Mixing for dynamical systems driven by stationary noises

Sergei Kuksin, Armen Shirikyan

Mixing for dynamical systems driven by stationary noises

Sergei Kuksin* Armen Shirikyan†

February 16, 2025

Abstract

The paper deals with the problem of long-time asymptotic behaviour of solutions for classes of ODEs and PDEs, perturbed by stationary noises. The latter are not assumed to be δ-correlated in time, so that the evolution in question is not necessarily Markovian. We first prove an abstract result which imply the mixing for random dynamical systems satisfying appropriate dissipativity and controllability conditions. It is applicable to a large class of evolution equations, and we illustrate it on the examples of a chain of anharmonic oscillators coupled to heat reservoirs, the 2d Navier–Stokes system, and a complex Ginzburg–Landau equation. Our results also apply to the general theory of random processes on the 1d lattice and allow one to get for them results related to Dobrushin's theorems on reconstructing processes via their conditional distributions. The proof is based on an iterative construction with quadratic convergence. It uses the method of Kantorovich functional, introduced in [KPS02, Kuk02, Kuk06] in the context of randomly forced PDEs, and some ideas suggested in [Shi15, KNS20] to prove mixing with the help of controllability properties of an associated system.

AMS subject classifications: 35Q30, 35Q56, 37H30, 37L40, 60G10, 60H25, 60J05, 76D06, 76F25

Keywords: Navier–Stokes system, complex Ginzburg–Landau equation, chain of oscillators, stationary noise, exponential mixing

Contents

0 Introduction **2**

*Université Paris Cité and Sorbonne Université, CNRS, IMJ-PRG, 75013 Paris, France & Peoples Friendship University of Russia (RUDN University) & Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia; Email: Sergei.Kuksin@imj-prg.fr

†Department of Mathematics, CY Cergy Paris University, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France; Email: Armen.Shirikyan@cyu.fr

1 A class of random dynamical systems **12**
1.1 Description of the model 12
1.2 Dynamical systems with stationary noises 14
1.3 A criterion for mixing 16
1.4 Proof of Theorem 1.3 18

2 Main results **24**
2.1 Formulation of the main theorem 24
2.2 Proof of Theorem 2.4 27
2.3 Concluding remarks 31
2.4 Proof of Lemma 2.5 32

3 Stationary processes satisfying the mixing hypothesis **33**
3.1 Existence and uniqueness 33
3.2 Strong recurrence, strong irreducibility, and mixing 36
3.3 Examples of processes satisfying the hypotheses of Section 2.1 . . 37
3.3.1 Discrete-time Markovian noises 38
3.3.2 Discrete-time moving averages 38
3.3.3 Continuous-time noise 39

4 Application to randomly forced ODE and PDE **41**
4.1 Chain of anharmonic oscillators coupled to heat reservoirs 41
4.2 Two-dimensional Navier–Stokes system 46
4.3 Complex Ginzburg–Landau equation 48

5 Appendix: the image of transition probabilities under Lipschitz maps **49**

Bibliography **50**

0 Introduction

In a Hilbert space H of finite or infinite dimension, we consider smooth discrete-time and continuous-time random dynamical systems, stirred by bounded stationary stochastic processes. Our goal is to study the long-time asymptotic behaviour of their trajectories. That is, we examine random dynamical systems (RDS) of the form

$$u_k = S(u_{k-1}, \eta_k), \qquad k \geq 1; \qquad u_0 = v; \qquad u_k \in H; \tag{0.1}$$

$$\dot{u}(t) = F(u(t), \eta(t)), \qquad t \geq 0; \qquad u(0) = v; \qquad u(t) \in H. \tag{0.2}$$

Here $\{\eta_k^\omega, k \in \mathbb{Z}\}$ and $\{\eta^\omega(t), t \in \mathbb{R}\}$ are bounded stationary processes in a Banach space E, and the mappings $S : H \times E \to H$ and $F : D \times E \to H$ are C^2-smooth (in the sense of Fréchet), where $D \subset H$ is a suitable dense subspace. We begin with reducing Equation (0.2) to a discrete-time system of the form (0.1). To do it, we generalise the setting a bit and assume that the

random input $\eta^\omega(t)$ entering (0.2) is a bounded process in E that is stochastically T-periodic for some $T>0$; that is,

$$\mathcal{D}\big(\eta(\cdot)\big)=\mathcal{D}\big(\eta(\cdot+T)\big)$$

(so T is any positive number if the process η is stationary). We also assume that, for any $v\in H$ and $\eta\in L^2_{\rm loc}(\mathbb{R}_+;E)$, Equation (0.2) has a unique solution $u(t)$ in an appropriate functional class, and that the mapping

$$S_T: H\times L_2(0,T;E)\to H,\quad (v,\eta_{(0,T)})\mapsto u(T)$$

is C^2-smooth, and its first- and second-order derivatives are bounded on bounded subsets. Let us define intervals $J_k=[(k-1)T,kT)$ for $k\in\mathbb{Z}$ and denote

$$\eta_k:=\eta\big|_{J_k}\in L_2(J_k;E). \tag{0.3}$$

We naturally identify intervals J_k with $J_T:=[0,T)$ and spaces $L_2(J_k;E)$ with $L_2(J_T;E)=:\hat E$. Then $\{\eta_k,k\in\mathbb{Z}\}$ becomes a stationary process in $\hat E$, and if we set $u_k=u(kT)$ for $k\ge0$, then the sequence $\{u_k,k\ge0\}$ satisfies Equation (0.1) with $S=S_T$ and E replaced by $\hat E$. Thus, we have reduced the problem of long-time behaviour of trajectories of (0.2) to that for some system (0.1). So below in Introduction, we shall talk about systems (0.1), and only at the end shall discuss some applications to nonlinear PDEs which can be written as systems (0.2).

System (0.1)

We assume that $\mathcal{K}:=\operatorname{supp}\mathcal{D}(\eta_k)$ is a compact subset of E and that some compact set $X\subset H$ is invariant for system (0.1); that is, $S(X\times\mathcal{K})\subset X$. In practice the compactness of $\mathcal{K}$ is a mild restriction after we assumed that the process η is bounded, while an invariant compact set $X\subset H$ exists if (0.1) is a dissipative finite-dimensional system, or if it comes from a system (0.2) which is a well-posed nonlinear parabolic PDE; see the applications discussed at the end of Introduction. Our goal is to prove that system (0.1) is exponentially mixing in the dual-Lipschitz distance (also known as the Kantorovich–Rubinstein distance).[1] That is, there exists a probability measure μ on X such that, for any initial state $v\in X$, the trajectory $\{u_k\}$ of (0.1) satisfies

$$\|\mathcal{D}(u_k)-\mu\|_L^*\le Ce^{-\gamma k},\quad k\ge0, \tag{0.4}$$

for some positive constants C and γ, independent of v.[2] If the random variables $\{\eta_k\}$ are i.i.d., then system (0.1) defines a Markov process in H. The

[1] Recall that the distance between two (Borel) probability measures μ and ν equals the dual-Lipschitz norm of $\mu-\nu$, defined as $\|\mu-\nu\|_L^*=\sup|\langle f,\mu\rangle-\langle f,\nu\rangle|$, where the supremum is taken over all Lipschitz functions on X such that $|f|\le1$ and $\mathrm{Lip}(f)\le1$. This distance metrises the weak convergence of measures; in particular, relation (0.4) implies that $\mathcal{D}(u_k)\rightharpoonup\mu$ as $k\to\infty$; e.g., see [KS12, Section 1.2.3].

[2] See below the first item of Remark 0.1 for a discussion of relation of this notion with the traditional mixing for stationary random processes.

mixing in Markov systems has a long history and goes back to the thirties of the last century. Staring from the pioneering works of Kolmogorov [Kol37] and Doeblin [Doe38, Doe40] various methods were developed to prove it in one or another form, similar to (0.4); see the books [Kha12, MT93, Bor98, KS12]. But if the random variables $\{\eta_k\}$ and $\{\eta(t)\}$ are not independent, then the dynamics in systems (0.1) and (0.2) are not Markovian, and their long time behaviour remains essentially an open problem, which was studied only for some specific models. When system (0.2) is a chain of anharmonic oscillators, coupled to heat baths, Eckmann–Pillet–Rey-Bellet [EPR99] used a Markovian reduction to prove, in some special cases, the convergence in law of its solutions to a unique stationary measure (see also Jakšić–Pillet [JP97, JP98] for general Hamiltonian systems coupled to one heat bath). A similar idea was used by Hairer [Hai05] and Hairer–Ohashi [HO07] for dissipative SDEs driven by a fractional Brownian motion. On the other hand, for some classes of non-Markovian systems that arise in applied probability and are rather different from (0.1) and (0.2), the mixing is known, probably, starting with Sevastianov's work [Sev57]; see also the paper [Ver17] and the references therein.

Our goal in this work is to obtain a theorem which establishes the mixing (0.4) for a large class of systems (0.1) in a phase space H of finite or infinite dimension (which includes the systems that Equations (0.2) may be reduced to). To do that, apart from the compactness assumption on the support $\mathcal{K}$ of $\mathcal{D}(\eta_k)$ and on the invariant set X, we now impose two more restrictions on the process $\{\eta_k\}$ and two restrictions on the mapping S. They are simplified versions of Hypotheses (GD), (ALC), (DLP′) and (SRZ), stated in Section 2.1.

Assumptions on the process $\{\eta_k\}$

We recall that $\boldsymbol{\eta} := \{\eta_k\}_{k\in\mathbb{Z}}$ is a stationary process in the Banach space E such that $\operatorname{supp}\mathcal{D}(\eta_k) = \mathcal{K}$ is a compact subset of E. For $l \in \mathbb{Z}$, we denote by $\boldsymbol{\eta}_l = (\dots, \eta_{l-1}, \eta_l)$ the *l-past* of the process, so that $\boldsymbol{\eta}_l$ is a random element of $\mathcal{K}^{\mathbb{Z}_-}$, where $\mathbb{Z}_- = \mathbb{Z}\backslash\mathbb{N}$. We provide the product space $E^{\mathbb{Z}_-}$ with the Tikhonov topology, which can be metrised on its compact subset $\boldsymbol{\mathcal{K}} := \mathcal{K}^{\mathbb{Z}_-}$ with the help of the distance

$$\mathrm{d}(\boldsymbol{\xi}, \boldsymbol{\xi}') = \sum_{k=-\infty}^{0} \iota^k \|\xi_k - \xi'_k\|_E, \quad \boldsymbol{\xi} = (\xi_k, k \in \mathbb{Z}_-), \quad \boldsymbol{\xi}' = (\xi'_k, k \in \mathbb{Z}_-), \quad (0.5)$$

where $\iota > 1$ is an arbitrary fixed number. For any $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$, consider the conditional distribution $\mathbb{P}\{\eta_1 \in \cdot \,|\, \boldsymbol{\eta}_0 = \boldsymbol{\xi}\} =: Q(\boldsymbol{\xi}; \cdot)$ of η_1 under the condition that the 0-past is fixed: $\boldsymbol{\eta}_0 = \boldsymbol{\xi}$. This is a probability measure on E, depending on $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ and supported by $\mathcal{K}$ (e.g., see [Dud02, Section 10.2]). For any integer $m \geq 1$, the measures $Q(\boldsymbol{\xi}; \cdot)$ allow one to calculate the conditional distribution $Q_m(\boldsymbol{\xi}; \Gamma) = \mathbb{P}\big\{(\eta_1, \dots, \eta_m) \in \Gamma \,|\, \boldsymbol{\eta}_0 = \boldsymbol{\xi}\big\}$, where $\Gamma \in \mathcal{B}(E^m)$ (see relation (2.6) in Section 2.1). For $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$, let $\{\eta_j(\boldsymbol{\xi}), j = 1, 2, \dots\}$ be a process in $\mathcal{K}$ such that, for any $m \geq 1$,

$$\mathcal{D}\big(\eta_1(\boldsymbol{\xi}), \dots, \eta_m(\boldsymbol{\xi})\big) = Q_m(\boldsymbol{\xi}; \cdot).$$

We impose on the process η the following two restrictions:

($\boldsymbol{\eta}$1) *If* $\dim E < \infty$, *then the measures* $\{Q(\boldsymbol{\xi};\cdot), \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}\}$ *have densities, so that*

$$Q(\boldsymbol{\xi}; \mathrm{d}x) = p_{\boldsymbol{\xi}}(x)\,\mathrm{d}x, \tag{0.6}$$

where the function $p : \boldsymbol{\mathcal{K}} \times X \to \mathbb{R}$ *is Lipschitz-continuous. If* E *is a Hilbert space with an orthonormal basis basis* $\{f_l\}$, *then* η_k *is representable in the form*

$$\eta_k = \sum_{j=1}^{\infty} r_l \eta_k^l f_l, \tag{0.7}$$

where the numbers $r_l > 0$ *are such that* $\sum |r_l| < \infty$, *and* $\{\eta_k^l, k \in \mathbb{Z}\}$ *are independent real-valued stationary processes (indexed by the integer* $l \geq 1$*) such that* $|\eta_k^l| \leq 1$, *and the corresponding conditional laws* $Q^l(\boldsymbol{\xi};\cdot)$ *are representable in the form* (0.6) *with some Lipschitz-continuous functions.*

Let us emphasise that no uniformity in l is required for the Lipschitz constants for the densities of the conditional laws $Q^l(\boldsymbol{\xi};\cdot)$. For a general form of this condition when E is a Banach space, see Hypothesis (DLP′) in Section 2.1.

The second assumption on the process η requires that it stays close to zero with positive probability, for every past:

($\boldsymbol{\eta}$2) *For any* $n \in \mathbb{N}$ *and* $\delta > 0$ *there exist* $s \in \mathbb{N}$ *and* $\varepsilon > 0$ *such that*

$$\mathbb{P}\{\|\eta_j(\boldsymbol{\xi})\| < \delta,\ j = s+1, \dots, s+n\} \geq \varepsilon \quad \text{for every } \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}.$$

Crucial assumption (η1) means not only regularity of the conditional distributions of the process η, but also that it forgets the past exponentially fast. Indeed, definition (0.5) of the distance on $\boldsymbol{\mathcal{K}}$ implies that $p_{\boldsymbol{\xi}}(x)$ regarded as a function of ξ_{-l} is Lipschitz-continuous with a constant of order ι^{-l}. Moreover, in Section 3.2, we show that the results of our work imply that every process satisfying (η1) and (η2) is mixing in some sense that is weaker than it is customary in ergodic theory, and which we call *feeble mixing*.

Condition (η1) is close to those used by Dobrushin in his work [Dob68, Dob70] on reconstructing the distribution of a lattice random field from its conditional distributions. The results of the above papers immediately imply that every family of measures $\{Q(\boldsymbol{\xi};\cdot), \boldsymbol{\xi} \in \mathcal{K}\}$ as in (η1) is a system of conditional distributions of a discrete-time stationary process η. However, they do not ensure the uniqueness of the law of η.[3] In this regard, let us note that, in Theorem 3.2, we derive from our results that, if in addition to (η1) a family of measures $\{Q(\boldsymbol{\xi};\cdot)\}$ satisfies (η2), then the law of the process η is defined uniquely.

[3] Dobrushin addressed the uniqueness problem in his later works devoted to Gibbs systems (e.g., see [Dob74]). To the best of our knowledge, those results were not extended to random processes.

Assumptions on the mapping S

We recall that the map $S : H \times E \to H$ is C^2-smooth and bounded on bounded sets, together with its derivatives of the first and second order. We assume in addition that it possesses the following two properties:

(S1) *There is $k \geq 1$ and a number $a \in (0,1)$ such that the trajectory of the free system (0.1) with $\eta_1 = \eta_2 = \cdots = 0$ satisfies the inequality*

$$\|u_k\| \leq a\|v\| \quad \text{for all } v \in H.$$

Assumption (S1) implies, in particular, that the origin is a stable equilibrium for the free system. We state the second assumption in a weaker and simplified form, sufficient for the validity of the main theorem, referring the reader to (ALC) in Section 2.1 for its complete version.

(S2) *There exist a Hilbert space V, compactly and densely embedded into H, and an open set $O \subset H \times E$ containing $X \times \mathcal{K}$ such that S defines a C^1-smooth mapping from $H \times E$ to V, and for any $(u, \eta) \in O$ the linear operator $D_\eta S(u, \eta) : E \to H$ has a dense image.*

Note that if $\dim H < \infty$, then Assumption (S2) means that, for $(u, \eta) \in O$, the mapping $D_\eta S(u, \eta) : E \to H$ is surjective. Both in finite- and infinite-dimensional cases, this type of properties of system (0.1) often appears in the control theory and may be verified by well-known arguments; see examples in Section 4.

The main result

The statement below is a simplified version of the main result of this paper, proved in Sections 1 and 2:

Main Theorem. *Let us assume that Hypotheses (η1), (η2), (S1), and (S2) are fulfilled. Then there is a Borel probability measure μ on H such that* $\operatorname{supp}\mu \subset X$ *and, for any initial condition v which is a random variable in X independent from $\{\eta_k, k \geq 1\}$, inequality (0.4) holds.*

In Section 2.3, we show that μ is a stationary measure for system (0.1) in the sense that there is an X-valued stationary process $\{\hat{u}_k, k \geq 0\}$, which is a *weak solution* of the equation in (0.1) [4] such that $\mathcal{D}(\hat{u}_k) \equiv \mu$. We also prove that the attraction property for μ established in the Main Theorem extends to the convergence in distribution of the whole trajectories: if v is as in the theorem and $\{u_k\}$ is the corresponding trajectory of (0.1), then for any $s \in \mathbb{N}$

$$\mathcal{D}(u_k, \dots, u_{k+s}) \rightharpoonup \mathcal{D}(\hat{u}_0, \dots, \hat{u}_s) =: \mu^{s+1} \quad \text{as} \quad k \to \infty. \tag{0.8}$$

[4] See Section 1 for a definition of this notion.

Remark 0.1. 1. Let $\{\hat{u}_k\}$ be a weak stationary solution of (0.1) and let $\mathcal{F}_0$ be the σ-algebra generated by the initial state $\hat{u}_0$. Take any $s \in \mathbb{N}$ and any bounded continuous function $g : X^{s+1} \to \mathbb{R}$. Then, by (0.8) and the Markov property, we have

$$\mathbb{E}\big\{g(\hat{u}_k, \dots, \hat{u}_{k+s}) \,\big|\, \mathcal{F}_0\big\} = \int_{E^{s+1}} g(\vec{w})\mu^{s+1}(\mathrm{d}\vec{w}) + o(1) \quad \text{as } k \to \infty.$$

It follows that if $f : X \to \mathbb{R}$ is another bounded continuous function, then

$$\mathbb{E}\big(f(\hat{u}_0)g(\hat{u}_k, \dots, \hat{u}_{k+s})\big) = \mathbb{E}f(\hat{u}_0)\,\mathbb{E}g(\hat{u}_k, \dots, \hat{u}_{k+s}) + o(1) \quad \text{as } k \to \infty.$$

If this relation was valid for the indicator functions f and g of all Borel sets in X and X^{s+1}, then the stationary process $\{\hat{u}_k\}$ would have been mixing in the sense of classical ergodic theory. Thus, the convergence relations (0.4) and (0.8) can be characterised as a *mixing with a smaller set of observables*. This type of feeble mixing has been actively studied over the last thirty years; e.g., see [DDL+07]. It is known, for example, that similarly to the classical mixing, it implies both SLLN and CLT for the process $\{\hat{u}_k\}$; see [DDL+07], Section 4.1 in [KS12], and Subsection 3.2.

2. In fact, our proofs do not use the stationarity of the process $\{\eta_k\}$ entering (0.1). They also apply in the case when the conditional distribution $Q(\boldsymbol{\xi}; \cdot)$ of η_{l+1}, given the past $\boldsymbol{\eta}_l = \boldsymbol{\xi}$, depends on l, but the bounds on characteristics of decompositions (0.6) and (0.7) in (η1) and the choice of constants s and ε in (η2) may be made uniform in l.

3. If $\{\eta_k\}$ are independent random variables, then the Main Theorem is close to the result of paper [KZ20] (which is a simplified version of the work [KNS20] in which a stronger statement is proved).

4. In the case when the phase space H is finite-dimensional, convergence (0.4) can be strengthen to a similar inequality for the total variation metric. This result can be derived from the Main Theorem and will be presented in the subsequent article [KS25].

5. When applied to systems (0.2), our results allow one to treat nonlinear PDEs which depend nonlinearly on stationary processes. This is in contrast with the theory of stochastic PDEs, where the nonlinear PDEs are driven by random forces that as functions of time are white noises and therefore should enter the equation linearly (since nonlinear functions of such noises are not well defined).

6. In the interesting and important case of stochastic PDEs with conservative polynomial nonlinearities, existing techniques permit one to establish exponential mixing only for equations with quadratic nonlinearities. They also may apply to an equation with cubic conservative nonlinearity if it allows for very strong a priori estimates. In the latter case, the existing techniques give only mixing with an algebraic rate. The reason for this restriction comes

from the Foiaş–Prodi inequality, used to prove the mixing in all such systems.[5] Consider, for example, paper [KN13]. It establishes the mixing for the cubic CGL equation (0.18) with $\mu = 0$ and any d, when $\eta(t,x)$ is a random field, smooth in x and white in t. There, a crucial step of the proof is a Foias–Prodi type estimate at the bottom of p. 415, which involves quadratic exponential moments of the norms of solutions. If the nonlinearity in equation was of degree $r > 3$ rather than cubic, then the estimate would involve exponential moments of solutions of degree $r - 1 > 2$. They are "hopelessly unbounded" since it is impossible to believe that norms of solutions for an SPDE with a conservative nonlinearity allow for better estimates then those for norms of the Wiener process whose time-derivative drives the equation.

About the proof

It is well known that any random process becomes Markovian if we add to the process its past. In our case this thesis implies the following (cf. [BF92, EPR99, Hai05, HO07] and [Bor98, Section 13]). Recalling that $\boldsymbol{\eta}_k = (\eta_l, l \le k) \in E^{\mathbb{Z}_-}$, for a trajectory $\{u_k, k \ge 0\}$ of (0.1) we denote

$$U_k = (u_k, \boldsymbol{\eta}_k) \in H \times E^{\mathbb{Z}_-} =: \mathfrak{H}.$$

Then $\{U_k, k \ge 0\}$ is a trajectory of the following lifting of system (0.1) in H to a system in $\mathfrak{H}$:

$$U_k = \mathscr{S}(U_{k-1}, \eta_k), \quad k \ge 1, \tag{0.9}$$

$$U_0 = (v, \boldsymbol{\eta}_0), \tag{0.10}$$

where

$$\mathscr{S} : \mathfrak{H} \times E \to \mathfrak{H}, \qquad \big((u, \boldsymbol{\eta}), \xi\big) \mapsto \big(S(u,\xi), (\boldsymbol{\eta}, \xi)\big).$$

Since $\eta_k \in \mathcal{K}$ a.s., Equation (0.9) defines a stochastic system in $\mathfrak{X} = X \times \boldsymbol{\mathcal{K}}$, and trajectories of (0.1) in X define trajectories of (0.9), (0.10) in $\mathfrak{X}$. The natural projection $\Pi_X : \mathfrak{X} \to X$ sends these trajectories of (0.9), (0.10) back to trajectories of (0.1). It is not hard to see how to define properly trajectories of (0.9) with any initial data $U_0 \in \mathfrak{H}$, and that the corresponding dynamics in $\mathfrak{H}$ is Markovian; see Section 1.1. So it remains to show that the obtained Markov process in $\mathfrak{H}$ is exponentially mixing with some stationary measure $\boldsymbol{\mu}$ since then (0.4) holds with $\mu = (\Pi_X)_* \boldsymbol{\mu}$. But the Markov process, defined by (0.9), is rather complicated because the phase space $\mathfrak{H}$ is "much bigger" than H, and the resulting process is not strong Feller. Accordingly, existing approaches do not imply the mixing property for system (0.9). Below, to establish the mixing, we use coupling, combined with the method of Kantorovich functional suggested in [KPS02, Kuk02] (also see in [Kuk06, KS12]), and enrich it with some ideas from our works [Shi15, KNS20, KZ20, Shi21].

[5] with the exception of the stochastic Burgers equation, which possesses some very special extra properties.

The proof goes in two steps. Firstly, in Section 1, we prove Theorem 1.3, which establishes the mixing for a class of Markov systems that includes (0.9) as a particular case. The proof uses Newton's method of quadratic convergence. Due to the complexity of the situation, the method provides only exponential (rather than super-exponential) rate of convergence to a limit: the power of the method is used to cope with the fact that available estimates are rather weak. Secondly, in Section 2, we verify in Theorem 2.4 that if Hypotheses (η1), (η2), (S1), and (S2) hold for system (0.1), then the extended system (0.9) meets the conditions of Theorem 1.3. This proves the Main Theorem stated above.

Applications

The Main Theorem applies to Equations (0.1) obtained by the discrete-time reduction from various ODEs and PDEs depending on bounded stationary processes. To show this, we first construct in Section 3.3 a large class of T-periodic random processes $\eta(t)$, taking values in a Hilbert space E of finite or infinite dimension and having locally square-integrable trajectories, such that the corresponding path-valued stationary process $\{\eta_l\}$ in $L^2(0,T;E)$ defined by (0.3) satisfies (η1) and (η2). Next, in Section 4, we give various applications of our results to equations driven by additive stationary processes possessing the above properties.

In Section 4.1, we examine a chain of n anharmonic 1d oscillators with coordinates q_j, momenta p_j $(1 \le j \le n)$, and a smooth Hamiltonian

$$H(p,q) = \sum_{j=1}^{n} \frac{p_j^2}{2} + V(q), \quad (p,q) = (p_1,\dots,p_n,q_1,\dots,q_n) \in \mathbb{R}^{2n}, \tag{0.11}$$

where V is a C^2-smooth function. At sites $j=1$ and $j=n$, the chain is coupled with heat reservoirs which affect the 1^{st} and n^{th} particles by adding to the corresponding equations terms $-\gamma_1 p_1 + \zeta_1(t)$ and $-\gamma_n p_n + \zeta_n(t)$. Here $\gamma_1, \gamma_n > 0$ are damping coefficients and ζ_1, ζ_n are statistically 1-periodic processes as above with $E = \mathbb{R}$ and $T = 1$. We shall assume that

$$|\zeta_1(t)|, |\zeta_2(t)| \le c_* \quad \text{for almost every } t \ge 0, \tag{0.12}$$

where $c_* > 0$ is a number. Let us set $\zeta = (\zeta_1, \zeta_n)$ and denote by $\mathcal{K}$ the support of the process $\{\zeta(t), 0 \le t \le 1\}$ in the space $L_2(0,1;\mathbb{R}^2)$. Let $v(p,q)$ be the vector field of the chain of oscillators in the case when $\zeta_1 = \zeta_2 = 0$. Roughly speaking, we impose the two hypotheses below:

Global stability. *The vector field $v(p,q)$ has a globally asymptotically stable equilibrium at $(0,0)$ and admits a coercive*[6] *Lyapunov function $L(p,q)$ such that*

$$\big\langle \nabla L(p,q), v(p,q) \big\rangle + 2c_* |\nabla L(p,q)| \le -\delta L(p,q) \tag{0.13}$$

[6] This means that $L(p,q) \to +\infty$ as $|(p,q)| \to +\infty$.

outside a compact set in $\mathbb{R}^{2n}$, where $\langle\cdot,\cdot\rangle$ stands for the standard inner product, and $\delta > 0$ is a number.

This condition implies that there is a compact set $X \subset \mathbb{R}^{2n}$ which, for every ω, is invariant and absorbing for the inhomogeneous dynamics under the study.

Linearised controllability. *For every initial condition at $t = 0$ in X and every curve $\zeta(\cdot) \in \mathcal{K}$, the system linearised at the corresponding solution is linearly controllable at $t = 1$.*

In Section 4.1, we prove Theorem 4.2, which shows that the above properties are sufficient for the system in question to be mixing. We also provide an example when those conditions are satisfied.

It should be noted that chains and networks of anharmonic oscillators and rotators have been the focus of intensive research for the last thirty years; see the papers [EPR99, EH00, CE16, CEHR18, Raq19, NR21] and the references therein. The problem in question consists of a system of ODEs driven by Brownian or Poissonian noises, which are *unbounded* processes in time. In those works, a sufficient condition for mixing is the existence of a *Lyapunov function in the stochastic sense* and the *Hörmander bracket condition*. In our setting, we deal with a bounded noise and require the existence of a *Lyapunov function in the deterministic sense* (which may be easier to construct; see [DLS24]) and *controllability of a family of linearised equation*, which is different from, but related to the Hörmander condition.

Next, in Section 4.2, we deal with the 2d Navier–Stokes system, which is considered in a smooth bounded domain $D \subset \mathbb{R}^2$ and is subject to an external random force:

$$\partial_t u + \langle u, \nabla\rangle u - \nu\Delta u + \nabla p = \eta(t,x), \quad \operatorname{div} u = 0, \quad x \in D. \tag{0.14}$$

Here $u = (u_1, u_2)$ and p are unknown velocity field and pressure, $\nu > 0$ is the viscosity, and η is a random force, described below. Equations (0.14) are supplemented with the no-slip boundary condition

$$u\big|_{\partial D} = 0 \tag{0.15}$$

and with an initial condition for the velocity

$$u(0,x) = u_0(x). \tag{0.16}$$

Let us denote by H the standard space of square-integrable divergence-free vector field on D whose normal component vanishes on the boundary ∂D (see (4.12)). Then, under mild hypotheses on η, for any $u_0 \in H$, problem (0.14), (0.16) has a unique H-valued solution $u(t)$, and our goal is to investigate its long-time asymptotics.

Let $\{e_j, j \geq 1\}$ be an orthonormal basis in H, formed by the eigenfunctions of the corresponding Stokes operator. We assume that $\eta(t)$ is a random process in H of the form

$$\eta(t) = \sum_{j=1}^{\infty} b_j \eta^j(t) e_j,$$

where the sequence $\{b_j > 0\}$ converges to zero sufficiently fast and $\{\eta^j\}$ are i.i.d. statistically 1-periodic real process such that the corresponding stationary process in $L^2(0,1;\mathbb{R})$ satisfies (η1) and (η2); see Section 3.3.3. In Theorem 4.4, we show that the Main Theorem applies to (0.14) and therefore convergence (0.4) holds for the laws of its solutions.

The Main Theorem also applies if $\eta(t)$ in (0.14) is a kick force of the form

$$\eta(t,x) = \sum_{k=1}^{\infty} \eta_k(x)\delta(t-k), \quad \eta_k(x) = \sum_{j=1}^{\infty} b_j \eta_k^j e_j(x), \tag{0.17}$$

where the sequence of positive numbers $\{b_j\}$ converges to zero sufficiently fast, and for $j = 1, 2, \dots$ the real-valued random variables $\{\eta_k^j, k = 1, 2, \dots\}$ form i.i.d. bounded stationary processes. If they satisfy Hypotheses (η1) and (η2), then the Main Theorem applies in this situation as well.

Finally, in Section 4.3, we consider the following complex Ginzburg–Landau equation in a smooth bounded domain $D \subset \mathbb{R}^d$, with the Dirichlet boundary condition:

$$\partial_t u - (\nu + i\mu)\Delta u + i|u|^{2s}u = \eta(t,x), \quad x \in D. \tag{0.18}$$

Here ν and μ are positive numbers and $s \geq 1$ is an integer such that $s \leq 2/(d-2)$ if $d \geq 3$. If the external force η is a statistically 1-periodic random process in a suitable Sobolev space of functions of x, then as we show in Theorem 4.5, the Main Theorem also applies to the solutions of this equation.

Acknowledgments

The authors thank V. I. Bogachev for pertinent remarks. The research of AS was supported by the *CY Initiative of Excellence* through the grant *Investissements d'Avenir* ANR-16-IDEX-0008 and by the *ANR project DYNACQUS* through the grant ANR-24-CE40-5714-01.

Notation

We write $\mathbb{Z}$ ($\mathbb{Z}_+$, $\mathbb{Z}_-$) for the set of (non-negative, non-positive) integers, denote by $B_E(a,r)$ an open r-ball in a Banach space E, centred at a, and write $B_E(r)$ if $a = 0$. For an integer $m \geq 1$, a vector $\vec{\xi}_m = (\xi_1, \dots, \xi_m) \in E^m$, and a number $\delta > 0$, we denote by $\mathcal{O}_\delta(\vec{\xi}_m)$ the set $B_E(\xi_1,\delta) \times \dots \times B_E(\xi_m,\delta)$. When E is finite dimensional, we write ℓ for the Lebesgue measure on E. By $\mathcal{D}(\eta)$ we denote the law of a random variable η, and $\rightharpoonup$ signifies the weak convergence of measures.

If X is a Polish space, then we write $\mathcal{B}(X)$ for its Borel σ-algebra and $\mathcal{P}(X)$ for the set of probability measures on $(X, \mathcal{B}(X))$. Unless otherwise stated, spaces $\mathcal{P}(X)$ are provided with the weak topology. Borel-measurable maps from X to another measurable space are often called just *measurable*, and a random field $\{\zeta_x^\omega, x \in X\}$ is said to be measurable if it define a measurable map $(\omega, x) \to \zeta_x^\omega$ from $\Omega \times X$ to X. Given a subset $Q \subset X$, we denote by $Q^c = X \setminus Q$ its complement, by $\mathbf{1}_Q$ its indicator function, and by $\overline{Q}$ its closure.

1 A class of random dynamical systems

1.1 Description of the model

Let $\mathfrak{X}$ be a compact metric space with a distance $\mathrm{d}_{\mathfrak{X}}$, E be a separable Banach space, and $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Consider a transition probability $\{\mathscr{P}(U;\cdot)\}_{U\in\mathfrak{X}}$ from $\mathfrak{X}$ to E.[7] In what follows, $\mathscr{P}(U;\cdot)$ plays the role of the law of a random input (or noise) acting on a dynamical system in the space $\mathfrak{X}$, and we shall always assume that it satisfies the following hypothesis:

(DN) Driving noise. *The mapping $U \mapsto \mathscr{P}(U;\cdot)$ is continuous from $\mathfrak{X}$ to the space $\mathcal{P}(E)$ endowed with the weak topology. Moreover, there is a number $C > 0$ and a compact set $\mathcal{K} \subset E$ such that* $\operatorname{supp} \mathscr{P}(U;\cdot) \subset \mathcal{K}$ *for any $U \in \mathfrak{X}$, and*

$$\|\mathscr{P}(U;\cdot) - \mathscr{P}(U';\cdot)\|_L^* \le C\mathrm{d}_{\mathfrak{X}}(U,U') \quad \text{for } U, U' \in \mathfrak{X}. \tag{1.1}$$

Let $\{\zeta_k^U, U \in \mathfrak{X}\}_{k\ge1}$ be a sequence of independent $\mathcal{K}$-valued measurable random fields on $\mathfrak{X}$ defined on a complete probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$. Thus, the map $(U,\widehat{\omega}) \mapsto \zeta_k^U(\widehat{\omega})$ is measurable from $\mathfrak{X}\times\widehat{\Omega}$ to $\mathcal{K}$ for any $k \ge 1$. We shall always assume that

$$\mathcal{D}(\zeta_k^U) = \mathscr{P}(U;\cdot) \quad \text{for any } U \in \mathfrak{X} \text{ and } k \ge 1. \tag{1.2}$$

Let us fix a Lipschitz-continuous map $\mathscr{S} : \mathfrak{X}\times\mathcal{K} \to \mathfrak{X}$ and consider the following random dynamical system (RDS) in $\mathfrak{X}$:

$$U_k = \mathscr{S}(U_{k-1}, \zeta_k^{U_{k-1}}), \quad k \ge 1. \tag{1.3}$$

Equation (1.3) is supplemented with the initial condition

$$U_0 = V, \tag{1.4}$$

where V is an $\mathfrak{X}$-valued random variable that is defined on the same probability space and is independent of the family $\{\zeta_k^U, U \in \mathfrak{X}, k \ge 1\}$.

In what follows, we always assume that the underlying probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ for (1.3) is the *completion of the tensor product of countably many copies of a fixed complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, writing $\widehat{\omega} = (\omega_k, k \ge 0)$ for points of $\widehat{\Omega}$, we shall assume that

$$V \text{ depends only on } \omega_0 \text{ and } \zeta_k^U \text{ depends only on } \omega_k \text{ for } k \ge 1,\ U \in \mathfrak{X}. \tag{1.5}$$

It is well known that, given a transition probability $\{\mathscr{P}(U;\cdot)\}_{U\in\mathfrak{X}}$ on E that satisfies Hypothesis (DN), there is a measurable random field $\{\zeta^U, U \in \mathfrak{X}\}$ (where for the probability space one may take the interval $[0,1]$ endowed with the Lebesgue measure) such that $\mathcal{D}(\zeta^U) = \mathscr{P}(U;\cdot)$ for any $U \in \mathfrak{X}$; see [Dud02,

[7] Thus, for any $U \in \mathfrak{X}$, we have a Borel probability measure $\mathscr{P}(U;\cdot)$ on E such that, for any $\Gamma \in \mathcal{B}(E)$, the function $U \mapsto \mathscr{P}(U;\Gamma)$ is measurable from $\mathfrak{X}$ to $\mathbb{R}$; cf. [Par05, Chapter 5].

Theorem 11.7.5] and [KS12, Theorem 1.2.28]. Hence, setting $\zeta_k^U(\widehat{\omega}) = \zeta^U(\omega_k)$ for $U \in \mathfrak{X}$ and $k \geq 1$, we obtain a sequence of independent random fields satisfying (1.2). In what follows, we write $\{U_k(V), k \geq 0\}$ for a trajectory of (1.3), (1.4) and call it a *solution* of (1.3), (1.4) corresponding to the probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ and the random fields $\{\zeta_k^{\cdot}\}$. Very often, we shall drop a specification of the probability space and random fields.

Let $\{\mathcal{F}_k\}_{k\geq 0}$ be the natural filtration of $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$; that is, $\mathcal{F}_k$ consists of those elements of $\widehat{\mathcal{F}}$ that depend only on $\boldsymbol{\omega}_k := (\omega_0, \dots, \omega_k)$. In what follows, we call a space $\widehat{\Omega}$ as above with the filtration $\{\mathcal{F}_k\}$ a *suitable (filtered) probability space* constructed from the space $(\Omega, \mathcal{F}, \mathbb{P})$, always assuming (1.5). The following result is a consequence of (1.2), (1.5), and Fubini's theorem.

Proposition 1.1. *The family of trajectories $\{U_k(V)\}$ of (1.3), corresponding to all possible deterministic initial conditions $V \in \mathfrak{X}$, form a Markov process in $\mathfrak{X}$, corresponding to the filtration $\{\mathcal{F}_k\}$, with the continuous time-1 transition probability*

$$P_1(U;\cdot) = \mathscr{S}_*\big(U, \mathscr{P}(U;\cdot)\big), \quad U \in \mathfrak{X}. \tag{1.6}$$

Due to (1.6), the Markov semigroups defined by this process do not depend on the specific choice of a suitable probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ and random fields $\{\zeta_k^{\cdot}\}$. Therefore, the laws of solutions $\{U_k(V)\}$ are also independent of that choice.

Proof of Proposition 1.1. By construction, the random variable $U_k(V)$, $k \geq 0$, depends only on $\boldsymbol{\omega}_k$, so it is $\mathcal{F}_k$-measurable. We now prove that, for any integers $k, m \geq 0$ and any bounded measurable function $f : \mathfrak{X} \to \mathbb{R}$,

$$\mathbb{E}\{f(U_{k+m}(V)) \,|\, \mathcal{F}_k\} = \hat{f}_m(U_k(V)) \quad \mathbb{P}\text{-almost surely}, \tag{1.7}$$

where $\hat{f}_m : \mathfrak{X} \to \mathbb{R}$ is a bounded measurable function depending only on f and m. The relation is obvious for $m = 0$, and a simple induction argument shows that it suffices to establish it for $m = 1$. But for $m = 1$ and any $\mathcal{F}_k$-measurable function function $\varphi(\boldsymbol{\omega}_k)$, we have

$$\begin{aligned}\mathbb{E}\big(f(U_{k+1}(V))\varphi(\boldsymbol{\omega}_k)\big) &= \mathbb{E}^{\boldsymbol{\omega}_k}\Big[\mathbb{E}^{\omega_{k+1}} f\big(\mathscr{S}(U_k^{\boldsymbol{\omega}_k}, \zeta_{k+1}^{U_k^{\boldsymbol{\omega}_k}}(\omega_{k+1}))\big)\varphi(\boldsymbol{\omega}_k)\Big] \\ &= \mathbb{E}^{\boldsymbol{\omega}_k}\Big[\varphi(\boldsymbol{\omega}_k)\,\mathbb{E} f\big(\mathscr{S}(U_k^{\boldsymbol{\omega}_k}, \zeta_1^{U_k^{\boldsymbol{\omega}_k}})\big)\Big],\end{aligned}$$

since the law of ζ_k^U does not depend on k. Thus,

$$\mathbb{E}\{f(U_{k+1}(V)) \,|\, \mathcal{F}_k\} = \mathbb{E} f\big(\mathscr{S}(x, \zeta_1^x)\big)\big|_{x=U_k}.$$

It remains to note that the map $x \mapsto \mathbb{E} f(\mathscr{S}(x, \zeta_1^x))$ defines a bounded measurable function $\hat{f}_1(x)$, depending only on f, so (1.7) holds for $m = 1$. Relation (1.6) follows obviously from the last equality. □

In what follows, we denote by $\{P_k(U;\cdot), k \in \mathbb{Z}_+, U \in \mathfrak{X}\}$ the transition probability for the Markov process associated with the RDS (1.3) (it is given by (1.6) for $k = 1$). Often, when studying the long-time asymptotics of trajectories for (1.3), we need to replace the underlying probability space $\widehat{\Omega}$ with another *suitable* probability space. This will not change the Markovian character of the dynamics and the laws of solutions of (1.3). The next subsection describes an important example of a dynamical system driven by a stationary process that can be reduced to an RDS of the form (1.3).

1.2 Dynamical systems with stationary noises

Consider an RDS of the form (0.1):

$$u_k = S(u_{k-1}, \eta_k), \quad k \geq 1, \qquad u_0 = v. \tag{1.8}$$

Here $u_k \in H$, where H is a Hilbert space, $\{\eta_k\}_{k\in\mathbb{Z}}$ is a stationary process that takes values in a Banach space E and is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $S : H \times E \to H$ is a continuous map. We assume that $\operatorname{supp} \mathcal{D}(\eta_k) =: \mathcal{K}$ is a compact subset of E (so $\{\eta_k\}$ also may be regarded as a process in $\mathcal{K}$), and suppose that there is a compact subset $X \subset H$ such that $S(X \times \mathcal{K}) \subset X$. Then (1.8) defines an RDS in X. We are interested in the long-time behaviour of trajectories of this system.

In what follows, we distinguish between *weak* and *strong* solutions of (1.8): strong solutions are those that are constructed by relations (1.8), and will be denoted $\{u_k(v), k = 0, 1, \dots\}$, whereas weak solutions are, as usual, trajectories $\{\tilde{u}_k\}$ of system (1.8), in which $\{\eta_k\}$ and v are replaced with another process $\{\tilde{\eta}_k\}$ and a random variable $\tilde{v}$ having the same joint distribution and possibly defined on a different probability space. All weak solutions have the same distribution. This is in difference with system (1.3), where the random fields $\{\zeta_k\}$ must be defined on some suitable probability space, and ζ_k must be $\mathcal{F}_k$-measurable. Accordingly, a weak solution of the equation in (1.8) is that of problem (1.8) with some v. Weak solutions of (1.8) will be denoted as $(\{\tilde{u}_k\}, \{\tilde{\eta}_k\})$ or, for short, as $\{\tilde{u}_k\}$.

As before, we provide the space $E^{\mathbb{Z}_-}$ with the Tikhonov topology whose restriction to the compact subset $\boldsymbol{\mathcal{K}} := \mathcal{K}^{\mathbb{Z}_-}$ is metrised by distance (0.5). We define the measure

$$\sigma := \mathcal{D}(\boldsymbol{\eta}_-), \quad \boldsymbol{\eta}_- = \{\eta_k, k \in \mathbb{Z}_-\}, \tag{1.9}$$

which is an element of $\mathcal{P}(\boldsymbol{\mathcal{K}})$, and denote by $\mathcal{E} \subset \boldsymbol{\mathcal{K}}$ its support.

For any $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$, denote by $Q(\boldsymbol{\xi};\cdot)$ the conditional law $\mathbb{P}\{\eta_1 \in \cdot \,|\, \boldsymbol{\eta} = \boldsymbol{\xi}\}$ of η_1 given the past $\{\eta_k = \xi_k\}_{k\in\mathbb{Z}_-}$; see [Dud02, Theorem 10.2.1]. It is uniquely defined up to $\boldsymbol{\xi}$'s from a negligible set with respect to the measure σ. Now we assume in addition that the process $\{\eta_k\}$ satisfies the following continuity condition, similar to the Feller property for Markov processes.

(Fel) *The conditional measure $Q(\boldsymbol{\xi};\cdot)$ can be chosen to be a continuous mapping from $\mathcal{E}$ to $\mathcal{P}(\mathcal{K}) \subset \mathcal{P}(E)$. Moreover, there is $C > 0$ such that*

$$\|Q(\boldsymbol{\xi};\cdot) - Q(\boldsymbol{\xi}';\cdot)\|_L^* \le C\mathrm{d}(\boldsymbol{\xi},\boldsymbol{\xi}') \quad \textit{for any } \boldsymbol{\xi},\boldsymbol{\xi}' \in \mathcal{E}.$$

This condition allows us to (uniquely) define $Q(\boldsymbol{\xi};\cdot)$ for $\boldsymbol{\xi} \in \mathcal{E}$ and regard it as a transition probability, continuous in $\boldsymbol{\xi}$ (see [Par05, Section 5.1]).

We now fix a specific realisation of the process $\{\eta_k\}$ in (1.8), convenient for the purposes of this work. We construct it as a process defined on some suitable probability space $(\widehat{\Omega},\widehat{\mathcal{F}},\widehat{\mathbb{P}})$, where $\widehat{\Omega} = \{(\omega_0,\omega_1,\dots)\}$ with $\omega_j \in \Omega$ (see Section 1.1). For $k \le 0$, the random variables η_k depends on ω_0 as for the original process η_k in (1.8), while for $k \ge 1$ we take $\eta_k = \eta_k(\boldsymbol{\omega}_k) = \eta_k(\boldsymbol{\omega}_{k-1},\omega_k)$ with $\boldsymbol{\omega}_k := (\omega_0,\dots,\omega_k)$, where for any $\boldsymbol{\omega}_{k-1} \in \Omega^k$ the mapping $\omega \mapsto \eta_k(\boldsymbol{\omega}_{k-1},\omega)$ is a random variable $\omega \mapsto \zeta_k^{\boldsymbol{\eta}_{k-1}}(\omega)$, with law equal to $Q(\boldsymbol{\eta}_{k-1};\cdot)$, such that $\zeta_k^{\boldsymbol{\eta}}(\omega)$ is a measurable random field, for each $k \ge 1$. Here $\boldsymbol{\mathcal{K}} \ni \boldsymbol{\eta}_k := (\eta_l, l \le k)$. An induction argument in $k \ge 0$ readily shows that this modification of the process $\{\eta_k\}$ is distributed as the original one.

Let us introduce the product space $\mathfrak{X} = X \times \mathcal{E}$ endowed with the metric

$$\mathrm{d}_{\mathfrak{X}}(U,U') = L\,\|v - v'\|_H + \mathrm{d}(\boldsymbol{\xi},\boldsymbol{\xi}'), \tag{1.10}$$

where $L \ge 1$ is a parameter specified later, and denote the natural projections to the two components by

$$\Pi_X : \mathfrak{X} \to X, \quad \Pi_{\mathcal{E}} : \mathfrak{X} \to \mathcal{E}. \tag{1.11}$$

We consider a map $\mathscr{S} : \mathfrak{X} \times \mathcal{K} \to \mathfrak{X}$ defined by the relation

$$\mathscr{S}(U,\eta) = \big(S(v,\eta),(\boldsymbol{\xi},\eta)\big), \tag{1.12}$$

where $U = (v,\boldsymbol{\xi}) \in \mathfrak{X}$ and $\eta \in \mathcal{K}$. Finally, we introduce an RDS in $\mathfrak{X}$ by

$$U_k = \mathscr{S}(U_{k-1},\zeta_k^{\boldsymbol{\xi}_{k-1}}), \quad k \ge 1, \tag{1.13}$$

where $U_k = (v_k,\boldsymbol{\xi}_k)$. The system (1.13) is supplemented with the initial condition

$$U_0 = V := (v,\boldsymbol{\xi}),$$

where $(v,\boldsymbol{\xi})$ is a random variable depending only on ω_0. We thus obtain a special case of system (1.3). By Proposition 1.1, it defines a Markov process in the extended phase space $\mathfrak{X}$.

Let $\{u_k = u_k(v)\}$ be a trajectory of the RDS (1.8). For $k \ge 0$, we denote $\widehat{U}_k = (u_k,\boldsymbol{\eta}_k)$, where $\boldsymbol{\eta}_k = (\eta_l, l \le k)$. By the construction of the version of the process $\{\eta_k\}$ described above, for $k \ge 0$ we have $\eta_k = \eta_k(\boldsymbol{\omega}_k) = \eta_k(\boldsymbol{\omega}_{k-1},\omega_k)$ (so also $\boldsymbol{\eta}_k = \boldsymbol{\eta}_k(\boldsymbol{\omega}_k)$), where $\eta_k(\boldsymbol{\omega}_{k-1},\cdot)$ is the random variable $\omega_k \mapsto \zeta_k^{\boldsymbol{\eta}_{k-1}(\boldsymbol{\omega}_{k-1})}(\omega_k)$ with $\mathcal{D}(\zeta_k^{\boldsymbol{\xi}}) = Q(\boldsymbol{\xi};\cdot)$. It follows that

$$\mathscr{S}\big(\widehat{U}_k,\zeta_k^{\boldsymbol{\eta}_{k-1}}\big) = \big(S(u_{k-1},\eta_k),(\boldsymbol{\eta}_{k-1},\eta_k)\big) = \widehat{U}_k.$$

So $\{\widehat{U}_k, k \ge 0\}$ is a trajectory of (1.13) with $\widehat{U}_0 = (v_0,\boldsymbol{\eta}_0)$. We have thus proved the following result.

Lemma 1.2. *Let an $\mathcal{F}_0$-measurable random initial condition $V = (v, \boldsymbol{\xi})$ be such that $v \in H$ and $\boldsymbol{\xi}$ is a $\boldsymbol{\mathcal{K}}$-valued random variable whose law is equal to σ. Let $\{U_k = (v_k, \xi_k), k \geq 0\}$ be a solution of (1.13), (1.4), where for $U = (v, \boldsymbol{\xi}) \in \mathfrak{X}$ we have $\mathcal{D}(\zeta_k^U) = Q(\boldsymbol{\xi}; \cdot)$ (so $\mathscr{P}((v, \boldsymbol{\xi}), \cdot) = Q(\boldsymbol{\xi}; \cdot)$). Then, for any $k \geq 0$, the distribution of the vector $[U_0, \dots, U_k]$ is such that*

$$\mathcal{D}([v_0, \dots, v_k]) = \mathcal{D}\big([u_0(v), \dots, u_k(v)]\big), \quad \mathcal{D}(\boldsymbol{\xi}_k) = \sigma. \tag{1.14}$$

Since all weak solutions have the same law, relation (1.14) remains true if we replace there the strong solution $\{u_k(v)\}$ with any weak solution of (1.8).

In conclusion, let us note that both spaces H and E can be assumed to be separable. Indeed, denoting by $\widetilde{H}$ and $\widetilde{E}$ the closures of the vector spans of X and $\mathcal{K}$, respectively, we define the map

$$\widetilde{S} : \widetilde{H} \times \widetilde{E} \to \widetilde{H}, \quad (u, \eta) \mapsto \mathsf{P}_{\widetilde{H}} S(u, \eta),$$

where $\mathsf{P}_{\widetilde{H}} : H \to H$ stands for the orthogonal projection to $\widetilde{H}$. Since the image of $X \times \mathcal{K}$ under S is contained in X, we see that $\widetilde{S}$ and S coincide on $X \times \mathcal{K}$. Thus, replacement of S with $\widetilde{S}$ in (1.8) will not change the trajectories, and we obtain an RDS of the same form with separable spaces.

1.3 A criterion for mixing

We now prove a result providing a sufficient condition for the property of exponential mixing for the RDS (1.3). To this end, we impose the two hypotheses below on the map $\mathscr{S}$ and the transition probabilities $\mathscr{P}(U; \cdot)$

(GCP) Global controllability to points. *For any $\varepsilon > 0$, there is an integer $m \geq 1$ and a point $\widehat{U} \in \mathfrak{X}$ (both depending on ε) for which the following property holds: for any $U \in \mathfrak{X}$ there are $\xi_1, \dots, \xi_m \in \mathcal{K}$ such that the points $U_0, \dots, U_m \in \mathfrak{X}$ defined by $U_0 = U$ and $U_k = \mathscr{S}(U_{k-1}, \xi_k)$ for $1 \leq k \leq m$ satisfy the relations*

$$\xi_k \in \operatorname{supp} \mathscr{P}(U_{k-1}, \cdot) \quad \text{for } 1 \leq k \leq m, \qquad \mathrm{d}_{\mathfrak{X}}(U_m, \widehat{U}) < \varepsilon. \tag{1.15}$$

Let us set

$$D_\delta = \{(U, U') \in \mathfrak{X} \times \mathfrak{X} : \mathrm{d}_{\mathfrak{X}}(U, U') \leq \delta\} \tag{1.16}$$

and define the Lipschitz seminorm of a map $\Phi : \mathcal{K} \to E$ by the relation

$$\operatorname{Lip}_\xi(\Phi) := \sup_{\xi_1, \xi_2} \frac{\|\Phi(\xi_1) - \Phi(\xi_2)\|_E}{\|\xi_1 - \xi_2\|_E},$$

where the supremum is taken over all $\xi_1, \xi_2 \in \mathcal{K}$ such that $0 < \|\xi_1 - \xi_2\|_E \leq 1$.

(LAC) Local approximate controllability. *There are positive numbers C_*, δ, and $q < 1$, a finite-dimensional subspace $F \subset E$, and a continuous map*

$$\Phi : D_\delta \times \mathcal{K} \to F$$

such that, for any $(U, U') \in D_\delta$,

$$\sup_{\xi \in \mathcal{K}} \|\Phi(U, U', \xi)\|_E + \mathrm{Lip}_\xi\big(\Phi(U, U', \cdot)\big) \le C_* \mathrm{d}_{\mathfrak{X}}(U, U'), \tag{1.17}$$

$$\sup_{\xi \in \mathcal{K}} \mathrm{d}_{\mathfrak{X}}\big(\mathscr{S}(U, \xi), \mathscr{S}(U', \xi + \Phi(U, U', \xi))\big) \le q\, \mathrm{d}_{\mathfrak{X}}(U, U'). \tag{1.18}$$

We shall also need a decomposability and regularity hypothesis on the driving noise in (1.3). We recall that a closed subspace $F \subset E$ is said to be *complemented* if there is another closed subspace $F^\dagger \subset E$ such that $E = F \dotplus F^\dagger$. Obviously, the existence of a *complementary subspace* is equivalent to the existence of a continuous projection $\mathsf{P}_F : E \to E$ onto F. A simple consequence of the Hahn–Banach theorem is that any finite-dimensional subspace is complemented. In what follows, if E is represented as the direct sum of closed subspaces F and $F^\dagger$, then we denote by P_F and $\mathsf{P}_{F^\dagger}$ the associated projections.

(DLP) Decomposability and Lipschitz property. *The subspace F entering (LAC) possesses a complementary subspace $F^\dagger \subset E$ such that the measure $\mathscr{P}(U; \cdot)$ admits a disintegration of the form*

$$\mathscr{P}(U; \mathrm{d}\xi_F, \mathrm{d}\xi_{F^\dagger}) = \mathscr{P}_{F^\dagger}(U; \mathrm{d}\xi_{F^\dagger}) \mathscr{P}_F(U, \xi_{F^\dagger}; \mathrm{d}\xi_F), \quad U \in \mathfrak{X}, \tag{1.19}$$

where $\mathscr{P}_{F^\dagger}(U; \cdot)$ denotes the projection of $\mathscr{P}(U; \cdot)$ to $F^\dagger$, and $\mathscr{P}_F(U, \xi_F; \cdot)$ stands for the regular conditional probability of $\mathscr{P}(U; \cdot)$ when the projection to $F^\dagger$ is fixed and is equal to $\xi_{F^\dagger}$. Moreover, there is a function $p_F : \mathfrak{X} \times E \to \mathbb{R}_+$, Lipschitz continuous in both arguments, i.e.

$$|p_F(U, \xi) - p_F(U', \xi')| \le M\big(\mathrm{d}_{\mathfrak{X}}(U, U') + \|\xi - \xi'\|_E\big) \text{ for } U, U' \in \mathfrak{X},\ \xi, \xi' \in E, \tag{1.20}$$

such that, for any $U \in \mathfrak{X}$ and $\xi_{F^\dagger} \in F^\dagger$, we have

$$\mathscr{P}_F(U, \xi_{F^\dagger}; \mathrm{d}\xi_F) = p_F(U, \xi_{F^\dagger}, \xi_F)\, \ell_F(\mathrm{d}\xi_F), \tag{1.21}$$

where ℓ_F stands for the Lebesgue measure on F.

In view of the well-known formula for the total variation distance (e.g., see Proposition 1.2.7 in [KS12]), it follows from (1.20) and (1.21) that, for any $U, U' \in \mathfrak{X}$ and $\xi_{F^\dagger}, \xi'_{F^\dagger} \in F^\dagger$, we have

$$\|\mathscr{P}_F(U, \xi_{F^\dagger}; \cdot) - \mathscr{P}_F(U', \xi'_{F^\dagger}; \cdot)\|_{\mathrm{var}} \le M_1 \big(\mathrm{d}_{\mathfrak{X}}(U, U') + \|\xi_{F^\dagger} - \xi'_{F^\dagger}\|_E\big),$$

where $M_1 > 0$ does not depend on U, U' and $\xi_{F^\dagger}, \xi'_{F^\dagger}$.

Recall that, by Proposition 1.1, the trajectories of (1.3) form a discrete-time Markov process. The following theorem is the main result of this section.

Theorem 1.3. *Suppose that Hypotheses (GCP), (LAC), and (DLP) are fulfilled. Then the Markov process associated with the RDS (1.3) has a unique stationary measure $\boldsymbol{\mu} \in \mathcal{P}(\mathfrak{X})$, and there are positive numbers C and γ such that*[8]

$$\|P_k(U;\cdot) - \boldsymbol{\mu}\|_L^* \le Ce^{-\gamma k} \quad \text{for any } U \in \mathfrak{X},\ k \ge 0. \tag{1.22}$$

Moreover, if $\{V_k, k \ge 0\}$ is a stationary trajectory of this Markov process, $\mathcal{D}(V_k) \equiv \boldsymbol{\mu}$, and $\boldsymbol{\mu}_s = \mathcal{D}(V_1, \dots, V_s)$, then

$$\|\mathcal{D}(U_k, \dots, U_{k+s}) - \boldsymbol{\mu}_{s+1}\|_L^* \le C_s e^{-\gamma k} \quad \text{for any } U \in \mathfrak{X},\ k \ge 0, \tag{1.23}$$

where $\{U_k\}$ stands for the trajectory of (1.3) issued from U.

A proof of this result is presented in the next subsection. It uses Theorem 3.1.1 of [KS12] and some ideas from [Shi15, KNS20, KZ20, Shi21].

1.4 Proof of Theorem 1.3

We begin with a description of the scheme of the proof. It is well known that, to prove (1.22) under some minimal compactness assumption (which trivially holds now since the space $\mathfrak{X}$ is compact), it suffices to establish that, for any probability measures λ, λ' on $\mathfrak{X}$, we have

$$\|\mathfrak{P}_k^*\lambda - \mathfrak{P}_k^*\lambda'\|_L^* \le Ce^{-\gamma k} \quad \forall\, k \ge 0, \tag{1.24}$$

where $\mathfrak{P}_k^* : \mathcal{P}(\mathfrak{X}) \to \mathcal{P}(\mathfrak{X})$ denote the Markov operators, corresponding to the kernels $P_k(U;\Gamma)$. To establish (1.24), we use the method of Kantorovich functional (suggested in [KPS02, Kuk02, Kuk06] in the context of randomly forced PDEs), following its presentation in [KS12, Section 3.1.1]. Let $F : \mathfrak{X} \times \mathfrak{X} \to \mathbb{R}_+$ be a measurable symmetric function minorised by $\mathrm{d}_{\mathfrak{X}}$. We define the *Kantorovich functional associated with F* by the formula

$$K_F : \mathcal{P}(\mathfrak{X}) \times \mathcal{P}(\mathfrak{X}) \to \mathbb{R}_+, \quad K_F(\mu, \mu') = \inf_{\mathfrak{M}} \int_{\mathfrak{X}\times\mathfrak{X}} F(V, V')\, \mathfrak{M}(\mathrm{d}V, \mathrm{d}V'),$$

where the infimum is taken over all couplings $\mathfrak{M}$ of μ and μ', i.e., over all measures $\mathfrak{M} \in \mathcal{P}(\mathfrak{X} \times \mathfrak{X})$ whose marginals coincide with μ and μ'. In view of Theorem 3.1.1 of [KS12], relation (1.24) (and so also (1.22)) will be established if we find a function F satisfying the above properties, a number $\varkappa \in (0,1)$, and an integer $s \ge 1$ such that

$$K_F(\mathfrak{P}_s^*\lambda, \mathfrak{P}_s^*\lambda') \le \varkappa K_F(\lambda, \lambda') \quad \text{for any } \lambda, \lambda' \in \mathcal{P}(\mathfrak{X}). \tag{1.25}$$

To prove (1.25), it suffices to find a coupling $(\Phi_s^\lambda, \Phi_s^{\lambda'})$ for the pair of measures $(\mathfrak{P}_s^*\lambda, \mathfrak{P}_s^*\lambda')$ such that

$$\mathbb{E}\, F(\Phi_s^\lambda, \Phi_s^{\lambda'}) \le \varkappa K_F(\lambda, \lambda'), \tag{1.26}$$

[8] Recall that the norm $\|\cdot\|_L^*$ was defined in footnote 1.

since the left-hand side of (1.25) is bounded by $\mathbb{E}\, F(\Phi_s^\lambda, \Phi_s^{\lambda'})$. The following construction of a coupling satisfying this property was suggested in the references mentioned above.

Suppose that, for some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we have constructed measurable maps $\mathcal{R}, \mathcal{R}' : \mathfrak{X} \times \mathfrak{X} \times \Omega \to \mathfrak{X}$ such that $\mathcal{R} = \mathcal{R}(U, U'; \omega)$ and $\mathcal{R}' = \mathcal{R}'(U, U'; \omega)$ form a coupling for $(P_1(U; \cdot), P_1(U'; \cdot))$ for any $U, U' \in \mathfrak{X}$, i.e.,

$$\mathcal{D}(\mathcal{R}) = P_1(U; \cdot), \quad \mathcal{D}(\mathcal{R}') = P_1(U'; \cdot). \tag{1.27}$$

Then, denoting by $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}, \mathcal{F}_k)$ the suitable probability space as in Section 1.1, constructed from the space $(\Omega, \mathcal{F}, \mathbb{P})$ above and formed of points $\widehat{\omega} = (\omega^0, \omega^1, \dots)$ with $\omega^j \in \Omega$, and setting $\widehat{\omega}^k = (\omega^0, \dots, \omega^k)$, we define iterations of $(\mathcal{R}, \mathcal{R}')$ by the formulas[9]

$$\left.\begin{aligned} &\Phi_0(U, U'; \omega^0) = U, \quad \Phi'_0(U, U'; \omega^0) = U', \\ &\Phi_k(U, U'; \widehat{\omega}^k) = \mathcal{R}\big(\Phi_{k-1}(U, U'; \widehat{\omega}^{k-1}), \Phi'_{k-1}(U, U'; \widehat{\omega}^{k-1}), \omega^k\big), \\ &\Phi'_k(U, U'; \widehat{\omega}^k) = \mathcal{R}'\big(\Phi_{k-1}(U, U'; \widehat{\omega}^{k-1}), \Phi'_{k-1}(U, U'; \widehat{\omega}^{k-1}), \omega^k\big). \end{aligned}\right\} \tag{1.28}$$

The construction readily implies that

$$\begin{aligned} \mathcal{D}\big(\{\Phi_k(U, U')\}_{k\ge 0}\big) &= \mathcal{D}\big(\{U_k(U)\}_{k\ge 0}\big), \\ \mathcal{D}\big(\{\Phi'_k(U, U')\}_{k\ge 0}\big) &= \mathcal{D}\big(\{U_k(U')\}_{k\ge 0}\big); \end{aligned} \tag{1.29}$$

we recall that $\{U_k(V), k \ge 0\}$ is a solution of (1.3), (1.4). Besides, it is easy to see that

$$\begin{aligned} &\text{the family } \big\{\big(\Phi_k(U, U')\big), \Phi'_k(U, U')\big), k \ge 0\big\} \text{ is a Markov} \\ &\text{process in } \mathfrak{X} \times \mathfrak{X} \text{ with respect to the filtration } \{\mathcal{F}_k\}. \end{aligned} \tag{1.30}$$

It is shown in [KS12] that if the random variables $\mathcal{R}$ and $\mathcal{R}'$ satisfy certain properties, then for $\lambda = \delta_U$ and $\lambda' = \delta_{U'}$ relation (1.26) with $\Phi_s^\lambda = \Phi_s(U, U')$, $\Phi_s^{\lambda'} = \Phi'_s(U, U')$ holds for some $\varkappa < 1$, and that this implies the validity of (1.26) for any measures λ and λ', with suitable Φ_s^λ and $\Phi_s^{\lambda'}$ constructed from Φ_s and Φ'_s.

In accordance with this programme, we now construct a pair of maps $\mathcal{R}$, $\mathcal{R}'$ and show that they possess the required properties. Recall that the sets D_δ were defined by (1.16). In the following two propositions, we assume that the hypotheses of Theorem 1.3 are fulfilled.

Proposition 1.4. *There is a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, measurable maps $\mathcal{R}, \mathcal{R}' : \mathfrak{X} \times \mathfrak{X} \times \Omega \to \mathfrak{X}$, and positive numbers θ, N, and $q < 1$ satisfying the inequality*

$$N\theta < 1 - q, \tag{1.31}$$

such that (1.27) holds for any $(U, U') \in \mathfrak{X} \times \mathfrak{X}$, and the following properties are fulfilled.

[9] When we write a random variable $\eta(\widehat{\omega})$ as $\eta(\widehat{\omega}^k)$, it means that η depends only on the components $\widehat{\omega}^k$ of a point $\widehat{\omega} \in \widehat{\Omega}$.

Independence. *For $(U,U')\in D_\theta^c$, the random variables $\mathcal{R}(U,U')$ and $\mathcal{R}'(U,U')$ are independent.*

Squeezing. *For $(U,U')\in D_\theta$, we have*

$$\mathbb{P}\big\{\mathrm{d}_{\mathfrak{X}}\big((\mathcal{R}(U,U'),\mathcal{R}'(U,U')\big)\le q\,\mathrm{d}_{\mathfrak{X}}(U,U')\big\}\ge 1-N\mathrm{d}_{\mathfrak{X}}(U,U'). \tag{1.32}$$

As was mentioned above, iterations of $\mathcal{R}$ and $\mathcal{R}'$ allows one to construct a Markov process (Φ_k,Φ'_k) on a suitable filtered probability space $(\widehat{\Omega},\widehat{\mathcal{F}},\widehat{\mathbb{P}},\mathcal{F}_k)$. The following result establishes two key properties of that process.

Proposition 1.5. *There is a number $\theta>0$ satisfying (1.31) such that, for any $(U,U')\in D_\theta$, the following inequality holds for the Markov process (1.30):*

$$\widehat{\mathbb{P}}\big\{\mathrm{d}_{\mathfrak{X}}\big(\Phi_k(U,U'),\Phi'_k(U,U')\big)\le q^k\mathrm{d}_{\mathfrak{X}}(U,U')\ \textit{for all}\ k\ge 0\big)\big\}\ge 1-N_1\,\mathrm{d}_{\mathfrak{X}}(U,U'), \tag{1.33}$$

where $N_1=N(1-q)^{-1}$. Besides, for any $\delta>0$, there exist a number $p>0$ and an integer $l\ge 1$ such that, for any $(U,U')\in\mathfrak{X}\times\mathfrak{X}$,

$$\widehat{\mathbb{P}}\big\{\mathrm{d}_{\mathfrak{X}}\big(\Phi_l(U,U'),\Phi'_l(U,U')\big)\le\delta\big\}\ge p. \tag{1.34}$$

Once these two propositions are established, direct application of the results in [KS12, Section 3.2.3] will imply the validity of (1.26) and prove Theorem 1.3.

Proof of Proposition 1.4. In what follows, we denote $d:=\mathrm{d}_{\mathfrak{X}}(U,U')$. The proof is divided into three steps.

Step 1: Estimate for a total variation distance. Let Φ be the map entering Hypothesis (LAC) and let $\Psi(U,U',\xi)=\xi+\Phi(U,U',\xi)$. We claim that there is $\theta\in(0,\delta)$ such that, for any $(U,U')\in D_\theta$,

$$\Delta(U,U'):=\big\|\mathscr{P}(U';\cdot)-\Psi_*(U,U',\mathscr{P}(U';\cdot))\big\|_{\mathrm{var}}\le C_1 d. \tag{1.35}$$

Here and below C_i, $i=1,2,\dots$, stand for some positive numbers not depending on U and U'.

To prove (1.35), note that, for any $(U,U')\in D_\delta$, the map $\Phi(U,U',\cdot):\mathcal{K}\to E$ satisfies the hypotheses mentioned in the beginning of Section 5. In particular, in view of (1.17) and (1.18), inequalities (5.1) hold with $\varkappa=C_*d$. Applying Proposition 5.1, we see that $\Phi(U,U',\cdot)$ must satisfy inequality (5.2), which coincides with (1.35).

Step 2: Construction of $\mathcal{R}$ and $\mathcal{R}'$. Let us fix any $(U,U')\in D_\theta$ and consider two pairs of random variables in E, $\big(\zeta^U,\Psi(U,U',\zeta^U)\big)$ and $\big(\Psi(U,U',\zeta^U),\zeta^{U'}\big)$, where $\mathcal{D}(\zeta^U)=\mathcal{P}(U,\cdot)$ and $\mathcal{D}(\zeta^{U'})=\mathcal{P}(U',\cdot)$. We denote by $\lambda(U,U')$ the law of the first pair and define $\lambda'(U,U')$ to be the law of a maximal coupling for the second pair, i.e., for the pair

$$\big(\mathcal{D}(\Psi(U,U',\zeta^U))=\Psi_*(U,U',\mathscr{P}(U';\cdot)),\ \mathcal{D}(\zeta^{U'})=\mathscr{P}(U';\cdot)\big).$$

In view of Theorem 1.2.28 in [KS12], we can assume that λ and λ' are transition probabilities from D_θ to $E \times E$. Then $\Pi_{2*}\lambda(U,U') = \Pi_{1*}\lambda'(U,U')$ for any $(U,U') \in D_\theta$, where $\Pi_i : E \times E \to E$, $i = 1,2$ denotes the natural projection from the direct product to its i^{th} component. In view of the measurable version of the gluing lemma (see [Vil09, Chapter 1] and [BM20, Corollary 3.7]), there exists a triplet of E-valued random variables $\tilde\zeta(U,U')$, $\xi(U,U')$, $\tilde\zeta'(U,U')$, defined on the same probability space, such that the maps $\tilde\zeta, \xi, \tilde\zeta' : D_\theta \times \Omega \to E$ are measurable, and for any $(U,U') \in D_\theta$, we have

$$\mathcal{D}\big(\tilde\zeta(U,U'), \xi(U,U')\big) = \lambda(U,U'), \quad \mathcal{D}\big(\xi(U,U'), \tilde\zeta'(U,U')\big) = \lambda'(U,U').$$

Enlarging, if necessary, the probability space, we extend $\tilde\zeta(U,U')$ and $\tilde\zeta'(U,U')$ from $(U,U') \in D_\theta$ to random fields on $\mathfrak{X} \times \mathfrak{X}$ in such a way that

$$\tilde\zeta(U,U') \text{ and } \tilde\zeta'(U,U') \text{ are independent if } (U,U') \in D_\theta^c, \tag{1.36}$$

and their laws are $\mathscr{P}(U;\cdot)$ and $\mathscr{P}(U';\cdot)$, respectively. Finally, for any $U, U' \in \mathfrak{X}$, we set

$$\mathcal{R}(U,U',\omega) = \mathscr{S}\big(U, \tilde\zeta(U,U')\big), \quad \mathcal{R}'(U,U',\omega) = \mathscr{S}\big(U', \tilde\zeta'(U,U')\big). \tag{1.37}$$

This implies, in particular, that $(\mathcal{R}, \mathcal{R}')$ is a coupling for the measures $P_1(U;\cdot)$ and $P_1(U';\cdot)$.

Step 3: Properties of $\mathcal{R}$ and $\mathcal{R}'$. We claim that the maps defined by (1.37) satisfy the properties of Independence and Squeezing. Indeed, the independence of $\mathcal{R}$ and $\mathcal{R}'$ for $(U,U') \in D_\theta^c$ follows from (1.36), so we only need to establish (1.32) for $(U,U') \in D_\theta$.

Since $\mathcal{D}(\xi, \tilde\zeta') = \lambda'(U,U')$ is the law of a maximal coupling for $\mathcal{D}(\Psi(U,U',\zeta^U))$ and $\mathcal{D}(\zeta^{U'})$, it follows from (1.35) that

$$\mathbb{P}\{\xi(U,U') \neq \tilde\zeta'(U,U')\} \le C_1\, d. \tag{1.38}$$

On the other hand, since the law of $(\tilde\zeta, \xi)$ coincides with that of $(\zeta^U, \Psi(U,U',\zeta^U))$, the definition of Ψ and inequality (1.18) imply that, with probability 1,

$$\mathrm{d}_{\mathfrak{X}}\big(\mathscr{S}(U, \tilde\zeta(U,U')), \mathscr{S}(U', \xi(U,U'))\big) \le q\, d.$$

Therefore, if $\omega \in \Omega$ is such that $\xi(U,U') = \tilde\zeta'(U,U')$, then

$$\mathrm{d}_{\mathfrak{X}}\big(\mathscr{S}(U, \tilde\zeta(U,U')), \mathscr{S}(U', \tilde\zeta'(U,U'))\big) \le q\, d.$$

Combining this with (1.38) and (1.37), we arrive at (1.32) with $N = C_1$. □

Proof of Proposition 1.5. Step 1: Proof of (1.33). Given an integer $k \ge 1$, we define the event

$$G_k = \Big\{ \mathrm{d}_{\mathfrak{X}}\big(\Phi_k(U,U'), \Phi'_k(U,U')\big) \le q\, \mathrm{d}_{\mathfrak{X}}\big(\Phi_{k-1}(U,U'), \Phi'_{k-1}(U,U')\big) \Big\}.$$

We claim that, for any $(U, U') \in D_\theta$,

$$\widehat{\mathbb{P}}\big(\widetilde{G}_k\big) \geq 1 - Nd \sum_{l=0}^{k-1} q^l =: \delta_k, \quad \text{where} \quad \widetilde{G}_k = \bigcap_{l=1}^{k} G_l. \tag{1.39}$$

If this inequality is proved, then the probability of the intersection $\widetilde{G}$ of all G_k can be minorised by

$$1 - Nd \sum_{l=0}^{\infty} q^l = 1 - N_1 d,$$

and the required inequality (1.33) will follow from the observation that $\widetilde{G}$ is a subset of the event in the left-hand side of (1.33).

To prove (1.39), we argue by induction. For $k = 1$, inequality (1.39) coincides with (1.32). Assuming that inequality (1.39) holds for $k = m$, we now prove it for $k = m + 1$.

By (1.30) and the Markov property, we have

$$\widehat{\mathbb{P}}\big\{G_{m+1} \,|\, \mathcal{F}_m\big\} = \widehat{\mathbb{P}}\Big\{\mathrm{d}_{\mathfrak{X}}\big(\mathcal{R}(V, V'), \mathcal{R}'(V, V')\big) \leq q\, \mathrm{d}_{\mathfrak{X}}(V, V')\Big\},$$

where $V = \Phi_m(U, U')$ and $V' = \Phi'_m(U, U')$. Now note that if $\widehat{\omega} \in \widetilde{G}_m$, then $\mathrm{d}_{\mathfrak{X}}(\Phi_m(U, U'), \Phi'_m(U, U')) \leq q^m d \leq \theta$, so by (1.32) the probability on the right-hand side is minorised by $1 - Nq^m d$. Using the induction hypothesis, we thus obtain

$$\begin{aligned}\widehat{\mathbb{P}}\big(\widetilde{G}_{m+1}\big) &= \widehat{\mathbb{E}}\Big(\mathbf{1}_{\widetilde{G}_m}\, \widehat{\mathbb{P}}\big\{G_{m+1} \,|\, \mathcal{F}_m\big\}\Big) \geq \big(1 - Nq^m d\big)\, \widehat{\mathbb{P}}\big(\widetilde{G}_m\big) \\ &\geq \big(1 - Nq^m d\big)\, \delta_m > \delta_{m+1}.\end{aligned}$$

This completes the induction step and the proof of (1.33).

Step 2: Transition to a ball of small radius. Let us note that all our constructions and conclusions above remain true if we replace the number θ, defining the set D_θ^c (cf. (1.31)), with any smaller positive constant, depending only on the system (1.3). Thus, without loss of generality, we shall assume in what follows that

$$\theta \leq \frac{1}{2N_1}, \tag{1.40}$$

where $N_1 > 0$ is the number in (1.33). In this case, the right-hand side of (1.33) is minorised by $1/2$ if $d = \mathrm{d}_{\mathfrak{X}}(U, U') \leq \theta$.

As a first step in the proof of (1.34), we establish a simpler result concerning single trajectories $\{U_k, k \geq 0\}$. Namely, we claim that, for any $\delta > 0$, there is $U_\delta \in \mathfrak{X}$, a number $p_\delta > 0$, and an integer $m_\delta \geq 1$ such that

$$\widehat{\mathbb{P}}\big\{U_{m_\delta} \in B_{\mathfrak{X}}(U_\delta, \delta)\big\} \geq p_\delta \quad \text{for any initial point } U \in \mathfrak{X}. \tag{1.41}$$

Indeed, in view of Hypothesis (GCP), in which we take $\varepsilon = \delta/2$, there is a point $\widehat{U} \in \mathfrak{X}$ and an integer $m \geq 1$ such that, for any $U \in \mathfrak{X}$ and some appropriately

chosen vectors $\xi_k \in E$, $1 \le k \le m$ relations (1.15) hold with $\varepsilon = \delta/2$. By the continuity of $\mathscr{S}$, there is $\nu > 0$ such that, for any $\tilde{\xi}_k \in B_E(\xi_k, \nu)$, $1 \le k \le m$, the corresponding trajectory of the RDS (1.3) belongs to the ball $B_{\mathfrak{X}}(\widehat{U}, \delta)$ at time $k = m$. Since $\xi_k \in \operatorname{supp} \mathcal{D}(\zeta^{U_{k-1}})$ for $1 \le k \le m$, it follows that

$$\widehat{\mathbb{P}}\big\{\zeta_k^{U_{k-1}} \in B_E(\xi_k, \nu) \text{ for } 1 \le k \le m\big\} > 0.$$

We have thus proved that the probability of transition from any point $U \in \mathfrak{X}$ to the δ-neighbourhood of $\widehat{U}$ at time $k = m$ is positive. By the Feller property and the well-known characterisation of the weak convergence of measures, the function $U \mapsto P_m(U, B_{\mathfrak{X}}(\widehat{U}, \delta))$ is lower semicontinuous. Since it is positive everywhere, we conclude that

$$\inf_{U \in \mathfrak{X}} P_m\big(U, B_{\mathfrak{X}}(\widehat{U}, \delta)\big) > 0.$$

This shows that (1.41) holds with $m_\delta = m$ and $U_\delta = \widehat{U}$.

Step 3: Proof of (1.34). In view of (1.29), if the processes $\Phi_k(U, U')$ and $\Phi'_k(U, U')$ were independent, then the required result would follow immediately from (1.41). However, they are not, and we have to proceed differently. In what follows, we can assume without loss of generality that $\delta < \theta$, where $\theta > 0$ is satisfies (1.40). Let $k_* > 0$ be the smallest integer such that $\theta q^{k_*} \le \delta$. We claim that (1.34) holds with $l = m_{\delta/2} + k_*$, where $m_{\delta/2}$ is the integer entering (1.41) in which the radius δ of the ball is replaced with $\delta/2$.

To prove this, we abbreviate $\Phi_k(U, U') = \Phi_k$, $\Phi'_k(U, U') = \Phi'_k$, $m = m_{\delta/2}$. We set $\zeta_l = (\Phi_l, \Phi'_l) \in \mathfrak{X} \times \mathfrak{X}$, $d(\zeta_l) = \mathrm{d}_{\mathfrak{X}}(\Phi_l, \Phi'_l)$ and define the stopping times

$$\tau = \tau(U, U') = \min\big\{k \ge 0 : d(\zeta_k) \le \theta\big\}, \quad \tau^m = \tau \wedge (m+1).$$

Without loss of generality, we may assume that, in construction (1.28), when $(U, U') \in D_\theta$ we take random variables $\mathcal{R}(U, U', \omega_k)$, $\mathcal{R}'(U, U', \omega_k)$, $k \ge 1$, from some fixed collection of pairs of independent random fields

$$(\mathcal{R}(\cdot, \omega_k), \mathcal{R}'(\cdot, \omega_k), k = 1, 2, \dots).$$

Then we define processes $\{\hat{\Phi}_k, k \ge 1\}$ and $\{\hat{\Phi}'_k, k \ge 1\}$ by relation (1.28), where always the random fields $\mathcal{R}$ and $\mathcal{R}'$ are independent and are taken from the collection above. We set $\hat{\zeta}_l = (\hat{\Phi}_l, \hat{\Phi}'_l)$ and define $d(\hat{\zeta}_l)$, $\hat{\tau}$, $\hat{\tau}^m$ similarly to $d(\zeta_l)$, τ, τ^m. Note that

$$\tau^m = \hat{\tau}^m \quad \text{and} \quad \zeta_l = \hat{\zeta}_l \quad \text{for } l \le \tau^m.$$

In view of (1.41),

$$\widehat{\mathbb{P}}\{\hat{\tau}^m = m+1\} \le \widehat{\mathbb{P}}\{d(\hat{\zeta}_m) > \theta\} \le 1 - (p_{\theta/2})^2. \tag{1.42}$$

Let us set $A_m = \{\mathrm{d}_{\mathfrak{X}}(\Phi_k(\zeta_{\tau^m}), \Phi'_k(\zeta_{\tau^m})) \le q^k\theta \text{ for } k \ge 0\}$. By the strong Markov property, applied to the process $\{\zeta_l\}$, we have

$$\begin{aligned}\widehat{\mathbb{P}}\{d(\zeta_{\tau^m+k}) \le q^k\theta \text{ for } k \ge 0\} &= \widehat{\mathbb{P}}(A_m) \ge \widehat{\mathbb{P}}\big(\{\tau^m \le m\} \cap A_m\big) \\ &= \mathbb{E}\big(\mathbf{1}_{\{\tau^m \le m\}}\, \mathbb{P}\{A_m \,|\, \mathcal{F}_{\tau^m}\}\big).\end{aligned} \tag{1.43}$$

But for every $\omega \in \{\tau^m \le m\}$, we have $\tau^m = \tau$ and $d(\zeta_{\tau^m}) \le \theta$. So for every such ω it holds that $\mathbb{P}\{A_m \,|\, \mathcal{F}_{\tau^m}\} \ge \frac{1}{2}$ in view of (1.33) and (1.40). Since $\tau^m = \hat{\tau}^m$, then by (1.42) the right-most term in (1.43) is no less than $\frac{1}{2}(p_{\theta/2})^2$. As the event in (1.34) contains the event $\{d(\zeta_{\tau^m+k}) \le q^k\theta \text{ for } k \ge 0\}$ if $l = m+1$, then (1.34) is proved with $l = m_{\delta/2} + 1$ and $p = \frac{1}{2}(p_{\theta/2})^2$.

We have thus established (1.22). To establish (1.23), we first note that it suffices to find $C_s > 0$ such that

$$\|\mathcal{D}(U_0, \dots, U_s) - \mathcal{D}(U'_0, \dots, U'_s)\|_L^* \le C_s \mathrm{d}_{\mathfrak{X}}(U, U') \quad \text{for } U, U' \in \mathfrak{X}, \tag{1.44}$$

where $\{U_k\}$ and $\{U'_k\}$ stand for the trajectories of (1.3) issued from U and U', respectively. Once this inequality is proved, (1.23) will easily follow due to the Markov property.

To prove (1.44), we first consider the case $s = 1$. In view of (1.3) and (1.6), for any continuous function $f : \mathfrak{X} \times \mathfrak{X} \to \mathbb{R}$, we have

$$\big\langle f, \mathcal{D}(U_0, U_1)\big\rangle = \int_{\mathcal{K}} f\big(U, \mathscr{S}(U, y)\big)\mathscr{P}(U; \mathrm{d}y).$$

Assuming that f is 1-Lipschitz and satisfies the inequality $\|f\|_\infty \le 1$, and using the Lipschitz property of $\mathscr{S}$ and inequality (1.1), it is easy to show that

$$\big|\big\langle f, \mathcal{D}(U_0, U_1)\big\rangle - \big\langle f, \mathcal{D}(U'_0, U'_1)\big\rangle\big| \le C_1 \, \mathrm{d}_{\mathfrak{X}}(U, U'),$$

where $C_1 > 0$ does not depend on f. Taking the supremum with respect to f, we arrive at (1.44) for $s = 1$. A similar argument combined with induction allows one to establish (1.44) for any integer $s \ge 2$. The proof of Theorem 1.3 is complete. □

2 Main results

2.1 Formulation of the main theorem

Let us recall the setting of Subsection 1.2. As was mentioned there after Lemma 1.2, we assume without loss of generality that H is a separable Hilbert space, E is a separable Banach space, $S : H \times E \to H$ is a continuous map, and $X \subset H$ and $\mathcal{K} \subset E$ are compact subsets such that $S(X \times \mathcal{K}) \subset X$. We consider the stochastic system in (1.8),

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

supplemented with an initial condition

$$u_0 = u, \tag{2.1}$$

where $u \in X$ is a given point, and $\{\eta_k\}$ is a stationary sequence satisfying the properties mentioned in Subsection 1.2. In particular, the law of η_k has the compact support $\mathcal{K} \subset E$.

Definition 2.1. We shall say that stochastic system (1.8) is *exponentially mixing in the dual-Lipschitz norm* [10] if there is a measure $\mu \in \mathcal{P}(X)$ and positive numbers C and γ such that, for any initial state $u \in X$, the corresponding trajectory $\{u_k\}_{k\in\mathbb{Z}_+}$ of (1.8), (2.1) satisfies the inequality

$$\|\mathcal{D}(u_k) - \mu\|_L^* \le C\, e^{-\gamma k}, \quad k \ge 0. \tag{2.2}$$

In what follows, we drop for short *in the dual-Lipschitz norm* and call stochastic systems as in Definition 2.1 just *exponentially mixing*. Let us note that if the initial state u_0 is an X-valued random variable independent of $\{\eta_k\}_{k\in\mathbb{Z}}$, then (2.2) remains valid. To see this, it suffices to remark that if $\lambda = \mathcal{D}(u_0)$, then

$$\mathcal{D}(u_k) = \int_X P_k(v,\cdot)\lambda(\mathrm{d}v),$$

where $P_k(v,\cdot)$ stands for the law at time k of the trajectory of (1.8), (2.1) with $u = v$. We recall that the laws $\mathcal{D}(u_k)$, $k \ge 1$, depend not on the process $\{\eta_k\}$, but only on its distribution.

Below, we shall prove that, under certain restrictions, system (1.8) is mixing, and at the end of Section 2.2, shall show that the measure μ in (2.2) is stationary in the following sense: there exists a random variable v, measurable with respect to $\mathcal{F}_0$, such that $\mathcal{D}(v) = \mu$ and a weak solution $\{u_k, k \ge 0\}$ for (1.8) such that $\mathcal{D}(u_k) \equiv \mu$. It is obvious from (2.2) that μ is the only measure with this property.

To ensure the property of exponential mixing for (1.8), we assume that the map $S : H \times E \to H$ is *twice continuously differentiable and is bounded on bounded subsets together with its derivatives up to the second order*, and impose the following four hypotheses on the dynamics and driving noise. For an integer $k \ge 1$, we denote $\mathbf{0}_k = (0,\dots,0)$, where $0 \in E$ is repeated k times, and for a vector $\vec{\xi}_n = (\xi_1,\dots,\xi_n) \in \operatorname{supp}\mathcal{D}(\eta_1,\dots,\eta_n) \subset \mathcal{K}^n$, we denote

$$S_n(v;\vec{\xi}_n) = u_n, \text{ where } \{u_1,\dots,u_n\} \text{ is a trajectory of (1.8) with } \eta_k = \xi_k. \tag{2.3}$$

(GD) Global dissipation. *There is an integer $k \ge 1$ and a number $a \in (0,1)$ such that*

$$\|S_k(u;\mathbf{0}_k)\|_H \le a\,\|u\|_H \quad \text{for any } u \in H.$$

Remark 2.2. This hypothesis implies that the unperturbed dynamics has one globally stable equilibrium. In a subsequent paper, we shall relax this condition to include stochastic perturbations of more general systems. The result will be applicable, for instance, to a reaction-diffusion equation possessing finitely many hyperbolic equilibria and a Lyapunov function.

We shall say that a closed subspace $G \subset H$ is *determining* if there is a number $\varkappa \in (0,1)$ such that

$$\big\|(I - \mathsf{P}_G)D_u S(u,\eta)\big\|_{\mathcal{L}(H)} \le \varkappa \quad \text{for any } u \in X,\ \eta \in \mathcal{K}, \tag{2.4}$$

[10] See the first footnote.

where $\mathsf{P}_G : H \to H$ denote the orthogonal projection to G. This definition is related to the usual concept of *determining modes* going back to Foiaş–Prodi [FP67]. Indeed, assuming that X is convex and applying the mean value theorem, we see that

$$\big\|(I - \mathsf{P}_G)\big(S(u,\eta) - S(u',\eta)\big)\big\|_H \le \varkappa \|u - u'\|_H.$$

In particular, if $\{u_k\}$ and $\{u'_k\}$ are two trajectories of (1.8) with the same driving noise such that $\mathsf{P}_G u_k = \mathsf{P}_G u'_k$ for $k \ge 1$, then the difference $u_k - u'_k$ goes to zero exponentially fast.

Remark 2.3. If H is a finite-dimensional space, then (2.4) trivially holds with $G = H$. It also holds if the map S is smoothing in the sense that for some Hilbert space V, compactly and densely embedded in H, the map S is continuously differentiable from $H \times E$ to V. Indeed, in this case H admits a Hilbert basis $\{e_j, j \ge 1\}$ which also is an orthogonal basis of V such that the norms $\|e_j\|_V$ go to infinity with j (e.g., see [LM72, Section 2.1]). Then (2.4) is fulfilled if we take for G the vector span of $\{e_j, 1 \le j \le N\}$ with a sufficiently large N.

(ALC) Approximate linearised controllability. *There exists an open set $O \supset X \times \mathcal{K}$ and a finite-dimensional determining subspace $G \subset H$ such that the closure of the image of the operator $D_\eta S(u,\eta) : E \to H$ contains G for any $(u,\eta) \in O$.*

Recall that we denote by $Q(\boldsymbol{\xi};\cdot)$ the conditional law of η_1 given the past $\{\eta_k, k \in \mathbb{Z}_-\} = \boldsymbol{\xi} \in \mathcal{E}$ and that Assumption (Fel) holds for it. The next hypothesis is an analogue of (DLP) in the current setting.

(DLP′) Decomposability and Lipschitz property. *There is an increasing sequence of finite-dimensional subspaces $F_n \subset E$, with complementary subspaces $F_n^\dagger \subset E$, and projections $\mathsf{P}_n : E \to F_n$, such that the union $\cup_n F_n$ is dense in E, the operator norms of P_n are bounded uniformly in $n \ge 1$, and the following property holds for any $n \ge 1$: for each $\boldsymbol{\xi} \in \mathcal{E}$, the measure $Q(\boldsymbol{\xi};\cdot)$ admits a disintegration*

$$Q(\boldsymbol{\xi}; \mathrm{d}y) = Q_n^\dagger(\boldsymbol{\xi}; \mathrm{d}y_n^\dagger) Q_n(\boldsymbol{\xi}, y_n^\dagger; \mathrm{d}y_n), \tag{2.5}$$

where the decomposition $y = (y_n, y_n^\dagger)$ is associated with the direct sum $E = F_n \dotplus F_n^\dagger$, and we write $Q_n^\dagger(\boldsymbol{\xi};\cdot) \in \mathcal{P}(F_n^\dagger)$ for the image of $Q(\boldsymbol{\xi};\cdot)$ under the projection $I - \mathsf{P}_n$ onto $F_n^\dagger$. Moreover, there is a Lipschitz continuous function $\rho_n : \mathcal{E} \times E \to \mathbb{R}_+$ supported by $\mathcal{E} \times \mathcal{K}$ such that

$$Q_n(\boldsymbol{\xi}, y_n^\dagger; \mathrm{d}y_n) = \rho_n(\boldsymbol{\xi}, y_n^\dagger, y_n) \ell_n(\mathrm{d}y_n),$$

where ℓ_n stands for the Lebesgue measure on F_n.

This hypothesis is a version of a condition often imposed in Dobrushin's work on reconstructing random fields on lattices from their conditional distributions (e.g., Assumption 2 in [Dob70, Theorem 1] or [Dob68, Theorems 5 and 6]). If

$\dim F < \infty$, we take $F_n = F$ for all $n \geq 1$, so that $F_n^\dagger = \{0\}$, and condition (DLP′) simplifies to the assumption that $Q(\boldsymbol{\xi}; \mathrm{d}y) = \rho(\boldsymbol{\xi}, y)\ell(\mathrm{d}y)$, where ρ is a Lipschitz function.

For any $\boldsymbol{\xi} \in \mathcal{E}$ and any integer $k \geq 1$, the conditional law of the vector $(\eta_1, \dots, \eta_k)$ given that the past $\boldsymbol{\eta}_- := \{\eta_l, l \leq 0\}$ equals $\boldsymbol{\xi}$ is the measure

$$\boldsymbol{Q}_k(\boldsymbol{\xi}; \mathrm{d}y_1, \dots, \mathrm{d}y_k) = Q(\boldsymbol{\xi}; \mathrm{d}y_1)Q(\boldsymbol{\xi}_1; \mathrm{d}y_2) \cdots Q(\boldsymbol{\xi}_{k-1}; \mathrm{d}y_k), \tag{2.6}$$

where $\boldsymbol{\xi}_l = (\boldsymbol{\xi}, y_1, \dots, y_l)$ is the sequence obtained by concatenation of $\boldsymbol{\xi}$ and $(y_1, \dots, y_l)$. Given integers $k \geq 0$ and $m \geq 1$, we denote by $Q_k^m(\boldsymbol{\xi}; \cdot) \in \mathcal{P}(\mathcal{K}^m)$ the measure

$$Q_k^m(\boldsymbol{\xi}; \cdot) = \langle \text{projection of } \boldsymbol{Q}_{k+m}(\boldsymbol{\xi}; \cdot) \text{ to the last } m \text{ components} \rangle. \tag{2.7}$$

Thus, $Q_k^m(\boldsymbol{\xi}; \cdot)$ is the conditional law of the vector $(\eta_{k+1}, \dots, \eta_{k+m})$ given the past $\boldsymbol{\eta}_- = \boldsymbol{\xi}$. The following hypothesis is similar to the property of *strong recurrence* in the context of Markov processes.

(SRZ) Strong recurrence to zero. *For any $n \in \mathbb{N}$ and $\delta > 0$, there is an integer $s \geq 0$ such that*

$$\inf_{\boldsymbol{\xi} \in \mathcal{E}} Q_s^n\bigl(\boldsymbol{\xi}; \mathcal{O}_\delta(\mathbf{0}_n)\bigr) > 0. \tag{2.8}$$

The following theorem is the main result of this paper. Its proof is based on Theorem 1.3 and is presented in the next subsection.

Theorem 2.4. *Suppose that Hypotheses (GD), (ALC), (DLP′), and (SRZ) are fulfilled. Then the RDS (1.8) is exponentially mixing.*

In conclusion of this subsection, let us mention that the question of existence of processes for which Hypotheses (DLP′) and (SRZ) are satisfied will be discussed in Section 3. Here we only point out that if an E-valued stationary process $\boldsymbol{\eta} := \{\eta_k\}_{k \in \mathbb{Z}}$ is such that (DLP′) holds for its conditional probabilities $Q(\boldsymbol{\xi}; \cdot)$, then we have the following implications:

$$\text{strong feeble mixing for } \boldsymbol{\eta} \implies \text{strong recurrence} \implies \text{feeble mixing for } \boldsymbol{\eta}; \tag{2.9}$$

see Subsection 3.2 for exact statements. Thus, our class of admissible random forces $\boldsymbol{\eta}$ can be informally characterised as *mixing stationary processes with Lipschitz continuous densities for conditional laws*.

2.2 Proof of Theorem 2.4

Let us recall that the law of a trajectory u_k for (1.8) issued from $u \in X$ depends not on the concrete realisation of the process $\{\eta_k, k \geq 1\}$, but only on its law. It coincides with that of the first component of the trajectory $U_k = (v_k, \boldsymbol{\xi}_k)$ for (1.13), with the phase space $\mathfrak{X} = X \times \mathcal{E}$, where $\mathscr{S}$ is defined by (1.12), and $\mathcal{D}(U_0) = \delta_u \otimes \sigma$; see (1.9). Below we call the corresponding Markov process in $\mathfrak{X}$

the *U-process*. If we prove that the U-process satisfies (1.22), then taking its projection Π_X to the X-component, we can conclude that $\mathcal{D}(u_k)$ satisfies (2.2) with

$$\mu = (\Pi_X)_* \boldsymbol{\mu}\,; \tag{2.10}$$

see (1.11). Moreover, applying (1.22) to a solution of (1.3) with initial condition U_0 as above and using (1.14), we get

$$(\Pi_{\mathcal{E}})_* \boldsymbol{\mu} = \sigma. \tag{2.11}$$

Thus, to prove Theorem 2.4, it suffices to check that the three hypotheses of Theorem 1.3 are fulfilled. The proof of this fact is divided into three steps. To simplify notation, we denote by $\boldsymbol{\sigma} \in \mathcal{P}(E^{\mathbb{Z}})$ the law of $\boldsymbol{\eta} = \{\eta_k, k \in \mathbb{Z}\}$ and write $\boldsymbol{\sigma}_m$ for the law of $(\eta_1, \dots, \eta_m)$, so that $\boldsymbol{\sigma}_m$ is the projection of $\boldsymbol{\sigma}$ to E^m.

Step 1: Transition probability. We first express the law of the noise in the RDS (1.13) and the transition probability of the U-process in terms of the objects entering (1.8). To this end, note that since the law of $\zeta_1^{\boldsymbol{\xi}}$ equals $Q(\boldsymbol{\xi}; \cdot)$ for any $\boldsymbol{\xi} \in \mathcal{E}$, we have

$$\mathscr{P}(U; \Gamma) = Q(\boldsymbol{\xi}; \Gamma) \quad \text{for } \Gamma \in \mathcal{B}(E),\ U = (v, \boldsymbol{\xi}) \in \mathfrak{X}. \tag{2.12}$$

This relation and Hypothesis (DLP′) imply that $\mathscr{P}(U; \cdot)$ satisfies Hypothesis (DLP), in which one can take $F = F_n$ with an arbitrary $n \geq 1$.

We now describe the transition probability of the U-process. Recalling (1.12) and (1.6), we see that the time-1 transition probability can be written as

$$P_1\big(U, B \times (\boldsymbol{\Gamma} \times \Gamma)\big) = \delta_{\boldsymbol{\xi}}(\boldsymbol{\Gamma}) \int_{\Gamma} \mathbf{1}_B\big(S(v, y)\big) Q(\boldsymbol{\xi}; \mathrm{d}y), \tag{2.13}$$

where U and Γ are the same as in (2.12), $\boldsymbol{\Gamma} \in \mathcal{B}(\mathcal{E})$, and $B \in \mathcal{B}(X)$. In other words, for $U = (v, \boldsymbol{\xi}) \in \mathfrak{X}$, denoting $\Psi_U : \mathcal{K} \to X \times \mathcal{E}$, $\eta \mapsto \mathscr{S}(U, \eta)$, we can write $P_1(U; \cdot) = (\Psi_U)_* Q(\boldsymbol{\xi}; \cdot)$. Then, using (2.13) and arguing by induction, it is straightforward to see that the time-k transition probability has the form

$$P_k\big(U, B \times (\boldsymbol{\Gamma} \times \Gamma_k)\big) = \delta_{\boldsymbol{\xi}}(\boldsymbol{\Gamma}) \int_{\Gamma_k} \mathbf{1}_B\big(S_k(v; y_1, \dots, y_k)\big) \boldsymbol{Q}_k(\boldsymbol{\xi}; \mathrm{d}y_1, \dots, \mathrm{d}y_k),$$

where $U = (v, \boldsymbol{\xi}) \in \mathfrak{X}$, $\boldsymbol{\Gamma} \in \mathcal{B}(\mathcal{E})$, $\Gamma_k \in \mathcal{B}(E^k)$, $B \in \mathcal{B}(X)$, and we recall (2.3) and (2.6).

Step 2: Global controllability to points (GCP). We claim that the required property holds with $\widehat{U} = (0, \mathbf{0})$, where $\mathbf{0} \in \mathcal{E}$ is the sequence whose elements are all zero. Indeed, let us fix any $\varepsilon > 0$ and choose an integer $l \geq 1$ and a number $\delta > 0$ so that, for any element $\hat{\boldsymbol{\xi}} = (\hat{\xi}_j, j \leq 0) \in \mathcal{E}$ satisfying the inequality $\|\hat{\xi}_j\|_E \leq \delta$ for $1 - l \leq j \leq 0$, we have $\mathrm{d}(\hat{\boldsymbol{\xi}}, \mathbf{0}) < \varepsilon/3$. Let us take any $U = (u, \boldsymbol{\xi}) \in \mathfrak{X}$. By (GD), there is an integer $n \geq l$ not depending on U such that $\|S_n\big(u; \mathbf{0}_n\big)\|_H < \frac{\varepsilon}{3L}$. Since S is continuous, we can find $\delta > 0$ such that

$$\big\|S_n\big(u; \xi_1', \dots, \xi_n'\big)\big\|_H < \frac{2\varepsilon}{3L} \quad \text{for } u \in X,\ (\xi_1', \dots, \xi_n') \in \mathcal{O}_\delta(\mathbf{0}_n). \tag{2.14}$$

By (SRZ), there is $s \in \mathbb{N}$ such that (2.8) holds. It follows that there exist vectors $\xi_1, \dots, \xi_{s+n} \in \mathcal{K}$ such that

$$(\xi_{s+1}, \dots, \xi_{s+n}) \in \mathcal{O}_\delta(\mathbf{0}_n), \tag{2.15}$$

$$\xi_{j+1} \in \operatorname{supp} Q\big((\boldsymbol{\xi}; \xi_1, \dots, \xi_j); \cdot\big) \quad \text{for } 0 \le j < s+n. \tag{2.16}$$

We now set $v_k = S(v_{k-1}, \xi_k)$, where $1 \le k \le s+n$ and $v_0 = u$. In view of (2.14) and (2.15), we have $\|v_{s+n}\|_H < \frac{2\varepsilon}{3L}$. Furthermore, inequality (2.15) and the choice of l imply that $\mathrm{d}(\boldsymbol{\xi}_{s+n}, \mathbf{0}) < \frac{\varepsilon}{3}$, where $\boldsymbol{\xi}_k = (\boldsymbol{\xi}, \xi_1, \dots, \xi_k)$. Recalling that the U-process is given by $U_k = (v_k, \boldsymbol{\xi}_k)$ and using (1.10), we see that

$$\mathrm{d}_{\mathfrak{X}}(U_{s+n}, \widehat{U}) = L\|v_{s+n}\|_H + \mathrm{d}(\boldsymbol{\xi}_{s+n}, \mathbf{0}) < \varepsilon.$$

Finally, it follows from (2.16) and (2.12) that $\xi_k \in \operatorname{supp} \mathscr{P}(U_{k-1}; \cdot)$ for $1 \le k \le s+n$. This completes the verification of (1.15) with $m = s+n$.

Step 3: Local approximate controllability (LAC). To prove (LAC), we define the set

$$\widetilde{D}_\delta = \{(v, v') \in X \times X : \|v - v'\|_H \le \delta\}$$

and suppose that, for a small $\delta > 0$, we have constructed a continuous map

$$\tilde{\Phi} : \widetilde{D}_\delta \times \mathcal{K} \to E, \quad (v, v', \xi) \mapsto \xi',$$

such that, for any $(v, v') \in \widetilde{D}_\delta$, the map $\tilde{\Phi}(v, v', \cdot) : \mathcal{K} \to E$ is continuously differentiable and satisfies the inequalities (cf. (1.17) and (1.18))

$$\sup_{\xi \in \mathcal{K}} \Big(\|\tilde{\Phi}(v, v', \xi)\|_E + \operatorname{Lip}_\xi\big(\tilde{\Phi}(v, v', \xi)\big) \Big) \le C' \|v - v'\|_H, \tag{2.17}$$

$$\sup_{\xi \in \mathcal{K}} \big\| S(v, \xi) - S\big(v', \xi + \tilde{\Phi}(v, v', \xi)\big) \big\|_H \le q' \|v - v'\|_H, \tag{2.18}$$

where C' and $q' < 1$ are some positive numbers. In this case, recalling the constant $L \ge 1$ in the distance (1.10) and defining $\Phi : D_\delta \times \mathcal{K} \to E$ by the relation

$$\Phi(U, U', \xi) = \tilde{\Phi}(v, v', \xi), \quad U = (v, \boldsymbol{\xi}), \quad U' = (v', \boldsymbol{\xi}'),$$

we see that (1.17) is trivially satisfied with $C_* = C' L^{-1}$. To prove (1.18), we use (2.17) and (2.18) to write

$$\begin{aligned} \mathrm{d}_{\mathfrak{X}}\big(\mathscr{S}(U, \xi), \mathscr{S}(U', \xi + \Phi(U, U', \xi)\big) &\le \iota^{-1}\, \mathrm{d}(\boldsymbol{\xi}, \boldsymbol{\xi}') + \|\Phi(U, U', \xi)\|_E \\ &\quad + L\, \|S(v, \xi) - S(v', \xi + \tilde{\Phi}(v, v', \xi)\|_H \\ &\le \iota^{-1}\, \mathrm{d}(\boldsymbol{\xi}, \boldsymbol{\xi}') + L\, (L^{-1} C' + q') \|v - v'\|_H. \end{aligned}$$

Up to this point, the number $L \ge 1$ was arbitrary. We now choose it so large that $L^{-1} C' + q' < 1$. Then the above estimate implies inequality (1.18) with $q = \max\{\iota^{-1}, L^{-1} C' + q'\} < 1$. Thus, it remains to construct $\tilde{\Phi}$ satisfying (2.17) and (2.18).

To this end, let us denote by $\Delta(v, v', \xi)$ the expression under the norm sign in the left-hand side of (2.18), where $\tilde{\Phi}$ is a map to be defined. Using the Taylor formula and the C^2-smoothness of S, we write

$$\Delta(v, v', \xi) = (D_v S)(v, \xi)(v - v') - (D_\xi S)(v, \xi)\tilde{\Phi}(v, v', \xi) + r(v, v', \xi), \quad (2.19)$$

where the remainder term r satisfies the inequality

$$\|r(v, v', \xi)\|_H \le C_1\big(\|v' - v\|_H^2 + \|\tilde{\Phi}(v, v', \xi)\|_E^2\big) \quad (2.20)$$

with some constant C_1 not depending on v, v', and ξ. We now need the following lemma whose proof is given in Section 2.4 and is based on the *Moore–Penrose formula* for pseudo-inverse; cf. [KNS20, Theorem 2.8].

Lemma 2.5. *Let H and E be separable Hilbert and Banach spaces, $G \subset H$ be a finite-dimensional vector subspace, O be an open subset of a Banach space $\mathcal{H}$, and $A : O \to \mathcal{L}(E, H)$ be a C^1-smooth map such that, for any $y \in O$, the closure of the image of the linear application $A(y)$ contains G. Let finite-dimensional subspaces $F_n \subset E$ be as in (DLP′). Then, for any $\varepsilon > 0$ and any compact set $Y \subset O$, there is an integer $n \ge 1$ and a C^1-smooth map*[11] *$B_\varepsilon : O \to \mathcal{L}(G, F_n)$ such that*

$$\|A(y)B_\varepsilon(y)f - f\|_H \le \varepsilon \|f\|_H \quad \text{for any } f \in G,\ y \in Y. \quad (2.21)$$

By Hypothesis (ALC), for any $(v, \xi) \in O \subset H \times E$, the closure of the image of the operator $D_\xi S(v, \xi) : E \to H$ contains the finite-dimensional subspace G. Applying Lemma 2.5 to $A(v, \xi) = D_\xi S(v, \xi)$, for any $\varepsilon > 0$ we can construct C^1-smooth mapping $B_\varepsilon : O \to \mathcal{L}(G, F_n)$ such that

$$\|D_\xi S(v, \xi)B_\varepsilon(v, \xi)f - f\|_H \le \varepsilon \|f\|_H \quad \text{for } (v, \xi) \in X \times \mathcal{K},\ f \in G. \quad (2.22)$$

We now set

$$\tilde{\Phi}(v, v', \xi) = B_\varepsilon(v, \xi)\, \mathsf{P}_G (D_v S)(v, \xi)(v - v'). \quad (2.23)$$

The validity of (2.17) with some constant $C' = C'(\varepsilon)$ follows from the compactness of $X \times \mathcal{K}$ and the continuity of the functions entering (2.23). Furthermore, substituting $\tilde{\Phi}$ into (2.19) and using (2.4), (2.22), and (2.20), we derive

$$\|\Delta(v, v', \xi)\|_H \le (\varkappa + C_2\varepsilon)\|v - v'\|_H + C_3\|v - v'\|_H^2,$$

where we set

$$C_2 = \sup\big\{\big\|\mathsf{P}_G(D_v S)(v, \xi)\big\|_{\mathcal{L}(H,G)} : (v, \xi) \in X \times \mathcal{K}\big\}, \quad C_3 = C_1(1 + C').$$

Up to this point, the number $\varepsilon > 0$ was arbitrary. We now choose $\varepsilon > 0$ and $\delta > 0$ so that $q' := \varkappa + C_2\varepsilon + C_3\delta < 1$. In this case, inequality (2.18) holds, and this completes the proof of the theorem. □

[11] Let us emphasise that the map B_ε is defined on O, but it depends on the choice of Y.

2.3 Concluding remarks

The proof given above implies two additional properties: the stationarity of the limiting measure μ and convergence in the space of trajectories. Namely, we have the two propositions below, provided that the conditions of Theorem 2.4 are fulfilled.

Proposition 2.6 (Stationarity of μ)**.** *Equation* (1.8) *possesses a weak solution* $\{(\hat{u}_k, \hat{\xi}_k), k \in \mathbb{Z}_+\}$ *such that* $\{(\hat{u}_k, \hat{\xi}_k)\}$ *is a stationary process in* $X \times \mathcal{E}$*. Moreover,* $\mathcal{D}(\hat{u}_k) \equiv \mu$*.*

Proof. Let us consider again the extended system (1.13), which generates a Markov U-process in the space $\mathfrak{X} = X \times \mathcal{E}$. The latter has a unique stationary measure $\boldsymbol{\mu} \in \mathcal{P}(\mathfrak{X})$ whose projection to X is the limiting measure $\mu \in \mathcal{P}(X)$ for (1.8); see (2.10). Let $\{V_k = (v_k, \boldsymbol{\xi}_k)\}$ be a stationary trajectory of (1.13), so that $\mathcal{D}(V_k) \equiv \boldsymbol{\mu}$. It suffices to take $\hat{u}_k = v_k$ for $k \geq 0$. Indeed, by (2.11), $\mathcal{D}(\boldsymbol{\xi}_k) = \sigma$. So, by Lemma 1.2, $\{v_k, k \geq 0\}$ is a weak solution of (1.8), and $\mathcal{D}(v_k) = (\Pi_X)_* \boldsymbol{\mu} = \mu$ for each $k \geq 0$. The construction readily implies that $\{\hat{u}_k, k \geq 0\}$ is an X-valued stationary process whose law is equal to μ at any time $k \geq 0$. □

Relation (1.8) implies that the process $\{\hat{u}_k\}_{k \in \mathbb{Z}_+}$ is feebly mixing in the sense of Section 3.2; cf. the proof of Proposition 3.4.

Proposition 2.7 (Convergence of trajectories)**.** *Let* u *be an* X*-valued random variable independent of* $\{\eta_k\}$ *and let* $\{u_k, k \in \mathbb{Z}_+\}$ *be the trajectory defined by* (1.8), (2.1)*. Then, for any integer* $s \geq 1$*, the process* $[u_k, \dots, u_{k+s}]$ *converges in law in the dual-Lipschitz metric to* $[\hat{u}_0, \dots, \hat{u}_s]$*, where* $\hat{u}_k$ *is the stationary trajectory described above. Moreover, the convergence is uniform with respect to all initial conditions.*

Proof. Let us consider again the extended system (1.13) and its stationary trajectory $\{V_k, k \in \mathbb{Z}_+\}$ as in Theorem 1.3. Let U_k be the trajectory of (1.13) issued from an initial state whose law equals $\lambda \otimes \sigma \in \mathcal{P}(\mathfrak{X})$, where $\lambda = \mathcal{D}(u) \in \mathcal{P}(X)$. Projecting estimate (1.23) to the first component, we obtain the required convergence. □

Finally, let us notice that, by Theorem 2.4, Theorem 1.3 applies to the U-process in the space $\mathfrak{X} = X \times \mathcal{E}$ and, hence, the latter is exponentially mixing. So, for any $V = (v, \boldsymbol{\xi}) \in X \times \mathcal{E}$, the corresponding trajectory U_k defined by (1.13) converges to $\boldsymbol{\mu}$ in the dual-Lipschitz metric. The following proposition, whose proof may be obtained by literal repetition of the arguments in Section 2.2, shows that the initial point V can be taken from the larger set $X \times \boldsymbol{\mathcal{K}}$, provided that the strong recurrence to zero holds for $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$; cf. Theorem 3.2 below.

Proposition 2.8. *Let* $\{Q(\boldsymbol{\xi}; \cdot), \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}\} \subset \mathcal{P}(\mathcal{K})$ *be a transition probability on* $\mathcal{K}$ *that coincides with the conditional probability introduced in Subsection 1.2 for*

$\boldsymbol{\xi} \in \mathcal{E}$. *Suppose that the hypotheses of Theorem 2.4 are fulfilled, with inequality (2.8) replaced by the stronger condition*

$$\inf_{\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}} Q_s^n\big(\boldsymbol{\xi}; \mathcal{O}_\delta(\mathbf{0}_n)\big) > 0, \tag{2.24}$$

where $Q_s^n(\boldsymbol{\xi}; \cdot)$ is defined in (2.7). Then there are positive numbers γ and C such that

$$\sup_{V \in X \times \boldsymbol{\mathcal{K}}} \|\mathcal{D}(U_k) - \boldsymbol{\mu}\|_L^* \le C e^{-\gamma k}, \quad k \ge 0,$$

where $\boldsymbol{\mu} \in \mathcal{P}(\mathfrak{X})$ is the stationary measure of the U-process on $X \times \mathcal{E}$ constructed in the proof of Theorem 2.4, and $\|\cdot\|_L^*$ stands for the dual-Lipschitz metric over $X \times \boldsymbol{\mathcal{K}}$.

2.4 Proof of Lemma 2.5

The result would follow immediately from Theorem 2.8 in [KNS20] if E was a Hilbert space. Since this is not the case, we need to proceed differently.

Let us fix a compact set $Y \subset O$ and a number $\varepsilon > 0$, and denote by $\mathrm{d}_H(f, H')$ the distance of f from a subspace $H' \subset H$. Since the union $\cup_n F_n$ is dense in E, then the continuous functions $(y, f) \mapsto \mathrm{d}_H\big(f, A(y)F_n\big)$ pointwise monotonically converge to zero as $n \to \infty$. By the Dini theorem, this convergence is uniform on compact sets. So there exists an integer $n \ge 1$ such that

$$\sup_{y \in Y} \sup_{f \in \overline{B_G(1)}} \mathrm{d}_H\big(f, A(y)F_n\big) \le \varepsilon. \tag{2.25}$$

Let us endow F_n with an inner product and denote by $A_n(y)$ the restriction of $A(y)$ to F_n. Then we can consider the adjoint operator $A_n(y)^* : H \to F_n$ and, given $\delta > 0$, define the family of operators

$$B(y, \delta) = A_n(y)^* \big(A_n(y) A_n(y)^* + \delta I\big)^{-1} : H \to F_n,$$

which form a C^1 function of $(y, \delta) \in O \times (0, +\infty)$ with range in $\mathcal{L}(H, F_n)$. Suppose we have proved that, for any $y \in Y$ and $f \in \overline{B_G(1)}$,

$$\limsup_{\delta \to 0^+} \|D(y, \delta) f - f\|_H \le 2\varepsilon, \tag{2.26}$$

where we set $D(y, \delta) = A(y)B(y, \delta)$. Noting that $\|D(y, \delta)\|_{\mathcal{L}(H)} \le 1$ for any (y, δ) and denoting by $f_1, \dots, f_N \in G$ an ε-net in $\overline{B_G(1)}$, we can write

$$\|D(y, \delta) f - f\|_H \le 2\varepsilon + \|D(y, \delta) f_j - f_j\|_H,$$

where $j \in \{1, \dots, N\}$ is such that $\|f - f_j\|_H \le \varepsilon$. Combining this with (2.26), we see that, for any $y \in Y$ and a sufficiently small $\delta_y > 0$, we have

$$\sup_{f \in \overline{B_G(1)}} \|D(y, \delta) f - f\|_H \le 3\varepsilon \quad \text{for } 0 < \delta \le \delta_y.$$

Since the function $\delta \mapsto \|D(y,\delta)f - f\|_H$ decreases with $\delta > 0$, using the compactness of Y and the continuity of the function $y \to D(y,\delta)$, we can find $\delta_\varepsilon > 0$ such that

$$\sup_{y\in Y} \sup_{f\in\overline{B_G(1)}} \|D(y,\delta_\varepsilon)f - f\|_H \le 4\varepsilon.$$

By homogeneity, this implies the required inequality (2.21) with $B_\varepsilon(y) = B(y,\delta_\varepsilon)$, in which ε on the right-hand side is replaced with 4ε. Thus, it remains to establish (2.26) for fixed y and f.

In view of (2.25), there is $f_\varepsilon \in A_n(y)F_n$ such that $\|f - f_\varepsilon\|_H \le \varepsilon$. Since the norm of $D(y,\delta)$ does not exceed 1, we obtain

$$\|D(y,\delta)f - f\|_H \le 2\varepsilon + \|D(y,\delta)f_\varepsilon - f_\varepsilon\|_H. \tag{2.27}$$

Using the spectral theorem for self-adjoint operators, it is straightforward to check that $\|D(y,\delta)f_\varepsilon - f_\varepsilon\|_H \to 0$ as $\delta \to 0^+$; see the proof of Lemma 2.7 in [KNS20]. Combining this with (2.27), we arrive at (2.26). This completes the proof of the lemma.

3 Stationary processes satisfying the mixing hypothesis

In this section, we discuss Hypotheses (DLP′) and (SRZ) and construct stationary processes for which they are satisfied. This type of questions were intensively studied in the middle of last century; see the papers [Dob68, Dob70, Dob74] and the references therein. Here we prove that the existence and uniqueness (in law) of a stationary process with prescribed conditional laws follows from Theorem 1.3. We also discuss the relation of properties (DLP′) and (SRZ) with various concepts of mixing for stationary processes.

3.1 Existence and uniqueness

Let E be a separable Banach space, $\mathcal{K} \subset E$ be a compact subset, and $\boldsymbol{\mathcal{K}} = \mathcal{K}^{\mathbb{Z}_-}$ be the direct product of countably many copies of $\mathcal{K}$ with the distance d as in (0.5). We begin with the following result due to Dobrushin [Dob70].

Proposition 3.1. *Let $\{Q(\boldsymbol{\xi};\cdot), \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}\} \subset \mathcal{P}(E)$ be a transition probability from $\boldsymbol{\mathcal{K}}$ to E for which $\operatorname{supp} Q(\boldsymbol{\xi};\cdot) \subset \mathcal{K}$ for any $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ and Hypotheses (Fel) and (DLP′) hold with $\mathcal{E} = \boldsymbol{\mathcal{K}}$, and let $\boldsymbol{Q}_k(\boldsymbol{\xi};\cdot) \in \mathcal{P}(E^k)$ be defined by (2.6) for each $k \ge 1$. Then there is an E-valued stationary random process $\{\eta_l\}_{l\in\mathbb{Z}}$ such that, for every $k \ge 1$, the conditional law of $(\eta_1,\dots,\eta_k)$ given the past $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ is equal to $\boldsymbol{Q}_k(\boldsymbol{\xi};\cdot)$ for $\boldsymbol{\xi} \in \operatorname{supp}\sigma$, where σ is defined by (1.9).*

This proposition follows from Theorem 1 in [Dob70, Section 2] (applied to Case 2 discussed there). We do not give further details, referring the reader to the original paper and noting that the law of the constructed process may not be unique, and that Hypothesis (SRZ) is not necessarily satisfied for it.

The following result (whose proof does not use Proposition 3.1) shows that the uniqueness holds, provided that the conditional measures, in addition to (Fel) and (DLP′), satisfy a *strong recurrence* property, which is slightly more general than (SRZ); cf. Proposition 2.8. Moreover, the result and its proof remain true in the non-stationary case, as is mentioned in Remark 0.1 (2).

Theorem 3.2. *In addition to the conditions of Proposition 3.1, let us assume that* $\{Q(\boldsymbol{\xi};\cdot), \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}\}$ *possesses the following property:*

(SR) Strong recurrence. *For any integer* $n \geq 1$ *and any* $\delta > 0$, *there is a vector* $\vec{\xi}_n = (\bar{\xi}_1, \dots, \bar{\xi}_n) \in E^n$ *and an integer* $s \geq 0$ *such that*

$$\inf_{\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}} Q_s^n\big(\boldsymbol{\xi}; \mathcal{O}_\delta(\vec{\xi}_n)\big) > 0, \tag{3.1}$$

where the measures $Q_s^n\big(\boldsymbol{\xi};\cdot\big)$ *are constructed from the transition probability* $Q\big(\boldsymbol{\xi};\cdot\big)$ *using* (2.6) *and* (2.7).

Then there is an E*-valued stationary random process* $\{\eta_k\}_{k\in\mathbb{Z}}$ *such that the conditional measure of* η_1 *given the past* $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ *is equal to* $Q(\boldsymbol{\xi};\cdot)$ *for any* $\boldsymbol{\xi} \in \mathcal{E} := \operatorname{supp}\sigma$, *where* $\sigma = \mathcal{D}(\{\eta_k, k \in \mathbb{Z}_-\})$. *Moreover, the law of the process* $\{\eta_k\}$ *is uniquely defined, and there are positive numbers* C *and* γ *such that, for any* $m, k \geq 0$,

$$\sup_{\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}} \big\| Q_k^m(\boldsymbol{\xi};\cdot) - \boldsymbol{\sigma}_m \big\|_L^* \leq C\, e^{-\gamma k}, \tag{3.2}$$

where $\boldsymbol{\sigma}_m \in \mathcal{P}(E^m)$ *is the law of* $(\eta_1, \dots, \eta_m)$.

Proof. Step 1: Proof of (3.2). Let us set $\mathfrak{X} = \boldsymbol{\mathcal{K}}$ and denote by $\{\zeta^{\boldsymbol{\xi}}, \boldsymbol{\xi} \in \mathfrak{X}\}$ a random field on $\mathcal{K}$ such that $\mathcal{D}(\zeta^{\boldsymbol{\xi}}) = Q(\boldsymbol{\xi};\cdot)$ for any $\boldsymbol{\xi} \in \mathfrak{X}$. Consider the following RDS on $\mathfrak{X}$, defined on some suitable probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$:

$$\boldsymbol{\xi}_k = \big(\boldsymbol{\xi}_{k-1}, \zeta_k^{\boldsymbol{\xi}_{k-1}}\big), \quad k \geq 1. \tag{3.3}$$

Here $\zeta_k^{\boldsymbol{\xi}}(\widehat{\omega}) = \zeta^{\boldsymbol{\xi}}(\omega_k)$ for any $\boldsymbol{\xi} \in \mathfrak{X}$ and $\widehat{\omega} \in \widehat{\Omega}$; cf. (1.13). We note that (3.3) can be written in the form (1.3), where the mapping $\mathscr{S} : \mathfrak{X} \times \mathcal{K} \to \mathfrak{X}$ is given by $\mathscr{S}(\boldsymbol{\xi}, \eta) = (\boldsymbol{\xi}, \eta)$. By Proposition 1.1, it defines a Markov process in $\mathfrak{X}$. It follows from (3.3) that

$$P_1(\boldsymbol{\xi};\cdot) = \delta_{\boldsymbol{\xi}} \otimes Q(\boldsymbol{\xi};\cdot) \in \mathcal{P}(\boldsymbol{\mathcal{K}}) \otimes \mathcal{P}(\mathcal{K}) \simeq \mathcal{P}(\mathfrak{X}), \tag{3.4}$$

so that

$$P_l(\boldsymbol{\xi};\cdot) = \delta_{\boldsymbol{\xi}} \otimes \boldsymbol{Q}_l(\boldsymbol{\xi};\cdot) \in \mathcal{P}(\boldsymbol{\mathcal{K}}) \otimes \mathcal{P}(\mathcal{K}^l) \simeq \mathcal{P}(\mathfrak{X}) \tag{3.5}$$

for $l \geq 1$. In particular, if we write $\boldsymbol{\xi}_k$ as $(\xi_k^i, i \leq k)$, then taking $l = k + m$ in (3.5) and projecting to the last m components, we see that

$$\mathcal{D}\big(\{\xi_k^i, k+1 \leq i \leq k+m\}\big) = Q_k^m(\boldsymbol{\xi};\cdot). \tag{3.6}$$

We claim that the RDS (3.3) satisfies all the conditions of Theorem 1.3. Indeed, Hypothesis (GCP) follows from (SR) and an argument used in Step 2

of Section 2.2, with $\mathbf{0}_n$ replaced by $\vec{\xi}_n = (\bar{\xi}_1, \dots, \bar{\xi}_n)$. Furthermore, Hypothesis (LAC) with $F = \{0\}$ is trivial since we can take $\Phi \equiv 0$, and (DLP) follows from (DLP′). Thus, the conclusion of Theorem 1.3 holds for (3.3), so that (3.3) has a unique stationary measure $\sigma \in \mathcal{P}(\mathfrak{X})$, and there are positive numbers C and γ such that

$$\sup_{\boldsymbol{\xi} \in \mathfrak{X}} \|\mathcal{D}(\boldsymbol{\xi}_k) - \sigma\|_L^* \le Ce^{-\gamma k}, \quad k \ge 0. \tag{3.7}$$

Denoting by $\boldsymbol{\sigma}_m \in \mathcal{P}(\mathcal{K}^m)$ the projection of σ to the last m components and using (3.6), we obtain (3.2) as a direct consequence of (3.7).

Step 2: Stationarity and construction of the process. We now prove that $\sigma \in \mathcal{P}(\mathfrak{X})$ is shift invariant. To this end, it suffices to show that

$$\boldsymbol{\sigma}_{m+1}(\Gamma \times \mathcal{K}) = \boldsymbol{\sigma}_m(\Gamma) \quad \text{for any } m \ge 1 \text{ and } \Gamma \in \mathcal{P}(\mathcal{K}^m). \tag{3.8}$$

The definition of Q_k^m (see (2.7)) implies that

$$Q_k^{m+1}(\boldsymbol{\xi}; \Gamma \times \mathcal{K}) = Q_k^m(\boldsymbol{\xi}; \Gamma) \quad \text{for any } \Gamma \in \mathcal{B}(\mathcal{K}^m).$$

Passing to the (weak) limit as $k \to \infty$ in this relation (regarded as equality of measures) and using (3.2), we arrive at (3.8).

Thus, σ is shift invariant, so that, by the Kolmogorov theorem, it can be extended to a shift invariant measure on $\mathcal{K}^{\mathbb{Z}}$, which is denoted by $\boldsymbol{\sigma}$. We now define a stationary sequence $\{\eta_k, k \in \mathbb{Z}\}$ as the canonical process corresponding to $\boldsymbol{\sigma}$. In the next step, we show that the conditional measures of $\{\eta_k\}$ are given by $Q(\boldsymbol{\xi}; \cdot)$, establishing thus the existence claim of the theorem.[12]

Step 3: Conditional measures. Let us consider a trajectory of (3.3) issued from a random initial condition $\boldsymbol{\xi}_0$ that depends only on ω_0 and is distributed as $\sigma \in \mathcal{P}(\mathfrak{X})$. In view of the stationarity of σ, the Markov property, and relation (3.4), for any bounded continuous function $F : \mathfrak{X} \times \mathcal{K} \to \mathbb{R}$, we have

$$\begin{aligned}\int_{\mathfrak{X}} F(\boldsymbol{\xi})\sigma(\mathrm{d}\boldsymbol{\xi}) &= \mathbb{E}F(\boldsymbol{\xi}_1) = \mathbb{E}\big(\mathbb{E}\{F(\boldsymbol{\xi}_0, \zeta_1^{\boldsymbol{\xi}_0})|\, \mathcal{F}_0\}\big) \\ &= \mathbb{E}\int_{\mathcal{K}} F(\boldsymbol{\xi}_0, \xi) P_1(\boldsymbol{\xi}_0; \mathrm{d}\xi) = \int_{\mathfrak{X}} \sigma(\mathrm{d}\boldsymbol{\xi}) \int_{\mathcal{K}} F(\boldsymbol{\xi}, \xi) Q(\boldsymbol{\xi}; \mathrm{d}\xi).\end{aligned}$$

We see that $Q(\boldsymbol{\xi}; \cdot)$ is indeed the conditional measure of η_1, given the past $\boldsymbol{\eta} = \boldsymbol{\xi}$.

Step 4: The uniqueness. Suppose there is another $\mathcal{K}$-valued stationary process $\{\eta'_k\}_{k \in \mathbb{Z}}$ whose conditional law is given by $Q(\boldsymbol{\xi}; \cdot)$ for $\boldsymbol{\xi} \in \mathcal{E}'$, where $\mathcal{E}' \subset \boldsymbol{\mathcal{K}}$ is the support of the law of $\{\eta'_k, k \in \mathbb{Z}_-\}$, and denote by $\boldsymbol{\sigma}' \in \mathcal{P}(\mathcal{K}^{\mathbb{Z}})$ its law. We wish to prove that $\boldsymbol{\sigma}' = \boldsymbol{\sigma}$.

Given an integer $m \ge 1$, we denote by $\boldsymbol{\sigma}'_m \in \mathcal{P}(\mathcal{K}^m)$ the law of the random vector $(\eta'_1, \dots, \eta'_m)$. For any integers $k, m \ge 1$ and any function $f \in C_b(E^m)$, we can write

$$\int_{\mathcal{K}^m} f(\boldsymbol{y})\boldsymbol{\sigma}'_m(\mathrm{d}\boldsymbol{y}) = \int_{\boldsymbol{\mathcal{K}}} \int_{\mathcal{K}^m} f(\boldsymbol{y}) Q_k^m(\boldsymbol{\xi}; \mathrm{d}\boldsymbol{y})\sigma'(\mathrm{d}\boldsymbol{\xi}),$$

[12] By Proposition 3.1, the existence is valid under weaker hypotheses. Our proof does not use Dobrushin's result and is given to make the presentation self-contained.

where $\sigma' \in \mathcal{P}(\boldsymbol{\mathcal{K}})$ stands for the projection of $\boldsymbol{\sigma}'$ to $\boldsymbol{\mathcal{K}}$. Passing to the limit as $k \to \infty$ and using (3.2), we obtain $\langle f, \boldsymbol{\sigma}'_m \rangle = \langle f, \boldsymbol{\sigma}_m \rangle$. Since this is true for any $f \in C_b(E^m)$, we see that $\boldsymbol{\sigma}'_m = \boldsymbol{\sigma}_m$ for all $m \geq 1$, whence we conclude that $\boldsymbol{\sigma} = \boldsymbol{\sigma}'$. The proof of the theorem is complete. □

Corollary 3.3. *Let $Q(\boldsymbol{\xi}; \cdot)$ be a transition probability from $\boldsymbol{\mathcal{K}}$ to $\mathcal{K}$ that satisfies Hypotheses (Fel), (DLP′), and (SRZ). Then there is a unique in distribution stationary random process $\{\eta_k\}_{k\in\mathbb{Z}}$ whose conditional measure coincides with $Q(\boldsymbol{\xi}; \cdot)$ for $\boldsymbol{\xi} \in \operatorname{supp} \sigma$, where σ is given by (1.9).*

Since it is easy to construct numerous examples of transition probabilities $Q(\boldsymbol{\xi}; \cdot)$ from $\boldsymbol{\mathcal{K}}$ to $\mathcal{K}$ which meet assumptions (Fel), (DLP′) and (SRZ), then there are plenty of stationary processes $\{\eta_k\}$ with compact support that satisfy the hypotheses imposed on the random input in Theorem 2.4.

3.2 Strong recurrence, strong irreducibility, and mixing

The aim of this section is to prove implications (2.9). We shall say that an E-valued stationary process $\boldsymbol{\eta} := \{\eta_k\}_{k\in\mathbb{Z}}$ (or $\boldsymbol{\eta} := \{\eta_k\}_{k\in\mathbb{Z}_+}$) is *feebly mixing* if for any integers $n, m \geq 1$ and any functions $f \in C_b(E^n)$ and $g \in C_b(E^m)$ we have

$$\lim_{k\to\infty} \mathbb{E}\big(f(\boldsymbol{\eta}) g(\theta_k \boldsymbol{\eta})\big) = \mathbb{E}\, f(\boldsymbol{\eta})\, \mathbb{E}\, g(\boldsymbol{\eta}), \tag{3.9}$$

where $\theta_k \boldsymbol{\eta} = \{\eta_{j+k}, j \in \mathbb{Z}\}$ is the shifted process, and for a function $h : E^l \to \mathbb{R}$ we set $h(\boldsymbol{\eta}) = h(\eta_1, \dots, \eta_l)$. We denote by $\boldsymbol{\sigma}_m$ and σ the laws of $(\eta_1, \dots, \eta_m)$ and $\{\eta_k, k \in \mathbb{Z}_-\}$, respectively.

Proposition 3.4. *Let $\boldsymbol{\eta} = \{\eta_k\}_{k\in\mathbb{Z}}$ be an E-valued stationary process whose conditional probabilities satisfy Hypotheses (Fel), (DLP′), and (SR). Then $\boldsymbol{\eta}$ is feebly mixing.*

Proof. The main idea of the proof is well known, so that we skip the details. Let $f \in C_b(E^n)$ and $g \in C_b(E^m)$. For any $k > n$ and $\boldsymbol{\xi} \in \mathcal{E}$, the conditional expectation given the past can be written as

$$\mathbb{E}\big\{g(\theta_k \boldsymbol{\eta}) \,\big|\, \boldsymbol{\eta}^n = \boldsymbol{\xi}\big\} = \int_{E^m} Q^m_{k-n}(\boldsymbol{\xi}; \mathrm{d}\vec{y}_m) g(\vec{y}_m),$$

where $\vec{y}_m = (y_1, \dots, y_m)$ and $\boldsymbol{\eta}^n = \{\eta_j, j \leq n\}$. Denoting by $\{\mathcal{F}_k\}_{k\in\mathbb{Z}}$ the natural filtration associated with $\boldsymbol{\eta}$ and taking the conditional expectation with respect to $\mathcal{F}_n$, we obtain

$$\begin{aligned} \mathbb{E}\big(f(\boldsymbol{\eta}) g(\theta_k \boldsymbol{\eta})\big) &= \mathbb{E}\big(f(\boldsymbol{\eta})\, \mathbb{E}\big\{g(\theta_k \boldsymbol{\eta}) \,\big|\, \mathcal{F}_n\big\}\big) \\ &= \mathbb{E}\bigg\{ f(\boldsymbol{\eta}) \int_{E^m} Q^m_{k-n}(\boldsymbol{\eta}^n, \mathrm{d}\vec{y}_m) g(\vec{y}_m) \bigg\}. \end{aligned} \tag{3.10}$$

In view of Theorem 3.2, the integral over E^m converges (uniformly in $\boldsymbol{\eta}^n \in \mathcal{E}$) as $k \to \infty$ to $\langle g, \boldsymbol{\sigma}_m \rangle = \mathbb{E}\, g(\boldsymbol{\eta})$. Passing to the limit in (3.10) as $k \to \infty$, we arrive at (3.9). □

Now let $\boldsymbol{\eta}$ be a stationary process whose conditional probabilities satisfy Hypothesis (Fel). We shall say that $\boldsymbol{\eta}$ is *strongly feebly mixing* if, for each $m \geq 1$,

$$\lim_{k\to\infty} \sup_{\boldsymbol{\xi},g} \left| \int_{E^m} Q_k^m(\boldsymbol{\xi}; \mathrm{d}\vec{y}_m) g(\vec{y}_m) - \mathbb{E}\, g(\boldsymbol{\eta}) \right| = 0, \tag{3.11}$$

where the supremum is taken over all points $\boldsymbol{\xi} \in \mathcal{E} = \operatorname{supp}\sigma$ and all the functions $g \in C_b(E^m)$ with $\|g\|_\infty \leq 1$. In this case, it follows from (3.10) that, for fixed m and n, the limit in (3.9) holds uniformly in continuous functions f and g whose absolute values are bounded by 1. Thus, when property (Fel) is satisfied, the concept of *strong feeble mixing* defined above is close to the usual concept of *strong mixing*, for which the validity of (3.11) is required uniformly in $m \geq 1$; see [IL71, Definition 17.2.2].

Let us note that, in view of Proposition 2.7, a stationary weak solution of (1.8) is strongly feebly mixing. The following result prove an irreducibility property for such processes.

Proposition 3.5. *Any strongly feebly mixing E-valued stationary process $\boldsymbol{\eta}$ satisfying Hypothesis (DLP′) is strongly irreducible in the following sense:*

(SI) *For any $n \in \mathbb{N}$ and $\delta > 0$ there is an integer $s \geq 1$ such that*

$$\inf_{\boldsymbol{\xi}\in\operatorname{supp}\sigma} Q_s^n\big(\boldsymbol{\xi}; \mathcal{O}_\delta(\vec{\xi}_n)\big) > 0$$

for any $\vec{\xi}_n = (\xi_1, \dots, \xi_n) \in \operatorname{supp}\boldsymbol{\sigma}_n$, where $\sigma \in \mathcal{P}(\mathcal{E})$ is defined in (1.9).

Proof. Let us fix an integer $n \geq 1$, a point $\vec{\xi}_n \in \operatorname{supp}\boldsymbol{\sigma}_n$, and any number $\delta > 0$. We take any function $g \in C_b(E^n)$ supported by $\mathcal{O}_\delta(\vec{\xi}_n)$ such that $0 \leq g \leq 1$ and $g(\vec{y}_n) = 1$ for $\vec{y}_n \in \mathcal{O}_{\delta/2}(\vec{\xi}_n)$. In this case, we have

$$p := \mathbb{E}\, g(\boldsymbol{\eta}) > 0, \quad Q_s^n\big(\boldsymbol{\xi}; \mathcal{O}_\delta(\vec{\xi}_n)\big) \geq \int_{E^n} Q_s^n(\boldsymbol{\xi}; \mathrm{d}\vec{y}_n) g(\vec{y}_n).$$

Combining this with (3.11), we see that

$$\inf_{\boldsymbol{\xi}\in\operatorname{supp}\sigma} Q_s^n\big(\boldsymbol{\xi}; \mathcal{O}_\delta(\vec{\xi}_n)\big) \geq p/2,$$

provided that $s \geq 1$ is sufficiently large. This completes the proof. □

3.3 Examples of processes satisfying the hypotheses of Section 2.1

Corollary 3.3 implies the existence of many stationary processes satisfying (Fel), (DLP′), and (SRZ). In Subsections 3.3.1 and 3.3.2 below, we describe two classes of stationary processes possessing those properties.

3.3.1 Discrete-time Markovian noises

Let E be a finite-dimensional space and let $\boldsymbol{\eta} = \{\eta_k\}_{k\geq 0}$ be a stationary Markov process in E corresponding to transition probabilites $\{P_k(\xi;\cdot), k \geq 0\}$ that possess the following properties:

- *the law of η_1 has a compact support, and there is a Lipschitz continuous function $\rho : E\times E \to \mathbb{R}_+$ such that $P_1(\xi;\mathrm{d}y) = \rho(\xi,y)\ell(\mathrm{d}y)$ for any $\xi \in E$;*
- *there is a ball $B \subset E$, a point $a \in E$, an integer $m \geq 1$, and a number $p > 0$ such that $\rho(\xi,a) > 0$ for any $\xi \in B$, and $P_m(\xi;B) \geq p$ for any $\xi \in E$.*

In this case, Hypothesis (DLP′) is satisfied for $\boldsymbol{\eta}$. Moreover, using Theorem 3.2, it is not difficult to prove that (3.11) holds with $\boldsymbol{\xi} \in \mathcal{E}$, so that $\boldsymbol{\eta}$ is strongly feebly mixing. Then, by Proposition 3.5, the process $\boldsymbol{\eta}$ satisfies the property of strong irreducibility (SI). Moreover, if the second condition mentioned above is fulfilled for the point $a = 0$ and some ball $B = B_E(\delta)$ with $\delta > 0$, then a simple calculation based on the Kolmogorov–Chapman relation shows that Hypothesis (SRZ) is also satisfied.

3.3.2 Discrete-time moving averages

Let $\boldsymbol{\zeta} = \{\zeta_k\}_{k\in\mathbb{Z}}$ be a sequence of i.i.d. random variables in finite-dimensional space E whose law possesses a Lipschitz continuous density $\rho(y)$. We assume that the support of ρ is compact. Let us fix any sequence $\{a_l\}_{l\geq 1}$ exponentially going to zero and define

$$\eta_k = \zeta_k + \sum_{l=1}^{\infty} a_l \zeta_{k-l}, \quad k \in \mathbb{Z}. \tag{3.12}$$

Obviously $\boldsymbol{\eta} = \{\eta_k\}$ is a stationary process in E. To simplify the presentation, we further assume that $\sum_l |a_l| < 1$. In this case, denoting by $\ell_\infty(E)$ the space of bounded E-valued sequences indexed by $\mathbb{Z}_-$, we see that the linear operator $A : \ell_\infty(E) \to \ell_\infty(E)$ taking $\{\zeta_j, j \in \mathbb{Z}_-\}$ to the sequence $\{\eta_k, k \in \mathbb{Z}_-\}$ defined by (3.12) is invertible, and the inverse operator B has the form

$$(B\boldsymbol{\xi})_k = \xi_k + \sum_{l=1}^{\infty} b_l \xi_{k-l}, \quad k \in \mathbb{Z}_-,$$

where $\{b_l, l \in \mathbb{N}\}$ is another exponentially decaying sequence of real numbers. It follows that the conditional law of $\boldsymbol{\eta}$ can be written as

$$Q(\boldsymbol{\xi};\mathrm{d}y) = \rho(y - h(\boldsymbol{\xi})), \quad h(\boldsymbol{\xi}) = \sum_{l=1}^{\infty} a_l (B\boldsymbol{\xi})_{1-l} = -\sum_{l=1}^{\infty} b_l \xi_{1-l},$$

where $\boldsymbol{\xi} = (\xi_l, l \in \mathbb{Z}_-) \in \mathcal{E}$. Since ρ is Lipschitz continuous, it follows that $Q(\boldsymbol{\xi};\mathrm{d}y)$ has a Lipschitz continuous density, and so $\{\eta_k\}$ satisfies (DLP′). Furthermore, since an i.i.d. random process satisfies (SI), using the fact that the

processes $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$ can be reconstructed from each other with the help of the continuous linear operators A and B, it is not difficult to conclude that (SI) holds also for $\boldsymbol{\eta}$. If, in addition, the support of the law of ζ_k contains $0 \in E$, then (3.12) readily implies that Hypothesis (SRZ) is satisfied.

3.3.3 Continuous-time noise

In this section, we denote by J the interval $[0,1) \subset \mathbb{R}$ and by E the Lebesgue space $L^2(J)$. Let $\{\varphi_l\}_{l\geq 1}$ be an orthonormal basis in E and $\{\zeta_l\}_{l\geq 1}$ be a sequence of independent scalar random variables whose laws possess Lipschitz continuous densities $d_l : \mathbb{R} \to \mathbb{R}_+$ with respect to Lebesgue measure on $\mathbb{R}$. We assume that $|\zeta_l| \leq a_l$ almost surely for any $l \geq 1$, where the numbers $a_l > 0$ are such that $\sum_l a_l^2 < \infty$. Thus, the function d_l vanishes outside the interval $[-a_l, a_l]$, and the random series

$$\zeta(t) := \sum_{l=1}^{\infty} \zeta_l \varphi_l(t), \quad t \in J$$

converges almost surely in E. We denote by $\nu \in \mathcal{P}(E)$ the law of the random variable ζ. Let us write $\mathcal{K}$ for the support of ν. This is a compact subset of E. Next, we fix any number $\iota > 1$ and define a distance $\mathrm{d}(\boldsymbol{\xi}, \boldsymbol{\xi}')$ on $\boldsymbol{\mathcal{K}} = \mathcal{K}^{\mathbb{Z}_-}$ by relation (0.5).

Proposition 3.6. *Let the above hypotheses be fulfilled, let $g : \boldsymbol{\mathcal{K}} \times E \to \mathbb{R}_+$ be a Lipschitz continuous function such that $g(\boldsymbol{\xi}, y) \geq c > 0$ for all $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ and $y \in E$, and let $m(\boldsymbol{\xi}) = \int_E g(\boldsymbol{\xi}, y)\nu(\mathrm{d}y)$. Then the transition probabilities*

$$Q(\boldsymbol{\xi}; \mathrm{d}y) = \rho(\boldsymbol{\xi}, y)\nu(\mathrm{d}y), \quad \rho(\boldsymbol{\xi}, y) = m(\boldsymbol{\xi})^{-1} g(\boldsymbol{\xi}, y), \tag{3.13}$$

satisfy Hypothesis (DLP′), in which F_n is the vector span of $\varphi_1, \dots, \varphi_n$, and $F_n^\dagger$ is $F_n^\perp$ – the orthogonal complement to F_n in E. Moreover, the densities of the conditional measures $Q_n(\boldsymbol{\xi}, y_n^\dagger; \cdot)$ entering (2.5) are given by the formula

$$\rho_n(\boldsymbol{\xi}, y_n^\dagger, y_n) = \frac{g(\boldsymbol{\xi}, y_n, y_n^\dagger) D_n(y_n)}{\int_{F_n} g(\boldsymbol{\xi}, z, y_n^\dagger)\nu_n(\mathrm{d}z)}, \tag{3.14}$$

where ν_n is the projection of ν to F_n, and D_n stands for the product of the functions $d_1, \dots, d_n$.

Proof. In view of (3.13), for any integer $n \geq 1$ and arbitrary bounded continuous functions $f : F_n \to \mathbb{R}$ and $g : F_n^\perp \to \mathbb{R}$, we can write

$$\begin{aligned} &\int_E f(y_n) g(y_n^\dagger) Q(\boldsymbol{\xi}; \mathrm{d}y) \\ &\qquad = \int_{F_n^\perp} g(y_n^\dagger) \biggl\{ \int_{F_n} f(y_n) \rho(\boldsymbol{\xi}, y_n, y_n^\dagger) D_n(y_n) \ell_n(\mathrm{d}y_n) \biggr\} \nu_n^\perp(\mathrm{d}y_n^\dagger), \end{aligned} \tag{3.15}$$

where $\nu_n^\perp$ is the projection of ν to $F_n^\perp$, and ℓ_n stands for the Lebesgue measure on F_n. On the other hand, the projection of $Q(\boldsymbol{\xi};\cdot)$ to $F_n^\perp$ is given by

$$Q_n^\dagger(\boldsymbol{\xi};\mathrm{d}y_n^\dagger) = \biggl\{\int_{F_n} \rho(\boldsymbol{\xi},z,y_n^\dagger)\nu_n(\mathrm{d}z)\biggr\}\nu_n^\perp(\mathrm{d}y_n^\dagger).$$

Combining this with (3.15), we arrive at relation (3.14) for the density of the measure $Q(\boldsymbol{\xi};y_n^\dagger,\cdot)$. The Lipschitz continuity of ρ_n follows from similar property of the function g and the inequality $g \ge c$. Finally, the density of the union $\cup_n F_n$ is a consequence of the construction. □

Corollary 3.7. *In addition to the conditions of Proposition 3.6, let us assume that $0 \in \operatorname{supp} d_l$ for any $l \ge 1$. Then Hypothesis (SRZ) holds for the family $\{Q(\boldsymbol{\xi};\cdot), \boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}\}$. Moreover, for any $m \ge 1$ and $\delta > 0$, inequality (2.24) holds with $s = 0$.*

Proof. Since the support of d_l contains zero, we see that $Q(\boldsymbol{\xi};B_E(0,\delta)) > 0$ for any $\boldsymbol{\xi} \in E$. Combining this with (2.6), for any $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ and $m \ge 1$, we derive

$$\boldsymbol{Q}_m\bigl(\boldsymbol{\xi};\mathcal{O}_\delta(\mathbf{0}_m)\bigr) > 0.$$

By the portmanteau theorem, the function taking $\boldsymbol{\xi} \in \boldsymbol{\mathcal{K}}$ to the left-hand side of this relation is lower semicontinuous and, hence, separated from zero on any compact set. This implies the required inequality (3.1). □

We now construct a continuous-time process whose restrictions to the intervals $J_k := [k-1,k)$, $k \in \mathbb{Z}$ satisfy Hypotheses (DLP′) and (SRZ). To this end, we first use Corollary 3.3 to construct an E-valued stationary process $\{\eta_k\}_{k\in\mathbb{Z}}$ such that the conditional law of η_1 given the past $\{\eta_k = \xi_k, k \in \mathbb{Z}_-\}$ is equal to $Q(\boldsymbol{\xi};\cdot)$ on the support of the law of $\{\eta_k\}_{k\in\mathbb{Z}_-}$. Then we define a process $\eta = \{\eta(t)\}_{t\in\mathbb{R}}$ by the relation

$$\eta(t) = \eta_k(t-k) \quad \text{for } k \le t < k+1,\ k \in \mathbb{Z}.$$

The following result follows immediately from the construction.

Proposition 3.8. *Almost every trajectory of η is a bounded function of time, and the following two properties hold:*

Periodicity. The law of η regarded as a probability measure on the separable Fréchet space $L^2_{\mathrm{loc}}(\mathbb{R})$ is invariant under the time-1 shift.

Regularity and recurrence. The restrictions of η to the intervals $\{J_k\}_{k\in\mathbb{Z}}$ form a stationary process in E that satisfy Hypotheses (DLP′) and (SRZ).

Finally, we construct a continuous-time process that takes values in an arbitrary separable Hilbert space H, has a dense image, and satisfies Hypotheses (DLP′) and (SRZ). To this end, we denote by $\{e_j\}$ an orthonormal basis

in H, choose independent identically distributed processes $\{\eta^j(t), t \in \mathbb{R}\}$ satisfying the conclusions of Proposition 3.8, and consider the process

$$\eta(t) = \sum_{j=1}^{\infty} b_j \eta^j(t) e_j, \quad t \in \mathbb{R}, \tag{3.16}$$

where $\{b_j\}$ are non-zero numbers satisfying the inequality $\sum_j b_j^2 < \infty$. The following result is a consequence of Proposition 3.8 and the observation that a conditional measure of a tensor product is the tensor product of conditional measures.

Proposition 3.9. *Almost every trajectory of the process $\eta(t)$ constructed above is a bounded function of time with range in H. Moreover, the two properties formulated in Proposition 3.8 remain valid if we replace $L^2_{\mathrm{loc}}(\mathbb{R})$ with $L^2_{\mathrm{loc}}(\mathbb{R}; H)$ and $E = L^2(J)$ with $L^2(J; H)$.*

4 Application to randomly forced ODE and PDE

Theorem 2.4 directly applies to systems (0.1). In this section, we give three examples which illustrate its applications to systems (0.2).

4.1 Chain of anharmonic oscillators coupled to heat reservoirs

Our first example deals with an application of the theorem to a chain of anharmonic oscillators coupled to heat reservoirs. Following [JP98, EPR99], we fix a smooth Hamiltonian of the form (0.11) for a "small system" and, after excluding the variables describing the reservoirs, write the equations of motion for the coupled system in the form[13]

$$\dot q = p, \quad \dot p = -\nabla V(q) + (-\gamma_1 p_1 + \zeta_1(t))e_1 + (-\gamma_n p_n + \zeta_n(t))e_n, \tag{4.1}$$

where $\{e_j, 1 \le j \le n\}$ is the standard basis in $\mathbb{R}^n$, γ_1, γ_n are positive numbers, and ζ_1, ζ_n are stochastic processes with locally square-integrable trajectories. Thus, the dissipation $\gamma_i p_i$ and stochastic force ζ_i enter only the equations for p_1 and p_n. We shall prove that the question of mixing for (4.1) can be reduced to a pure control problem for a family of linear ODEs with variable coefficients. The latter can be studied with the help of well-known methods of control theory and will be investigated in a subsequent publication (see also Example 4.3).

In what follows, we assume that V is twice continuously differentiable and denote by $v(p, q)$ the vector field entering (4.1):

$$v(p, q) = (p, -\nabla_q V(q) - \gamma_1 p_1 e_1 - \gamma_n p_n e_n).$$

[13] We confine ourselves to a "linear" chain coupled to reservoirs at the endpoints. It is not difficult to extend our analysis to more complicated geometries, similar to those studied in [CEHR18, Raq19].

A simple calculation shows that if $(p(t), q(t))$ is a trajectory for (4.1), then

$$\frac{\mathrm{d}}{\mathrm{d}t}H\big(p,q\big) = -\gamma_1 p_1^2 - \gamma_n p_n^2 + \big(\zeta_1 p_1 + \zeta_n p_n\big) \le C\big(\zeta_1^2 + \zeta_n^2\big), \tag{4.2}$$

where $H(p,q) = \sum_j p_j^2/2 + V(q)$. Since ζ_1 and ζ_n are locally square-integrable, we see that $H(p(t), q(t))$ does not explode in finite time. Therefore, assuming that

$$V(q) \to +\infty \quad \text{as} \quad |q| \to +\infty,$$

we can conclude that all the solutions are defined for all $t \in \mathbb{R}$. We shall denote by $\Phi^t(\cdot;\zeta) : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ the map taking an initial condition (p^0, q^0) to the value of the corresponding solution of (4.1) at time $t \ge 0$, and by $\overline{B}(r)$ the closed ball in $\mathbb{R}^{2n}$ of radius r centred at zero. Let us impose, in addition, the following two hypotheses on V.

(BAS) Bounded absorbing set. *For any $M > 0$, there is $\rho > 0$ such that, for every $r > 0$, we can find $T_r > 0$ satisfying the following property: if a measurable function $\zeta = (\zeta_1, \zeta_n) : \mathbb{R}_+ \to \mathbb{R}^2$ is such that*

$$\sup_{t\ge 0} \int_t^{t+1} \big(|\zeta_1(s)|^2 + |\zeta_n(s)|^2\big)\,\mathrm{d}s \le M, \tag{4.3}$$

then for each $(p^0, q^0) \in \overline{B}(r)$ we have $\Phi^t(p^0, q^0; \zeta) \in \overline{B}(\rho)$ if $t \ge T_r$.

This property is certainly satisfied if we impose the stronger condition (0.12) on the functions ζ_1, ζ_n and require the vector field $v(p,q)$ to possesses a Lyapunov function for which (0.13) holds. As will be shown in Proposition 4.1 below, Hypothesis (BAS) implies the existence of a compact set $X \subset \mathbb{R}^{2n}$ that absorbs the trajectories of (4.1) in finite time and is invariant under the flow with an arbitrary ζ satisfying (4.3).

(SE) Stable equilibrium point. There is $c \in \mathbb{R}^n$ such that

$$\sup_{(p^0,q^0)\in X} \big|\Phi^t(p^0, q^0; 0) - (0, c)\big| \to 0 \quad \text{as} \quad t \to \infty. \tag{4.4}$$

Since all trajectories of system (4.1) with $\zeta \equiv 0$ are absorbed by X in finite time, this condition implies that $(0, c)$ is a unique globally stable equilibrium for the flow $\Phi^t(\cdot, 0)$. A sufficient condition under which (BAS) and (SE) are fulfilled will be discussed in Example 4.3. The following result describes some further properties of trajectories for (4.1) with a deterministic driving force ζ.

Proposition 4.1. *Let us assume that Hypotheses (BAS) and (SE) are fulfilled, and deterministic measurable functions ζ_1, ζ_n in (4.1) satisfy (4.3). Then the following properties hold.*

Invariant set. There is a compact set $X \subset \mathbb{R}^{2n}$ depending only on M such that, for any $(p^0, q^0) \in X$, the trajectory of (4.1) issued from (p^0, q^0) satisfies the inclusion

$$\big(p(t), q(t)\big) \in X \quad \textit{for any } t \ge 0.$$

Absorption. For any $r > 0$, there is an integer $T_r > 0$ such that the trajectory of (4.1) issued from an initial condition $(p^0, q^0) \in \overline{B}(r)$ belongs to X for $t \geq T_r$.

Proof. This result is established using a well-known argument, and we only outline it. Let us recall that $\rho > 0$ is the radius of the absorbing ball defined in (BAS). We denote $T \geq 1$ the minimal time such that $\Phi^t(p^0, q^0; \zeta) \in \overline{B}(\rho)$ for any $t \geq T$, $(p^0, q^0) \in \overline{B}(\rho)$, and any pair $\zeta = (\zeta_1, \zeta_n)$ satisfying (4.3). Let E_M be the set of functions $\zeta \in L^2(0, T; \mathbb{R}^2)$ for which inequality (4.3) holds with the supremum taken over $t \in [0, T-1]$. We endow E_M with the weak topology and consider the map $(t, p^0, q^0, \zeta) \mapsto \Phi^t(p^0, q^0; \zeta)$ acting from the compact set $[0, T] \times \overline{B}(\rho) \times E_M$ to the space $\mathbb{R}^{2n}$. Standard theory of ODEs implies that this is a continuous map, so that the set

$$X := \bigcup_{t=0}^{T} \bigcup_{\zeta \in E_M} \Phi^t\big(\overline{B}(\rho); \zeta\big)$$

is also compact. Moreover, it contains the absorbing ball $\overline{B}(\rho)$, and the very definition of X easily implies that X is invariant under the flow $\Phi^t(\cdot, \zeta)$ for any choice of ζ satisfying (4.3). □

To formulate the main result of this subsection, we define $E = L^2(J, \mathbb{R}^2)$, where $J = [0, 1)$, and impose the following two hypotheses on the random process $\zeta = (\zeta_1, \zeta_n)$ (whose trajectories are assumed to be elements of $L^2_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^2)$ almost surely).

(CP) Compactness and periodicity. *The law of ζ is invariant under the time-1 shift, and the support of its projection to J is a compact subset $\mathcal{K} \subset E$.*

(DR) Decomposability and recurrence. *The stationary process $\{\eta_k\}$ in E defined by $\eta_k(t) = \zeta(t + k - 1)$ for $t \in J$ satisfies Hypothesis (SRZ), in which $\mathcal{E} \subset \boldsymbol{\mathcal{K}} := \mathcal{K}^{\mathbb{Z}_-}$ is the support of the law of $\{\eta_l, l \in \mathbb{Z}_-\}$. Moreover, the conditional law $Q(\boldsymbol{\xi}; \cdot)$ of η_1, given the past $\{\eta_l = \xi_l, l \in \mathbb{Z}_-\}$, is such that (DLP′) holds.*

We also need a condition on the linearisation of Eqs. (4.1) about its solutions $(p(t), q(t))$,

$$\dot{x} = y, \quad \dot{y} = -(D^2 V)(q(t))x + (-\gamma_1 y_1 + \xi_1(t))e_1 + (-\gamma_n y_n + \xi_n(t))e_n, \quad (4.5)$$

supplemented with the zero initial condition:

$$x(0) = 0, \quad y(0) = 0. \quad (4.6)$$

Here $(p(t), q(t))$ is a solution issued from the set X, which was constructed in Proposition 4.1. We thus obtain a family of linear ODEs, parametrised by the initial condition (p^0, q^0) and the (deterministic) right-hand side ζ entering (4.1). Writing $\xi = (\xi_1, \xi_n)$, we impose the following hypothesis.

(LC) Linearised controllability. *There is an open set $O \subset \mathbb{R}^{2n} \times E$ containing $X \times \mathcal{K}$ such that, for any $((p^0, q^0), \zeta) \in O$, the associated linear system (4.5), (4.6) in $\mathbb{R}^{2n}$ is controllable at time $t = 1$ with control $\xi \in E$.*

Before formulating the main result of this subsection, let us mention that Hypotheses (CP) and (DR) are satisfied, for example, for random forces described in Subsection 3.3.3 (see Corollary 3.7 and Proposition 3.8), whereas the controllability property (LC) requires a separate study. A simple system for which the latter holds is given in Example 4.3.

Theorem 4.2. *Let Hypotheses (BAS), (SE), (CP), (DR), and (LC) be fulfilled. Then the random flow of (4.1) restricted to integer times is exponentially mixing in the dual-Lipschitz metric. More precisely, there is a measure $\mu \in \mathcal{P}(\mathbb{R}^{2n})$ with compact support and a number $\gamma > 0$ such that, for any $r > 0$ and a sufficiently large $C_r > 0$, we have*

$$\sup_{(p^0,q^0)\in\overline{B}(r)} \|\mu_t(p^0, q^0) - \mu\|_L^* \le C_r e^{-\gamma t}, \quad t \in \mathbb{Z}_+, \tag{4.7}$$

where we denote by $\mu_t(p^0, q^0)$ the law at time t of the trajectory of (4.1), issued from $(p^0, q^0) \in \mathbb{R}^{2n}$.

Proof. We claim that Theorem 2.4 is applicable to the random flow of (4.1), restricted to integer times and the invariant subset X. Indeed, (GD) is a consequence of (4.4), and (ALC) and (SRZ) are postulated in (LC) and (DR), respectively. The latter condition also implies that the properties of Hypothesis (DLP′) are satisfied for $F_n = E$. Thus, all the conditions of Theorem 2.4 holds, so that we have inequality (4.7), in which $\overline{B}(r)$ is replaced with X. Finally, the validity of (4.7) for an arbitrary $r > 0$ follows from the fact that X absorbs all the trajectories issued from $\overline{B}(r)$ at some positive time T_r; see the second point of Proposition 4.1. This completes the proof of the theorem. □

Theorem 4.2 ensures the convergence of the restrictions of trajectories to integer times. It is not difficult to prove that, for any $s \in (0, 1)$, the restrictions of trajectories to the times $s+\mathbb{Z}_+$ also converge to a limiting measure $\mu_s \in \mathcal{P}(\mathbb{R}^{2n})$. Since the corresponding argument is well known and is exactly the same as in the case of the Navier–Stokes system, we discuss it briefly in the next subsection. We conclude with an example for which Hypotheses (BAS) (CL), and (SE) are satisfied.

Example 4.3. Let us assume that the potential V entering (0.11) is representable in the form

$$V(q) = \sum_{i=1}^{n} \bigl(a_i q_i^2 + F_i(q_i)\bigr) + \sum_{i=1}^{n-1} b_i (q_i - q_{i+1})^2, \quad q \in \mathbb{R}^n, \tag{4.8}$$

where a_i and b_i are positive numbers, and $F_i \in C^2(\mathbb{R})$ are some functions with bounded derivatives. We claim that (BAS) holds. To see this, we denote

by $V_0(q)$ the sum of the quadratic terms in (4.8) (that is, the expression in the right-hand side of (4.8) with $F_i \equiv 0$) and write $H_0(p,q)$ for the corresponding Hamiltonian. The damped Hamiltonian system associated with H_0 has the form

$$\dot q = p, \quad \dot p = -Dq - \gamma_1 p_1 e_1 - \gamma_n p_n e_n, \tag{4.9}$$

where $D > 0$ is the symmetric matrix corresponding to the quadratic form V_0. Using the inequalities $\gamma_1, \gamma_n > 0$, it is easy to check that $(p,q) = (0,0)$ is a stable equilibrium for (4.9); see the argument below applied in the verification of Hypothesis (LC). Therefore there is a positive-definite quadratic form $L_0(p,q)$ such that

$$\langle \nabla L_0(p,q), v_0(p,q)\rangle \le -\delta(|p|^2 + |q|^2), \quad (p,q) \in \mathbb{R}^{2n},$$

where $\delta > 0$ is a number, and v_0 is the linear vector field in (4.9). Since the difference $V - V_0$ has bounded first-order partial derivatives, it follows that

$$\langle \nabla L_0(p,q), v(p,q)\rangle \le -\frac{\delta}{2}(|p|^2 + |q|^2)$$

outside a sufficiently large ball. This readily implies the existence of an absorbing ball and of an absorbing invariant set $X \subset \mathbb{R}^{2n}$.

To ensure the validity of (SE) and (LC), we further assume that

$$F_i'(0) = 0, \quad \inf_{s\in\mathbb{R}} F_i''(s) > -2a_i \quad \text{for any } 1 \le i \le n. \tag{4.10}$$

To check (SE), note that if (p,q) is a trajectory of (4.1) with $\zeta_1 \equiv \zeta_n \equiv 0$,

$$\dot q = p, \quad \dot p = -\nabla V(q) - \gamma_1 p_1 e_1 - \gamma_n p_n e_n, \tag{4.11}$$

then $H(p(t),q(t))$ is a non-decreasing function of time; see (4.2). It follows that the Hamiltonian H must be constant on the ω-limit set of a solution. In particular, for any trajectory (p,q) of (4.11) lying on such an ω-limit set, we have $p_1 \equiv p_n \equiv 0$. We claim that $p \equiv q \equiv 0$ is the only such trajectory. This will imply the validity of (SE) by a simple compactness argument.

The relation $\dot q_1 = p_1$ (which is the first equation in (4.11)) implies that q_1 is constant in time. The equation

$$0 \equiv \dot p_1 = -\partial_{q_1} V(q) = 2b_1 q_2 - 2(a_1 + b_1)q_1 - F_1'(q_1)$$

implies that q_2 is also constant in time, so that $p_2 \equiv 0$. Arguing by induction, we see that all p_i vanish identically and all q_i are constant. It remains to prove that $q_i = 0$. The second equation in (4.11) imply that

$$2a_i q_i + 2b_{i-1}(q_i - q_{i-1}) + 2b_i(q_i - q_{i+1}) + F_i'(q_i) = 0, \quad 1 \le i \le n,$$

where $b_0 = q_0 = b_n = q_{n+1} = 0$. A simple analysis based on (4.10) shows that $q_1 = \cdots = q_n = 0$ is the only solution of this system.

Finally, the verification of the linearised controllability (LC) is based on well-known techniques in the control theory of ODEs, using the propagation of perturbation via the derivative of the potential V. Since the corresponding argument is technically more involved, we shall give it in a subsequent publication devoted to a study of more general chains of oscillators and rotators.

Thus, Theorem 4.2 applies to the chain of anharmonic oscillators (4.1) with a potential of the form (4.8), provided that the random forces ζ_1, ζ_n are independent and belong to the class of processes described in Subsection 3.3.3.

4.2 Two-dimensional Navier–Stokes system

Let us consider the Navier–Stokes system (0.14) in a bounded domain $D \subset \mathbb{R}^2$ with a smooth boundary ∂D. Let us impose the no-slip boundary condition (0.15), define the usual functional space

$$H = \{u \in L^2(D, \mathbb{R}^2) : \operatorname{div} u = 0 \text{ in } D,\ \langle u, \boldsymbol{n}\rangle = 0 \text{ on } \partial D\}, \tag{4.12}$$

and denote by Π the orthogonal projection in $L^2(D, \mathbb{R}^2)$ onto H. Excluding the pressure from (0.14) with the help of the standard argument, we write the problem in question as a nonlocal PDE in H,

$$\partial_t u + \nu L u + B(u) = \eta(t), \quad u(t) \in H, \tag{4.13}$$

supplemented with the initial condition (0.16). Here $L = -\Pi\Delta$ is the Stokes operator, $B(u) = \Pi(\langle u, \nabla\rangle u)$ is a quadratic map, $u_0 \in H$ is an initial function, and η is a bounded stationary stochastic process in H of the form (3.16), in which $\{e_j\}$ is an orthonormal basis in H composed of the eigenfunctions of L, $b_j \neq 0$ are some numbers such that $\sum_j b_j^2 < \infty$, and $\eta^j = \{\eta^j(t), t \in \mathbb{R}\}_{j\geq 1}$ are i.i.d. stochastic processes possessing the properties stated in Proposition 3.8.

Let us denote $J = [0, 1)$ and introduce random functions $\eta_k^j : J \to \mathbb{R}$ by the formula $\eta_k^j(t) = \eta^j(t + k - 1)$. Writing

$$\boldsymbol{\eta}_k = \sum_{j=1}^{\infty} b_j \eta_k^j e_j,$$

we obtain a random variable in $E := L^2(J, H)$, and the construction implies that $\{\boldsymbol{\eta}_k\}_{k\geq 1}$ is a stationary sequence in E. Denoting by $\mathcal{K}$ the Hilbert cube $\{v \in H : |(v, e_j)| \leq b_j \text{ for } j \geq 1\}$, we see that $\boldsymbol{\eta}_k \in \mathcal{K}$ with probability 1.

For any initial condition $u_0 \in H$, Equation (4.13) has a unique solution $u(t; u_0)$ that belongs to the space $L^2_{\mathrm{loc}}(\mathbb{R}_+, V) \cap C(\mathbb{R}_+, H)$ almost surely and satisfies the initial condition (0.16). Let us denote by $\{u_k(v), k \in \mathbb{Z}_+\}$ the restrictions of $u(\cdot; v)$ to integer times. The following theorem describes the asymptotic behaviour of u_k as $k \to \infty$.

Theorem 4.4. *Under the above hypotheses, there is a measure $\mu \in \mathcal{P}(H)$ with compact support and a number $\gamma > 0$ such that, for any $R > 0$, we have*

$$\sup_{v \in B_H(R)} \bigl\|\mathcal{D}\bigl(u_k(v)\bigr) - \mu\bigr\|_L^* \leq C_R e^{-\gamma k}, \quad k \geq 0, \tag{4.14}$$

where $C_R > 0$ depends only on R.

Proof. Since the norms of trajectories of $\eta(t)$ are bounded by a universal number for all $t \in \mathbb{R}$, a standard argument shows that there is a compact absorbing set $X \subset H$ that is invariant under the random flow; cf. [BV92, Section 1.6]. Thus, it suffices to prove inequality (4.14) in which the ball $B_H(R)$ is replaced with X. To this end, we shall use Theorem 2.4. Namely, denoting by $S : H \times E \to H$ the time-1 shift along trajectories of Equation (4.13), we see that

$$u_k = S(u_{k-1}, \boldsymbol{\eta}_k), \quad k \geq 1. \tag{4.15}$$

Hence, it suffices to check the validity of Hypotheses (GD), (ALC), (DLP′), and (SRZ). The first of them follows from a well-known dissipation property of the Navier–Stokes system and is valid with $k = 1$. The validity of (DLP′) and (SRZ) is the content of Proposition 3.9. Finally, to check (ALC), it suffices to prove that the linearised Navier–Stokes system

$$\partial_t v + \nu L v + B(u, v) + B(v, u) = \boldsymbol{\zeta}(t), \quad v(0) = 0 \tag{4.16}$$

is approximately controllable at time 1. Here $u \in \mathcal{H} := L^2(J, V) \cap C(J, H)$ is a reference trajectory of (4.13) associated with an initial condition $u_0 \in X$ and a right-hand side $\eta \in \boldsymbol{\mathcal{K}}$. The required property follows from the observations that the control $\boldsymbol{\zeta}$ has a full range in H and that the reference trajectories u are taken from a bounded ball in $\mathcal{H}$. This completes the proof of the theorem. □

Theorem 4.4 proves the convergence of solutions for the Navier–Stokes system as the times goes to $+\infty$ along integer values. Let us show that, for any $s \in (0, 1)$, the measures $\mathcal{D}(u(s + k; v))$ exponentially converge to a limit μ_s in the dual-Lipschitz metric. Indeed, let $\{(v_k, \boldsymbol{\xi}_k), k \in \mathbb{Z}_+\}$ be a stationary weak solution of (4.15) (see Section 1.2). Then $\mathcal{D}(v_k) \equiv \mu$, and the process $\{\boldsymbol{\xi}_k, k \in \mathbb{Z}_+\}$ is distributed as $\{\boldsymbol{\eta}_k, k \in \mathbb{Z}_+\}$. Let $\xi(t)$ be the corresponding H-valued process (so that $\xi(t) = \boldsymbol{\xi}_k(t - k + 1)$ for $t \in [k - 1, k)$) and let $w(t)$, $t \geq 0$ be a solution of (4.13) with η replaced by ξ such that $w(0) = v_0$. Then $w(k) = v_k$ for $k \in \mathbb{Z}_+$. For $v \in H$, let $u^\xi(t; v)$ be a solution of (4.13) with $\eta \equiv \xi$ such that $u^\xi(0; v) = v$. Then $\mathcal{D}(u^\xi(t; v)) = \mathcal{D}(u(t; v))$, where, as before, $u(t; v)$ solves (4.13) and equals v at $t = 0$. Relation (4.14) combined with a simple result on preservation of the weak convergence of measures under Lipschitz maps (e.g., see [Shi21, Lemma 3.2]) implies that

$$\sup_{v \in B_H(R)} \big\| \mathcal{D}(u(t; v)) - \mathcal{D}(w(t)) \big\|_L^* \leq C'_R e^{-\gamma t}, \quad t \geq 0,$$

where $C'_R > 0$ depends only on R. Obviously, for $t = s + k$ with $k \in \mathbb{Z}_+$ and $0 < s < 1$, the measures $\mathcal{D}(w(s + k))$ do not depend on k. Denoting that measure by $\mu_s \in \mathcal{P}(H)$, we obtain the exponential convergence of $\mathcal{D}(u(s+k; v))$ to it as $k \to \infty$.

In conclusion, we note that Theorem 4.4 remains valid for a kick force of the form (0.17), provided that the stationary random process $\{\eta_k\}_{k \in \mathbb{Z}}$ satisfies appropriate non-degeneracy and decomposability conditions. The corresponding

result is similar to that discussed in the next subsection for the Ginzburg–Landau equation.

4.3 Complex Ginzburg–Landau equation

To simplify the presentation, we shall confine ourselves to the case of a kick force. However, the result remains valid for continuous-time forces of the form (3.16) under the same conditions as in Section 4.2.

Let $D \subset \mathbb{R}^d$, $1 \le d \le 4$ be a bounded domain with a smooth boundary. We consider the PDE

$$\partial_t u - (\nu + i)\Delta u + ic|u|^{2s}u = \eta(t,x), \quad x \in D, \tag{4.17}$$

supplemented with the boundeaty condition (0.15). Here ν and c are positive numbers, and $s \ge 1$ is an integer such that $s \le \frac{2}{d-2}$ for $d = 3$ or 4. In this case, problem (4.17), (0.15) with $\eta \equiv 0$ is well posed in the Sobolev space $\mathscr{H} := H^1_0(D, \mathbb{C})$, regarded as a real Hilbert space with the inner product

$$(u, v) = \operatorname{Re} \int_D \langle \nabla u, \nabla \bar{v} \rangle \mathrm{d}x.$$

The corresponding phase flow in $\mathscr{H}$ possesses a Lyapunov function,

$$\mathcal{H}(u) = \int_D \Big(\frac{1}{2}|\nabla u|^2 + \frac{c}{2s+2}|u|^{2s+2} \Big) \mathrm{d}x, \tag{4.18}$$

which satisfies the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(u(t)) \le -\nu\|\Delta u\|^2 - \nu c\big(|u|^{2s}, |\nabla u|^2\big), \tag{4.19}$$

where $(\cdot,\cdot)$ and $\|\cdot\|$ stand for the L^2 real inner product and the corresponding norm in $L^2(D;\mathbb{C})$. Concerning the random force η, we assume that it has the form (0.17), in which $\{e_j\}$ is the complete set of L^2-normalised eigenfunctions of the Dirichlet Laplacian $-\Delta$ corresponding to the eigenvalues α_j indexed in the increasing order, η_k^j are real-valued random variables such that $|\eta_k^j| \le 1$ with probability 1, and $\{b_j\}$ are positive numbers such that

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

These conditions ensure that $\{\eta_k\}$ are $\mathscr{H}$-valued random variables whose law is supported by a compact subset $\mathcal{K} \subset \mathscr{H}$. Thus, denoting by $S : \mathscr{H} \to \mathscr{H}$ the time-1 shift along trajectories of (4.17) with $\eta \equiv 0$ and writing $u_k = u(k)$, we obtain the relation

$$u_k = S(u_{k-1}) + \eta_k, \quad k \ge 1. \tag{4.20}$$

This is a particular case of the RDS (1.8), in which $H = E = \mathscr{H}$. The following theorem provides sufficient conditions under which the results of Section 2 can be applied to (4.20).

Theorem 4.5. *In addition to the above conditions, assume that* $\{\eta_k\}_{k\in\mathbb{Z}}$ *is a stationary random process that satisfies Hypotheses* (SRZ) *and* (DLP′) *with* $F_n = \operatorname{span}\{e_j, 1 \le j \le n\}$ *and* $F_n^\dagger = F_n^\perp$, *where the orthogonal complement is taken in* $\mathscr{H}$. *Then there is a measure* $\mu \in \mathcal{P}(\mathscr{H})$ *with compact support and a number* $\gamma > 0$ *such that, for any* $R > 0$, *inequality* (4.14) *holds with* $H = \mathscr{H}$ *and a constant* C_R *depending only on* R.

Proof. We apply Theorem 2.4. The validity of Hypotheses (DLP′) and (SRZ) are postulated in the statement of the theorem, and (GD) follows immediately from (4.19) and the fact that $\mathcal{H}(u)$ can be estimated from above and from below by a power of the H^1-norm. To check (ALC), note that the derivative of the operator $S(u,\eta) = S(u)+\eta$ with respect to $\eta \in \mathscr{H}$ is the identity operator in $\mathscr{H}$, so that its image is equal to $\mathscr{H}$. This completes the proof of the theorem. □

5 Appendix: the image of transition probabilities under Lipschitz maps

Let E be a separable Banach space represented as the direct sum of its closed subspaces F and $F^\dagger$ such that $\dim F < \infty$, and let $\mathcal{K} \subset E$ be a compact subset. We consider a map $\Psi : \mathcal{K} \to E$ of the form $\Psi(\eta) = \eta + \Phi(\eta)$, where $\Phi : \mathcal{K} \to E$ is a continuous map whose image is included in F, and for any $\eta, \eta' \in \mathcal{K}$, we have

$$\|\Phi(\eta)\|_E \le \varkappa, \quad \|\Phi(\eta) - \Phi(\eta')\|_E \le \varkappa \|\eta - \eta'\|_E, \tag{5.1}$$

where $\varkappa > 0$ is a number not depending on η and η'. Let $\mathscr{P}(U;\cdot)$ be a transition probability with an underlying compact metric space $\mathfrak{X}$.

Proposition 5.1. *Suppose that* $\mathscr{P}(U;\cdot)$ *satisfies Hypothesis* (DLP). *Then there is* $C > 0$ *not depending on* U *and* $\varkappa$ *such that*

$$\big\|\mathscr{P}(U;\cdot) - \Psi_*\big(\mathscr{P}(U;\cdot)\big)\big\|_{\mathrm{var}} \le C\varkappa \quad \text{for any } U \in \mathfrak{X}. \tag{5.2}$$

Proof. This result is essentially a simpler version of Theorem 2.4 in [KNS20], and we confine ourselves to outlining the main steps.

Since the total variation distance is bounded by 1, there is no loss of generality in assuming that $\varkappa \le \frac{1}{2}$. We need to prove that if f is the indicator function of a Borel set in E, then the expression $|\langle f\circ\Psi, \mathscr{P}(U;\cdot)\rangle - \langle f, \mathscr{P}(U;\cdot)\rangle|$ is bounded above by the right-hand side of (5.2) with a number C not depending on U and f.

To this end, we use (1.19) to write

$$\begin{aligned}
\langle f\circ\Psi, \mathscr{P}(U;\cdot)\rangle &= \int_{F^\dagger} \mathscr{P}_{F^\dagger}(U, \mathrm{d}\eta) \int_F f\big(\xi+\eta+\Phi(\xi+\eta)\big) p_F(U,\eta;\xi)\, \ell_F(\mathrm{d}\xi),\\
\langle f, \mathscr{P}(U;\cdot)\rangle &= \int_{F^\dagger} \mathscr{P}_{F^\dagger}(U, \mathrm{d}\eta) \int_F f(\xi+\eta) p_F(U,\eta;\xi)\, \ell_F(\mathrm{d}\xi).
\end{aligned}$$

Note that the integrations are carried out over the projections of $\mathcal{K}$ to $F^\dagger$ and F (which are compact subsets), since the supports of the measures $\mathscr{P}(U;\cdot)$ are contained in $\mathcal{K}$. The required estimate will be established if we prove that

$$\left|\int_F f(\xi+\eta+\Phi(\xi+\eta))p_F(U,\eta;\xi)\,\ell_F(\mathrm{d}\xi)-\int_F f(\xi+\eta)p_F(U,\eta;\xi)\,\ell_F(\mathrm{d}\xi)\right|\le C\varkappa. \tag{5.3}$$

To this end, we wish to make the change of variable $\xi'=\xi+\Phi(\xi+\eta)$ in the first integral. However, the map Φ is defined only on $\mathcal{K}$. To overcome this difficulty, we first use the Kirzsbraun–McShane theorem (see [Dud02, Theorem 6.1.1]) to extend Φ to the whole space E in such a way that the extended map (for which we keep the same notation) is also $\varkappa$-Lipschitz. In this case, for any $\eta\in F^\dagger$, the map $\xi\mapsto\xi+\Phi(\xi+\eta)$ is a bi-Lipschitz homeomorphism of the space F onto itself, and we denote by $\Theta_\eta(\cdot)$ its inverse. The latter and its differential (which exists almost everywhere in view of the Rademacher theorem; see [Fed69, Section 3.1.6]) can be written in the form

$$\Theta_\eta(\xi')=\xi'-\Phi(\Theta_\eta(\xi')+\eta),\quad D\Theta_\eta(\xi')=\bigl(\mathrm{Id}_F+D\Phi(\Theta_\eta(\xi')+\eta)\bigr)^{-1}. \tag{5.4}$$

Using Theorem 3.2.5 in [Fed69] to perform the change of variable $\xi=\Theta_\eta(\xi')$, we can write the first integral in (5.3) as

$$\int_F f(\xi+\eta+\Phi(\xi+\eta))p_F(U,\eta;\xi)=\int_F f(\xi'+\eta)\frac{p_F(U,\eta;\Theta_\eta(\xi'))}{\det\bigl(\mathrm{Id}_F+D\Phi(\Theta_\eta(\xi')+\eta)\bigr)},$$

where we dropped the Lebesgue measure ℓ_F to shorten the formula. Thus, the left-hand side $\delta(U)$ of (5.3) satisfies the inequality

$$\delta(U)\le\int_F\left|\frac{p_F(U,\eta;\Theta_\eta(\xi'))}{\det\bigl(\mathrm{Id}_F+D\Phi(\Theta_\eta(\xi')+\eta)\bigr)}-p_F(U,\eta;\xi)\right|\ell_F(\mathrm{d}\xi), \tag{5.5}$$

where we used the inequality $0\le f\le 1$. It follows from (5.4) and (5.1) that Θ_η is a 2-Lipschitz function such that $|\Theta_\eta(\xi')-\xi'|\le\varkappa$. Combining this with the Lipschitz property of p_F (see (1.20)), we can easily show that

$$\begin{aligned}p_F(U,\eta;\Theta_\eta(\xi))&=p_F(U,\eta;\xi)+q(U,\xi,\eta),\\ \det\bigl(\mathrm{Id}_F+D\Phi(\Theta_\eta(\xi)+\eta)\bigr)^{-1}&=1+r(\xi,\eta),\end{aligned}$$

where q and r are some functions satisfying the inequality

$$|q(U,\xi,\eta)|+|r(\xi,\eta)|\le C_1\varkappa$$

with a constant $C_1>0$ not depending on U, ξ, η, and $\varkappa$. Combining this with (5.5), we arrive at the required inequality (5.3). □

References

[BF92] A. A. Borovkov and S. G. Foss, *Stochastically recursive sequences and their generalizations*, Siberian Adv. Math. **2** (1992), no. 1, 16–81.

[BM20] V. I. Bogachev and I. I. Malofeev, *Kantorovich problems and conditional measures depending on a parameter*, J. Math. Anal. Appl. **486** (2020), no. 1, 123883, 30.

[Bor98] A. A. Borovkov, *Ergodicity and Stability of Stochastic Processes*, John Wiley & Sons, Ltd., Chichester, 1998.

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

[CE16] N. Cuneo and J.-P. Eckmann, *Non-equilibrium steady states for chains of four rotors*, Comm. Math. Phys. **345** (2016), no. 1, 185–221.

[CEHR18] N. Cuneo, J.-P. Eckmann, M. Hairer, and L. Rey-Bellet, *Non-equilibrium steady states for networks of oscillators*, Electron. J. Probab. **23** (2018), Paper No. 55, 28.

[DDL+07] J. Dedecker, P. Doukhan, G. Lang, J. R. León, S. Louhichi, and C. Prieur, *Weak Dependence: with Examples and Applications*, Springer, New York, 2007.

[DLS24] A. Dymov, L. Lokutsievskiy, and A. Sarychev, *Strict Lyapunov functions and energy decay in Hamiltonian chains with degenerate damping*, Proc. Steklov Inst. Math. **327** (2024), to appear.

[Dob68] R. L. Dobrušin, *Description of a random field by means of conditional probabilities and conditions for its regularity*, Teor. Verojatnost. i Primenen **13** (1968), 201–229.

[Dob70] R. L. Dobrušin, *Prescribing a system of random variables by conditional distributions*, Teor. Verojatnost. i Primenen. **15** (1970), no. 3, 458–486.

[Dob74] R. L. Dobrušin, *Conditions for the absence of phase transitions in one-dimensional classical systems*, Math. USSR-Sb. **22** (1974), no. 1, 28–48.

[Doe38] W. Doeblin, *Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états*, Rev. Math. Union Interbalkan **2** (1938), 77–105.

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

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

[EH00] J.-P. Eckmann and M. Hairer, *Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators*, Comm. Math. Phys. **212** (2000), no. 1, 105–164.

[EPR99] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures*, Comm. Math. Phys. **201** (1999), no. 3, 657–697.

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

[FP67] C. Foiaş and G. Prodi, *Sur le comportement global des solutions non-stationnaires des équations de Navier–Stokes en dimension* 2, Rend. Sem. Mat. Univ. Padova **39** (1967), 1–34.

[Hai05] M. Hairer, *Ergodicity of stochastic differential equations driven by fractional Brownian motion*, Ann. Probab. **33** (2005), no. 2, 703–758.

[HO07] M. Hairer and A. Ohashi, *Ergodic theory for SDEs with extrinsic memory*, Ann. Probab. **35** (2007), no. 5, 1950–1977.

[IL71] I. A. Ibragimov and Yu. V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff Publishing, Groningen, 1971.

[JP97] V. Jakšić and C.-A. Pillet, *Ergodic properties of the non-Markovian Langevin equation*, Lett. Math. Phys. **41** (1997), no. 1, 49–57.

[JP98] ———, *Ergodic properties of classical dissipative systems. I*, Acta Math. **181** (1998), no. 2, 245–282.

[Kha12] R. Khasminskii, *Stochastic Stability of Differential Equations*, Springer, Heidelberg, 2012.

[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.

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

[Kol37] A. N. Kolmogorov, *Markov chains with a countable state space*, Moscow Univ. Math. Bulletin **1** (1937), no. 3, 1–16 (in Russian).

[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.

[KS12] S. Kuksin and A. Shirikyan, *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, Cambridge, 2012.

[KS25] ———, *Markovian reduction and exponential mixing in total variation for random dynamical systems*, in preparation.

[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.

[Kuk06] ———, *Randomly Forced Nonlinear PDEs and Statistical Hydrodynamics in 2 Space Dimensions*, European Mathematical Society (EMS), Zürich, 2006.

[KZ20] S. Kuksin and H. Zhang, *Exponential mixing for dissipative PDEs with bounded non-degenerate noise*, Stochastic Process. Appl. **130** (2020), no. 8, 4721–4745.

[LM72] J.-L. Lions and E. Magenes, *Non-Nomogeneous Boundary Value Problems and Applications. Vol. I*, Springer-Verlag, New York, 1972.

[MT93] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, London, 1993.

[NR21] V. Nersesyan and R. Raquépas, *Exponential mixing under controllability conditions for SDEs driven by a degenerate Poisson noise*, Stochastic Process. Appl. **138** (2021), 26–55.

[Par05] K. R. Parthasarathy, *Introduction to Probability and Measure*, Hindustan Book Agency, New Delhi, 2005.

[Raq19] R. Raquépas, *A note on Harris' ergodic theorem, controllability and perturbations of harmonic networks*, Ann. Henri Poincaré **20** (2019), no. 2, 605–629.

[Sev57] B. A. Sevastjanov, *An ergodic theorem for Markov processes and its application to telephone systems with refusals*, Teor. Veroyatnost. i Primenen. **2** (1957), 106–116.

[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.

[Shi21] ———, *Controllability implies mixing II. Convergence in the dual-Lipschitz metric*, J. Eur. Math. Soc. **23** (2021), no. 4, 1381–1422.

[Ver17] A. Yu. Veretennikov, *On convergence rate for Erlang-Sevastyanov type models with infinitely many servers*, Theory Stoch. Process. **22** (2017), no. 1, 89–103.

[Vil09] C. Villani, *Optimal Transport*, Springer-Verlag, Berlin, 2009.