HAL
open science

EXPONENTIAL MIXING FOR DISSIPATIVE PDES WITH BOUNDED NON-DEGENERATE NOISE

Sergei Kuksin, Huilin Zhang

To cite this version:

Sergei Kuksin, Huilin Zhang. EXPONENTIAL MIXING FOR DISSIPATIVE PDES WITH BOUNDED NON-DEGENERATE NOISE. Stochastic Processes and their Applications, 2020, 130, pp.4721-4745. 10.1016/j.spa.2020.01.014 . hal-02922318

HAL Id: hal-02922318

https://hal.science/hal-02922318v1

Submitted on 26 Aug 2020

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 DISSIPATIVE PDES WITH BOUNDED NON-DEGENERATE NOISE

SERGEI KUKSIN AND HUILIN ZHANG

Abstract. We prove that well posed quasilinear equations of parabolic type, perturbed by bounded nondegenerate random forces, are exponentially mixing for a large class of random forces.

Contents

0. Introduction	1
1. Mixing for a class of Markov chains	4
1.1. Settings and assumptions	4
1.2. Main results	5
1.3. Main lemma and proof of Theorem 1.3	8
1.4. Proof of the main lemma	11
2. Applications	14
2.1. Random Haar series ("colored noises")	15
2.2. The 2D Navier-Stokes system	15
2.3. Quasilinear parabolic systems on $\mathbb{T}^d$	19
2.4. Appendix.	22
References	23

0. Introduction

In this paper we consider nonlinear PDEs, perturbed by random forces, which we write as

$$\partial_t u_t + \nu L u_t + F(u_t) = \eta_t, \quad u_t \in H. \tag{1}$$

Here H is a Hilbert space of functions of x, $\nu \in (0,1]$, L is a self-adjoint operator with compact resolvent, F is a nonlinearity and $\eta_t = \eta_t^\omega$ is a random process in H. We are concerned with the question when a solution u_t of (1) is a mixing random process, i.e. when its distribution $\mathcal{D}(u_t)$ converges to a unique measure in H, independent from the initial data u_0, while $t \to \infty$. The problem of mixing in equations (1) is well motivated by modern physics, and it has received much attention during the last two decades, see in [5]. In the corresponding papers (discussed in [5]) the authors prove the mixing for various classes of equations (1), assuming that the random force η_t has the structure which we will now discuss.

2010 *Mathematics Subject Classification.* 35Q59, 35Q56, 35K59, 37A25, 37L55, 60F05, 60H15, 76D05.

Key words and phrases. exponential mixing, 2d Navier-Stokes system, Ginzburg-Landau equation, dissipative PDEs, Markovian dynamic system, Haar colored noise.

We suppose that

$$\eta_t = \sum_{i=1}^{\infty} b_i \eta_t^i \phi_i, \qquad \sum b_i^2 < \infty,$$

where $\{\phi_i\}_i$ is an orthonormal basis of H and $\{\eta_t^i, i \geq 1\}$ are i.i.d. real processes, distributed as a certain etalon process η_t^0. Concerning the latter it was assumed that either this is
a) a kick-process, $\eta_t^0 = \sum_{k=1}^{\infty} \delta(t - kT)\xi_k$, where $T > 0$ and $\{\xi_k\}$ are i.i.d. real random variables; or that
b) η_t^0 is a white noise; or
c) η_t^0 is a compound Poisson process (see in [5]).
Concerning the coefficients b_i it was usually assumed that either all of them are non-zero, or that $b_i \neq 0$ for all $i \leq N_\nu$, where N_ν grows to infinity as $\nu \to 0$. In the paper [2] the mixing was established for the case when (1) is the 2D Navier-Stokes system on the two-dimensional torus, perturbed by a white in time random force (see below eq. (50)), and only a ν–independent finite system of these coefficients do not vanish. The proof of [2] uses an infinite–dimensional version of the Malliavin calculus, and all attempts to generalize it to equations with kick-forces or compound Poisson processes have failed.
Instead, in [6] equations of the form (1) were considered, where $\eta^0(t)$ is a random Haar series "of time-width one."[1] This means the following:
The processes $\eta^0 |_{[k-1,k)}$, $k = 1, 2, \dots$, are i.i.d., so it suffices to define the process $\eta^0 |_{[0,1]}$. The latter is a random Haar series

$$\eta_t^0 = \sum_{j=0}^{\infty} c_j \sum_{l=0}^{2^j-1} \xi_{jl} h_{jl}(t) \quad \forall\, 0 \leq t \leq 1, \qquad c_j \neq 0 \ \ \forall\, j,$$

where $\{h_{jl}\}$ is the Haar base in $L_2[0,1]$ (see in [7] and see (47) below), and ξ_{jl} are independent random variables, $|\xi_{jl}| \leq 1$, whose density functions $\rho_{jl}(x)$ are Lipschitz-continuous and do not vanish at $x = 0$. Let us denote $E = L_2([0,1], H)$ and consider the mapping

$$S : H \times E \to H, \quad S(u_0, \eta) = u_1, \tag{2}$$

where u_t is a solution of (1), equal u_0 at $t = 0$.
Assuming that
(B0) $\sum_j 2^{j/2}|c_j| < \infty$, *so the process* η_t *is bounded uniformly in* t *and* ω *(see below in Section 2.1);*
(B1) the mapping S *is well defined. Moreover, there exists a compactly embedded Banach subspace* $V \subset H$ *such that* $S(H \times E) \subset V$, *the mapping* $S : H \times E \to V$ *is analytic and its derivatives are bounded on bounded sets;*
(B2) $0 \in H$ *is an asymptotically stable equilibrium for eq.* (1) *with* $\eta = 0$, *and relation* (9) *holds;*
(B3) $b_j \neq 0$ *for* $j \in J \subset \mathbf{N}$, *where the set* J, *finite or infinite, is such that the linearised equation*

$$\partial_t v + \nu L v + dF(u_t)v = \sum_{j \in J} b_j \xi_t^j \phi_j, \quad 0 \leq t \leq 1,$$

is approximately controllable by controls ξ_t^j, $j \in J$, *provided that* u_t *is a solution of* (1) *with arbitrary* u_0 *and with* $\eta = \eta^\omega$, *where* ω *does not belong to a certain null-set depending on* u_0,
it was proved in [6] that eq. (1) is exponentially mixing. [2]

[1] the time-width one may be replaced by any positive width.
[2] our statement of this result is a bit imprecise and less general than the result of [6] is; see the original paper for the exact statement.

The condition (B3) has been intensively studied, and for many important equations it is now established that (B3) holds if J is a finite set, satisfying certain explicit conditions, independent from ν. See in [6].

Our work continues the research in [6]. Namely, making the assumptions (B0′) – (B3′), where
(B0′) $\sum_j c_j^2 2^j < \infty$;
(B1′) the mapping $S : H \times E \to V$ *is* C^2*–smooth and its derivatives up to second order are bounded on bounded sets;*
(B2′) = (B2);
(B3′) for any $u \in H$ *and* $\eta \in E$

$$\text{the operator } d_\eta S(u,\eta) : E \to H \text{ has dense image,} \tag{3}$$

we prove in Theorem 1.5, Section 1, the following result:

there exists a unique Borel measure μ_ν *in* H *such that if* u_t *is a solution of* (1) *with the initial data* u_0^ω*, satisfying* $u_0^\omega \in \{u \in H : \|u\| \le R\}$ *almost surely for some* $R > 0$*, then*

$$\|\mathcal{D}u_k - \mu_\nu\|_L^* \le C\kappa^k, \quad \forall\, k \in \mathbb{N}. \tag{4}$$

Here $\|\cdot\|_L^*$ *is the Lipschitz-dual distance in the space of Borel measures in* H *(see* (5)*), and* $C = C(R) > 0$*,* $\kappa = \kappa(R) \in (0,1)$*. Moreover, if* u_0^ω *is any random variable in* H*, then* $\mathcal{D}u_k$ *weakly converges to* μ_ν *as* $k \to \infty$*.*

A special case of systems (2) appears when E is a subspace of H and the mapping S has the form

$$S(u_0,\eta) = S_1(u_0) + \eta.$$

Such systems correspond to equations (1) with kick-forces η as in a) above (see Remark 2.7 in Sectuion 2.2). They are easier than general systems (2) and for them the mixing can be proved if the map S_1 is Lipschitz–continuous on bounded sets, see [5].

In Section 2 we show that the result applies to the 2D Navier-Stokes system on torus as well as to equations (1) who are random perturbations of quasilinear parabolic systems which are
– well posed for any bounded force $\eta(t,x)$, sufficiently smooth is x;
– satisfy the dissipativity assumption (B2).
The proof uses some ideas from [5, 6] and is significantly shorter than that in [6] since the nondegeneracy assumption (B3′) which we assume now is stronger than the assumption (B3); essentially it holds if the force η_t is non-degenerate, i.e. all coefficients $b_i \neq 0$, (unfortunately, we are not aware of a weaker sufficient condition which would guarantee (B3′)). As in [6], the exponential convergence (4) follows from Doeblin's coupling, enhanced with the quadratic convergence in the form, close to that in the KAM–theory.[3] In particular, our work shows that the analyticity assumption in (B1) is not an intrinsic feature of the approach of [6], but is needed to work under the very weak nondegeneracy assumption (B3) and may be replaced by the C^2–smoothness if the equation is non-degenerate in the sense (B3′). Furthermore, our result is easier to apply then that in [6] since it is easier to check assumption (B3′) than (B3) in [6].

Notation. As usual, by $C, C_1, \dots$ we denote various constants which change from line to line. By $B_F(R)$ we denote the closed R-ball in a Banach space F, centered at the origin; by $\mathcal{D}(\xi)$ – a law of a r.v. ξ; $\langle \mu, f\rangle = \langle f, \mu\rangle$ stands for the integral of a measurable function f against a measure μ, and $\mathcal{L}(X,Y)$ – for the space of bounded linear operators between Banach spaces X and Y. A

[3]The peculiarity of the used KAM–scheme is that now, in difference with the "traditional" KAM, the rate of convergence is not super-exponential, but only exponential. Cf. [6], Section 1.3.

complete separable metric space (i.e. a Polish space) X always is equipped with its Borel σ-algebra $\mathcal{B}_X$. By $\mathcal{P}(X)$ we denote the space of probability measures on $(X, \mathcal{B}_X)$. We provide it with the Lipschitz-dual distance

$$\|\mu-\nu\|_L^* = \|\mu-\nu\|_{L(X)}^* = \sup |\langle \mu-\nu, f\rangle|, \tag{5}$$

where the supremum is taken over all Lipschitz functions f on X such that

$$|f|_{L(X)} := \sup_{x\in X} |f(x)| \vee \mathrm{Lip}(f) \le 1,$$

with the Kantorovich distance

$$\|\mu-\nu\|_K = \sup_{\mathrm{Lip}(f)\le 1} |\langle \mu-\nu, f\rangle|, \tag{6}$$

and the distance of total variation $\|\cdot\|_{var}$. If X, Y are Polish spaces and $f : X \to Y$ is a Borel–measurable mapping, then $f_* : \mathcal{P}(X) \to \mathcal{P}(Y)$ denotes the corresponding push-forward of measures. A pair of random variables ξ, ζ defined on a probability space is *a coupling* for given measures $\mu, \nu \in \mathcal{P}(X)$ if $\mathcal{D}(\xi) = \mu, \mathcal{D}(\zeta) = \nu$.

Acknowledgements: S.K. thanks the Russian Science Foundation for support through the project 18-11-00032. H.Z. is supported by the National Postdoctoral Program for Innovative Talents, No: BX20180075 and China Postdoctoral Science Foundation.

1. Mixing for a class of Markov chains

1.1. **Settings and assumptions.** Let H, E be two separable Hilbert spaces and V be a Banach space, compactly and densely embedded in H. Let $S : H \times E \to V$ be a continuous mapping. We consider the following random dynamical system (RDS) in the space H:

$$u_k = S(u_{k-1}, \eta_k), \quad k \ge 1, \tag{7}$$

where $\{\eta_k\}_k$ is a sequence of i.i.d. random variables in E. We denote by ℓ the law of η_k,

$$\ell = \mathcal{D}(\eta_k) \in \mathcal{P}(E) \qquad \forall\, k,$$

and suppose that ℓ is supported by a compact set $\mathcal{K} \subset E$,

$$\mathcal{K} := \mathrm{supp}\, \ell \Subset E.$$

By $(u_k(v), k \ge 0)$, we denote a trajectory of (7) such that $u_0(v) = v$. The process $\{u_k(v)\}$ is a Markov chain in H, whose transition probability function after k steps is $P_k(u, \Gamma) = \mathbb{P}(u_k(u) \in \Gamma)$. It defines a semigroup of Markov operators in the space of functions

$$\mathcal{P}_k : C_b(H) \to C_b(H), \quad f(\cdot) \mapsto \mathbb{E}\big(f(u_k(\cdot))\big), \quad k \ge 0,$$

and a semigroup of operators in the space of measures

$$\mathcal{P}_k^* : \mathcal{P}(H) \to \mathcal{P}(H), \quad \mu \mapsto \mathcal{D}(u_k(v)), \quad k \ge 0,$$

where v is a r.v. in H, independent from the noise η, such that $\mathcal{D}(v) = \mu$ (e.g. see in [5]).

We make the following assumptions concerning regularity of our system:

(A1) (**Regularity**). *The mapping $S : H \times E \to V$ is twice continuously differentiable, and its derivatives up to second order are bounded on bounded sets.*

Concerning the noise η we assume the following:

(H1) (**Decomposability and non-degeneracy**). *There exists an orthonormal basis* $\{e_j, j \geq 1\}$ *of* E *such that*

$$\eta_k = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j, \quad b_j \neq 0 \, \forall \, j.$$

Here ξ_{jk} *are independent random variables and* b_j *are real numbers, satisfying*

$$|\xi_{jk}| \leq 1 \ \ a.s., \qquad \mathcal{D}(\xi_{jk}) = \rho_j(r)dr, \qquad R_\eta^2 := \sum_{j=1}^{\infty} b_j^2 < \infty, \tag{8}$$

where $\rho_j : \mathbb{R} \to \mathbb{R}$ *are Lipschitz functions and* $\rho_j(0) \neq 0$ *for all* j.

(H2) (**Dissipativity**). *For any* $u \in H$, $\eta \in \mathcal{K}$ *we have*

$$\|S(u,\eta)\|_H \leq \gamma \|u\|_H + \beta, \qquad \|S(u,0)\|_H \leq \gamma \|u\|_H \ \text{with some}\ 0 < \gamma < 1, \beta > 0. \tag{9}$$

(H3) (**Non-degeneracy**). *For any* $u \in H$ *and* $\eta \in \mathcal{K}$ *the image of the operator* $D_\eta S(u,\eta) : E \to V$ *is dense in* H.

Note that by (8) $\operatorname{supp} \ell = \mathcal{K}$ belongs to the Hilbert brick

$$\{\eta = \sum \eta_k e_k : |\eta_k| \leq b_k\} =: \tilde{\mathcal{K}}, \tag{10}$$

so indeed $\mathcal{K}$ is a compact set, and $\mathcal{K} \subset B_E(R_\eta)$.
It is easy to see that since $\rho_j(0) \neq 0$ for all j, then according to relations (9) and assumption (A1), for any $a > \gamma$ and $\delta > 0$,

$$\mathbb{P}\left(\|S(u,\eta)\|_H < a\|u\|_H\right) > p_\delta > 0, \quad \text{if } \|u\|_H \geq \delta. \tag{11}$$

Remark 1.1. *It is not necessary to define the map* S *on the whole space* $H \times E$*: it suffices that it is defined on a neighbourhood* Q *of* $H \times \mathcal{K}$ *which contains each point* $(u_0 \times \eta) \in B_H(R) \times \mathcal{K}$ *with its vicinity in* $H \times E$ *of a positive radius* $r(R)$*. The map* S *should be* C^2*–smooth and its* C^2*–norm should depend only on* $\|u\|_H$.

Remark 1.2. *Assume that the system possesses the following additional property: there exists a function* $R_+(R)$ *such that if* $\|u_0\|_H \leq R$ *and* $\eta_1, \eta_2, \dots$ *are arbitrary points in* $\mathcal{K}$*, then* $\|u_k\|_H \leq R_+(R)$ *for all* $k \geq 0$*. Then in order to study solutions of* (7) *with* $\|u_0\|_H \leq R$ *we can do the following:*
– define $O \subset H$ *as a union of all trajectories of* (7) *with* $\|u_0\|_H \leq R$ *and* $\eta_k \in \mathcal{K}$ *for all* k*;*
– consider the closure $\bar{O}$*. This is a closed subset of* $B_H(R_+)$*, invariant for* (7) *a.s.*
Then to study solutions with $\|u_0\|_H \leq R$ *we can work with system's restriction to* $\bar{O}$*. In particular it suffices to verify assumptions (A1) and (H3) for* $u \in B_H(R_+)$*. Even more, it suffices to check (A1) and (H3) for* $u \in B_H(R_+)$ *with the norm* $\|\cdot\|_H$ *replaced by any equivalent Hilbert norm* $\|\cdot\|'_H$ *in the space* H*, depending on* R_+ *(i.e., depending on* R*). Applying Remark 1.1 we observe that it suffices to verify (A1) and (H3) on the set* $Q \cap (B_H(R_+) \times E)$.

1.2. **Main results.** Here we formulate the main results of this paper. In what follows, by some confusion of notations, η stays for an element of $\mathcal{K}$ or for a random variable with the law ℓ, depending on the context. We will use a modified distance in the space H which depends on a parameter $d_0 \in (0,1]$:

$$\|\xi_1 - \xi_2\|_{d_0} := \|\xi_1 - \xi_2\|_H \wedge d_0.$$

Clearly $\|\xi_1-\xi_2\|_{d_0}\le\|\xi_1-\xi_2\|_H$ and

$$\|\xi_1-\xi_2\|_H\le\frac{2R_*}{d_0}\|\xi_1-\xi_2\|_{d_0}\quad\text{if}\quad \xi_1,\xi_2\in B_H(R_*),\text{with } 2R_*\ge d_0. \tag{12}$$

Theorem 1.3. *Under the assumptions (A1), (H1), (H2) and (H3), for any $R_*>0$ and any $u,u'\in B_H(R_*)$,*

$$\|P_k(u,\cdot)-P_k(u',\cdot)\|_L^*\le C\|u-u'\|_H\,\kappa^k, \tag{13}$$

where $0<\kappa<1$ and $C>0$ depend on R_*.

Corollary 1.4. *If $\mu,\nu\in\mathcal{P}(B_H(R_*))$, then*

$$\|\mathcal{P}_k^*\mu-\mathcal{P}_k^*\nu\|_L^*\le C_1\|\mu-\nu\|_L^*\,\kappa^k\quad\forall\,k\ge1, \tag{14}$$

where $0<\kappa<1$ and $C_1>0$ depend on R_*.

Proof. Let $(u^{\omega'},u'^{\omega'})$ be a coupling for (μ,ν). Then $\mathcal{P}_k^*(\mu)=\mathcal{D}u_k(u)$, $\mathcal{P}_k^*(\nu)=\mathcal{D}u_k(u')$. Take any $f\in C_b(H)$ with $|f|_{L(H)}\le1$. Then

$$|\langle f,\mathcal{P}_k^*(\mu)-\langle f,\mathcal{P}_k^*(\nu)\rangle|=\left|\mathbb{E}^{\omega'}\mathbb{E}^{\omega}(f(u_k(u^{\omega'}))-f(u_k(u'^{\omega'})))\right|\le C\kappa^k\mathbb{E}^{\omega'}\|u^{\omega'}-u'^{\omega'}\|_H,$$

where the inequality follows from the theorem. This inequality remains true if in the r.h.s. we take infimum over all couplings (u,u') for (μ,ν). By the Kantorovich–Rubinshtein theorem (see in [5])

$$\inf_{u,u'}\mathbb{E}^{\omega'}\|u^{\omega'}-u'^{\omega'}\|_H=\|\mu-\nu\|_K,$$

where $\|\cdot\|_K$ is the Kantorovich distance (6). Since ν,μ supported on $B_H(R_*)$, then $\|\mu-\nu\|_K\le R_*\|\mu-\nu\|_L^*$, and the result follows. □

Clearly without loss of generality we may assume that in Theorem 1.3

$$R_*\ge\frac{\beta}{1-\gamma}=:R_*^0\,; \tag{15}$$

this relation is always assumed from now on. Then in view of (H2)

$$S(u,\eta)\in B_H(R_*)\quad\text{if}\quad u\in B_H(R_*),\eta\in\mathcal{K}.$$

By $(A1)$ if $u\in B_H(R_*)$ and $\eta\in\mathcal{K}$, then $\|S(u,\eta)\|_V\le C(\mathcal{K},R_*)$, where $C(\mathcal{K},R_*)$ is a constant depending only on $\mathcal{K},R_*$. We see that the set

$$X=X_{\mathcal{K},R_*}=\text{completion in } H \text{ of } B_H(R_*)\cap B_V(C(\mathcal{K},R_*)),$$

is a compact subset of H such that

$$S:B_H(R_*)\times\mathcal{K}\mapsto X.$$

In particular,

$$S:X\times\mathcal{K}\mapsto X. \tag{16}$$

That is, the RDS (7) defines a Markov chain in X and

$$S(u,\eta)\in X\quad\text{a.s.}\quad\text{if } u\in B_H(R_*).$$

From here for u,u' as in the theorem's assumption we have

$$u_k(u),\ u_k(u')\in X\quad\text{for}\ \ k\ge1,\quad\text{a.s.}$$

The proof of Corollary 1.4 and this relation show that (14) holds if (13) is established only for $u, u' \in X$. But (14) implies the validity of the assertion of Theorem 1.3 for all $u, u' \in B_H(R_*)$. So

$$\text{proving the theorem we may assume that } u, u' \in X \text{ and regard (7) as a system in the compact set } X. \tag{17}$$

Choosing in (14) k sufficiently big we find that the operator $\mathcal{P}_k^*$ defines a contraction of the complete metric space $(\mathcal{P}(X), \|\cdot\|_L^*)$. So it has a unique fixed point $\mu_* \in \mathcal{P}(X)$. Then $\mathcal{P}_k^*\mathcal{P}_1^*\mu_* = \mathcal{P}_1^*\mu_*$, so $\mathcal{P}_1^*\mu_* = \mu_*$ by the uniqueness, i.e. μ_* is a stationary measure for the Markov chain, defined by (7) on X. Due to (14) it is unique. Since $X = X_{\mathcal{K},R_*}$, then formally this measure depends on R_*, $\mu_* = \mu_*(R_*)$. But the measure $\mu_*(R_*^0)$ (see (15)) is stationary for the system on $X_{\mathcal{K},R_*}$ for any $R_* \geq R_*^0$, so $\mu_*(R_*^0) = \mu_*(R_*)$ by the uniqueness. Denoting this measure by μ_* we see that it is stationary for the system (7), considered on H, and derive from (14) (with a suitable R_*) that

$$\|P_k(u,\cdot) - \mu_*\|_L^* \to 0 \quad \text{as} \quad k \to \infty, \tag{18}$$

for every $u \in H$. Assume that μ' is another stationary measure for the system on H. Then for any bounded Lipschitz function f on H we have

$$\langle f, \mu' \rangle = \langle f, \mathcal{P}_k^*\mu' \rangle = \langle \mathcal{P}_k f, \mu' \rangle = \int_H \langle f, P_k(u,\cdot) \rangle \, \mu'(du) \to \langle f, \mu_* \rangle$$

by the Lebesgue theorem and (18). So $\mu' = \mu_*$. We finalize our analysis of eq. (7) in the following theorem:

Theorem 1.5. *Under the assumptions (A1), (H1), (H2) and (H3) equation* (7) *defines in H a Markov chain which has a unique stationary measure μ_*. This measure is supported by the ball $B_H(R_*^0)$, and if $\mu \in \mathcal{P}(H)$ is supported by a ball $B_H(R_*)$, $R_* \geq R_*^0$, then*

$$\|\mathcal{P}_k^*\mu - \mu_*\|_L^* \leq C\kappa^k, \quad \forall\, k \geq 1, \tag{19}$$

where $C > 0$ and $\kappa \in (0,1)$ depend on R_*. If ν is any measure in $\mathcal{P}(H)$, then

$$\mathcal{P}_k^*\nu \rightharpoonup \mu_* \quad \text{as } k \to \infty, \tag{20}$$

where $\rightharpoonup$ signifies the weak convergence of mesures.

Proof. The first part of the theorem follows from what was said above if we note that the measure $\mathcal{P}_1^*\mu$ is supported by the compact set $X_{\mathcal{K},R_* \vee R_*^0}$, so (19) is a consequence of (14) with $\mu := \mathcal{P}_1^*\mu$ and $\nu := \mu_*$. Now consider any $\nu \in \mathcal{P}(H)$. For R large, define a probability measure supported by $B_H(R)$:

$$\nu_R(\cdot) := \frac{1}{\nu(B_H(R))}\nu(\cdot \cap B_H(R)).$$

In view of Ulam's theorem ν_R converges to ν in the total variation distance. We have

$$\begin{aligned} \|\mathcal{P}_k^*\nu - \mu_*\|_L^* &\leq \|\mathcal{P}_k^*\nu - \mathcal{P}_k^*\nu_R\|_L^* + \|\mathcal{P}_k^*\nu_R - \mu_*\|_L^* \\ &= \sup_{|f|_{L(H)} \leq 1} \langle \mathcal{P}_k f, \nu - \nu_R \rangle + C_R\kappa_R^k \leq 2\|\nu - \nu_R\|_{var} + C_R\kappa_R^k. \end{aligned}$$

Choosing first R so big that the first term in the r.h.s. goes to zero and then k so large such that the second term goes to null we see that the l.h.s. converges to zero with k. This implies (20). □

Note that since the theorem deals with initial data, supported by a ball $B_H(R_*)$, then Remark 1.2 applies and it suffices to check Assumption (H2) in the weaker form, specified in that remark.

1.3. **Main lemma and proof of Theorem 1.3.** In this section we state the main technical lemma and derive from it Theorem 1.3. Based on (17) we regard (7) as a system on the compact set $X = X_{\mathcal{K},R_*}$.

Lemma 1.6. *Under the assumptions (A1), (H1), (H2), (H3), there exist constants $C > 0$, $\delta \in (0,1]$ and a continuous mapping $\Phi : X \times H \times \mathcal{K} \to E$ such that $\Phi(u,u';\eta) = 0$ if $u = u'$, and the mapping*

$$\Psi(u,u';\eta) := \eta + \Phi(u,u';\eta)$$

satisfies

$$\|\ell - \Psi_*(u,u';\cdot)\ell\|_{var} \leq C\|u-u'\|_H, \tag{21}$$

$$\|S(u,\eta) - S(u',\Psi(u,u';\eta))\|_H \leq \tfrac{1}{2}\|u-u'\|_H, \tag{22}$$

for all $u,u' \in D_\delta := \{(u,u') \in X \times H : \|u-u'\|_H \leq \delta\}$ and all $\eta \in \mathcal{K}$.

Lemma 1.6 will be proved in the next section. Now, supposing that we have this result, we will first establish a coupling lemma and then derive from it Theorem 1.3.

Lemma 1.7. *For δ as in Lemma 1.6 there exists $C_1 = C_1(\delta)$, a probability space $(\Omega_0, \mathcal{F}_0, P_0)$ and a measurable mapping*

$$D_\delta \times \Omega_0 \to \mathcal{K} \times \mathcal{K}, \quad (u,u',\omega_0) \mapsto (\eta_1^{\omega_0}(u,u'), {\eta_1'}^{\omega_0}(u,u'))$$

such that $\mathcal{D}(\eta_1(u,u')) = \mathcal{D}(\eta_1'(u,u')) = \ell$, and $u_1 := S(u,\eta_1)$, $u_1' := S(u',\eta_1')$ satisfy

$$P\big(\|u_1 - u_1'\|_H \leq \tfrac{1}{2}d\big) > 1 - C_1 d, \quad d = \|u-u'\|_H \leq \delta. \tag{23}$$

Proof. Recall that η has the law ℓ. Denote $\eta' = \Psi(u,u',\eta)$ and $\ell' = \mathcal{D}\eta'$. Then according to Lemma 1.6,

$$\|S(u,\eta) - S(u',\eta')\|_H \leq \tfrac{1}{2}d, \ a.s. \tag{24}$$

$$\|\ell - \ell'\|_{var} \leq C_1 d. \tag{25}$$

The law ℓ' of η' need not to be ℓ. To improve this, note that in view of (22) and the Dobrushin lemma (see in [5]), there exists a coupling $(\tilde\eta, \tilde\eta')$ for (ℓ, ℓ'), defined on a probability space $(\Omega_1, \mathcal{F}_1, P_1)$, satisfying

$$P_1(\tilde\eta \neq \tilde\eta') = \|\ell - \ell'\|_{var} \leq C_1 d. \tag{26}$$

Moreover, the mapping $(\omega_1, u, u') \mapsto (\tilde\eta, \tilde\eta')$ may be chosen to be measurable, see [5], Section 1.2.4. For the pair of measures (ℓ, ℓ') we have obtained two couplings – (η, η') and $(\tilde\eta, \tilde\eta')$.
According to the *Gluing lemma* (see [10]), there exist random variables $(\zeta_1, \zeta_2, \zeta_3)$ defined on a probability space $(\Omega_0, \mathcal{F}_0, P_0)$ such that

$$\mathcal{D}(\zeta_1,\zeta_2) = \mathcal{D}(\eta,\eta'), \quad \mathcal{D}(\zeta_2,\zeta_3) = \mathcal{D}(\tilde\eta',\tilde\eta).$$

The triplet $(\zeta_1, \zeta_2, \zeta_3)$ may be chosen to be a measurable function of (ω_0, u, u'), see [6], Appendix 3. In particular, $\mathcal{D}(\zeta_1) = \mathcal{D}(\zeta_3) = \ell$ and by inequality (24)

$$\|S(u,\zeta_1) - S(u',\zeta_2)\|_H \leq \tfrac{1}{2}d, \quad P_0 - a.s.$$

Furthermore, in view of inequality (26),

$$P_0(\zeta_2 \neq \zeta_3) = P_1(\tilde\eta' \neq \tilde\eta) \leq C_1 d,$$

which implies that

$$P_0\big(\|S(u,\zeta_1) - S(u',\zeta_3)\|_H \leq \tfrac{1}{2}d\big) \geq 1 - C_1 d.$$

To complete the proof it remains to choose $\eta_1 = \zeta_1$ and $\eta_1' = \zeta_3$.

□

To obtain the exponential mixing, claimed by Theorem 1.3, we will apply the method of Kantorovich functional (see in [5]). A Kantorovich functional for measures on X is a symmetric function $K : \mathcal{P}(X) \times \mathcal{P}(X) \to \mathbb{R}_+$, defined via its density f_K. The latter is a symmetric measurable function f_K on $X \times X$, satisfying

$$f_K(\xi_1, \xi_2) \geq \|\xi_1 - \xi_2\|_{d_0} \quad \forall\, \xi_1, \xi_2 \in X,$$

(see (12)), and K is defined in terms of f_K by the relation

$$K(\mu_1, \mu_2) = \inf\{\mathbb{E} f_K(\xi_1, \xi_2)\},$$

where the infimum is taken over all couplings (ξ_1, ξ_2) for (μ_1, μ_2).

For R_* as above and $u, u' \in B_H(R_*)$ we denote $d = \|u - u'\|_H$ and set

$$(u_1, u_1')(\omega_0) = \begin{cases} (S(u,\eta), S(u',\eta)), & d > d_0\,, \\ (S(u,\eta_1), S(u',\eta_1')), & d \leq d_0\,, \end{cases} \tag{27}$$

where $\eta_1(\omega_0)$, $\eta_1'(\omega_0)$ are as in Lemma 1.7, and η is a r.v. on Ω_0, distributed as ℓ, and independent from η_1, η_1'. Clearly (u_1, u_1') is a measurable function of (u, u', ω_0) and $\mathcal{D}(u_1) = P_1(u, \cdot)$ and $\mathcal{D}(u_1') = P_1(u', \cdot)$.

Lemma 1.8. *There exist $d_0 > 0$ and a function*

$$f : (\tfrac{1}{2} d_0, R_*] \to (2d_0, 3d_0] \quad \textit{non-decreasing}, \tag{28}$$

such that for any $\xi_1, \xi_2 \in X$,

$$f_K(\xi_1, \xi_2) = \begin{cases} \|\xi_1 - \xi_2\|_H, & \textit{if } \|\xi_1 - \xi_2\|_H \leq d_0, \\ f(\|\xi_1\|_H \vee \|\xi_2\|_H), & \textit{if } \|\xi_1 - \xi_2\|_H > d_0, \end{cases} \tag{29}$$

is a Kantorovich density. It satisfies

$$\mathbb{E}\,[f_K(u_1, u_1')] \leq \kappa f_K(u, u') \quad \textit{for any } u, u' \in B_H(R_*), \tag{30}$$

where $\kappa \in (0,1)$ is a constant, depending on R_*.

Everywhere below in the proof f_K always stays for the Kantorovich density, defined in (29).

Proof. We may assume that $d_0 \leq \delta$, where δ is introduced in Lemma 1.6. The fact that f_K is a Kantorovich density is obvious, as well as that it satisfies

$$f_K(\xi_1, \xi_2) \leq 3\|\xi_1 - \xi_2\|_H. \tag{31}$$

It remains to verify (30). Let us denote $R_1(\omega_0) := \|S(u,\eta)\|_H \vee \|S(u',\eta')\|_H$, $R := \|u\|_H \vee \|u'\|_H$. Recall that $d = \|u - u'\|_H$.

Case 1: $d > d_0$.

Now $\frac{1}{2} d_0 \leq R \leq R_*$. Let as take $a = \frac{1+\gamma}{2}$. According to (11) and (9),

$$P[R_1 \leq aR] > p_{d_0} =: p > 0, \qquad R_1 \leq \gamma R + \beta, \quad a.s.$$

Therefore

$$\mathbb{E}[f_K(u_1, u_1')] \leq p f(aR) + (1-p) f(\gamma R + \beta).$$

To obtain inequality (30), we need to find f of the form (28), satisfying

$$p f(aR) + (1-p) f(\gamma R + \beta) \leq \kappa_1 f(R) \quad \text{for} \quad R \in (\tfrac{1}{2} d_0, R_*], \tag{32}$$

with some $\kappa_1 \in (0,1)$.

Let us first consider $f(R)$ when $R \in J := (\frac{1}{2}d_0, R_*^0]$. We define the segments

$$I_1 = (aR_*^0, R_*^0],\ I_2 = (a^2R_*^0, aR_*^0], \dots, I_n = (a^nR_*^0, a^{n-1}R_*^0], \dots$$

Then $\cup_{j=1}^{N_0} I_j \supset J$, where $N_0 = \left[\frac{\ln 2R_0 + \ln d_0^{-1}}{\ln a^{-1}}\right] + 1$. We define f on J by the relation

$$f(x) = a_j, \quad x \in I_j, \quad 1 \le j \le N_0,$$

where

$$2.5d_0 = a_1 > a_2 > \dots > a_{N_0} > 0 \tag{33}$$

is a sequence to be defined. If $R \in I_1$, then (32) is valid if

$$a_1 > pa_2 + (1-p)a_1,$$

which holds true since $a_1 > a_2$. If $R \in I_n$, $n \ge 2$, then (32) holds if

$$a_{n-1} > pa_n + (1-p)a_1. \tag{34}$$

Let us take some $p_1 \in (0, p)$ and define the sequence $\{a_j\}$ as follows:

$$a_n := a_1 - \frac{a_1 - a_2}{p_1^n}.$$

If $a_2 < a_1$ is chosen sufficiently close to a_1, then (33) holds. Relation (34) holds as well and (30) is proved for $d > d_0$ with some $\kappa = \kappa_1 < 1$. Choosing $a_1 - a_2 > 0$ sufficiently small we achieve that also $2d_0 \le f(x) \le 2.5d_0$ on J.
If $R_* = R_*^0$, we are done. If $R_* > R_*^0$, we extend $f(x)$ to the interval $(R_*^0, R_*]$ as an affine function $l(x)$ such that $l(R_0^*) = 2.6d_0$ and $l(R_*) = 3d_0$. Then for $R > R_*^0$ we have $f(R) \ge 2.6d_0$. From the other hand, $f(\gamma R + \beta) \le f(R) = l(R)$, while $f(aR)$ is $\le 2.5d_0$ if $aR \le R_*^0$ and is $l(aR)$ if $aR > R_*^0$. In both cases (32) holds with a suitable $\kappa_1 < 1$.
We have constructed a function f_K, satisfying (28), (31) and (32), if $d > d_0$.

Case 2. $d \le d_0$.
Denote $D = \|u_1 - u_1'\|_H$. Since $d_0 \le \delta$, then according to Lemma 1.7,

$$\mathbb{P}\big(D \le \tfrac{1}{2}d\big) \ge 1 - C_1 d \quad \text{if} \quad d \le d_0.$$

In view of (28),

$$\mathbb{E}[f_K(u_1, u_1')] \le C_1 d(3d_0) + (1 - C_1 d)(\tfrac{1}{2}d) \le \tfrac{1}{2}d + (\tfrac{5}{2}C_1 d)d_0.$$

Choosing $d_0 \le 1/(10C_1)$ we achieve that (30) holds with $\kappa = 3/4$.
Now the lemma is proved by choosing $d_0 \le 1/(10C_1) \wedge 2\beta/(1-\gamma)$ with $\kappa = \kappa_1 \vee (3/4)$. □

Proof of Theorem (1.3). Note that for any $f \in C_b(H)$ with $|f|_{L(H)} \le 1$ (see (5)) and any measures $\mu, \nu \in \mathcal{P}(H)$ we have

$$\Big|\int_H f(x)(\mu - \nu)(dx)\Big| = |\mathbb{E}\,[f(\xi) - f(\zeta)]| \le \mathbb{E}\|\xi - \zeta\|_H,$$

where (ξ, ζ) is a coupling for μ, ν. Therefore

$$\|\mu - \nu\|_L^* \le \inf_{(\xi,\zeta)} \mathbb{E}\|\xi - \zeta\|_H,$$

where the infimum is taken over all couplings (ξ,ζ) for (μ,ν). If μ,ν are supported by X, then the couplings are valued in $B_H(R_*)$ a.s., and due to (29) and (28) we have that

$$\|\xi^\omega-\zeta^\omega\|_H\le\frac{R_*}{d_0}f_K(\xi^\omega,\eta^\omega)\quad\text{a.s.}$$

For $u,u'\in X$ and $k\ge1$ let (u_k,u_k') be any fixed coupling for $(P_k(u,\cdot),P_k(u',\cdot)$. Then

$$\|P_k(u,\cdot)-P_k(u',\cdot)\|_L^*\le\inf_{\xi,\zeta}\mathbb{E}\|\xi-\zeta\|_H\le\frac{R_*}{d_0}\inf_{\xi,\zeta}\mathbb{E}f_K(\xi,\zeta)\le\frac{R_*}{d_0}\mathbb{E}f_K(u_k,u_k'),\tag{35}$$

where, as above, (ξ,ζ) is a coupling for the two measures.
Denote $\mu_k=P_k(u,\cdot)$ and $\mu_k'=P_k(u',\cdot)$. We claim that for any $k\ge1$ there exists a coupling (u_k,u_k') for the measures (μ_k,μ_k'), defined on a probability space $(\Omega^k,\mathcal{F}^k,P^k)$, such that

$$\mathbb{E}^k\left[f_K(u_k,u_k')\right]\le\kappa^kf_K(u,u'),\tag{36}$$

where $\kappa\in(0,1)$.
For $k=1$ this is the assertion of Lemma 1.8. Now let $k\ge2$. Consider a coupling $(u_{k-1}(\omega_{k-1}),u_{k-1}'(\omega_{k-1}))$ for (μ_{k-1},μ_{k-1}'), defined on a probability space $(\Omega_{k-1},\mathcal{F}_{k-1},P_{k-1})$ and satisfying (36) with $k:=k-1$, existing by the base of induction. Let (Ω_0,F_0,P_0) be the probability space from Lemma 1.7. We take for Ω_k the direct product of these two spaces: $\Omega_k=\{\omega_k=(\omega_{k-1},\omega_0)\}$, and set $(u_k,u_k')=(u_1,u_1')(\omega_{k-1},\omega_0)$, where the pair (u_1,u_1') is defined as in (27) with $u=u_{k-1}(\omega_{k-1}),u'=u_{k-1}'(\omega_{k-1})$.
By (30) and the base of induction we have:

$$\mathbb{E}f_K(u_k,u_k')=\mathbb{E}^{k-1}\Big[\mathbb{E}^0f_K\big(u_1,u_1'\big)(u_{k-1}(\omega_{k-1}),u_{k-1}'(\omega_{k-1}))\Big]\le\kappa\mathbb{E}^{k-1}f_K(u_{k-1},u_{k-1}')\le\kappa^kf_K(u,u'),$$

so (36) is proved. By (35), (36) and (31)

$$\|P_k(u,\cdot)-P_k(u',\cdot)\|_L^*\le\frac{R_*}{d_0}f_K(u,u')\kappa^k\le\frac{3R_*}{d_0}\|u-u'\|_H\kappa^k,$$

which proves Theorem 1.3. □

1.4. **Proof of the main lemma.** We start with a result on almost-inverse linear operators, depending on a parameter, following closely Section 2.2 of [6]. Let H, E and V be the spaces as above and X, $\mathcal{K}$ be the compact sets as above. For the Hilbert base $\{e_j\}$ of E as in (H1) and any $M\ge1$ we denote by E_M the subspace of E, generated by the first M vectors of the base, and denote by P_M the orthogonal projection $E\to E_M$.

Proposition 1.9. *Let $A:X\times\mathcal{K}\to\mathcal{L}(E,H)$ be a continuous mapping. Assume that for any $u\in X$ and $\eta\in\mathcal{K}$ the operator $A(u,\eta)$ has dense image in H. Then for any $\varepsilon>0$ there exists $M_\varepsilon\in\mathbb{N}$, $C(\varepsilon)>0$ and a continuous mapping $R_\varepsilon:X\times\mathcal{K}\to\mathcal{L}(H,E)$ such that for any $u\in X,\eta\in\mathcal{K}$,*

$$Im\,(R_\varepsilon(u,\eta))\subseteq E_{M_\varepsilon},\quad\|R_\varepsilon(u,\eta)\|_{\mathcal{L}(H,E)}\le C(\varepsilon);\tag{37}$$

$$\|A(u,\eta)R_\varepsilon(u,\eta)f-f\|_H\le\varepsilon\|f\|_V\quad\forall f\in V.\tag{38}$$

Proof. For any $u\in X$ and $\eta\in\mathcal{K}$ define $G(u,\eta):=A(u,\eta)A^*(u,\eta)$, where A^* is the adjoint operator. This is a non-negative self-adjoint operator. Its kernel equals to that of A^* and is trivial since A has dense image. So $G=G^*>0$, and the operator $(G+rI)^{-1}$ is well-defined and smooth in $\{r>0\}$. For $M\in\mathbb{N}$ denote

$$G_r(u,\eta):=A^*(u,\eta)(G(u,\eta)+rI)^{-1},\quad G_{r,M}(u,\eta):=P_M\circ G_r(u,\eta).$$

For any r and M the operator $G_{r,M}$ is continuous in u and η. Below we will choose $M = M_\varepsilon$ and $r = r_\varepsilon$ in such a way that $R_\varepsilon := G_{r,M}$ satisfies (37),(38).
Since $\sup_{X\times\mathcal{K}} \|A(u,\eta)\| < \infty$ and $\|(G+rI)^{-1}\| \le r^{-1}$, then $\sup_{X\times\mathcal{K}} \|G_{r,M}(u,\eta)\| \le Cr^{-1}$, which implies (37).
Now we will prove (38). Doing that we may assume that $f \in B_V(1)$. Since $AG_r - AG_{r,M} = A(I-P_M)G_r$, then

$$\|AG_r f - AG_{r,M} f\|_E \le C\|(I-P_M)G_r f\|_H =: CF_M(u,\eta,f),$$

for all $(u,\eta,f) \in X\times\mathcal{K}\times B_V(1) =: \mathfrak{A}$. The functions F_M are continuous on $\mathfrak{A}$, and $F_M \searrow 0$ as $M\to\infty$, pointwise. Since the set $\mathfrak{A}$ is compact, this convergence is uniform by Dini's theorem. That is,

$$\sup_{\mathfrak{A}} \|AG_r f - AG_{r,M} f\| \to 0 \quad \text{as} \quad M\to\infty,$$

for any $r>0$. So to obtain inequality (38) it remains to show that

$$\sup_{\mathfrak{A}} \|AG_r f - f\| \to 0 \quad \text{as} \quad r\to 0. \tag{39}$$

To prove this note that since $AG_r = AA^*(G+rI)^{-1} = G(G+rI)^{-1}$, then

$$\|AG_r f - f\|_E = r\|(G+rI)^{-1} f\|_E =: \Phi_r(u,\eta,f).$$

Consider the family of continuous functions

$$\Phi_r : \mathfrak{A} \to \mathbb{R}_+, \quad r>0.$$

Writing the operator $r(G+rI)^{-1}$ and the vector f in terms of the spectral decomposition for the positive self-adjoint operator $G(u,\eta)$, we get that the functions Φ_r pointwise monotonically converge to zero as $r\to 0$ (see [6], Lemma 2.4, for details).
Since $\mathfrak{A}$ is a compact set, then evoking again Dini's theorem [4] we see that the convergence is uniform. So (39) follows and the proposition is proved. □

Now we are ready to prove the main lemma.

Proof of Lemma 1.6. Step 1: construction of the mapping Φ. By the regularity assumption (A1), for any $\zeta \in E$

$$S(u',\eta+\zeta) = S(u,\eta) + D_u S(u,\eta)(u'-u) + D_\eta S(u,\eta)\zeta + O(\|u'-u\|_H^2 + \|\zeta\|_E^2). \tag{40}$$

To obtain (22), consider the relation above with $\zeta = \Phi(\eta) = \Phi(u,u';\eta)$. If we can construct a measurable mapping $\Phi : D_\delta \times \mathcal{K} \to E_{M_\varepsilon}$ such that

$$\|D_\eta S(u,\eta)\Phi(\eta) + (D_u S(u,\eta))(u'-u)\|_H \le \tfrac{1}{4}\|u'-u\|_H, \tag{41}$$

$$\|\Phi(\eta)\|_E \le C\|u'-u\|_H, \tag{42}$$

for all (u,u',η) with a uniform constant C, then (22) should follow if $\delta \ll 1$.
For a fixed $(u,u',\eta) \in D_\delta \times \mathcal{K}$ consider the following equation on ζ :

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

[4] A version of the theorem for families of functions easily follow from the usual Dini's theorem for sequences of functions.

By assumption $(H3)$ and Proposition 1.9, for any $\varepsilon > 0$ there exists an $M = M_\varepsilon \in \mathbb{N}$ and a linear operator $R_\varepsilon(u,\eta) : H \to E_{M_\varepsilon}$ whose norm is bounded uniformly in $u, \eta \in X \times \mathcal{K}$, such that for any $f \in V$,

$$\|(D_\eta S)R_\varepsilon f - f\|_H \le \varepsilon \|f\|_V.$$

Take $f = -(D_u S)(u' - u)$. By $(A1)$ f belongs to the space V and $\|f\|_V \le C\|u' - u\|_H$, uniformly in $(u,\eta) \in X \times \mathcal{K}$. Define

$$\Phi : D_\delta \times \mathcal{K} \to E_{M_\varepsilon}, \quad \Phi(u,u';\eta) = R_\varepsilon(u,\eta)f = -R_\varepsilon(u,\eta)(D_u S(u,\eta))(u' - u).$$

For any $\varepsilon > 0$ this is a continuous mapping, and due to assumption $(A1)$

$$\|\Phi(u,u';\eta)\|_E \le C_\varepsilon \|u' - u\|_H, \tag{43}$$

$$\|\Phi(u,u';\eta) - \Phi(u,u';\eta')\|_E \le C_\varepsilon \|u' - u\|_H \|\eta - \eta'\|_E, \tag{44}$$

for any $(\eta,\eta') \in B_E(R_\eta) \times B_E(R_\eta)$ and $(u,u') \in D_\delta$.

Step 2: estimate (22). According to identity (40) and estimate (43), for any $\eta \in \mathcal{K}$ and $u, u' \in D_\delta$ we have

$$\begin{aligned} \|S(u',\eta + \Phi(u,u';\eta)) - S(u,\eta)\|_H &\le C\left(\|u' - u\|_H^2 + \|\Phi(u,u';\eta)\|_E^2\right) + \varepsilon\|(D_u S)(u' - u)\|_V \\ &\le \|u' - u\|_H \left(C\varepsilon + (1 + C_\varepsilon^2)C\|u' - u\|_H\right) \\ &\le \tfrac{1}{2}\|u' - u\|_H, \end{aligned}$$

where to get the last inequality we choose ε so small that $C\varepsilon < \frac{1}{4}$ next choose $\delta \le \delta(\varepsilon)$ so small that $(1 + C_\varepsilon^2)C\delta < \frac{1}{4}$. This proves (22).

Step 3: estimate (21). Recall that $E_M = \text{span}\{e_1, \dots, e_M\}$. Denote by $E_M^\perp$ the complementary space of E_M in E. For any $A \in \mathcal{B}_E$ we will denote by A_M and $A_M^\perp$ its projections to E_M and $E_M^\perp$, respectively. Consider the Hilbert brick $\tilde{\mathcal{K}}$, defined in (10). Then $\tilde{\mathcal{K}}_M$ is a parallelepiped, $\tilde{\mathcal{K}}_M^\perp$ is a Hilbert brick in $E_M^\perp$,

$$\tilde{\mathcal{K}}_M = \{\sum_{i=1}^M x_i e_i \Big| \ |x_i| \le b_i, i = 1, ..., M\},$$

and $\tilde{\mathcal{K}} = \tilde{\mathcal{K}}_M \times \tilde{\mathcal{K}}_M^\perp$.

By assumption (H1) the measure ℓ is supported by $\tilde{\mathcal{K}}$ and $\ell = \ell_M \otimes \ell_M^\perp$ where ℓ_M is supported by $\tilde{\mathcal{K}}_M$ and $\ell_M^\perp$ – by $\tilde{\mathcal{K}}_M^\perp$. Besides, $\rho_M = \rho_M(dv)$, where $E_M = \{v\}$ and ρ_M is a Lipschitz function, supported by $\tilde{\mathcal{K}}_M$.

For any fixed $A \in \mathcal{B}_E$, denote by h_A the indicator function of A. According to Fubini's theorem,

$$\ell(A) = \int h_A(e)\ell(de) = \int_{\tilde{\mathcal{K}}_M^\perp} \ell_M^\perp(dw) \int_{\tilde{\mathcal{K}}_M} h_A(w+v)\ell_M(dv) = \int_{\tilde{\mathcal{K}}_M^\perp} \ell_M^\perp(dw) \int_{\tilde{\mathcal{K}}_M} h_A(w+v)\rho_M(v)dv.$$

Similar,

$$\Psi_*(\ell)(A) = \int h_A(e + \Phi(u,u';e))\ell(de) = \int_{\tilde{\mathcal{K}}_M^\perp} \ell_M^\perp(dw) \int_{\tilde{\mathcal{K}}_M} h_A(w + v + \Phi(u,u';w+v))\rho_M(v)dv.$$

Since $\|\ell - \Psi_*\ell\|_{var} = \sup_{A \in \mathcal{B}_E} |\ell(A) - \Psi_*(\ell)A|$, then (21) would follow if we prove that

$$|\int_{\tilde{\mathcal{K}}_M} h_A(w+v)\rho_M(v)dv - \int_{\tilde{\mathcal{K}}_M} h_A(w+v+\Phi(u,u';w+v))\rho_M(v)dv| \le C\|u - u'\| \quad \forall w, u, u', \tag{45}$$

where C does not depend on A as well as on w, u, u'. To do this let us consider the mapping

$$\Phi_w : \tilde{\mathcal{K}}_M \ni v \mapsto \Phi(u, u'; v + w) \in E_M.$$

By (43) and (44) the map's norm and the Lipschitz constant are bounded by $C_\varepsilon \|u - u'\|_H \le C_\varepsilon \delta$. We will assume that δ is so small that

$$C_\varepsilon \delta \le 1/2 \tag{46}$$

(where ε was fixed at Step 2 of the proof). By the Kirszbraun theorem (see [1]), for each fixed u, u' and w, Φ_w extends to a mapping $\tilde{\Phi}_w : E_M \to E_M$ with the same Lipschitz constant. Now consider the mapping

$$\Xi : E_M \to E_M, \quad v \mapsto \xi = v + \tilde{\Phi}_w(v).$$

Due to (46) its Lipschitz constant is $\le 1 + C_\varepsilon \delta \le 3/2$. The inverse mapping $\Xi^{-1} : E_M \to E_M$ exists and

$$\mathrm{Lip}\,(\Xi^{-1} - \mathrm{id}) \le C'_\varepsilon \|u - u'\|.$$

Let us consider the second integral in (45) and write it as an integral over E_M with an integrand which vanishes outside $\tilde{\mathcal{K}}_M$:

$$\int_{\tilde{\mathcal{K}}_M} h_A(w + v + \Phi(u, u'; w + v))\rho_M(v)dv = \int_{E_M} h_A(w + v + \Phi(u, u'; w + v))\rho_M(v)dv.$$

Passing there from the variable v to ξ we write the integral as

$$\int_{E_M} h_A(w + \xi)\rho_M(v(\xi))\Big|\det \frac{\partial \xi}{\partial v}\Big|^{-1} d\xi = \int_{\Xi(\tilde{\mathcal{K}}_M)} h_A(w + \xi)\rho_M(v(\xi))\Big|\det \frac{\partial \xi}{\partial v}\Big|^{-1} d\xi$$

(concerning Lipschitz changes of variables in integrals over $\mathbb{R}^N$ see e.g. Theorem 3.2.5 in [1]). Writing the first integral in (45) as $\int_{\mathcal{K}_M} h_A(w+\xi)\rho_M(\xi)\, d\xi$ and using that the mapping $v \mapsto \xi$ and its inverse both are $C_\varepsilon \|u - u'\|$-close to the identity in the Lipschitz norm, we get (45). The main lemma is proved. □

Remark 1.10. *If, as in Remark 1.1, the mapping S is defined only on a subdomain Q of $H \times E$, containing $H \times \mathcal{K}$, and the assumption of Remark 1.2 holds, then assumption* (46) *should be strengthened to $C_\varepsilon \delta \le \min(1/2, r(R_+))$, where r is the positive functions from Remark 1.1. Then the extension $\tilde{\Phi}_w$ can be chosen to be bounded in norm by $r(R_+)$, so the points (u', Ψ) stay in the domain Q, where S is well defined.*

2. Applications

In this section we apply the abstract theorems above to the two-dimensional Navier-Stokes system on the torus $\mathbb{T}^2 = \mathbb{R}^2/(2\pi)\mathbb{Z} \oplus (2\pi)\mathbb{Z}$ and to well posed quasilinear parabolic systems on $\mathbb{T}^d$, perturbed by random forces, and prove that these systems are exponentially mixing. To do that we will pass from the corresponding PDE (1) to a discrete time system (2), will show that the latter satisfies the assumptions (A1), (H1), (H2), (H3) and then will apply Theorem 1.5. We will do that in the situation when the space H is a Hilbert space of functions of x and E is the space $L_2([0, 1]; H)$. Accordingly, the Hilbert base $\{e_j\}$ of E will be of the form $\{h_r \otimes \phi_i\}$, where $\{\phi_i\}$ is a base of H and $\{h_r\}$ – a base of $L_2(0, 1)$. To apply the abstract theorems to the 2d Navier-Stokes system the base $\{\phi_i\}$ may be arbitrary, while to apply them to the quasilinear parabolic systems its elements should be bounded functions. Below we restrict ourselves to the case when $\{\phi_i\}$ is the Haar base.

2.1. **Random Haar series ("colored noises").** We will apply Theorem 1.5 to equations (1), perturbed by random forces η_t of the form

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

where $\{\eta_k(\tau), 0 \le \tau < 1\}_k$ is a sequence of i.i.d. random variables in $L^2([0,1],H)$. It suffices to define the process $\eta_1 = \eta\,|_{[0,1)}$. To do this, let us denote by $\{h_{jl} : j \in \mathbb{N}, 0 \le l \le 2^j - 1\}$ the orthonormal Haar basis on $[0,1]$:

$$h_{00}(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & t < 0 \text{ or } t \ge 1, \end{cases} \qquad h_{jl}(t) = \begin{cases} 0 & t < l2^{-j} \text{ or } t \ge (l+1)2^{-j}, \\ 2^{\frac{j}{2}} & l2^{-j} \le t < (l+\frac{1}{2})2^{-j}, \\ -2^{\frac{j}{2}} & (l+\frac{1}{2})2^{-j} \le t < (l+1)2^{-j}. \end{cases} \tag{47}$$

Then $\{h_{jl}\phi_i\}_{ijl}$ is an orthonormal basis of $E = L_2([0,1],H)$ (we recall that $\{\phi_i\}_i$ is a Hilbert basis of H).
Now we define the process $\eta_1 = \eta\,|_{[0,1)}$ in H as

$$\eta_1(t) = \sum_{i=1}^{\infty} b_i \eta_t^i \phi_i, \quad b_i \ne 0\ \forall\, i. \tag{48}$$

Here $\sum_i b_i^2 < \infty$ and for $i \ge 1$, η_t^i is the random process

$$\eta_t^i = \sum_{j=0} c_j^i \sum_{l=0}^{2^j-1} \xi_{jl}^i h_{jl}, \quad c_j^i \ne 0 \ \ \forall\, i,j, \tag{49}$$

where

$$\sum_{ijl} (c_j^i b_i)^2 = \sum_i b_i^2 \sum_j 2^j (c_j^i)^2 =: R_\eta^2 < \infty,$$

and $\{\xi_{jl}^i\}_{i,j,l}$ are independent real random variables with $|\xi_{jl}^i| \le 1$, whose density functions ρ_{ijl} are Lipschitz–continuous and $\rho_{ijl}(0) \ne 0$. Denote the basis $\{h_{jl}\phi_i\}_{i,j,l}$ as $\{e_k, k \ge 1\}$, where

$$j(k) \to \infty, \ \ i(k) \to \infty \quad \text{as } k \to \infty.$$

Then assumption $(H1)$ obviously holds.

2.2. **The 2D Navier-Stokes system.** Consider the 2D Navier-Stokes system

$$\partial_t u - \nu \triangle u + \langle u, \nabla \rangle u + \nabla p = \eta(t,x), \quad \text{div } u = 0, \quad \int u\, dx = 0, \tag{50}$$

where $x \in \mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2, u = (u_1, u_2)$, p is an unknown scalar function (the pressure), $\nu > 0$ is the viscosity and η is a random force as in the previous section. Denote by $\mathcal{H}$ the L_2–space

$$\mathcal{H} = \Big\{ u \in L_2(\mathbb{T}^2, \mathbb{R}^2) : \ \text{div } u = 0, \ \int_{\mathbb{T}^2} u(x)dx = 0 \Big\}$$

with the norm $\|\cdot\|$. For $m \in \mathbb{N}$ we set $\mathcal{H}^m := \mathcal{H} \cap H^m(\mathbb{T}^2, \mathbb{R}^2)$, where H^m is the Sobolev space on the torus, equipped with the Sobolev norm $\|\cdot\|_m$, and for $m \in -\mathbb{N}$ denote by $\mathcal{H}^m$ the closure of $\mathcal{H}$ in the Sobolev norm $\|\cdot\|_m$. Let $\{\phi_j, j \ge 1\}$ be the usual trigonometric Hilbert basis of $\mathcal{H}$, (e.g. see Chapter 2 in [5]). This also is an orthogonal basis of every space $\mathcal{H}^m$.

Suppose that $\eta(t,\cdot) \in \mathcal{H}$ for all t, a.s. Applying the Leray projection $\Pi : L_2(\mathbb{T}^2, \mathbb{R}^2) \to \mathcal{H}$ to equation (50) we obtain the following nonlocal PDE:

$$\partial_t u + \nu L u + B(u) = \eta(t,x), \quad u(0) = u_0, \tag{51}$$

where $L = -\Pi\Delta$ and $B(u) = \Pi(\langle u, \nabla\rangle u)$.
For $j \in \mathbb{Z}$ denote

$$E_j = L_2([0,1], \mathcal{H}^j).$$

By the classical results (see e.g. Chapter 2 in [5]), if $m \geq 1$, then for any $u_0 \in \mathcal{H}^{m-1}$ and $\eta \in E_{m-2}$ eq. (50) with $0 \leq t \leq 1$ has a unique solution

$$u \in \mathcal{U}_m := \{u_\cdot \in E_m : \partial_t u_\cdot \in E_{m-2}\} \subseteq C([0,1]; \mathcal{H}^{m-1}).$$

We equip the space $\mathcal{U}_m$ with the Hilbert norm $\|u\|^2_{\mathcal{U}_m} = \int_0^1 \|u_t\|^2_m dt + \int_0^1 \|\partial_t u_t\|^2_{m-2} dt$.
Consider the dynamical system

$$S(u_{k-1}, \eta_k) = u(1), \tag{52}$$

where $u(t)$ is a solution of equation (51) with $u(0) = u_{k-1}$ and $\eta = \eta_k$. We are going to apply to system (52) Theorem 1.5 with $H = \mathcal{H}^1$ and $E = E_1$. Accordingly we rescale the basis $\{\phi_i\}$ to be an orthonormal basis of the space $\mathcal{H}^1$. The validity of assumption (H1) was already checked in Section 2.1. In order to verify the the remaining assumptions we need two auxiliary results:

Lemma 2.1. *For any $u \in \mathcal{U}_2$ and $j = 0$ or 1, consider the mapping*

$$A_2(u) : \mathcal{U}_{j+2} \longmapsto E_j \times \mathcal{H}^{j+1}; \quad v \longmapsto (\partial_t v - \nu\Delta v + \Pi(\langle u, \nabla\rangle v + \langle v, \nabla\rangle u), v(0)).$$

Then $A_2(u)$ is a linear isomorphism. Furthermore, the norm of its inverse depends only on $\|u\|_{\mathcal{U}_2}$.

The proof easily follows by Galerkin' method.

Proposition 2.2. *For $u_0 \in \mathcal{H}^1$ and $\eta \in E_1$, we have $S(u_0, \eta) \in \mathcal{H}^2$. Furthermore, the mapping $S : \mathcal{H}^1 \times E_1 \mapsto \mathcal{H}^2$ is analytic.*

Proof. Consider equation (51) with $u_0 \in \mathcal{H}^1$ and $\eta \in E_0$. It has a unique solution $u \in \mathcal{U}_2$. By Lemma 2.1 with $j = 0$ and the implicit function theorem, u analytically depends on (u_0, η) (see [3] for details), so the mapping

$$A_1 : \quad \mathcal{H}^1 \times E_0 \longmapsto \mathcal{U}_2; \quad (u_0, \eta) \longmapsto u,$$

is analytic, as well as the mapping $\mathcal{H}^1 \times E_1 \longmapsto \mathcal{H}^1$, $(u_0, \eta) \mapsto u_1$. It remains to improve the regularity and show that the map

$$S : \quad \mathcal{H}^1 \times E_1 \longmapsto \mathcal{H}^2; \quad (u_0, \eta) \longmapsto u_1,$$

is analytic. Note that

$$S(u_0, \eta) = \int_0^1 (d/dt) S(tu_0, t\eta)\, dt = \int_0^1 (D_{u_0} S(tu_0, t\eta) u_0 + D_\eta S(tu_0, t\eta)\eta) dt, \tag{53}$$

so it suffices to show that $D_{u_0} S(u_0, \eta) h$ and $D_\eta S(u_0, \eta)\xi$ as mappings $\mathcal{H}^1 \times E_1 \times \mathcal{H}^1 \to \mathcal{H}^2$ and $\mathcal{H}^1 \times E_1 \times E_1 \to \mathcal{H}^2$, respectively, are analytic.
Denote by $u = u(u_0, \eta) \in \mathcal{U}_2$ a solution of (51) with $(u_0, \eta) \in \mathcal{H}^1 \times E_1$. Then $D_\eta S(u_0, \eta)\xi = v_1(0, \xi), \xi \in E_1$ where $v_t(v_0, \xi)$ stands for a solution of the following linear equation

$$\partial_t v + \nu L v + \Pi(\langle u, \nabla\rangle v + \langle v, \nabla\rangle u) = \xi, \quad v(0) = v_0. \tag{54}$$

The map $D_\eta S$: By Lemma 2.1 with $j = 1$, we have $v_\cdot(0,\xi) \in \mathcal{U}_3$. It remains to show $v_\cdot(0,\xi)$ analytically depends on u, which would imply that $D_\eta S(u_0,\eta)\xi = v_1$ is an analytic mapping. To this end consider the mapping

$$F : \mathcal{U}_2 \times \mathcal{U}_3 \longrightarrow E_1 \times \mathcal{H}^2, \quad (u,v) \longrightarrow A_2(u)v = \left(\partial_t v - \nu\Delta v + \Pi\left(\langle u,\nabla\rangle v + \langle v,\nabla\rangle u\right), v(0)\right).$$

To prove the required analyticity it suffices to apply Lemma 2.1 and the implicit function theorem to the equation $F(u,v) = (\xi, 0)$.
The map $D_{u_0}S$: Similarly for any $h \in \mathcal{H}^1$, $D_{u_0}S(u_0,\eta)h = \bar v_1$, where $\bar v_t$ solves (54) with $\xi = 0$, $\bar v_0 = h$. By Lemma 2.1 with $j = 0$, $\bar v \in \mathcal{U}_2$; so $\bar v(1) \in \mathcal{H}^1$. To improve the smoothness and show that $\bar v(1) \in \mathcal{H}^2$ we consider the function $w = t\bar v$. Calculating $\partial_t w$ and using that $\bar v$ satisfies (54) with $\xi = 0$, we find that in its turn, w satisfies (54) with $v_0 = 0$, $\xi = \bar v \in \mathcal{U}_2$. Then by Lemma 2.1, $w \in \mathcal{U}_3$. It follows that $\bar v_1 = w_1 \in \mathcal{H}^2$ analytically depends on (u_0,η) for the same reason as above. This proves the required analyticity of the mapping S. □

The last proposition implies (A1). To verify (H2) we note that since

$$\frac{1}{2}\frac{d}{dt}\|u\|_1^2 = \langle Lu, \partial_t u\rangle = -\nu\|u\|_2^2 + \langle Lu,\eta\rangle \le -\frac{\nu}{2}\|u\|_2^2 + \frac{1}{2\nu}\|\eta\|^2,$$

then $\|u_1\|_1^2 \le e^{-\nu}\|u_0\|_1^2 + \beta$ for $\eta \in \mathcal{K}$, with some constant β. So (H2) follows.
It remains to check assumption (H3):

Lemma 2.3. *For any $(u_0,\eta) \in \mathcal{H}^1 \times E_1$ the mapping $D_\eta S(u_0,\eta) : E_1 \mapsto \mathcal{H}^1$ has dense image.*

Proof. Denote by $u = u(u_0,\eta) \in \mathcal{U}_2$ the solution of equation (51) and by $v_t(v_0,\xi)$ the solution of equation (54). Define $\mathcal{L} : E_1 \mapsto \mathcal{H}^1$ by $\mathcal{L}(\xi) = v_1(0,\xi)$. Then $D_\eta S(u,\eta)\xi = \mathcal{L}(\xi)$, so we should show that the mapping $\mathcal{L} : E_1 \mapsto \mathcal{H}^1$ has dense image. According to Fredholm's alternative, we only need to verify that the adjoint operator $\mathcal{L}^* : \mathcal{H}^1 \mapsto E_1$ has trivial kernel.
In order to do that let us consider the dual problem

$$\begin{cases} -\partial_t w + \nu L w - \Pi(\langle u,\nabla\rangle w) + \Pi(du)^* w = 0, \\ w_1 = w, \end{cases} \tag{55}$$

where $((du)^*w)^j = \sum_{l=1}^2 \frac{\partial u^l}{\partial x^j} w^l$. It is dual to the problem (54) with $\xi = 0$ in the sense that if v solves (54) with $\xi = 0$ and w solves (55), then

$$\langle v_t, w_t\rangle = \text{const}.$$

For a fixed $\tau \in [0,1]$, let $S_\tau^t : v \mapsto v_t$, $\tau \le t \le 1$, be the resolving operator for equation (54) with $\xi = 0$ and the initial value $v_\tau = v$. Similar, let $\tilde S_1^t : w \mapsto w_t$ be the resolving operator for eq. (55) with terminal value $w_1 = w$. Then $\tilde S_1^t$ is the dual operator for S_t^1 with respect to the L_2–scalar product. Accordingly, for any $\eta \in E_1$ and $w \in \mathcal{H}^2$, by Duhamel's principle,

$$\int_0^1 \langle \eta_t, (\mathcal{L}^* w)_t\rangle_1 dt = \langle \eta, \mathcal{L}^* w\rangle_{E_1} = \langle \mathcal{L}(\eta), w\rangle_1 = \langle \int_0^1 S_t^1 \eta_t dt, w\rangle_1 = \int_0^1 \langle \eta, \tilde S_1^t L w\rangle dt. \tag{56}$$

If $w \in \mathcal{H}^1$, then we approximate w in $\mathcal{H}^1$ by smooth functions $w^n, n \ge 1$, and substitute

$$\tilde S_1^t L w^n = L^{1/2}\xi_t^n. \tag{57}$$

Then $\xi_1^n = L^{-1/2}Lw^n = L^{1/2}w^n$ and

$$-\partial_t \xi^n + \nu L\xi^n + L^{-1/2}\Pi\big(-\langle u,\nabla\rangle L^{1/2}\xi^n + (du)^* L^{1/2}\xi^n\big) = 0, \quad \xi_1^n = L^{1/2}w^n.$$

Taking a scalar product of this equation with ξ^n in $\mathcal{H}$ and using that the vectors $\xi_1^n = L^{1/2}w^n$ are bounded uniformly in $\mathcal{H}$ we find that $|\xi^n|_{\mathcal{U}_1} \le C(|u|_{\mathcal{U}_2}, \|w\|_1)$, uniformly in n. So $\xi^{n_j} \rightharpoonup \xi \in \mathcal{U}_1$ weakly in $\mathcal{U}_1$, for a suitable sequence $n_j \to \infty$. From this convergence, (57) and (56) we get that

$$(L\mathcal{L}^*w)(t) = L^{1/2}\xi_t \in C([0,1], \mathcal{H}^{-1}), \quad (L\mathcal{L}^*w)(1) = L^{1/2}\xi_1 = Lw.$$

Now let $\bar{w}_1 \in \mathcal{H}^1$ be such that $\mathcal{L}^*\bar{w}_1 = 0$. Then $L\bar{w}_1 = (L\mathcal{L}^*\bar{w}_1)(1) = 0$, which implies $\bar{w}_1 = 0$ and completes the proof. □

It follows that Theorem 1.5 applies and implies that

Theorem 2.4. *Suppose that the noise η has the form* (48), (49), *where $\{\phi_i\}$ is the trigonometric Hilbert basis of $\mathcal{H}^1$. Then for any $\nu > 0$, the Markov chain (u_k, P_k), defined by the 2D Navier-Stokes equation* (50) *in $\mathcal{H}^1$, has a unique stationary measure $\mu_\nu \in \mathcal{P}(\mathcal{H}^1)$. Furthermore, for any $R > 0$ and any measure μ supported by the ball $B_{\mathcal{H}^1}(R)$, there exist $C = C(R) > 0$ and $\kappa = \kappa(R) \in (0,1)$ such that*

$$\|\mathcal{P}_k^*\mu - \mu_\nu\|^*_{L(\mathcal{H}^1)} \le C\kappa^k, \quad \forall\, k \ge 1.$$

Let u_t^ν be a solution of (50) such that $\mathcal{D}(u_0^\nu) = \mu_\nu$. Then $\mathcal{D}(u_k^\nu) = \mu_\nu$ for $k \in \mathbb{N}$, while for $t = k + \tau$, $0 \le \tau < 1$, we have

$$\mathcal{D}(u_t^\nu) = \Sigma^\tau \mathcal{D}(u_k^\nu) = \Sigma^\tau \mu,$$

where Σ^τ is a Lipschitz operator in the space $(\mathcal{P}(X), \|\cdot\|_L^*)$. So we get

Corollary 2.5. *Under the assumptions of Theorem 2.4 let $u_t(x)$ be a solutions of* (50) *such that $\mathcal{D}(u_0) = \mu$. Then*

$$\|\mathcal{D}(u_t) - \mathcal{D}(u_t^\nu)\|^*_{L(\mathcal{H}^1)} \le C\kappa^t, \quad \forall\, t \ge 0.$$

Note that the mapping $S : \mathcal{H}^0 \times E_1 \to \mathcal{H}^1$ is also well defined (see e.g. [5]), and for any initial distribution $\mu \in \mathcal{P}(\mathcal{H}^0)$, $\mathcal{D}(u_1) \in \mathcal{P}(\mathcal{H}^1)$. By applying Theorem 2.4 and Theorem 1.5, we have the following result.

Corollary 2.6. *For any $\mu \in \mathcal{P}(\mathcal{H}^0)$, supported by the ball $B_{\mathcal{H}^0}(R)$, there exist $C = C(R) > 0$ and $\kappa = \kappa(R) \in (0,1)$ such that*

$$\|\mathcal{P}_k^*\mu - \mu_\nu\|^*_{L(\mathcal{H}^1)} \le C\kappa^k, \quad \forall\, k \ge 1.$$

Furthermore, if $\nu \in \mathcal{P}(\mathcal{H}^0)$ is any measure, then $\mathcal{P}_k^*\nu \rightharpoonup \mu_\nu$ as $k \to \infty$.

Remark 2.7. *Now consider the 2D NSE* (50), *where $\eta(t,x)$ is a random kick force (see* [5], *Section 2.3): $\eta = \sum_{k=1}^\infty \eta_k\, \delta(t-k)$. Here $\{\eta_k\}$ is a sequence of i.i.d. random variables in $\mathcal{H}^2$. Solutions $u_t, t \ge 0$, of the equation define the map $S : \mathcal{H}^1 \times \mathcal{H}^2 \to \mathcal{H}^1$ by the relation*

$$S(u_0, \eta_1) = u_1 = S_1(u_0) + \eta_1,$$

where S_1 is the time-one flow-map for the free equation $(50)_{\eta=0}$. According to Proposition 2.2, the map S satisfies the regularity assumption (A1) with $E := \mathcal{H}^2$. Now as in [5], *Section 3.2.4, set $\eta_k = \sum_{j=1}^\infty b_j \xi_{jk} \phi_j$, where $\{b_j, j \ge 1\}$, is an l_2–sequence of nonzero real numbers, $\{\phi_j, j \ge 1\}$, is the orthonormal trigonometric basis of the space $\mathcal{H}^2$, and $\{\xi_{jk}\}$ are i.i.d. real variables whose density function is Lipschitz–continuous and does not vanish at the origin. Then (H1), (H2) and (H3) hold trivially. Applying Theorem 1.5 we recover the well known result that the 2D NSE with a non-degenerate random kick force is exponentially mixing, see in* [5].

Similarly Theorem 1.5 applies to the CGL equation (62) as below in Example 2.10, where η is a nondegenerate kick-force.

2.3. **Quasilinear parabolic systems on $\mathbb{T}^d$.** In this part we consider quasilinear parabolic systems

$$\partial_t u = \Delta u + f(x, u, \nabla u) + \eta, \quad u(0) = u_0, \tag{58}$$

where $f(x, 0, 0) = 0$, $u = u_t(x) \in \mathbb{R}^n$, $x \in \mathbb{T}^d$ and $f : \mathbb{T}^d \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \mapsto \mathbb{R}^n$ is a C^∞–smooth function. We restrict ourselves to the case when the solution u is sought in the space

$$H^m = H^m(\mathbb{T}^d; \mathbb{R}^n), \quad m > \frac{d}{2} + 1,$$

and η_t is a process as in Section 2.1 with $H = H^m$.
We are going to apply Theorem 1.5 with $H = H^m$. To do that we will make two assumptions concerning the well-posedness and regularity of eq. (58). To formulate them we define the following spaces, where $k \in \mathbb{Z}$ and $T > 0$:

$$E_k^T = L_2([0, T]; H^k), \qquad \mathcal{U}_k^T = \{u \in E_k^T : \partial_t u \in E_{k-2}^T\} \subset C([0, T]), H^{k-1}),$$

and abbreviate $E_k^1 = E_k$, $\mathcal{U}_k^1 = \mathcal{U}_k$.

(A0) (**Well-posedness**). For any $M, T > 0$, $u_0 \in H^m$ and any $\eta \in E_{m-1}^T$ such that $\|\eta_t\|_{m-1} \le M$ for all t, the problem (58) has a unique solution $u \in \mathcal{U}_{m+1}^T$. It satisfies

$$\|u_t\|_m \le C(M, \|u_0\|_m) \quad \forall\, t. \tag{59}$$

(H1′) (**Random force**). The force η has the form (48), (49), where $\{\phi_i\}$ is a Hilbert base of H^m, $\{h_{ij}\}$ is the Haar base and

$$M^2 = (\sup_i \sum_j 2^{\frac{j}{2}} |c_j^i|)^2 (\sum_i b_i^2) < \infty.$$

(H2′) (**Dissipativity**). There exists $0 \le r' \le m$ and $\kappa > 0$, such that if u solves (58) with $\eta = 0$ and $u_0 \in H^m$, then

$$\|u_t\|_{r'} \le e^{-\kappa t} \|u_0\|_{r'}. \tag{60}$$

Note that due to (H1′)

$$\|\eta_t\|_m \le M \quad \forall\, t,\ \forall\, \omega. \tag{61}$$

Remark 2.8. *Here we assume that the random force η is bounded uniformly in ω and t, while in Section 2.2 we assumed that it is bounded in t in the L_2–sense (i.e. in the norm of the space E). This is needed since the class of equations* (58) *for which it is possible to prove well-posedness of the initial value problem for $u_0 \in H$ and $\eta \in E_m^T$ is smaller than the class of equations for which we can prove the well-posedness for the problem with bounded in time η (e.g. we cannot prove that assumption (A0) holds for the CGL equations as in Example 2.10 without the additional restriction $\|\eta_t\|_{m-1} \le M$). So working with bounded in time random forces as in* (61) *we can apply our main theorem to a larger class of quasilinear equation* (58).

Lemma 2.9. *Under the assumption (A0) there exists an open neighbourhood Q of $H^m \times \mathcal{K}$ in $H^m \times E_{m-1}$ which contains each point $(u_0, \eta) \in H^m \times \mathcal{K}$ with its vicinity in $H^m \times E_{m-1}$ of radius which depends only on $\|u\|_m$, such that for every $(u_0, \eta) \in Q$ the problem* (58) *has a unique solution u. Moreover, the mapping*

$$Q \ni (u_0, \eta) \mapsto u \in \mathcal{U}_{m+1}$$

is C^∞–smooth, and for any $k \ge 1$ its C^k–norm is bounded on bounded sets.

Proof. Consider the mapping

$$\Phi : \mathcal{U}_{m+1} \mapsto H^m \times E_{m-1}; \quad u \mapsto (u_0, \partial_t u - \Delta u - f(x, u, \nabla u)).$$

It is C^∞ smooth (the smoothness of the nonlinear component follows from a much more general result in [8], pp. 14, 381), and by (A0), its image contains $H^m \times \mathcal{K}$. For any $(u_0, \eta) \in H^m \times \mathcal{K}$ and for $u = \Phi^{-1}(u_0, \eta)$, the linear mapping

$$d\Phi(u) : \mathcal{U}_{m+1} \mapsto H^m \times E_{m-1}; \quad d\Phi(u)(v) = (v_0, \Delta v + D_u f(x, u, \nabla u)v + D_{\nabla u} f(x, u, \nabla u)\nabla v).$$

is an isomorpism of $\mathcal{U}_{m+1}$ and $H^m \times E_{m-1}$ (see e.g. in [9]), so by the inverse map theorem the point (u_0, η) has a neighbourhood $Q_{(u_0,\eta)} \subset H^m \times E_{m-1}$, where the inverse mapping Φ^{-1} exists and is C^∞. By the constructive nature of the inverse map theorem, the size of the neighbourhood and the norms of derivatives of the inverse mapping are bounded in terms of the norms $\|u_0\|_m$ and $|\eta|_{E_{m-1}}$. Since η belongs to the compact set $\mathcal{K}$, these quantities may be chosen η–independent. Taking for Q the open set $\cup_{(u_0,\eta)} Q_{(u_0,\eta)}$ we arrive at the conclusion. □

Due to the lemma and (61), Remarks 1.1, 1.10 apply to the equations (58) if the assumptions $(A0) - (H2')$ hold.

Example 2.10. *Consider the complex Ginzburg-Landau (CGL) equation:*

$$\partial_t u - (\nu_1 + i\nu_2)\Delta u + \gamma u + i|u|^{2r} u = \eta, \quad x \in \mathbb{T}^d, \quad u(0) = u_0, \tag{62}$$

where $\gamma > 0, \nu_1 > 0$ and $\nu_2 \geq 0$. This is an example of system (58) *with $n = 2$. Assumption (H2′) with $r' = 0$ obviously holds for all equations* (62)*. Assumption (A0) also is fulfilled for a large class of the equations. In particular, it holds if $\nu_2 > 0$ and $d \leq 2$, $r \in \mathbb{N}$ or $d = 3$ and $r = 1$; or if $\nu_2 = 0$ and $r \in \mathbb{N}$. See Appendix.*

We may assume that

$$u_0 \in B_{H^m}(R) \quad \text{a.s.},$$

for some $R > 0$. According to assumption (A0), the mapping $S : H^m \times E_{m-1} \supset Q_R \to H^m$ is C^2–smooth and its C^2–norm is bounded. Here Q_R is a neighbourhood of $B_{H^m}(R) \times \mathcal{K}$ in $H^m \times E_{m-1}$ which contains each point of $B_{H^m}(R) \times \mathcal{K}$ with its neighbourhood of radius $c_R > 0$. As above, $S(u_0, \eta) = u_1$ where u_t, $0 \leq t \leq 1$, is a solution of (58). So (58) defines in H^m the random dynamical system (52). By (61) and (A0) the trajectories of (52) with $u_0 \in B_{H^m}(R)$ satisfy $\|u_k\|_m \leq R_+$, $R_+ = R_+(R, M)$, for all k, a.s. In the following three steps we will check that the assumptions (A1), (H2) and (H3) hold with $H = H^m, E = E_m$ and $V = H^{m+1}$. Verifying assumption (H2), we only need to take into account Remark 1.2.

Step 1. (**Improvement of regularity**) In view of (A0) to prove (A1) it remains to show that S restricted to $Q_R \subset H^m \times E_m$ takes values in H^{m+1} and is smooth. Evoking equality (53) (where now S is the operator in (52)) we see that it suffices to examine the regularity of the maps $D_{u_0} S(u_0, \eta)h$ and $D_\eta S(u_0, \eta)\xi$. Let $u = u(u_0, \eta) \in \mathcal{U}_{m+1}$ solves (58) with $u_0 \in H^m$ and $\eta \in E_m$. Since $D_\eta S(u_0, \eta)\xi = v_1(0, \xi; u)$ and $D_{u_0} S(u_0, \eta)h = v_1(h, 0; u)$, where $v_t(v_0, \xi; u)$ solves the linear system

$$\partial_t v = \Delta v + D_u f(x, u, \nabla u)v + D_{\nabla u} f(x, u, \nabla u)\nabla v + \xi, \quad v(0) = v_0, \tag{63}$$

we only need to examine the regularity of the mapping $(v_0, \xi, u) \mapsto v_1(v_0, \xi; u)$. Due to the classical results from the linear parabolic theory (see [9], Part 7), if $u \in \mathcal{U}_{m+1}$, then for $j = m$ and $j = m+1$, for $v_0 \in H^j, \xi \in E_{j-1}$ eq. (63) has a unique solution $v \in \mathcal{U}_{j+1}$. By an argument similar to that in Proposition 2.2 (but using a smooth version of the implicit function theorem), we show that v

smoothly depends on $u \in \mathcal{U}_{m+1}$. So $v_1(0,\xi;u) \in H^{m+1}$ smoothly depends on $u \in \mathcal{U}_{m+1}$, $\xi \in E_m$. To analyse $v_t = v_t(h,0;u)$, as in Section 2.2 we consider $w = tv$. This vector-function satisfies (63) with $v_0 = 0, \xi = v(h,0;u)$. So $v_1 = w_1 \in H^{m+1}$ smoothly depends on $\xi \in \mathcal{U}_{m+1}$, i.e. on $h \in H^m$ and $u \in \mathcal{U}_{m+1}$. This implies the smoothness, required in (A1).

Remark 2.11. *By a similar argument we can deduce that the mapping S is analytic, if f is.*

Step 2. Below we follow [6], Section 4.2. Firstly we claim that for any $R > 0$ and $u_0 \in H^m$ with $\|u_0\|_m < R$,

$$\|S(u_0,0)\|_m \le C_R \|u_0\|_{r'}, \tag{64}$$

where $C_R > 0$ depends only on R. Indeed, since $f(x,0,0) = 0$, then

$$S(u_0,0) = \int_0^1 (\partial/\partial t) S(tu_0,0)u_0 \, dt = \int_0^1 D_u S(tu_0,0)u_0 \, dt,$$

and we only need to show that $\|D_u S(tu_0,0)u_0\|_m \le C_R\|u_0\|_{r'}$ if $\|u_0\|_m \le R$. Denote by $\tilde{u}_t(tu_0,0)$ the solution of equation (58) with u_0 replaced by tu_0 and $\eta = 0$. By (61) and (59), $\|\tilde{u}_t\|_m \le C(R,r) =: R_+$ for all t. Consider the following linear system

$$\partial_t v = \Delta v + D_u f(x,\tilde{u},\nabla\tilde{u})v + D_{\nabla u} f(x,\tilde{u},\nabla\tilde{u})\nabla v, \quad v(0) = u_0; \tag{65}$$

then $D_u S(tu_0,0)u_0 = v_1$. Choosing any points

$$0 < t_{r'+1} < t_{r'+2} < \cdots < t_m = 1$$

and arguing as at Step 1 we find that $\|v_{t_{r'+1}}\|_{r'+1} \le C'_R\|u_0\|_{r'}, \dots, \|v_{t_m}\|_m \le \|v_{t_{m-1}}\|_{m-1}$. So $\|v_1\|_m \le C''_R\|u_0\|_{r'}$ and (64) follows.
Due to (64) and Assumption (H2′), in the space H^m there exists an equivalent norm $\|\cdot\|'_m$ such that for any $u_0 \in H^m$ with $\|u_0\|_m \le R$,

$$\|S(u_0,0)\|'_m \le q\|u_0\|'_m, \tag{66}$$

where $q \in (0,1)$ depends on R. Indeed, let $\|\cdot\|'_m := \|\cdot\|_{r'} + \varepsilon\|\cdot\|_m$, where $\varepsilon > 0$ is a parameter to be defined. We have

$$\|S(u_0,0)\|'_m = \|S(u_0,0)\|_{r'} + \varepsilon\|S(u_0,0)\|_m \le e^{-\kappa}\|u_0\|_{r'} + \varepsilon C_R\|u_0\|_{r'} \le (e^{-\kappa} + \varepsilon C_R)\|u_0\|'_m.$$

It remains to choose ε so small that $q := e^{-\kappa} + \varepsilon C_R < 1$. Since the C^2–norm of the map $S : Q_R \to H^m$ is bounded, then (66) implies (9) for $\|u\|_m \le R_+$ with the norm $\|\cdot\|_m$ replaced by $\|\cdot\|'_m$. This proves assumption (H2) in the weaker form, suggested in Remark 1.2.

Step 3. Now we verify (H3), i.e. check that for any $u_0 \in H^m, \eta \in E_m$, the linear operator $D_\eta S(u_0,\eta) : E_m \mapsto H^m$ has dense image. Define the operator $\mathcal{L} : E_m \mapsto H^m$ by $\mathcal{L}(\xi) = v_1(0,\xi)$, where $v_t(v_0,\xi)$ solves equation (63). By Fredholm's alternative it suffices to show that $\mathcal{L}^* : H^m \mapsto E_m$ has trivial kernel. Denote $S^1_t : v_t \mapsto v_1$ the resolving operator for equation (63) with $\xi = 0$, $v(t) = v_t$. Consider the adjoint system

$$-\partial_t w = \Delta w + \big(D_u f(x,u,\nabla u)\big)^* w - \operatorname{div}\big[\big(D_{\nabla u} f(x,u,\nabla u)\big)^* w\big], \quad w(1) = w_1,$$

and denote by $\tilde{S}^t_1 : w_1 \mapsto w_t$, $0 \le t \le 1$, its resolving operator with initial condition at $t = 1$. Then $\langle v_t, w_t\rangle \equiv$ constant, so the operator $\tilde{S}^t_1$ is the L^2–dual of S^1_t. It follows that, for $\eta \in E_m$ and

$w \in H^{2m}$,

$$\int_0^1 \langle \eta_t, (\mathcal{L}^* w)_t \rangle_m dt = \langle \eta, \mathcal{L}^* w \rangle_{E^1_m} = \langle \mathcal{L}(\eta), w \rangle_m = \langle \int_0^1 S^1_t \eta_t dt, w \rangle_m$$
$$= \int_0^1 \langle L^m w, S^1_t \eta_t \rangle dt = \int_0^1 \langle \tilde{S}^t_1 L^m w, \eta_t \rangle dt, \quad L = 1 - \Delta.$$

If $w \in H^m$, then arguing as in Section 2.2 we get that

$$L^m \mathcal{L}^* w = L^{m/2} \xi, \;\; \xi \in \mathcal{U}_m; \qquad \xi_1 = L^{m/2} w.$$

So $\mathcal{L}^* w = 0$ implies that $\xi_1 = 0$ and $w = 0$, i.e. $\mathcal{L}^*$ has trivial kernel.

Now an application of Theorem 1.5 implies the validity for eq. (58) of obvious reformulations of Theorem 2.4 and Corollary 2.5.

2.4. **Appendix.** *Case* $\nu_2 > 0$. We will only discuss equations with $d \geq 2$ and to simplify notation take $\nu_1 = \nu_2 = \gamma = 1$:

$$\partial_t u - i\Delta u + (u - \Delta u) + i|u|^{2r} u = \eta_t, \quad \|u_0\|_m =: R, \tag{67}$$

where $\|\eta_t\|_{m-1} \leq M$ for all η, with a fixed $m > d/2$. We start with apriori estimates, assuming that u is a smooth solution of (67).

Step 1. Noting that eq. (67) with $\eta = 0$ and with removed term $u - \Delta u$ is hamiltonian with the Hamiltonian $H(u) = \int \big(\frac{1}{2}|\nabla u|^2 + \frac{1}{2r+2}|u|^{2r+2}\big) dx$ and since $\nabla H = -\Delta u + |u|^{2r} u$, we find that

$$\partial_t \big(H(u_t) + \tfrac{1}{2}\|u_t\|_0^2\big) = \langle \nabla H + u, \Delta u - u \rangle + \langle \nabla H + u, \eta \rangle$$
$$\leq -C(H(u_t) + \tfrac{1}{2}\|u_t\|_0^2) + \|u_t\|_1 \|\eta_t\|_1 + \int |u|^{2r+1} |\eta| \, dx \leq -\tfrac{C}{2}(H(u_t) + \tfrac{1}{2}\|u_t\|_0^2) + C_1(\|\eta_t\|_1).$$

So

$$H(u_t) + \tfrac{1}{2}\|u_t\|_0^2 \leq C(R, M) \quad \forall\, t \geq 0. \tag{68}$$

Step 2. Now let us consider $\frac{1}{2}\partial_t \|u\|_m^2$, where $\|u\|_m^2 = \|u\|_0^2 + \|\nabla^m u\|_0^2$. We have:

$$\tfrac{1}{2}\partial_t \|u\|_m^2 + \|u\|_{m+1}^2 \leq |\langle \nabla^m |u|^{2m} u, \nabla^m u \rangle| + |\langle \eta_t, ((-\Delta)^m + 1) u \rangle|. \tag{69}$$

The first term in the r.h.s. is bounded by a finite sum of the terms of the form

$$U = C \int |v^1 \dots v^{2r+2}| \, dx, \quad v^j = \partial^{a_j} u \text{ or } v^j = \partial^{a_j} \bar{u},$$

where

$$|a_1| + \dots + |a_{2r+2}| = 2m, \quad |a_j| \leq m \;\forall\, j. \tag{70}$$

Then

$$U \leq \prod_1^{2r+2} |v^j|_{p_j}, \quad \sum \frac{1}{p_j} = 1. \tag{71}$$

By (68) and the Sobolev embedding,

$$|u_t|_{L_q} \leq C_1(R, M, \varepsilon) \;\; \forall\, t \geq 0, \qquad \frac{1}{q} = \frac{d - 2 + \varepsilon}{2d}.$$

Here and below $\varepsilon = 0$ if $d \geq 3$ and ε is any positive number if $d = 2$. So by the Gagliardo–Nirenberg inequality

$$|v_t^j|_{p_j} \leq C\|u_t\|_{m+1}^{\alpha_j}|u_t|_{L_q}^{1-\alpha_j} \leq C_1(R,M,\varepsilon)\|u_t\|_{m+1}^{\alpha_j}, \tag{72}$$

where

$$\frac{1}{p_j} = \frac{|a_j|}{d} + \left(\frac{1}{2} - \frac{m+1}{d}\right)\alpha_j + \frac{(1-\alpha_j)(d-2+\varepsilon)}{2d} = \frac{2|a_j| + (d-2+\varepsilon)}{2d} - \frac{\alpha_j(m-\varepsilon)}{d}.$$

Denoting $\alpha = \sum \alpha_j$ we get from the last relation and (70) that

$$1 = \sum \frac{1}{p_j} = \frac{4m + (d-2)(2r+2) + \varepsilon'}{2d} - \frac{\alpha(m-\varepsilon')}{d}.$$

So $\alpha = \big(4m + (d-2)(2r+2) - 2d + \varepsilon'\big)/2m$, which is bounded by 2 if ε' is sufficiently small and $d < 2\frac{r+1}{r}$; i.e. r is any if $d = 2$ and $r = 1$ if $d = 3$. Under this condition relations (69), (71) and (72) jointly imply that

$$\partial_t\|u\|_m^2 + 2\|u\|_{m+1}^2 \leq C(M,R)\|u\|_{m+1}^{\alpha} + CM\|u\|_{m+1}, \quad \alpha < 2.$$

From here

$$\|u_t\|_m \leq C_1(M,R) \quad \forall\, t \geq 0; \qquad |u|_{\mathcal{U}_{m+1}^T} \leq C_2(M,R,T) \quad \forall\, T > 0.$$

Step 3. The obtained estimates imply (A0) via Galerkin's method.

Case $\nu_2 = 0$, d is any. In this case the function $r(t,x) = |u(t,x)|$ satisfies a differential inequality with the maximum principle, e.g. see in [4], where the white in time stochastic force ζ_t has to be replaced by the easier for analysis force η_t. This implies that $|u_t|_{L_\infty} \leq C(R,M)$ for all t and any dimension d. Then (A0) follows by the same argument as at the Steps 2-3 above.

References

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

[2] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), 993–1032.

[3] 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.

[4] S. Kuksin, *Stochastic nonlinear Schrödinger equation 1. A priori estimates*, Proceedings of the Steklov Institute of Mathematics **225** (1999), 219–242.

[5] S. Kuksin and A. Shirikyan, *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, Cambridge, 2012.

[6] S. Kuksin and V. Nersesyan, A. Shirikyan, *Exponential mixing for a class of dissipative PDEs with bounded degenerate noise*, GAFA (2020), to appear; arXiv:1802.03250v2.

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

[8] T. Runst and W. Sickel, *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*, Walter de Gruyter, Berlin & New York, 1996.

[9] O. A. Ladyzhenskaja, V. A. Solonnikov and N. N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, "Nauka", Moscow, 1967; English trans., AMS, Providence, R. I., 1968.

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

S. Kuksin, Institut de Mathémathiques de Jussieu-Paris Rive Gauche, CNRS, Université Paris Diderot, UMR 7586, Sorbonne Paris Cité, F-75013, Paris, France; and School of Mathematics, Shandong University, Shanda Nanlu 27, 250100, PRC; and Saint Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, Russia,
E-mail address: sergei.kuksin@imj-prg.fr

H. Zhang, Institute of Mathematics, Fudan University, Handan Road 220, 200433, PRC,
E-mail address: huilinzhang2014@gmail.com