H \times E \to V$ is twice continuously differentiable, and its derivatives up to the second order are bounded on bounded subsets. Moreover, for any $u \in H$, the mapping $\eta \mapsto S(u,\eta)$ is analytic from E to H, and the derivatives $(D^j_\eta S)(u,\eta)$ are continuous functions of (u,η) that are bounded on bounded subsets of $H \times E$; cf. (0.7).*

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

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

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.[5]

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

(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 such that $|\xi_{jk}| \le 1$ a.s. and b_j are non-zero numbers satisfying

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

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

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

Before turning to formulating our main result on the discrete-time random dynamical system (1.1) satisfying (H$_1$)–(H$_4$), we describe how such systems arise in the analysis of nonlinear PDEs of parabolic type perturbed by a random force and explain why, in this situation, Hypotheses (H$_1$), (H$_2$) and (H$_4$) hold trivially, whereas (H$_3$) imposes some restrictions on both the *nonlinearity* and the *random force*. Namely, let us consider the following equation in a functional space H which, as before, is assumed to be a separable Hilbert space:

$$\dot{u} = -Lu + F(u) + \eta(t), \tag{1.6}$$

where $L = L^* > 0$ is an unbounded linear operator in H, F is an analytic nonlinearity, and η is a random force. In the examples we deal with, L is a positive degree of $-\Delta$, and we do not assume that the linear part of F at the origin (i.e., $(DF)(0)$) vanishes. Let $\mathcal{H}$ be a separable Hilbert space of finite or infinite dimension, continuously embedded into H, and let $\{\varphi_i\}_{i\in\mathcal{I}}$ be an orthonormal basis in $\mathcal{H}$. We assume that the random force η has the form

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

[5] To see this, consider a measurable space $(Y, \mathcal{Y})$ and a measurable map $A : Y \to \mathcal{L}(E,H)$ and denote by $\mathcal{G}_A$ the set of points $y \in Y$ for which the image of $A(y)$ is dense in H. Choosing countable dense subsets $\{f_i\} \subset E$ and $\{h_j\} \subset H$, it is easy to see that $\mathcal{G}_A = \big\{y \in Y : \inf_{i\ge1} \|A(y)f_i - h_j\|_H = 0 \text{ for any } j \ge 1\big\}$. 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 $\mathcal{G}^u = \{\eta \in E : (u,\eta) \in \mathcal{G}_A\}$.

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 $E := L^2(J,\mathcal{H})$ with $J=[0,1]$ such that

$$\mathcal{K} = \operatorname{supp} \mathcal{D}(\eta_k) \text{ is a compact subset of } E; \tag{1.7}$$

this implies, in particular, that the random variables η_k are uniformly bounded in E.

Denote by $S : H \times E \to H$ the mapping which takes a pair $(v, \eta|_{[0,1]})$ 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 $t \geq 0$ is a trajectory of (1.1). The existence of a compact set $X \subset H$ such that the mapping $v \mapsto S(v, \eta_1)$ preserves X a.s. often follows from the parabolic regularity and the boundedness of η.

Assumption (H_1) is another usual consequence of the parabolic regularity, while (H_2) holds, for example, if the origin of H is an exponentially stable equilibrium for Eq. (1.6) with $\eta \equiv 0$, and $0 \in \mathcal{K}$. (This condition can be relaxed, requiring only the existence of a point that is in the set of accessibility for an arbitrary initial state; see [KNS19] for details.) Hypothesis (H_4) is not restrictive either. In particular, it holds if $\{\eta_k(\cdot)\}$ are independent realisations of a process $\eta(\cdot)$ defined by the relations

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

where $\{\psi_l^i\}_{l\geq 1}$ are orthonormal bases in $L^2(J)$ for each i, $\{\xi_l^i\}$ are independent random variables with Lipschitz densities against the Lebesgue measure such that $|\xi_l^i| \leq 1$ almost surely, and $\{b_{il}\}$ decay to zero sufficiently fast as $i, l \to \infty$ (the latter property implies (1.7)); see Section 5.2 for specific examples.

Finally, Hypothesis (H_3) means certain non-degeneracy of the function F, and it holds if the latter satisfies a Hörmander-type condition and if η is observable in a suitable sense; see Sections 4 and 5 for more details. As we show in the proof, these properties, together with some analyticity argument, imply that, with high probability, the trajectories of (1.6) (and those of (1.1)) can be locally stabilised by a finite-dimensional modification of the driving force.

In conclusion of this subsection, 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 solutions 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.

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*

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \le C e^{-\gamma k} \quad \textit{for all } k \ge 0 \textit{ and } \lambda \in \mathcal{P}(X). \tag{1.8}$$

The proof of this result is based on an application of the Kantorovich functional method, described in Section 3.1.1 of [KS12] and repeated here as Theorem 2.1.[6] By that result, to prove the theorem, it suffices to check a contraction property in the space of probability measures on X. This requires subtle analysis based on ideas from the optimal control, measure theory, and theory of analytic functions. A proof of Theorem 1.1 is given in Section 2. Here we discuss briefly some examples to which our main result is applicable; the details will be given in Section 4.

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

$$H = \left\{ u \in L^2(\mathbb{T}_a^2, \mathbb{R}^2) : \operatorname{div} u = 0 \text{ in } \mathbb{T}_a^2, \int_{\mathbb{T}_a^2} u(x)\,\mathrm{d}x = 0 \right\}. \tag{1.9}$$

We consider the Navier–Stokes system (0.1) in $\mathbb{T}_a^2$ perturbed by a random process. Applying the Leray projection $\Pi : L^2(\mathbb{T}_a^2, \mathbb{R}^2) \to H$ to the equations, we reduce the system to the following nonlocal PDE:

$$\partial_t u + \nu L u + B(u) = \eta(t, x). \tag{1.10}$$

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

$$\eta(t, x) = \sum_{k=1}^{\infty} \mathbb{I}_{[k-1,k)}(t)\eta_k(t - k + 1, x), \tag{1.11}$$

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 $J = [0, 1]$ whose law is decomposable in the sense of Hypothesis (H_4). We shall prove that if, in addition, the law of η_k is *observable* (see Section 4.1), then the Markov process obtained by restricting solutions of (1.10) to integer times has a unique stationary distribution, which is exponentially mixing in the dual-Lipschitz metric. Examples of observable processes η_k are given in Section 5.

Example 1.3 (Ginzburg–Landau equation). Let us fix a vector $a = (a_1, a_2, a_3)$ with positive coordinates and denote by $\mathbb{Z}_a^3$ the integer lattice generated by the vectors $a_k\iota_k$, where $\{\iota_1, \iota_2, \iota_3\}$ is the standard basis in $\mathbb{R}^3$. We set $\mathbb{T}_a^3 = \mathbb{R}^3/2\pi\mathbb{Z}_a^3$

[6] In a different form, this method as a tool to prove the exponential mixing for nonlinear PDEs with stochastic forcing was suggested in [KPS02]. In the final form (used in our work) the method was presented in [Kuk02].

and consider Eq. (0.2), in which $u = u(t,x)$ is an unknown complex-valued function, ν, γ, and c are positive parameters, and $m \in \{1,2\}$. Equation (0.2) is well posed in the Sobolev space $H^1(\mathbb{T}^3_a, \mathbb{C})$. Assuming again that the random force η has the form (1.11) and satisfies the decomposability and observability hypotheses, we shall prove uniqueness and exponential mixing of stationary measure for (0.2).

Example 1.4 (3D stochastic primitive equations). Let us consider the following equations for a vector function $(u,p) = (u_1, u_2, u_3, p)$ in the horizontally periodic box $D = \mathbb{T}^2 \times (0,1)$:

$$\partial_t v + \langle u, \nabla \rangle v - \nu \Delta v + \nabla p = \eta(t,x), \quad \operatorname{div} u = 0, \tag{1.12}$$

where $v = (u_1, u_2)$, the pressure p is assumed to be independent of x_3, and the buoyancy is taken to be zero for simplicity. This is a popular model in the theory of climate, and its mathematical treatment attracted a lot of attention in the last two decades (e.g., see the paper [CT07] and the references therein). In the stochastic setting with a spatially regular white noise η, the well-posedness of (1.12) and the existence of a stationary measure were established in [DGTZ12, GKVZ14]. On the other hand, the uniqueness of a stationary solution in that setting seems to be out of reach because of rather weak tail estimates for strong solutions (so their moments in Sobolev spaces are not know to be bounded); cf. Theorems 1.7 and 1.9 in [GKVZ14]. In the case of a bounded observable noise, the problem of mixing can be treated with the help of Theorem 1.1 and will be carried out in a subsequent publication. Thus, even in the case of quadratic nonlinearities, the class of noises we deal with provides a new framework for proving the mixing behaviour of dynamics.

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, based on a suitable coupling construction, and discuss its relation to the technique of fast convergence due to Newton–Kantorovich–Kolmogorov. To prove convergence (1.8), it suffices to verify that any two trajectories $\{u_k\}$ and $\{u'_k\}$ with random initial data u and u' converge in distribution when $k \to \infty$. Often this property is proven with the help of the 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 $\{(\eta_k, \eta'_k)\}$ is a sequence of independent random variables such that

$$\mathcal{D}(\eta'_k) = \mathcal{D}(\eta_k) = \ell \quad \text{for all } k \geq 1. \tag{1.13}$$

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 (1.13) such that, with probability 1, $\{v_k\}$ is asymptotically close to $\{u_k\}$ as $k \to \infty$. To construct $\{\eta'_k\}_{k\geq 1}$, we fix a small parameter $d > 0$ and distinguish between the following two cases:

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

(b) The case $\delta = \|u_{k-1} - v_{k-1}\| \leq d$ contains the main difficulty. To simplify notation, assume that $k = 1$; so $u_0 = u$ and $v_0 = u'$. We seek η'_1 in the form $\eta'_1 = \Psi^{u,u'}(\eta_1)$, where the transformation $\Psi^{u,u'}$ satisfying

$$\Psi_*^{u,u'}(\ell) = \ell \tag{1.14}$$

has to be constructed. Relation (1.14) implies (1.13) (with $k = 1$), and the goal is to find $\Psi^{u,u'}$ such that the inequality $\|u_1 - v_1\| \ll \delta$ holds with high probability.

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

$$S(u, \eta_1) - S(u', \eta'_1) = 0, \quad \text{a.s.} \tag{1.15}$$

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

$$H_\delta(p, q) = h_1(p) + \delta f_1(p, q), \qquad (p, q) \in B^n \times \mathbb{T}^n, \tag{1.16}$$

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

$$H_\delta \circ S_\delta(p, q) = h_\delta(p). \tag{1.17}$$

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 (1.17) only up to a higher order term:

$$H_\delta \circ S_1(p, q) = h_2(p) + \delta^2 f_2(p, q). \tag{1.18}$$

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

$$\{h_1(p), g_1(p, g)\} = f_1(p, q) - \langle f_1 \rangle(p), \tag{1.19}$$

where $\langle\cdot\rangle$ stands for the averaging in $q\in\mathbb{T}^n$. If h_1 meets a mild non-degeneracy condition, then (1.19) 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 (1.18) which is much closer to be integrable than H_δ. Iterating this argument for $j=1,2,\dots$, Kolmogorov constructed symplectomorphisms $\{S_1\circ\cdots\circ S_j\}_{j\ge1}$ transforming H_δ to some Hamiltonian $H_{\delta,j}$ that is 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 (1.17) 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 [Kol54]. (That paper contained only a scheme of the argument, 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].)

Going back to our problem, 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 (1.15) 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

$$D_\eta S(u,\eta_1)\Phi(\eta_1)=S'(u,u',\eta_1). \tag{1.20}$$

This is the homological equation of our proof, analogous to Eq. (1.19) from the KAM theory. Equation (1.20), as well as (1.19), 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 3.1 that, for any $\varepsilon>0$, an approximate solution $\Phi(\eta_1)$ solving (1.20) 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

$$\|R_\varepsilon(u,\eta_1)\|\le C(\varepsilon,\ell(\mathcal{N}^u))<\infty \tag{1.21}$$

with some function $C(\varepsilon,r)$ going to $+\infty$ as $\varepsilon\to0$ or $r\to0$. 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 (1.15) of order

$$\delta_1=O(\delta\varepsilon)+\delta^2C(\varepsilon,\ell(\mathcal{N}^u))^2\quad\text{for}\quad\eta_1\notin\mathcal{N}^u. \tag{1.22}$$

In the KAM theory, the usual strategy is to choose $\varepsilon=\delta^{\gamma_1}$ with some $\gamma_1>0$. Then, if we knew that $\delta^2C(\delta^{\gamma_1},\ell(\mathcal{N}^u))^2\le\delta^{1+\gamma_2}$ with some $\gamma_2>0$, this would

lead to a super-exponential convergence $\|v_k - u_k\| \to 0$, typical for the theory. 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 (1.22) that

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

and choose $\varepsilon = (4C_1)^{-1}$ and $d = \big(4C_1 C(\varepsilon, \ell(\mathcal{N}^u))^2\big)^{-1}$. 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[7] 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 (1.14) 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 (1.21). 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. In reality, the corresponding argument is a bit more complicated and evokes the gluing lemma; see in Section 7.1.

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. Their detailed discussion is presented in Section 2, which is devoted to the proof of Theorem 1.1. In conclusion, let us mention that

[7] The first difficulty does not have an analogue in the KAM theory, whereas the second is usually present and manifests itself in the fact that the homological equation (1.19) cannot be solved for all actions $p \in B$.

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; see [KNS19].

2 Proof of the main result

2.1 General scheme

Let us describe the scheme of the proof of Theorem 1.1. It is based on a sufficient condition for exponential mixing from [KS12, Section 3.1.1], stated below as Theorem 2.1. Namely, we shall define a majorant of the metric on the space $\mathcal{P}(X)$ for which the Markov operator $\mathfrak{P}_1^*$ is a contraction. The construction of that majorant uses an auxiliary result about a suitable coupling for the pair of measures $(P_1(u,\cdot), P_1(u',\cdot))$; see Theorem 2.2. The latter reduces, with the help of a proposition about transformation of measures under smooth maps, to an approximate solution of a homological equation. The construction of solution is based on the existence of an approximate right inverse for a family of operators depending on a parameter that act between two Hilbert spaces and have a dense image, for almost all values of the parameter. Before turning to the accurate proof, we present the above scheme in more details. A reader not interested in an informal discussion may safely skip the text coming after Theorem 2.1 and go to Subsection 2.2.

Sufficient condition for exponential mixing

For reader's convenience, we formulate the result we need in an abstract setting. 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

$$F(u_1,u_2) \geq c\, d_X(u_1,u_2)^\beta \quad \text{for any } u_1,u_2 \in X, \tag{2.1}$$

where $0 < \beta \leq 1$ and c are positive numbers not depending on u_1 and u_2. Recall that the *Kantorovich functional* associated with F is defined by

$$\mathcal{K}_F(\mu_1,\mu_2) = \inf_{\xi_1,\xi_2} \mathbb{E}\, F(\xi_1,\xi_2), \tag{2.2}$$

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 (2.1) and the definition of the dual-Lipschitz distance that

$$\|\mu_1 - \mu_2\|_L^* \leq C\mathcal{K}_F(\mu_1,\mu_2) \quad \text{for any } \mu_1,\mu_2 \in \mathcal{P}(X), \tag{2.3}$$

where $C = c^{-1}\operatorname{diam}(X)^{1-\beta}$. A proof of the following theorem[8] can be found in Section 3.1.1 of [KS12].

[8] In [KS12], the result is proved with $\beta = 1$. However, using (2.3), it is straightforward to check that the proof remains valid for any $\beta \in (0,1)$.

Theorem 2.1. *Suppose there is a number $\varkappa \in (0,1)$ and a bounded Borel-measurable function $F : X \times X \to \mathbb{R}_+$ satisfying (2.1) such that*

$$\mathcal{K}_F(\mathfrak{P}_1^*\mu_1, \mathfrak{P}_1^*\mu_2) \le \varkappa \mathcal{K}_F(\mu_1, \mu_2) \quad \text{for any } \mu_1, \mu_2 \in \mathcal{P}(X). \tag{2.4}$$

Then the Markov process $(u_k, \mathbb{P}_u)$ has a unique stationary measure $\mu \in \mathcal{P}(X)$, which satisfies inequality (1.8) with some positive numbers C and γ.

Thus, to prove Theorem 1.1, it suffices to construct a function F satisfying the hypotheses of the above result.

Reduction to a coupling

To construct the function F, we fix a small parameter $d > 0$ as in Section 1.3 (see there Steps (a) and (b)) and divide the product space $\boldsymbol{X} = X \times X$ into disjoint subsets $\boldsymbol{X}_n$, $-N \le n \le +\infty$, where $N \sim -\ln d$. Roughly speaking, for $n \ge 0$ the set $\boldsymbol{X}_n$ consists of those pairs $(u, u') \in \boldsymbol{X}$ for which $\|u - u'\| \sim q^n d$, where q^{-1} is the maximum between 2 and the Lipschitz constant of S on X, while for $k < 0$, $\boldsymbol{X}_k$ is the collection of the pairs $(u, u') \in \boldsymbol{X}$ for which $\|u - u'\| > d$ and $\|u\| \vee \|u'\| \sim Rb^{N+k}$ for some suitable $R > 0$ and $b \in (0,1)$. Then for $(u, u') \in \boldsymbol{X}_n$, $n \ge 0$, we should argue as at Step (a) of Section 1.3, while for $(u, u') \in \boldsymbol{X}_k$ with $-N \le k < 0$ the argument follows Step (b). We define the function F to be constant on the sets $\boldsymbol{X}_n$: if $n \ge 0$, then $F \le 1$ is $\sim q^{\gamma n/2}$ for some $\gamma \in (0,1)$, while if $n < 0$, then $F = M_n$ for a suitably chosen decreasing sequence $\{M_k, -N \le k \le -1\}$. Next, to prove (2.4) for any pair of measures $(P_1(u, \cdot), P_1(u', \cdot))$ with $(u, u') \in \boldsymbol{X}_n$, we construct a coupling $(V(u, u'), V'(u, u'))$ such that, no matter if $n \ge 0$ or $n < 0$, we have

$$\mathbb{P}\Big\{(V, V') \in \bigcup_{m \le n-2} \boldsymbol{X}_m\Big\} \lesssim \|u - u'\|^\gamma, \quad \mathbb{P}\Big\{(V, V') \in \bigcup_{m \ge n+1} \boldsymbol{X}_m\Big\} \ge 1 - \nu, \tag{2.5}$$

where $\nu \in (0,1)$ is sufficiently small. In other words, with a probability close to 1 the distance between the trajectories issued from u and u' squeezes by a factor of q at time $t = 1$, whereas the probability that distance increases by a factor of at least q^{-2} is estimated by $\|u - u'\|^\gamma$. Once these two inequalities are established, the validity of (2.4) follows from a simple computation; see Case 1 in Section 2.2.

The actual construction is more complicated since we need to take into account the points (u, u') that are far from each other (so that the first inequality in (2.5) does not give any information). The proof of (2.4), based on the existence of a coupling as above, is presented in Section 2.2.

Reduction to a homological equation

We now describe how to construct a coupling (V, V') satisfying (2.5), for any $(u, u') \in \boldsymbol{X}$. We know that $(u, u') \in \boldsymbol{X}_n$ for some $n \ge -N$. If $n \le -1$ or $n = \infty$, then $\|u - u'\| > d$ or $u = u'$, and we choose $V = S(u, \zeta)$, $V' = S(u', \zeta)$,

where $\mathcal{D}(\zeta) = \ell$. If $n \geq 0$, then $\|u - u'\| \leq d$, and to construct the coupling we use a general result about the transformation of probability measure under smooth finite-dimensional perturbations of the identical map. Roughly speaking, it says that if $\ell \in \mathcal{P}(E)$ is a decomposable measure with compact support, and $\Psi : E \to E$ is a smooth map of the form $\Psi(\zeta) = \zeta + \Phi(\zeta)$, where Φ has a finite-dimensional range and satisfies the inequalities

$$\|\Phi(\zeta)\| \leq \varkappa, \quad \|\Phi(\zeta) - \Phi(\zeta')\| \leq \varkappa\|\zeta - \zeta'\| \quad \text{for } \ell\text{-a.e. } \zeta, \zeta' \in E, \tag{2.6}$$

then the total variation distance between ℓ and $\Psi_*(\ell)$ can be estimated by $C\varkappa^\beta$, with some $\beta \in (0, 1)$. (We refer the reader to Theorem 2.4 for an exact statement.) Suppose now we have constructed finite-dimensional maps $\Phi^{u,u'} : E \to E$ that are perturbations of identity (in the sense that (2.6) holds for them with $\varkappa \simeq \|u - u'\|$) and satisfy the inequality

$$\|S(u, \eta) - S(u', \eta + \Phi^{u,u'}(\eta))\| \leq q\|u - u'\| \tag{2.7}$$

for $u, u' \in \boldsymbol{X}$ and ℓ-a.e. $\eta \in E$. Then the application of the above result to $\Psi^{u,u'} = \mathrm{Id} + \Phi^{u,u'}$ enables one to conclude that

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

Using now the gluing lemma (see Theorem 7.1), we can construct a coupling (V, V') satisfying (2.5).

We thus need to construct a small smooth map $\Phi^{u,u'}$ satisfying (2.7). To this end, we argue as at Step (b) in Section 1.3. Namely, setting $\zeta = \Phi^{u,u'}(\eta)$, using Taylor's expansion, and ignoring the second-order terms, we can rewrite (2.7) as

$$\|(D_u S)(u, \eta)(u' - u) + (D_\eta S)(u, \eta)\zeta\| \lesssim q\|u - u'\|. \tag{2.8}$$

If the map $D_\eta S(u, \eta)$ had a full image, we could annihilate the left-hand side of this inequality. However, this is not the case, and we arrive at the problem of solving (approximately) the *homological equation* (cf. (1.20))

$$(D_\eta S)(u, \eta)\zeta = -(D_u S)(u, \eta)(u' - u). \tag{2.9}$$

Solving the homological equation

It is exactly here where Hypothesis (H_3) comes into play. Namely, we know that the image of $(D_\eta S)(u, \eta)$ is dense in E for any $u \in X$ and ℓ-a.e. $\eta \in E$. Hence, when u is fixed, we can find an approximate solution of (2.9) for almost every η. However, this not sufficient, since the resulting function $\zeta = \Phi^{u,u'}(\eta)$ must satisfy a number of properties. More precisely, $\Phi^{u,u'}$ should have a finite-dimensional range, be a function of order $\|u - u'\|$, and solve approximately (2.9) for all close pairs $(u, u') \in \boldsymbol{X}$ and ℓ-almost all $\eta \in E$. This will be achieved with the help of a construction of approximate right inverses for a family of linear operators with dense images. The corresponding result is given in Theorem 2.8, which is one of the main technical novelties of this paper.

2.2 Reduction to a coupling

In this subsection, we construct a symmetric function $F : X \times X \to \mathbb{R}_+$ satisfying (2.1) and prove that (2.4) 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

$$\|S(u_1,\zeta) - S(u_2,\zeta)\| \le q^{-1}\|u_1 - u_2\| \quad \text{for any } u_1, u_2 \in X,\ \zeta \in \mathcal{K}; \tag{2.10}$$

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 for $n \ge 0$ and $-N \le k \le -1$ introduce the sets

$$\boldsymbol{X}_\infty = \big\{(u,u') \in \boldsymbol{X} : u = u'\big\}, \tag{2.11}$$

$$\boldsymbol{X}_n = \big\{(u,u') \in \boldsymbol{X} : q^{n+1}d < \|u-u'\| \le q^n d\big\}, \tag{2.12}$$

$$\boldsymbol{X}_k = \big\{(u,u') \in \boldsymbol{X} : \|u-u'\| > d, Rb^{N+k+1} < \|u\| \vee \|u'\| \le Rb^{N+k}\big\}. \tag{2.13}$$

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 (2.4) is the following result.

Theorem 2.2. *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*

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \ge n+1\big\} \ge 1-\nu, \tag{2.14}$$

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \le n-2\big\} \le C\,\|u-u'\|^\gamma, \tag{2.15}$$

where $n \ge 0$ in (2.14) and $n \ge 1$ in (2.15).

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

$$\mathbb{P}\big\{\big(V(u,u'),V'(u,u')\big) \in \boldsymbol{X}_m \text{ for some } m \ge k+1\big\} \ge p. \tag{2.16}$$

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 Theorem 2.2. We fix $\nu > 0$ so small that[9]

$$q^{\gamma/2} + q^{-\gamma/2}\nu < 1, \tag{2.17}$$

$$q^{\gamma/2} + 3\nu < 1. \tag{2.18}$$

Let $d_0 > 0$ be the constant constructed in Theorem 2.2 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 $-N \le n \le +\infty$ are uniquely determined. We define

$$F(u,u') = \begin{cases} 0 & \text{for } (u,u') \in \boldsymbol{X}_\infty, \\ (q^n d)^{\gamma/2} & \text{for } (u,u') \in \boldsymbol{X}_n, \\ M_k & \text{for } (u,u') \in \boldsymbol{X}_k, \end{cases} \tag{2.19}$$

where $n \ge 0$, $-N \le k \le -1$, and $M_k \ge 2d^{\gamma/2}$ is a decreasing sequence to be chosen below. It is straightforward to see that F satisfies (2.1). We shall prove that inequality (2.4) 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

$$\mathcal{K}_F(\mathfrak{P}_1^*\delta_u, \mathfrak{P}_1^*\delta_{u'}) \le \varkappa \mathcal{K}_F(\delta_u, \delta_{u'}) = \varkappa F(u,u') \quad \text{for any } (u,u') \in \boldsymbol{X}. \tag{2.20}$$

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

$$\mathcal{K}_F(\mu_1, \mu_2) \ge \mathbb{E}\, F(\xi_1, \xi_2) - \theta. \tag{2.21}$$

Now let $(V(u,u'), V'(u,u'))$ be the random variables constructed in Theorem 2.2. 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 (2.20) and (2.21), 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 (2.4).

To establish (2.20), 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, assuming that the parameters b, R, q, C, and γ are fixed (see (2.10) and Theorem 2.2).

Case 1: $(u,u') \in \boldsymbol{X}_n$ with $n \ge 1$, so $F = (q^n d)^{\gamma/2}$. Since (V, V') is a coupling for $(\mathfrak{P}_1^*\delta_u, \mathfrak{P}_1^*\delta_{u'})$, it follows from of (2.14), (2.15), and (2.19) that

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

[9] The role of inequalities (2.17) and (2.18) will be clarified below.

where G_n^1 and G_n^2 denote the events on the left-hand sides of (2.14) and (2.15), 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 (2.20) will be established if we prove that $\varkappa_1 < 1$, uniformly in $n \geq 1$. To this end, notice that $\mathbb{P}(G_n^3) \leq \nu$ in view of (2.14), so that

$$q^{\gamma/2}\,\mathbb{P}(G_n^1) + q^{-\gamma/2}\mathbb{P}(G_n^3) \leq q^{\gamma/2} + q^{-\gamma/2}\nu. \tag{2.22}$$

Furthermore, it follows from (2.15) that

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

Combining this with (2.22) and (2.17), we see that $\varkappa_1 < 1$, provided that

$$CM_{-N}d^{\gamma/2} < 1 - q^{\gamma/2} - q^{-\gamma/2}\nu. \tag{2.23}$$

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 $M_{-N} \leq 3d^{\gamma/2}$, so that (2.23) will be satisfied if

$$3Cd^{\gamma} < 1 - q^{\gamma/2} - q^{-\gamma/2}\nu. \tag{2.24}$$

In what follows, we fix $d \in (0, d_0)$ satisfying (2.24). Together with b, they determine $N \geq 1$ as the least positive integer satisfying $b^N R \leq d/2$.

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

$$M_k = 2d^{\gamma/2} + \varepsilon(B^{N-1} - B^{N+k}), \tag{2.25}$$

where $\varepsilon > 0$ and $B > 1$ will be chosen below. Arguing as above, using (2.14), and assuming that $\varepsilon \leq B^{1-N}d^{\gamma/2}$ (so that $M_{-N} \leq 3d^{\gamma/2}$), we derive

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

In view of (2.18), we have $\varkappa_2 < 1$.

Case 3: $(u,u') \in \boldsymbol{X}_{-1}$, so $F = M_{-1} = 2d^{\gamma/2}$. It follows from (2.16) with $m = -1$ that

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

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

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

Case 4: $(u,u') \in \boldsymbol{X}_k$ with $-N \leq k \leq -2$, so $F = M_k$. Using (2.16) and (2.25), we derive

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

Let us set $B = 2/p \geq 2$. 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 (2.20) holds with $\varkappa = \max\{\varkappa_i, 1 \leq i \leq 4\}$, provided that

$$\varepsilon = B^{-N} d^{\gamma/2} \min\bigl\{p(1-p)^{-1}, B\bigr\}. \tag{2.26}$$

Relation (2.20) implies inequality (2.4), and the exponential mixing (1.8) follows. Thus, to complete the proof of Theorem 1.1, it remains to establish Theorem 2.2, which is done in the next subsection.

2.3 Proof of Theorem 2.2

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). Recalling (2.13), we see that (2.16) will be established if we prove that

$$\mathbb{P}\bigl\{\|V\| \vee \|V\| \leq b(\|u\| \vee \|u'\|)\bigr\} \geq p, \tag{2.27}$$

where $p > 0$ is a number depending only on d (but not on the vectors $u, u' \in X$). To see this, note that $\|u\| \vee \|u'\| \geq r := d/2$. By Hypothesis ($\mathrm{H}_2$) and the Lipschitz property of $S(u, \cdot) : \mathcal{K} \to H$, we have

$$\begin{aligned} \|S(u,\zeta)\| &\leq \|S(u,0)\| + C\|\zeta\|_E \leq a\|u\| + C(r^{-1}\|u\| \vee \|u'\|)\|\zeta\| \\ &\leq \bigl(a + Cr^{-1}\|\zeta\|\bigr)\|u\| \vee \|u'\|. \end{aligned} \tag{2.28}$$

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 (2.27) 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.

Let us turn to the case $\|u - u'\| \leq d$, with a sufficiently small $d > 0$ to be specified below. Given any $\delta > 0$, we set $D_\delta = \{(u, u') \in X \times H : \|u - u'\| \leq \delta\}$. We shall need the following auxiliary result whose proof is given in Section 2.6.

Proposition 2.3. *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 Borel 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 $\Psi^{u,u'}(\eta) = \eta + \Phi^{u,u'}(\eta)$:*

$$\ell(\mathcal{K}^{u,\sigma,\theta}) \geq 1 - \sigma, \tag{2.29}$$

$$\|\ell - \Psi_*^{u,u'}(\ell)\|_{\mathrm{var}} \leq C\,\|u - u'\|^\beta, \tag{2.30}$$

$$\|S(u,\eta) - S(u', \Psi^{u,u'}(\eta))\| \leq \theta\,\|u - u'\|, \tag{2.31}$$

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

From now on, we assume that $d \le \delta$, where $\delta > 0$ is the number constructed in Proposition 2.3 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 2.3. 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

$$\mathfrak{N}_{u,u'}\big(\{(\zeta,\zeta') \in E \times E : \zeta \ne \zeta'\}\big) = \|\Psi_*^{u,u'}(\ell) - \ell\|_{\mathrm{var}}. \tag{2.32}$$

In view of Theorem 1.2.28 in [KS12], we can choose $\mathfrak{N}_{u,u'}$ to be measurable in (u,u'), i.e., 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 the gluing lemma (see Theorem 7.1 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 $(u,u') \in D_\delta$,

$$\mathcal{D}\big(\zeta(u,u',\cdot), \hat\zeta(u,u',\cdot)\big) = \mathfrak{M}_{u,u'}, \quad \mathcal{D}\big(\hat\zeta(u,u',\cdot), \zeta'(u,u',\cdot)\big) = \mathfrak{N}_{u,u'}. \tag{2.33}$$

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 (2.14) and (2.15).

To this end, for any $(u,u') \in D_\delta$ we introduce the events[10]

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

These events form a partition of the probability space Ω, and it follows from (2.29), (2.30), (2.32), and (2.33) that

$$\mathbb{P}(\Omega_1^{u,u'}) \ge 1 - \sigma - C_1\delta^\beta, \tag{2.34}$$

$$\mathbb{P}(\Omega_2^{u,u'}) \le C_1\|u-u'\|^\beta, \tag{2.35}$$

where C_1, β, and δ depend on σ and θ. Moreover, it follows from (2.31) and (2.10) that, on the event $\Omega_1^{u,u'}$, we have

$$\|V(u,u') - V'(u,u')\| = \|S(u,\zeta(u,u')) - S(u',\hat\zeta(u,u'))\| \le \theta\|u-u'\|, \tag{2.36}$$

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

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

$$\|V(u,u') - V'(u,u')\| = \|S(u,\zeta(u,u')) - S(u',\zeta(u,u'))\| \le q^{-1}\|u-u'\|. \quad (2.37)$$

We now specify our choice of the parameters. Let

$$\sigma = \frac{\nu}{2}, \quad \theta = q, \quad (2.38)$$

let $C =: C_1$, β and δ be the constants constructed in Proposition 2.3, and let $d_0 \in (0,\delta)$ be so small that $C_1 d_0^\delta \le \frac{\nu}{2}$. It follows from (2.36), (2.37), and (2.38) 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 (2.14) and (2.15) (with $\gamma = \beta$ and $C = C_1$) follow now from (2.34) and (2.35), respectively. This completes the proof of Theorem 2.2.

Thus, it remains to establish Proposition 2.3. To do it, we need two auxiliary results about transformation of measures under Lipschitz maps close to identity and existence of approximate right inverse for a family of operators with dense image. Since those results may have some independent interest, we formulate and prove them in an abstract setting.

2.4 Image of measures under piecewise Lipschitz maps

Let E be a separable Hilbert space, let $\mathcal{K} \subset E$ be a compact subset, and let $\Psi : \mathcal{K} \to E$ be a map of the form $\Psi(\eta) = \eta + \Phi(\eta)$, where $\Phi : \mathcal{K} \to E$ is a Borel-measurable "small" map 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 maps 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.

Suppose that E is 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 Lipschitz-continuous density ρ with respect to the Lebesgue measure on $\mathcal{E}$.

Theorem 2.4. *Let $\Phi : \mathcal{K} \to \mathcal{E} \subset E$ be a Borel-measurable map 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*

$$\|\Phi(\eta)\| \le \varkappa, \quad \|\Phi(\eta) - \Phi(\eta')\| \le \varkappa\|\eta - \eta'\| \quad \text{for } \eta, \eta' \in \mathcal{K}_1. \tag{2.39}$$

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

$$\mathrm{Leb}\big\{v \in \mathcal{E} : \mathrm{dist}\big(v, \partial_w \mathcal{K}_1\big) \le r\big\} \le c\, r^\gamma, \tag{2.40}$$

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*

$$\|\ell - \Psi_*(\ell)\|_{\mathrm{var}} \le C\varkappa^\beta. \tag{2.41}$$

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

$$\langle f \circ \Psi, \ell\rangle = \int_{\mathcal{K}'} \ell_{\mathcal{E}'}(\mathrm{d}w) \int_{\mathcal{K}(w)} f(v + w + \Phi(v+w))\rho(v)\mathrm{d}v, \tag{2.42}$$

$$\langle f, \ell\rangle = \int_{\mathcal{K}'} \ell_{\mathcal{E}'}(\mathrm{d}w) \int_{\mathcal{K}(w)} f(v + w)\rho(v)\mathrm{d}v. \tag{2.43}$$

Suppose we have shown that

$$\left| \int_{\mathcal{K}_1(w)} \big(f(v + w + \Phi(v+w)) - f(v+w)\big)\rho(v)\mathrm{d}v \right| \le C_1 \varkappa^\beta, \tag{2.44}$$

where C_1 and β do not depend on w, f, and $\varkappa$. In this case, taking the absolute value of the difference between (2.42) and (2.43), using that Φ is zero outside $\mathcal{K}_1$, and estimating the interior integral with the help of (2.44), 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 (2.41). Thus, we need to establish (2.44).

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 (2.39). 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 (2.44). 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 2.5 (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 $\leq \varkappa$, 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 $\leq C\varkappa$, where $C > 0$ does not depend on $\varkappa$.

To introduce truncations, we define the sets

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

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

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

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

$$\int_{\mathcal{K}_1(w)} f(v + w + \Phi(v + w))\rho(v)\mathrm{d}v \leq \int_{\mathcal{E}} f(v + w + \tilde{\Phi}(v + w))\chi_r^+(v, w)\rho(v)\mathrm{d}v,$$
$$\int_{\mathcal{K}_1(w)} f(v + w)\rho(v)\mathrm{d}v \geq \int_{\mathcal{E}} f(v + w)\chi_r^-(v, w)\rho(v)\mathrm{d}v$$
$$\geq \int_{\mathcal{E}} f(v + w)\chi_r^+(v, w)\rho(v)\mathrm{d}v - c\,\|\rho\|_\infty r^\gamma,$$

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 (2.40). Now note that, for $\varkappa \leq \frac{1}{2}$, 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 (2.44)

and using, for instance, Theorem 3.2.5 in [Fed69] to make a change of variable, we obtain

$$\begin{aligned}\delta(f) &\le \int_{\mathcal{E}} f(v+w)\frac{\chi_r^+(\Theta_w(v),w)\rho(\Theta_w(v))}{\det\bigl(I+(D\tilde{\Phi})(\Theta_w(v)+w)\bigr)}\,\mathrm{d}v \\ &\quad - \int_{\mathcal{E}} f(v+w)\chi_r^+(v,w)\rho(v)\mathrm{d}v + C_1 r^\gamma \\ &\le \int_B f(v+w)|\Delta(v,w)|\mathrm{d}v + C_1 r^\gamma, \end{aligned} \tag{2.45}$$

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

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

Since $v \mapsto \Theta_w(v)$ is a 2-Lipschitz map 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 (2.45), we derive

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

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

$$\delta(f) \le C_4\varkappa^{\gamma/(1+\gamma)}. \tag{2.46}$$

A similar argument shows that $\delta(f) \ge -C_5\varkappa^{\gamma/(1+\gamma)}$. Combining this with (2.46), we arrive at inequality (2.44) with $\beta = \frac{\gamma}{1+\gamma}$. This completes the proof of the theorem. □

2.5 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 2.6. *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*

$$\|AR_\varepsilon f - f\|_H \le \varepsilon\|f\|_V \quad \textit{for any } f \in V. \tag{2.47}$$

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 2.7. *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, then*

$$\lim_{\gamma \to 0} \|G(G + \gamma)^{-1} f - f\| = 0 \quad \textit{for any } f \in H. \tag{2.48}$$

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 2.7, 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

$$\|AA^*(G + \gamma)^{-1} g - g\| \le \frac{2\varepsilon}{3} \quad \text{for } 0 < \gamma \le \gamma_\varepsilon(f),\ \|g - f\| \le \delta_\varepsilon(f). \tag{2.49}$$

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 (2.49) that

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

Since the sequence $\{AR_{\gamma_\varepsilon,M}\}_{M \ge 1}$ 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 (2.47) with $R_\varepsilon = R_{\gamma_\varepsilon,M_\varepsilon}$. □

Proof of Lemma 2.7. 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 (2.48) for operators with a dense image, we first note that

$$G(G + \gamma)^{-1} - I = -\gamma(G + \gamma)^{-1}. \tag{2.50}$$

Since the norm of the operator $\gamma(G+\gamma)^{-1}$ is bounded by 1, it suffices to prove that the right-hand side of (2.50) 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 (2.48) and completes the proof. □

In what follows, we shall need a version of Proposition 2.6 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 (0.7) and the convention for analytic mappings with parameter). As in Proposition 2.6, 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 $f_1, \dots, f_N$. A proof of the following result will be given in Section 3.1.

Theorem 2.8. *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 $M_{\vec{\varepsilon}} \ge 1$, 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*

$$\mathcal{K}^u_{\vec{\varepsilon}} := \{\eta \in \mathcal{K} : \mathfrak{F}_{\vec{\varepsilon}}(u, \eta) \le \nu_{\varepsilon_2}\} \tag{2.51}$$

satisfies the inequality

$$\ell(\mathcal{K}^u_{\vec{\varepsilon}}) \ge 1 - \varepsilon_1 \quad \text{for } u \in X. \tag{2.52}$$

Right inverse. *Let us define the compact set*

$$\mathcal{D}_{\vec{\varepsilon}} = \{(u, \eta) \in X \times \mathcal{K} : \mathfrak{F}_{\vec{\varepsilon}}(u, \eta) \le 2\nu_{\varepsilon_2}\}. \tag{2.53}$$

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

$$\operatorname{Im}\big(R_{\vec{\varepsilon}}(u, \eta)\big) \subset F_{M_{\vec{\varepsilon}}}, \quad \|R_{\vec{\varepsilon}}(u, \eta)\|_{\mathcal{L}(H,F)} \le C_{\vec{\varepsilon}} \quad \text{for } (u, \eta) \in \mathcal{D}_{\vec{\varepsilon}}, \tag{2.54}$$

$$\|A(u, \eta)R_{\vec{\varepsilon}}(u, \eta)f - f\|_H \le \varepsilon_2\|f\|_V \quad \text{for } (u, \eta) \in \mathcal{D}_{\vec{\varepsilon}},\ f \in V. \tag{2.55}$$

2.6 Proof of Proposition 2.3

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

$$\|S(u, \eta) - S(u', \eta + \zeta)\| \le \theta\|u - u'\|, \tag{2.56}$$

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

$$S(u',\eta+\zeta)=S(u,\eta)+(D_uS)(u,\eta)(u'-u)+(D_\eta S)(u,\eta)\zeta+r(u,u',\eta,\zeta), \quad (2.57)$$

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

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

This equation is not necessarily solvable. However, by Hypothesis (H_3), the image of $(D_\eta S)(u,\eta)$ is dense in H. By Theorem 2.8, for any $\varepsilon>0$, there exists an approximate right inverse $R_\varepsilon(u,\eta):H\to E$ such that

$$\|(D_\eta S)(u,\eta)R_\varepsilon(u,\eta)f-f\|\le\varepsilon\|f\|_V \quad \text{for any } f\in V, \quad (2.58)$$

where V is the Banach space in (H_1). The mapping $(D_uS)(u,\eta)$ is continuous from H to V, so that in (2.58) we can take $f=-(D_uS)(u,\eta)(u'-u)$. We shall show that, for a sufficiently small $\varepsilon=\varepsilon(\theta)>0$, the mapping

$$\Phi^{u,u'}(\eta)=-R_\varepsilon(u,\eta)(D_uS)(u,\eta)(u'-u) \quad (2.59)$$

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 Theorem 2.8 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 (2.59). The main point is to choose them in such a way that inequality (2.40) is true with $\mathcal{K}_1=\mathcal{K}^{u,\sigma,\theta}$ (so that we can apply Theorem 2.4 to prove (2.30)).

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

$$\mathcal{K}^u_\varepsilon=\{\eta\in\mathcal{K}:\mathfrak{F}_\varepsilon(u,\eta)\le\nu_\varepsilon\} \quad \text{for } u\in X, \quad (2.60)$$

$$\mathcal{D}_\varepsilon=\{(u,\eta)\in X\times\mathcal{K}:\mathfrak{F}_\varepsilon(u,\eta)\le 2\nu_\varepsilon\}, \quad (2.61)$$

$$R_\varepsilon(u,\eta)=\mathsf{P}_M R_\gamma(u,\eta) \quad \text{for } (u,\eta)\in\mathcal{D}_\varepsilon, \quad (2.62)$$

where the operator R_γ is defined in (3.1), 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 $X\times E_M^\perp$. In other words,

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

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 2.9. *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

$$\mathcal{K}^{u,\sigma,\theta} = \bigcup_{l=1}^{m} \big\{\eta \in \mathcal{K} : \eta = v + w, (u,w) \in \mathcal{D}'_l, \mathfrak{F}_\varepsilon(u, w+v) \le \nu_l\big\} \tag{2.63}$$

and introduce a function $\Phi : X \times H \times E \to E$ as follows: $\Phi^{u,u'}(\eta)$ is defined by (2.59) 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 (2.29)–(2.31) hold for an appropriate choice of ε.

Step 2. Proof of (2.29). Up to now, the parameter $\varepsilon > 0$ was arbitrary. Let us choose it so small that $\varepsilon \le \sigma$. In view of (2.52), 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 (2.60) and (2.63) since $\nu_l > \nu_\varepsilon$.

Step 3. Proof of (2.31). We first note that (2.61) and (2.63) imply the inclusion $\mathcal{K}^{u,\sigma,\theta} \subset \{\eta \in \mathcal{K} : (u,\eta) \in \mathcal{D}_\varepsilon\}$. In view of relations (2.57) and (2.58), 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

$$\begin{aligned} &\|S(u,\eta) - S(u', \eta + \Phi^{u,u'}(\eta))\| \\ &\qquad \le \varepsilon \, \|(D_u S)(u,\eta)(u'-u)\|_V + \|r(u,u',\eta,\Phi^{u,u'}(\eta))\|. \end{aligned} \tag{2.64}$$

Hypothesis (H_1) implies that

$$\|(D_u S)(u,\eta)(u'-u)\|_V \le C_1 \|u'-u\|, \tag{2.65}$$

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

$$\|r(u,u',\eta,\zeta)\| \le C_2\big(\|u-u'\|^2 + \|\zeta\|_E^2\big). \tag{2.66}$$

Recalling the definition of $\Phi^{u,u'}$ (see (2.59)) and using the inequality in (2.54), we derive

$$\|\Phi^{u,u'}(\eta)\| \le C_3(\varepsilon)\|u-u'\| \quad \text{for } u \in X,\ \eta \in \mathcal{K}^{u,\sigma,\theta},\ u' \in H. \tag{2.67}$$

Substituting (2.65)–(2.67) into (2.64), 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 (2.31) for $(u,u') \in D_\delta$ and $\eta \in \mathcal{K}^{u,\sigma,\theta}$.

Step 4. Proof of (2.30). We shall use Theorem 2.4 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 (2.30) will be established

if we show that Properties (a) and (b) of Theorem 2.4 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 (2.67), the first inequality in (2.39) is satisfied. To prove[11] the second, let us fix any smooth 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

$$\Phi^{u,u'}(\eta) = -h\big(\mathfrak{F}_\varepsilon(u,\eta)\big) R_\varepsilon(u,\eta)(D_u S)(u,\eta)(u'-u) \quad \text{for } \eta \in \mathcal{K}^{u,\sigma,\theta}. \tag{2.68}$$

The right-hand side of (2.68) 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 (2.63). 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 (2.40) 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

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

Now note that the analytic function $O_M \ni v \mapsto \mathfrak{F}_\varepsilon(u, w+v) - \nu_l$ does not vanish identically for $(u,w) \in \overline{\mathcal{D}}_l$. An estimate for the neighbourhood of the set of zeros for such functions will be given in Corollary 3.3. Applying the latter to the Lebesgue measure on E_M, we conclude that (2.40) is valid for $\partial_w \mathcal{K}^l$. This completes the proof of Proposition 2.3.

Proof of Lemma 2.9. 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 $l = 1$. $\square$

3 Some auxiliary results

In this section, we prove the theorem about the existence of a right inverse for a family of operators with dense image and establish an estimate for the measure

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

of a tubular neighbourhood of a nodal set for holomorphic functions depending on a parameter. The latter was used in the proof of Proposition 2.8 and will also be needed in Section 5.2.

3.1 Proof of Theorem 2.8

We essentially repeat the proof of Proposition 2.6, 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

$$R_\gamma(u,\eta) = A(u,\eta)^*\big(G(u,\eta) + \gamma I\big)^{-1}, \quad R_{M,\gamma}(u,\eta) = \mathsf{P}_M R_\gamma(u,\eta), \tag{3.1}$$

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 (2.54) holds for any fixed M and γ. Thus, we need to construct a function $\mathfrak{F}_{\vec{\varepsilon}}$ and to prove (2.52) and (2.55).

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

$$\sup_{(u,\eta)\in\mathcal{D}_{\vec{\varepsilon}}} \|A(u,\eta)R_\gamma(u,\eta)f - f\| < \varepsilon_2 \quad \text{for } f \in B_V(1), \tag{3.2}$$

then inequality (2.55) 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 (3.2), we see that (2.55) 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 (2.52) and (3.2) hold.

To this end, note that, by Lemma 2.7, for any $u \in X$, $\eta \in \mathcal{K}^u$, and $f \in H$,

$$\lim_{\gamma\to 0} \|A(u,\eta)R_\gamma(u,\eta)f - f\| = 0. \tag{3.3}$$

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

$$\mathfrak{F}_\gamma(u,\eta) = \sum_{j=1}^{N} \|A(u,\eta)R_\gamma(u,\eta)f_j - f_j\|^2 \tag{3.4}$$

and notice that it is analytic in η and decreases with $\gamma > 0$ for each (u,η) in view of Lemma 2.7. Furthermore, setting $\nu_{\varepsilon_2} = \varepsilon_2^2/32$, we see that if $\mathfrak{F}_\gamma(u,\eta) \le 2\nu_{\varepsilon_2}$, then for any $f \in B_V(1)$ we have

$$\begin{aligned}\|A(u,\eta)R_\gamma(u,\eta)f - f\| &\le \min_{1\le j\le N}\big(\|AR_\gamma(f-f_j)\| + \|f-f_j\|\big) + \mathfrak{F}_\gamma(u,\eta)^{1/2}\\ &\le 3\varepsilon_2/4,\end{aligned} \tag{3.5}$$

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 3.1. *Under the hypotheses of Proposition 2.8, for any $\nu > 0$ we have*

$$\ell\big(\{\eta \in \mathcal{K} : \mathfrak{F}_\gamma(u,\eta) < \nu\}\big) \to 1 \quad \textit{uniformly in } u \in X \textit{ as } \gamma \to 0^+. \tag{3.6}$$

Applying this result with $\nu = \nu_{\varepsilon_2}$, we can find $\gamma = \gamma(\vec{\varepsilon}) > 0$ such that (2.52) holds for the set $\mathcal{K}^u_{\vec{\varepsilon}}$ defined by relation (2.51) with $\mathfrak{F}_{\vec{\varepsilon}} = \mathfrak{F}_{\gamma(\vec{\varepsilon})}$. It remains to note that inequality (3.2) follows immediately from (3.5). This completes the proof of Theorem 2.8.

Proof of Lemma 3.1. By assumption, the image of $A(u,\eta)$ is dense for any $u \in X$ and $\eta \in \mathcal{K}^u$. It follows from (3.3) 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

$$m_\gamma(u) = \ell\big(\{\eta \in \mathcal{K} : \mathfrak{F}_\gamma(u,\eta) < \nu\}\big) \tag{3.7}$$

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 2.7, 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. □

3.2 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 3.2. *In addition to the above hypotheses, let us assume that the one-dimensional projections of ℓ to the spaces* $\mathrm{span}(e_j)$ *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 $r \in [0,1]$,*

$$\ell\big(\{\eta \in \mathcal{O} : |\mathfrak{F}(u,\eta)| \le r\}\big) \le C\, r^c. \tag{3.8}$$

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 3.3. *Under the hypotheses of Lemma 3.2, there are numbers $C, c > 0$ such that, for any $u \in X$ and $r \in [0,1]$,*

$$\ell\big(\{\eta \in \mathcal{O} : \mathrm{dist}(\eta, \mathcal{N}(u)) \le r\}\big) \le C\, r^c. \tag{3.9}$$

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

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

where $C' > 0$ does not depend on $u \in X$ and $r \in [0,1]$. The required result is now implied by (3.8). □

Proof of Lemma 3.2. It suffices to establish (3.8) 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 (3.8) 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

$$|\mathfrak{F}(u,\eta)| > \sigma(u^0,\eta^0) > 0 \quad \text{for } u \in O_{u^0},\ \eta \in O_{\eta^0}. \tag{3.10}$$

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])

$$\mathrm{Leb}\big(\{t \in [-\delta,\delta] : |f(t)| \le r\}\big) \le C(u^0,\eta^0)\, r^{1/m} \tag{3.11}$$

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 (3.11) that

$$\ell\big(\mathcal{Z}(u,r) \cap O_{\eta^0}\big) \le C(u^0,\eta^0)\, r^{c(u^0,\eta^0)}. \tag{3.12}$$

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 (3.10) or (3.12) 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 (3.8) follows from (3.10) and (3.12). □

In Section 5.2, we shall need the following particular case of Lemma 3.2 when E is a finite-dimensional space and X is a singleton.

Corollary 3.4. *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*

$$\mathbb{P}\{|\mathfrak{F}(\xi_1,\dots,\xi_n)| \le r\} \le C(\rho)\, r^c \lambda \quad \textit{for any } r \in [0,1], \tag{3.13}$$

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

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 (1.10) 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 (1.11), 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 (1.10). 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 (1.10) 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 intersection with H. 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]. The restrictions $\{u_k\}$ of a solution $u(t)$ for (1.10) to integer times satisfy relation (1.1). 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 (1.10) 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 $(u_k, \mathbb{P}_u)$.

From now on, we assume that the measure $\ell = \mathcal{D}(\eta_k)$ has a compact support $\mathcal{K} \subset E$, which contains the origin, and satisfies[12] the decomposability hypothesis (H_4). To formulate the main result, we need the concepts of an *observable measure*[13] and a *saturating subspace*.

Definition 4.1. Let $\mathcal{H}$ be a Hilbert space with an inner product $(\cdot,\cdot)_{\mathcal{H}}$ 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

$$\sum_{i\in\mathcal{I}} \|a_i\|^2_{C(J)} < \infty \tag{4.1}$$

the equality[14]

$$\sum_{i\in\mathcal{I}} a_i(t)(\zeta(t), \varphi_i)_{\mathcal{H}} - b(t) = 0 \quad \text{in } L^1(J) \tag{4.2}$$

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

[12] We emphasise that (H_4) can be regarded as a condition on the law ℓ of the random variables η_k, since any other random variable with law ℓ has the same structure as η_k.

[13] 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 4.1, we have a similar property for functions: the left-hand side of (4.2) defines an affine function, and if it vanishes on ζ, then it must be zero.

[14] The convergence in $L^1(J)$ of the series in (4.2) (when $\mathcal{I}$ is infinite) follows easily from (4.1).

We now fix a finite- or infinite-dimensional Hilbert space $\mathcal{H} \subset H^3$ and assume that ℓ satisfies the following hypothesis, which is not very restrictive in view of the examples given in Section 5. In particular, as we show in Section 5.2, it holds for the coloured noise (0.3), (0.4).

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

Remark 4.2. If $\dim \mathcal{H} < \infty$, then property of Lipschitz-observability 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.

Given a subspace $\mathcal{H} \subset H^3$, we define a non-decreasing sequence of closed subspaces $\mathcal{H}_k \subset H^3$, $k \geq 0$, 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 linear space, formed by all vectors of the form

$$\eta + \sum_{l=1}^{n} Q(\zeta_l, \xi_l), \quad n = 1, 2, \dots, \tag{4.3}$$

  where $Q(\zeta, \xi) = \Pi(\langle \zeta, \nabla\rangle \xi + \langle \xi, \nabla\rangle \zeta)$, and for $l = 1, \dots, n$ the vectors $\eta, \eta_l \in \mathcal{H}_k$ and $\xi_l \in \mathcal{H}$ are such that $Q(\zeta_l, \xi_l) \in H^3$.

Definition 4.3. 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.[15]

Theorem 4.4. *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*

$$\sum_{i\in\mathcal{I}} \|\varphi_i\|_1^2 < \infty, \tag{4.4}$$

and let η be a random process of the form (1.11) whose law is decomposable and satisfies the observability hypothesis. 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 (1.8).

[15] 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.

Remark 4.5. Even though the convergence to the stationary measure is formulated in Theorem 4.4 for initial functions supported by X, the exponential convergence to μ_ν holds for any initial measure with a finite first moment. Namely, there are positive numbers C and γ such that, for any initial measure $\lambda \in \mathcal{P}(H)$ satisfying the condition $\mathfrak{m}(\lambda) := \int_H \|u\|\lambda(\mathrm{d}u) < \infty$, we have

$$\|\mathfrak{P}_k^*\lambda - \mu_\nu\|_L^* \le C_1 e^{-\gamma_1 k}\big(1 + \mathfrak{m}(\lambda)\big) \quad \text{for } k \ge 0. \tag{4.5}$$

The validity of (4.5) can be derived easily from (1.8) using the dissipativity and regularisation properties of the Navier–Stokes dynamics (e.g., see Section 3.2.4 in [KS12]), and we shall not dwell on it.

Furthermore, denoting by $\mu_\nu(t)$ the law of a solution for (1.10) issued from an initial condition u_0 that is independent of η and is distributed as μ_ν, we note that $\mu_\nu(t+1) = \mu_\nu(t)$ for any $t \ge 0$, and for any other solution $u(t)$ of (1.10) we have

$$\big\|\mathcal{D}(u(t)) - \mu_\nu(t)\big\|_L^* \le C_2 e^{-\gamma_2 t}\big(1 + \mathfrak{m}(u(0))\big), \quad t \ge 0.$$

Both these claims follow immediately from the fact that, for any integer $k \ge 0$ and any $s \in (0,1)$, the function $u(k+s)$ with can be written as a Lipschitz-continuous function of the independent random variables $u(k)$ and η_{k+1} (cf. Section 3 in [Shi19]).

The coloured noises (0.4) with $c_j = Cj^{-q}$ for some $q > 1$ are decomposable, and as we show below, they satisfy the observability hypothesis. Therefore the result above implies the assertion of Theorem 0.1, as well as inequality (4.5).

Proof. In view of Theorem 1.1, it suffices to check Hypotheses (H_1)–(H_4), where H is defined by (1.9), 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]); for the reader's convenience, we presnt a complete proof of analyticity of S in Section 7.3. Condition (H_2) with $\hat\eta = 0$ and $\hat u = 0$ follows immediately from the dissipativity of the homogeneous Navier–Stokes system. The decomposability condition (H_4) for ℓ holds by assumption. 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. (1.10) 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. (1.10) with the initial condition u at time $t = 0$ and let $R^{\tilde u}(t,s) : H \to H$ (with $0 \le s \le t \le 1$) be the two-parameter process solving the linearised problem

$$\dot v + \nu L v + Q(\tilde u(t))v = 0, \quad v(s) = v_0, \tag{4.6}$$

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 (4.6). Denote by $A(u,\eta)$ the operator $D_\eta S(u,\eta)$. This is the resolving operator for the non-homogeneous equation (4.6)

(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. $\eta \in \mathcal{K}$. We claim that the density of $\mathfrak{L}(u,\eta)$ is equivalent to the triviality of the kernel for the operator [16]

$$G^{\tilde{u}} := \int_0^1 R^{\tilde{u}}(1,t)\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(1,t)^* \mathrm{d}t, \tag{4.7}$$

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; cf. Theorem 2.5 in [Zab08, Part IV]. Indeed, let us introduce an operator $A_1(u,\eta) : L^2(J,H) \to H$ by the relation $A_1(u,\eta) = A(u,\eta)\mathsf{P}_{\mathcal{H}}$ and note that the image $\mathfrak{L}(u,\eta)$ of the operator $A(u,\eta)$ is dense in H if and only if so is the image of $A_1(u,\eta)$. The latter property is equivalent to the triviality of the kernel of the adjoint operator $A_1(u,\eta)^* : H \to L^2(J,H)$. Since $(A(u,\eta)^* g)(t) = \mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(1,t)^* g$ for $g \in H$, we see that $A_1(u,\eta)^* g = 0$ for some $g \in H$ if and only if $G^{\tilde{u}} g = 0$.

Step 2. Description of $R^{\tilde{u}}(1,t)^*$. Together with (4.6), let us consider the dual problem

$$\dot{w} - \nu L w - Q^*(\tilde{u}(t))w = 0, \quad w(1) = w_0, \tag{4.8}$$

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 (4.8) is a linear backward parabolic equation, and for any $w_0 \in H$, it possesses a unique solution $w \in L^2(J,H^1) \cap W^{1,2}(J,H^{-1})$. It is well known, and can be proved easily, that w can be written as

$$w(t) = R^{\tilde{u}}(1,t)^* w_0. \tag{4.9}$$

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, ℓ-almost surely, any element of $\mathrm{Ker}(G^{\tilde{u}})$ is orthogonal to $\mathcal{H}_k$ for all $k \geq 0$. The rest of the proof follows a rather common argument used in the Malliavin calculus to prove the almost sure non-degeneracy of Malliavin's matrix; see the proof of Theorem 2.3.2 in the book [Nua06] and the paper [MP06, Sections 3–5]. The main difference is that we use the observability of the measure ℓ, rather than fine properties of functionals of the Wiener process.

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 (4.7), we have

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

[16] The operator $G^{\tilde{u}}$ is called the *Gramian* of the linear control problem associated with (4.6). 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].

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

$$(\zeta, R^{\tilde{u}}(1,t)^* w_0) = 0 \quad \text{for } t \in J, \tag{4.10}$$

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 (4.10) holds for all $\zeta \in \mathcal{H}_k$ with 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 (4.10) holds for all $\zeta \in \mathcal{H}_{k+1}$.

For any given $\zeta \in \mathcal{H}_k$, differentiating (4.10) in time and using (4.8) and (4.9), we derive[17]

$$\big(\nu L\zeta + Q(\tilde{u}(t))\zeta, w(t)\big) = 0 \quad \text{for } t \in J, \tag{4.11}$$

where w is given by (4.9). If we set

$$y(t) = \tilde{u}(t) - \int_0^t \eta(s)\mathrm{d}s = \tilde{u}(t) - \sum_{i\in\mathcal{I}} \varphi_i(x) \int_0^t \eta^i(s)\,\mathrm{d}s, \tag{4.12}$$

where $\eta^i(t) = (\eta(t), \varphi_i)_{\mathcal{H}}$, then (4.11) 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

$$b(t) + \sum_{i\in\mathcal{I}} a_i(t)\eta^i(t) = 0 \quad \text{for all } t \in J. \tag{4.13}$$

The function b is continuous, and $\{a_i, i \in \mathcal{I}\}$ are differentiable. Moreover, condition (4.4) implies that (4.1) holds. By observability of ℓ, it follows that $a_i \equiv 0$ for all $i \in \mathcal{I}$, whence $w(1) = w_0$ is orthogonal to all vectors (4.3) with $\eta = 0$. Since by the above w_0 is orthogonal also to $\mathcal{H}_k$, we conclude that it is orthogonal to the whole space $\mathcal{H}_{k+1}$. This completes the proof of Theorem 4.4. □

[17] Let us check briefly that the regularity of $\tilde{u}$ and w is sufficient to justify (4.11). Indeed, as was mentioned above, the function w belongs to $W^{1,2}(J, H^{-1})$, so that the L^2-scalar product $(\zeta, w(t))$ can be differentiated for any $\zeta \in H^1$. Resolving the first equation in (4.8) with respect to $\dot{w}$ and substituting the resulting expression into the time derivative of (4.10), we obtain $(\zeta, \nu Lw + Q(\tilde{u})^* w) = 0$, where the equality is valid for almost every $t \in J$. Since $\tilde{u} \in L^2(J, H^3)$ and $w \in L^2(J, H^1)$, we can integrate by parts in the last equality, which gives the validity of equality (4.11) for almost every $t \in J$. Finally, recalling that $\tilde{u} \in C(J, H^2)$ and $w \in C(J, L^2)$, we see that the right-hand side is a continuous function of time and, hence, vanishes on J.

In conclusion of this subsection, let us note that Theorem 4.4 remains valid in the case when the Navier–Stokes system is studied on a two-dimensional sphere[18], and the space $\mathcal{H}$ contains sufficiently many spherical harmonics. Indeed, all the arguments of the proof are exactly the same as in the case of the torus, except the verification of the saturating property. However, the construction of the successive subspaces $\mathcal{H}_k$ coincides with that used by Agrachev and Sarychev in [AS08, Theorems 6.1] to study the controllability problem. The fact that sufficiently many spherical harmonics form a saturating subspace is established in [AS08, Theorem 10.4].

4.2 Complex Ginzburg–Landau equation

We now turn to Eq. (0.2). To simplify the presentation, we confine ourselves to the more complicated case $m = 2$. 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 consider the equation

$$\partial_t u - (\nu + i)\Delta u + \gamma u + ic|u|^4 u = \eta(t, x), \quad x \in \mathbb{T}^3_a. \tag{4.14}$$

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 (1.11), 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 (4.14) is supplemented with the initial condition

$$u(0) = u_0 \in H^1. \tag{4.15}$$

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

$$(u, v) = \operatorname{Re} \int_{\mathbb{T}^3_a} u\bar{v} \, \mathrm{d}x \tag{4.16}$$

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 (4.14)) possesses a compact absorbing set $X \subset H^1$, that is closed and bounded in H^2 (cf. Theorem 7.4).

Let us introduce a concept of a *saturating subspace* for (4.14). 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 $\mathcal{H}_k \subset H^2$, $k \geq 0$, by the following rule:

- $\mathcal{H}_0$ coincides with $\mathcal{H}$;

[18] A similar remarks applies to the case of a rectangular domain with Lions boundary condition; see Section 9 in [AS08]. However, we shall not elaborate on that point since the analysis is similar.

- 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 4.6. The subspace $\mathcal{H} \subset H^2$ is said to be *saturating* if the union of $\{\mathcal{H}_k\}_{k\geq 0}$ is dense in L^2.

Let us write $B(u) = ic|u|^4u$ for the nonlinear term in (4.14) 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:

$$Q(u)v = ic\big(3|u|^4v + 2|u|^2u^2\bar{v}\big).$$

Theorem 4.7. *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 (1.11) 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 hypothesis (H_4).*

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-observable 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 (1.8).

Remark 4.8. Theorem 4.7 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. Moreover, an analogue of Remark 4.5 is valid under the hypotheses of Theorem 4.7: the exponential convergence to the stationary measure holds for any initial measure with a finite moment. Finally, for equations with a higher-order nonlinearity for which the well-posedness can only be proved for essentially bounded (rather than square-integrable) functions $\eta : J \to H^2$, one can apply a slight modification of Theorem 1.1, requiring the map S to be defined and analytic in an L^2-open neighbourhood of $L^\infty(J, H^2)$. The latter can be obtained by applying the inverse mapping theorem at any point of the space $L^\infty(J, H^2)$; cf. Section 7.4 and Section 2.3 in [KZ18].

Proof of Theorem 4.7. As in the case of the Navier–Stokes system, it suffices to check Hypotheses (H_1)–(H_4), where $H = H^1$ and $E = L^2(J, H^2)$. The proof of regularity properties for the resolving operator is carried by well-known methods and is presented in Section 7.4. The decomposability condition (H_4) for ℓ holds by assumption. Thus, we only need to check Conditions (H_2) and (H_3).

Step 1: Dissipativity. Let us recall that $\|\cdot\|$ and $\|\cdot\|_1$ denote the standard norms in L^2 and H^1, respectively. Given a number $\delta > 0$, we define a new norm $|\cdot|_\delta$ on H^1 by the formula $|u|_\delta^2 = \|u\|^2 + \delta\|u\|_1^2$. We claim that, for any bounded subset $B \subset H^1$, there are $\delta > 0$ and $a \in (0,1)$ such that

$$|S(u,0)|_\delta \le a|u|_\delta \quad \text{for } u \in B. \tag{4.17}$$

This will imply the validity of (H_2) with $\hat{u} = 0$ and $\hat{\eta} = 0$, since we can assume from the very beginning that the space H^1 is endowed with the norm $|\cdot|_\delta$.

To establish (4.17), we first note that

$$\|S(u)\| \le q\|u\| \quad \text{for } u \in L^2, \qquad \|S(u)\|_1 \le C_B\|u\| \quad \text{for } u \in B, \tag{4.18}$$

where $S(u) = S(u,0)$, $q \in (0,1)$, and $C_B > 0$ depends only on B. Indeed, the first inequality is a standard estimate for the L^2-norm of solutions, while the second can easily be derived by taking the derivative of $t\|\nabla u\|^2$; cf. the proof of Theorem 7.4. Combining inequalities (4.18), we derive

$$|S(u)|_\delta^2 = \|S(u)\|^2 + \delta\,\|S(u)\|_1^2 \le q^2\|u\|^2 + \delta\, C_B^2\|u\|^2 \le \left(q^2 + \delta C_B^2\right)|u|_\delta^2.$$

Choosing $\delta > 0$ so small that $q^2 + \delta C_B^2 < 1$, we obtain inequality (4.17) with some $a < 1$.

Step 2: Approximate controllability. We need to prove 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 some functions $u \in X$ and $\eta \in E$, denote by $\tilde{u} \in L^2(J, H^3) \cap W^{1,2}(J, H^1)$ the solution of (4.14) issued from u, and consider the linearised problem

$$\dot{v} + Lv + Q(\tilde{u})v = g, \quad v(0) = v_0, \tag{4.19}$$

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 (4.19). 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 (4.19) 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 7.2), 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 (4.19) with $g \equiv 0$ and let

$$G^{\tilde{u}} := \int_0^T R^{\tilde{u}}(T,t)\mathsf{P}_{\mathcal{H}}R^{\tilde{u}}(T,t)^*\mathrm{d}t, \tag{4.20}$$

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 $k \geq 0$.

To this end, we fix $u \in X$ and take any $w_0 \in \mathrm{Ker}(G^{\tilde{u}})$. In view of (4.20), we have $\mathsf{P}_{\mathcal{H}} R^{\tilde{u}}(T,t)^* w_0 = 0$ for any $t \in J'$, so that

$$(\zeta, R^{\tilde{u}}(T,t)^* w_0) = 0 \quad \text{for } t \in J', \tag{4.21}$$

where $\zeta \in \mathcal{H}_0$ is an arbitrary vector. Assuming that relation (4.21) 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 (4.21) in time and using an analogue of relation (4.9) for (4.19), we derive

$$\big(L\zeta + Q(\tilde{u}(t))\zeta, w(t)\big) = 0 \quad \text{for } t \in J', \tag{4.22}$$

where $w(t)$ is given by (4.9). 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 (4.22) in time and using (4.14), we obtain

$$\big(L\zeta + Q(\tilde{u})\zeta, \dot{w}\big) - \big(B_2(\tilde{u}; \zeta, L\tilde{u} + B(\tilde{u})), w\big) + \sum_{i \in \mathcal{I}} \big(B_2(\tilde{u}; \zeta, \varphi_i), w\big)\eta^i(t) = 0, \tag{4.23}$$

where $B_k(u; \cdot)$ stands for the kth derivative of $B(u)$ (so that $Q = B_1$, and $B_k = 0$ for $k \geq 6$). We thus obtain relation (4.13), 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. (4.22))

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

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

$$\big(B_5(\zeta, \varphi_i, \varphi_j, \varphi_m, \varphi_n), w(t)\big) = 0 \quad \text{for } i, j, m, n \in \mathcal{I},\ t \in J', \tag{4.24}$$

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

$$B_5(\zeta, \varphi, 1, 1, 1) = 12ic\,(3\zeta\varphi + \bar{\zeta}\bar{\varphi} + 3\bar{\zeta}\varphi + 3\zeta\bar{\varphi}). \tag{4.25}$$

Since both $\mathcal{H}$ and $\mathcal{H}_k$ are invariant under complex conjugation, it follows from relation (4.25) 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. $\square$

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 4.1. 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 case of an infinite-dimensional space $\mathcal{H}$. Let a_i be Lipschitz-continuous functions and b a continuous function satisfying (4.1) and (4.2). 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 (4.2) 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*

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

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 (4.2) 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 (4.2) 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 $|d_\varepsilon(s)| \le C_1\varepsilon$. Applying the matrix $R_\varepsilon^{-1}(s)$ and recalling (5.2), we obtain $|a(s)| \le C_2\varepsilon^{1-\theta}$, 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 4.1

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 $\{\psi_l^i\}_{l\ge1}$ 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 $i_1 \ne i_2$.

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

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

It is obvious that the law ℓ of η is a decomposable measure[19] 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

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

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

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

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 their 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 $\{\varphi_i, 1 \le i \le N\} \subset \mathcal{H}$ and consider the following process in $\mathcal{H}$:

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

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 the form $C_1 \log \frac{1}{\varepsilon}$ is fixed, then there are N points

$$\tau^j_{l_1} =: \tau_1, \quad \ldots, \quad \tau^j_{l_N} =: \tau_N, \quad l_r = l_0 + r,$$

that belong to the ε-neighbourhood of s. Consider the matrix $R_\varepsilon(s)$ corresponding to the points $\tau_1, \ldots, \tau_N$. Its entries are $r^i_m = \Delta\eta^i(\tau_m) = \Delta\eta^i(\tau^j_{l_m})$, where

$$\begin{aligned}\Delta\eta^i(\tau^j_l) &= b_i\Big(-2c_j\xi^i_{jl} + \sum_{r=1}^{\infty} c_{j+r}\big(\xi^i_{j+r,(2l+1)2^{r-1}} - \xi^i_{j+r,(2l+1)2^{r-1}-1}\big)\Big)\\ &= b_i A^{-j}\zeta^i_l(j). \end{aligned} \tag{5.9}$$

Here $\{\zeta^i_l(j)\}$ 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 3.4 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 algebraically decaying coefficients

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

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

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

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

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 $(\zeta_{l_m}^i(j))$. A simple calculation shows that $\{\zeta_{l_m}^i(j)\}$ 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

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

where $\rho : \mathbb{R} \to \mathbb{R}$ is a continuous function with compact support. Applying Corollary 3.4, 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))

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

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 now 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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} 2Q(c_l^a,s_r^a) &= \Pi\big(\langle l^{\perp_a},r\rangle_a r^{\perp_a} + \langle r^{\perp_a},l\rangle_a l^{\perp_a}\big)\cos\langle l+r,x\rangle_a \\ &\quad + \Pi\big(\langle l^{\perp_a},r\rangle_a r^{\perp_a} - \langle r^{\perp_a},l\rangle_a l^{\perp_a}\big)\cos\langle l-r,x\rangle_a, \end{aligned} \tag{6.5}$$

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

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 $v \in L^2(\mathbb{T}_a^2,\mathbb{R}^2)$ 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:

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

To prove (6.4), we write

$$\langle c_l^a,\nabla\rangle c_r^a = \langle l^{\perp_a},r\rangle_a r^{\perp_a}\cos\langle l,x\rangle_a\sin\langle r,x\rangle_a.$$

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 $|l^{\perp_a}| \neq |r^{\perp_a}|$, $c_r^a, s_r^a \in \mathcal{H}(\mathcal{I})$, 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 $\mathcal{H}_{k+1}(\mathcal{I})$. Indeed, a direct verification shows that

$$\begin{aligned} &P((l-r)^{\perp_a})\big(\langle l^{\perp_a},r\rangle_a r^{\perp_a} - \langle r^{\perp_a},l\rangle_a l^{\perp_a}\big) \\ &\qquad = |(l-r)^{\perp_a}|^{-2}\langle l^{\perp_a},r\rangle_a\big(|l^{\perp_a}|^2 - |r^{\perp_a}|^2\big)(l-r)^{\perp_a}. \end{aligned} \tag{6.7}$$

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

$$\begin{aligned} Q(c_l^a, 2\alpha c_r^a) + Q(s_l^a, 2\beta s_r^a) &= C_a(l,r)(\alpha+\beta)(l-r)^{\perp_a}\sin\langle l-r,x\rangle_a \\ &\quad - C_a(l,-r)(\alpha-\beta)(l+r)^{\perp_a}\sin\langle l+r,x\rangle_a. \end{aligned} \tag{6.8}$$

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 $s_{l+r}^a, s_{l-r}^a \in \mathcal{H}_{k+1}(\mathcal{I})$. A similar argument using (6.2), (6.5), and (6.7) shows that $c_{l+r}^a, c_{l-r}^a \in \mathcal{H}_{k+1}(\mathcal{I})$.

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 $l \in \mathbb{Z}^2 \backslash \mathbb{Z}^2_{\mathcal{I}}$. 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

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

where $\langle l, x\rangle_a = \sum_i a_i^{-1} l_i x_i$. We now set $e^a_0 \equiv 1$ and

$$e^a_l(x) = \begin{cases} c^a_l(x) & \text{if } l_1 > 0 \text{ or } l_1 = 0,\ l_2 > 0 \text{ or } l_1 = l_2 = 0,\ l_3 > 0, \\ s^a_l(x) & \text{if } l_1 < 0 \text{ or } l_1 = 0,\ l_2 < 0 \text{ or } l_1 = l_2 = 0,\ l_3 < 0. \end{cases}$$

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

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

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 $\mathcal{I} = \{(0, 0, 0), (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)\}$ 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 $c^a_r, s^a_r \in \mathcal{H}(\mathcal{I})$ 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 $\mathcal{H}_{k+1}(\mathcal{I})$.

To see this, let us note that

$$c^a_l(x)c^a_r(x) = \frac{1}{2}\big(c^a_{l-r}(x) + c^a_{l+r}(x)\big), \quad s^a_l(x)s^a_r(x) = \frac{1}{2}\big(c^a_{l-r}(x) - c^a_{l+r}(x)\big).$$

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

7 Appendix

7.1 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 $\mathbb{R}$. A proof of the following lemma can be found in [BM19a, Corollary 3.6] (see also Lemma 7.6 in [Vil03] and Corollary 4 in [BM19b]).

Theorem 7.1. *Let X, Y, and Z be Polish spaces and let $\{\mathfrak{M}_w\}$ and $\{\mathfrak{N}_w\}$ be RPM 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

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

7.2 Density of sets of solutions for a linear parabolic PDE

Retaining the notation of Section 2.40, let us consider the following linear parabolic PDE on $J = [0, 1]$:

$$\dot{v} - \mu\Delta v + b(t, x)v + c(t, x)\bar{v} = 0, \quad x \in \mathbb{T}^3_a, \tag{7.1}$$

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 (7.1) 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 (7.1) by a starting control.

Proposition 7.2. *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. (7.1) 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 (7.1) 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

$$\dot{w} + \bar{\mu}\Delta w - L^{-1}(\bar{b}Lw) - L^{-1}(cL\bar{w}) = 0, \quad w(1) = w_1, \tag{7.2}$$

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

$$0 = \big(R(1,0)v_0, w_1\big)_1 = \big(v_0, R(1,0)^* w_1\big)_1 \quad \text{for any } v_0 \in H^1, \tag{7.3}$$

whence we conclude that $R(1,0)^* w_1 = 0$. Thus, the solution $w(t)$ of problem (7.2) vanishes at $t = 0$. By the backward uniqueness for (7.2) (e.g., see Section II.8 in [BV92]), we conclude that $w_1 = 0$. This completes the proof of the proposition. □

7.3 Resolving operator for the Navier–Stokes system

Let us consider the Navier–Stokes equations (1.10) on a 2D torus $\mathbb{T}^2_a$, where $a = (a_1, a_2)$ is an arbitrary vector with positive coordinates. Given any $T > 0$, we define the Hilbert spaces

$$\mathcal{X}_T = L^2(J_T, H^1_\sigma) \cap W^{1,2}(J_T, H^{-1}_\sigma), \quad \mathcal{Y}_T = L^2(J_T, H^{-1}_\sigma),$$

where $J_T = [0, T]$, and H^s_σ stands for the space of divergence-free vector functions on $\mathbb{T}^2_a$ whose components belong to the Sobolev space of order $s \in \mathbb{Z}$. It is a well-known fact that, for any $u_0 \in L^2_\sigma$ (where $L^2_\sigma = H^0_\sigma$) and $\eta \in \mathcal{Y}_T$, Eq. (1.10) has a unique solution $u \in \mathcal{X}_T$ satisfying the initial condition

$$u(0, x) = u_0(x). \tag{7.4}$$

We shall denote the solution of (1.10), (7.4) by $\mathcal{R}(u_0, \eta)$, so that $\mathcal{R}$ is a map from $L^2_\sigma \times \mathcal{Y}_T$ to $\mathcal{X}_T$. The following result, which is essentially established in [Kuk82] (see also [VF88, Chapter 1]), implies, in particular, that for any fixed $u_0 \in L^2_\sigma$, the map taking η to $u(1)$ is analytic from $\mathcal{Y}_T$ to L^2_σ.

Theorem 7.3. *The map $\mathcal{R} : L^2_\sigma \times \mathcal{Y}_T \to \mathcal{X}_T$ is analytic.*

Proof. We shall apply an analytic implicit function theorem (see Section 3.3B in [Ber77]). Let us denote by $z \in \mathcal{X}_T$ the solution of the (linear) Stokes problem

$$\partial_t z + \nu L z = \eta, \quad z(0) = u_0. \tag{7.5}$$

A solution of (1.10), (7.4) is sought in the form $u = z + A\zeta$, where $\zeta \in \mathcal{Y}_T$ is an unknown function and A stands for the resolving operator of problem (7.5) with $u_0 \equiv 0$. Substituting this expression into (1.10), we derive the following functional equation for ζ in the space $\mathcal{Y}_T$:

$$F(\zeta, z) := \zeta + B(z + A\zeta) = 0. \tag{7.6}$$

This is a functional equation in $\mathcal{Y}_T$ defined by the analytic function F. Due to the unique solvability of the initial value problem for the Navier–Stokes system, we know that (7.6) has a unique solution $\zeta \in \mathcal{Y}_T$, and we only need to prove that ζ is an analytic function of $z \in \mathcal{X}_T$. By the analytic implicit function theorem, this

will be established if we prove that the linear operator $(D_\zeta F)(z,\zeta) : \mathcal{Y}_T \to \mathcal{Y}_T$ is invertible for any $(z,\zeta) \in \mathcal{X}_T \times \mathcal{Y}_T$.

To this end, we first note that

$$(D_\zeta F)(z,\zeta)\xi = \xi + B(z + A\zeta, A\xi) + B(A\xi, z + A\zeta), \quad \xi \in \mathcal{Y}_T,$$

where $B(\cdot,\cdot)$ is the bilinear term in the Navier–Stokes system. By Banach's inverse mapping theorem, the required invertibility of $D_\zeta F$ will be established if we prove that, for any $f \in \mathcal{Y}_T$, the equation

$$\xi + B(z + A\zeta, A\xi) + B(A\xi, z + A\zeta) = f \tag{7.7}$$

has a unique solution $\xi \in \mathcal{Y}_T$. Setting $w = A\xi$, we obtain the following problem for w:

$$\partial_t w + \nu L w + B(z + A\zeta, w) + B(w, z + A\zeta) = f, \quad w(0) = 0. \tag{7.8}$$

This is a linear Navier–Stokes-type problem, and its unique solvability in $\mathcal{X}_T$ can be established by standard arguments. It remains to note that $\xi = (\partial_t + \nu L)w$ is an element of $\mathcal{Y}_T$, which satisfies (7.7). □

7.4 Resolving operator for the complex Ginzburg–Landau equation

Let us consider the Ginzburg–Landau equation (4.14), supplemented with the initial condition (4.15). For a given $T > 0$, we define the spaces

$$\mathcal{X}_T = L^2(J_T, H^2) \cap W^{1,2}(J_T, L^2), \quad \mathcal{Y}_T = L^2(J_T, H^1),$$

where H^s is the Sobolev space of order s on $\mathbb{T}^3_a$. Recall that, even though all the functions of this subsection are complex-valued, we regard H^s as a real Hilbert space, so that the nonlinear term in (4.14) defines a real-analytic map in H^2. The following theorem establishes the well-posedness of the Cauchy problem for the Ginzburg–Landau equation and the analyticity of its resolving operator.

Theorem 7.4. *For $u_0 \in H^1$ and $\eta \in \mathcal{Y}_T$, problem (4.14), (4.15) has a unique solution $u \in \mathcal{X}_T$, and the map $\mathcal{R}(u_0, \eta)$ acting from $H^1 \times \mathcal{Y}_T$ to $\mathcal{X}_T$ is analytic. Moreover, for any $\rho > 0$ there is $R > 0$ such that, if $\eta : \mathbb{R}_+ \to H^2$ is a locally square-integrable function, and*

$$\|\eta\|_{L^2([k-1,k],H^2)} \le \rho \quad \text{for all } k \ge 1, \tag{7.9}$$

then the corresponding solution $u(t)$ belongs to H^2 for $t > 0$ and satisfies the inequality

$$\|u(t)\|_{H^2} \le R \quad \text{for } t \ge T_0, \tag{7.10}$$

where $T_0 > 0$ is a number depending only on the H^1-norm of u_0.

Proof. We shall confine ourselves to the derivation of some apriori estimates for solutions and to the proof of uniqueness of a solution and its analytic dependence on data.

Step 1: Bound for the energy. Let us define the Hamiltonian

$$\mathscr{H}(u) = \int_{\mathbb{T}^3} \big(\tfrac{1}{2}|\nabla u|^2 + \tfrac{c}{6}|u|^6\big)\mathrm{d}x$$

and note that it is a Fréchet differentiable function on H^1 with the derivative

$$(D\mathscr{H})(u;\xi) = (-\Delta u + c|u|^4 u, \xi),$$

where the inner product $(\cdot,\cdot)$ is defined by (4.16). It follows that, for any solution $u \in \mathcal{X}_T$ of (4.14), we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(u(t)) &= (-\Delta u + c|u|^4 u, \partial_t u)\\ &= (-\Delta u + c|u|^4 u, (\nu + i)\Delta u - \gamma u - ic|u|^4 u + \eta)\\ &= -\nu\|\Delta u\|^2 + \nu c\,(|u|^4 u, \Delta u) - \gamma\|\nabla u\|^2 - \gamma c\|u\|_{L^6}^6\\ &\quad + (\nabla u, \nabla\eta) + c(|u|^4 u, \eta). \end{aligned} \tag{7.11}$$

Simple calculation combined with the continuity of the embedding $H^1 \subset L^6$ implies that

$$\begin{aligned}&(|u|^4 u, \Delta u) = -(3|u|^4\nabla u, \nabla u) - (2|u|^2 u^2 \nabla\bar{u}, \nabla u) \le -\big(|u|^4, |\nabla u|^2\big),\\ &(\nabla u, \nabla\eta) \le \|\nabla u\|\,\|\nabla\eta\| \le \frac{\gamma}{2}\|\nabla u\|^2 + \frac{1}{2\gamma}\|\eta\|_1^2,\\ &(|u|^4 u, \eta) \le \|u\|_{L^6}^5\|\eta\|_{L^6} \le C\mathscr{H}(u)^{5/6}\|\eta\|_1.\end{aligned}$$

Substituting these inequalities in (7.11) and carrying out some transformations, we derive

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(u) \le -\delta\big(\mathscr{H}(u) + \|\Delta u\|^2 + (|u|^4, |\nabla u|^2)\big) + K\big(\mathscr{H}(u)^{5/6}\|\eta\|_1 + \|\eta\|_1^2\big), \tag{7.12}$$

where δ and K are some positive numbers. We now need the following version of Gronwall lemma established in the end of this subsection.

Lemma 7.5. *Let $\varphi : J_T \to \mathbb{R}$ be an absolutely continuous function minorised by 1 such that*

$$\dot{\varphi} \le -a\varphi + b + C(t)\varphi^r \quad \textit{almost everywhere on } J_T, \tag{7.13}$$

where a, b, and $r < 1$ are some positive numbers, and $C : J_T \to \mathbb{R}_+$ is an integrable function. Then there is a number c_r depending only on r such that

$$\varphi(t) \le c_r e^{-at}\varphi(0) + c_r\left(\frac{b}{a} + (1-r)\int_0^t e^{-a(1-r)(t-s)}C(s)\mathrm{d}s\right)^{(1-r)^{-1}}. \tag{7.14}$$

Let us set $\varphi(t) = \mathscr{H}(u(t)) + 1$. It follows from (7.12) that φ satisfies inequality (7.13), in which $a = b = \delta$, $r = 5/6$, and $C(t) = K(\|\eta\|_1 + \|\eta\|_1^2)$. Applying Lemma 7.5, we derive

$$\mathscr{H}(u(t)) \le C_1 e^{-\delta t}\big(\mathscr{H}(u_0) + 1\big) + C_1\bigg(1 + \int_0^t e^{-\delta(t-s)/6}\big(\|\eta\|_1 + \|\eta\|_1^2\big)\mathrm{d}s\bigg)^6, \tag{7.15}$$

where $t \ge 0$ is arbitrary, and $C_1 > 0$ is a number not depending on the solution. In particular, if (7.9) is satisfied, then the energy $\mathscr{H}(u(t))$ of any solution is bounded on $\mathbb{R}_+$, and there is a number $R_* > 0$ such that

$$\mathscr{H}(u(t)) \le R_* \quad \text{for } t \ge T_*, \tag{7.16}$$

where $T_* > 0$ depends only on $\|u_0\|_1$. Moreover, integrating (7.12) in time and using (7.15), we can find $C_2 > 0$ depending only on ρ such that

$$\int_0^t \big(\|\Delta u\|^2 + (|u|^4, |\nabla u|^2)\big)\,\mathrm{d}s \le C_2(1+t)\mathscr{H}(u_0) \quad \text{for } t \ge 0. \tag{7.17}$$

Step 2: Uniqueness. If $u_1, u_2 \in \mathcal{X}_T$ are two solutions corresponding to the same data $(u_0, \eta) \in H^1 \times \mathcal{Y}_T$, then their difference $u = u_1 - u_2$ satisfies the equations

$$\partial_t u - (\nu + i)\Delta u + \gamma u + ic\big(|u_1|^4 u_1 - |u_2|^4 u_2\big) = 0, \quad u(0) = 0.$$

Taking the inner product in L^2 of the first equation with $2u$, noting that

$$\big||u_1|^4 u_1 - |u_2|^4 u_2\big| \le C_3\big(|u_1|^4 + |u_2|^4\big)|u|,$$

and carrying out some simple transformations, we derive

$$\partial_t \|u\|^2 \le -\gamma\|u\|^2 - \nu\|\nabla u\|^2 + C_4\big(\|u_1\|_{L^\infty}^4 + \|u_2\|_{L^\infty}^4\big)\|u\|^2. \tag{7.18}$$

Since $\|u_i\|_{L^\infty}^2 \le C_5\|u_i\|_1\|u_i\|_2$ and

$$\|u_i\|_{L^1(J_T,L^\infty)}^4 \le C_5\|u_i\|_{C(J_T,H^1)}^2\|u_i\|_{L^2(J_T,H^2)}^2,$$

we see that the coefficient in front of $\|u\|^2$ is an integrable function of time. Applying Gronwall's inequality to (7.18), we conclude that $u \equiv 0$.

Step 3: Analyticity. To prove the analyticity of the resolving operator for (4.14), (4.15), we repeat the scheme used in the case of the Navier–Stokes system. Namely, let us denote by $z \in \mathcal{X}_T$ the solution of the linear problem

$$\partial_t z + \gamma z - (\nu + i)\Delta z = \eta, \quad z(0) = u_0, \tag{7.19}$$

and denote by $A : \mathcal{Y}_T \to \mathcal{X}_T$ its resolving operator when $u_0 = 0$. We seek a solution of (4.14), (4.15) in the form $u = z + A\zeta$, where $\zeta \in \mathcal{Y}_T$ is an unknown function. Substituting this expression into (4.14) and setting $B(u) = ic|u|^4 u$, we

see that ζ must be a solution of (7.6). We know that (7.6) has a unique solution $\zeta \in \mathcal{Y}_T$ for any $z \in \mathcal{X}_T$; cf. Steps 1 and 2. If we prove that F is an analytic map in $\mathcal{Y}_T$ whose derivative with respect to ζ is an invertible linear operator in $\mathcal{Y}_T$, then the required result will follow from the analytic implicit function theorem (see Section 3.3B in [Ber77]). The analyticity is a straightforward consequence of the facts that B is a continuous multilinear map from $\mathcal{X}_T$ to $\mathcal{Y}_T$ and A is a continuous linear operator from $\mathcal{Y}_T$ to $\mathcal{X}_T$. We thus need to prove the invertibility of the derivative $D_\zeta F$.

Let us note that

$$(D_\zeta F)(\zeta, z)\xi = \xi + 3|z + A\zeta|^4 A\xi + 2|z + A\zeta|^2 (z + A\zeta)^2 \overline{A\xi}, \quad \xi \in \mathcal{Y}_T.$$

We now fix an arbitrary $f \in \mathcal{Y}_T$ and consider the equation

$$\xi + 3|z + A\zeta|^4 A\xi + 2|z + A\zeta|^2 (z + A\zeta)^2 \overline{A\xi} = f.$$

Setting $g = z + A\zeta \in \mathcal{X}_T$ and $w = A\xi$, so that w vanishes at $t = 0$, we rewrite this equation in the form

$$\partial_t w + \gamma w - (\nu + i)\Delta w + a(t,x)w + b(t,x)\bar{w} = f, \quad w(0) = 0, \tag{7.20}$$

where $a = 3|g|^4$ and $b = 2|g|^2 g^2$ are elements of the space $L^2(J_T, L^\infty)$. It remains to note that (7.20) is a zero-order perturbation of the linear problem (7.19), and its unique solvability can easily be established by standard arguments.

Step 4: Bound for the H^2 norm. We now prove (7.10). In view of (7.16), it suffices to prove that the H^2-norm of a solution of (4.14), (4.15) can be bounded at time $t = 1$ in terms of the H^1-norm of u_0. To this end, we calculate the derivative of the function $h(t) = t\|\Delta u(t)\|^2$. A simple calculation shows that

$$\partial_t h \le \|\Delta u\|^2 + t\|\Delta u\|\big(\|\Delta B(u)\| + \|\eta\|_2\big), \tag{7.21}$$

where $B(u)$ was defined in Step 3. Now note that

$$\|\Delta B(u)\| \le C_6\big(\|u\|_{L^\infty}^4 + 1\big)\|\Delta u\| \le C_7\big(U(t) + 1\big)\|\Delta u\|,$$

where $U(t) = \|u\|_1^2 \|u\|_2^2$ is a locally integrable function of time. Substituting this inequality into (7.21), we obtain

$$\partial_t h \le C_8\big(U(t) + 1\big)h + \|\Delta u\|^2 + \|\eta\|_2^2.$$

Applying Gronwall's inequality and using (7.17), we arrive at the required estimate. This completes the proof of the theorem. □

Proof of Lemma 7.5. Let us set $\alpha = (1 - r)^{-1}$ and $\psi = \varphi^{1/\alpha}$. Then ψ is an absolutely continuous function on J_T. Substituting $\varphi = \psi^\alpha$ into (7.13), we derive

$$\alpha\psi^{\alpha-1}\dot{\psi} \le -a\psi^\alpha + b + C(t)\psi^{\alpha r}.$$

Dividing this inequality by $\alpha\psi^{\alpha-1}$ and using the relations $\alpha r - \alpha + 1 = 0$ and $\psi \geq 1$, we obtain

$$\dot{\psi} \leq -a(1-r)\psi + \big(b + C(t)\big)(1-r).$$

Application of the usual Gronwall inequality results in

$$\psi(t) \leq \big(e^{-at}\varphi(0)\big)^{1-r} + \frac{b}{a} + (1-r)\int_0^t e^{-a(1-r)(t-s)}C(s)\,\mathrm{d}s.$$

This implies the required inequality (7.14) with $c_r = 2^{r/(1-r)}$. □

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.

[AS08] ———, *Solid controllability in fluid dynamics*, Instability in Models Connected with Fluid Flows. I, Int. Math. Ser. (N. Y.), vol. 6, Springer, New York, 2008, pp. 1–35.

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

[Ber77] M. S. Berger, *Nonlinearity and Functional Analysis*, Academic Press, New York–London, 1977.

[Bil99] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1999.

[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.

[BM19a] V. Bogachev and I. Malofeev, *Kantorovich problems and conditional measures depending on a parameter*, Preprint (2019).

[BM19b] ———, *On the Kantorovich Problem with a Parameter*, Dokl. Math. **100** (2019), no. 1, 349–353.

[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.

[CT07] C. Cao and E. S. Titi, *Global well-posedness of the three-dimensional viscous primitive equations of large scale ocean and atmosphere dynamics*, Ann. of Math. (2) **166** (2007), no. 1, 245–267.

[DGTZ12] A. Debussche, N. Glatt-Holtz, R. Temam, and M. Ziane, *Global existence and regularity for the 3D stochastic primitive equations of the ocean and atmosphere with multiplicative white noise*, Nonlinearity **25** (2012), no. 7, 2093–2118.

[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.

[GKVZ14] N. Glatt-Holtz, I. Kukavica, V. Vicol, and M. Ziane, *Existence and regularity of invariant measures for the three dimensional stochastic primitive equations*, J. Math. Phys. **55** (2014), no. 5, 051504, 34.

[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.

[HMS11] M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259.

[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.

[KNS19] S. Kuksin, V. Nersesyan, and A. Shirikyan, *Mixing via controllability for randomly forced nonlinear dissipative PDEs*, Preprint (2019).

[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.

[KPS02] S. Kuksin, A. Piatnitski, and A. Shirikyan, *A coupling approach to randomly forced nonlinear PDEs. II*, Comm. Math. Phys. **230** (2002), no. 1, 81–85.

[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.

[Kuk02] S. Kuksin, *On exponential convergence to a stationary measure for nonlinear PDEs perturbed by random kick-forces, and the turbulence limit*, Partial differential equations, Amer. Math. Soc. Transl. Ser. 2, vol. 206, Amer. Math. Soc., Providence, RI, 2002, pp. 161–176.

[KZ18] S. Kuksin and H. Zhang, *Exponential mixing for dissipative PDEs with bounded non-degenerate noise*, Preprint (2018).

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

[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.

[Shi19] ———, *Controllability implies mixing II. Convergence in the dual-Lipschitz metric*, J. Eur. Math. Soc. (2019), accepted for publication.

[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.

[VF88] M. I. Vishik and A. V. Fursikov, *Mathematical Problems in Statistical Hydromechanics*, Kluwer, Dordrecht, 1988.

[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.