A coupling approach to randomly forced nonlinear PDE's. I

Sergei Kuksin   Armen Shirikyan

Abstract

We develop a coupling approach to prove that a randomly forced dissipative PDE has a unique stationary measure and to study ergodic properties of this measure.

Contents

0 Introduction

Let H be a separable Hilbert space with a norm $\|\cdot\|$ and an orthonormal basis $\{e_j\}$ and let $S\colon H\to H$ be a locally Lipschitz operator such that $S(0)=0$. It is assumed that S satisfies some additional conditions which, roughly speaking, mean that S is compact and that $S^n(u)\to 0$ as $n\to\infty$ uniformly on bounded subsets of H. (For the exact statement, see conditions (A) – (C) in Section 2.) Let η_k, $k\in\mathbb{Z}$, be a sequence of i.i.d. random variables in H of the form

$$\eta_k=\sum_{j=1}^{\infty}b_j\xi_{jk}e_j. \tag{0.1}$$

Here $b_j\ge 0$ are some constants such that $\sum b_j^2<\infty$ and $\{\xi_{jk}\}$ are independent random variables such that $\mathcal{D}(\xi_{jk})=p_j(r)\,dr$, where p_j, $j=1,2,\dots$, are functions of bounded variation supported by the interval $[-1,1]$. (We denote by $\mathcal{D}(\xi)$ the distribution of a random variable ξ.)

Our goal is to study the random dynamical system (RDS)

$$u^k = S(u^{k-1}) + \eta_k. \tag{0.2}$$

For any $v \in H$, we denote by $u^k = u^k(v)$, $k \geq 0$, the solution of (0.2) such that $u^0 = v$. Let $C_b(H)$ be the space of bounded continuous functions on H and $\mathcal{P}(H)$ be the set of probability Borel measures on H. The RDS (0.2) defines a family of Markov chains in H. We shall denote by $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ the corresponding Markov semigroups acting in $C_b(H)$ and $\mathcal{P}(H)$, respectively:

$$\mathfrak{P}_k f(v) = \mathbb{E} f\bigl(u^k(v)\bigr), \quad f \in C_b(H),$$
$$\mathfrak{P}_k \mu(\Gamma) = \int_H \mathbb{P}\bigl\{u^k(v) \in \Gamma\bigr\}\mu(dv), \quad \mu \in \mathcal{P}(H).$$

Let us recall that $\mu \in \mathcal{P}(H)$ is called a *stationary measure* for (0.2) if $\mathfrak{P}_1^*\mu = \mu$.

The goal of this paper is to present a new, simple proof of the uniqueness and ergodicity of a stationary measure and to specify the rate of convergence to it. Namely, we prove the following result:

Theorem 0.1. *There is an integer $N \geq 1$ such that if*

$$b_j \neq 0 \quad \textit{for} \quad 1 \leq j \leq N, \tag{0.3}$$

then (0.2) *has a unique stationary measure μ. Moreover, for any $R > 0$ there is $C_R > 0$ such that*

$$\bigl|\mathfrak{P}_k f(u) - (\mu, f)\bigr| \leq C_R e^{-c\sqrt{k}}\bigl(\sup_H |f| + \mathrm{Lip}(f)\bigr) \quad \textit{for} \quad k \geq 0, \tag{0.4}$$

where $\|u\| \leq R$, f is an arbitrary bounded Lipschitz function on H, and $c > 0$ is a constant not depending on u, f, R, and k.

Example 0.2. Let us consider the 2D Navier–Stokes (NS) equations perturbed by a random kick-force:

$$\begin{aligned} &\dot{u} - \nu\Delta u + (u, \nabla)u + \nabla p = \eta(t, x) \equiv \sum_{k=-\infty}^{\infty} \eta_k(x)\delta(t-k), \\ &\operatorname{div} u = 0, \quad \langle u(t, \cdot)\rangle = 0, \end{aligned} \tag{0.5}$$

where $u = u(t, x)$, $x \in \mathbb{T}^2$, and $\langle u\rangle = \int_{\mathbb{T}^2} u(x)\,dx$. Let H be the space of divergence-free vector fields $u \in L^2(\mathbb{T}^2, \mathbb{R}^2)$ such that $\langle u\rangle = 0$ and let $\{e_j\}$ be the normalised trigonometric basis in H. Assuming that the kicks $\eta_k \in H$ have the form (0.1) and normalising solutions $u(t)$ for (0.5) to be continuous from the right, we observe that (0.5) can be written in the form (0.2), where $u^k = u(k, \cdot) \in H$ and $S\colon H \to H$ is the time-one shift along trajectories of the free NS system (i.e., of Equations (0.5) with $\eta \equiv 0$). As it is shown in [KS1], the operator S satisfies all the required assumptions, and therefore Theorem 0.1 applies to (0.5).

Theorem 0.1 can also be applied to many other dissipative nonlinear PDE's perturbed by a random kick-force, in particular, to the complex Ginzburg–Landau equation

$$\dot{u} - \nu(\Delta - 1)u + i|u|^2 u = \eta(t, x), \quad x \in \mathbb{T}^n,$$

where $u = u(t, x)$ and $\nu > 0$ (see [KS1, KS2]).

Uniqueness of a stationary measure for (0.2) was first established[1] in [KS1]. The proof in [KS1] is based on a Lyapunov–Schmidt type reduction of the system (0.2) to an N-dimensional RDS with delay (the integer N is the same as in Theorem 0.1). Due to this reduction, the problem of uniqueness of a stationary measure for (0.2) reduces to a similar question for an abstract 1D Gibbs system with an N-dimensional phase space. The uniqueness for the reduced Gibbs system is then established using a version of the Ruelle–Perron–Frobenius theorem.

E, Mattingly, Sinai [EMS] and Bricmont, Kupiainen, Lefevere [BKL] used later similar approaches to show that the NS system (0.5) perturbed by a white (in time) force of the form

$$\eta(t, x) = \sum_{j=1}^{N'} b_j \dot{\beta}_j(t) e_j(x), \quad N' < \infty,$$

also has a unique stationary measure $\mu \in \mathcal{P}(H)$, provided that $b_j \neq 0$ for $1 \leq j \leq N \leq N'$ with some sufficiently large $N = N(\nu)$. Moreover, it is shown in [BKL] that for the case of white noise the convergence in (0.4) is exponentially fast for μ-almost all $u \in H$.

In [KS3] the NS equations (0.5) with an unbounded kick-force $\eta(t, x)$ is studied and the scheme of [KS1] is used to prove the uniqueness and ergodicity of a stationary measure.

The approach presented in this work does not use a Lyapunov–Schmidt type reduction and the Gibbs measure technique. Instead it exploits some ideas from [KS2], interpreting them in terms of the coupling. The new approach gives rise to a shorter proof and is more flexible.

The coupling is a well-known effective tool for studying finite-dimensional Markov chains (e.g., see [Lin] and the Appendix in [V]) and dynamical systems (e.g., see [Y, BL]). In [EMS] a coupling is used to study the auxiliary finite-dimensional RDS with delay which arises as a result of the Lyapunov–Schmidt reduction. Our work shows that a form of coupling applies directly to infinite-dimensional Markov chains and randomly forced PDE's.

When a preprint of this paper was sent around, we learned from L.-S. Young that a similar approach to prove Theorem 0.1 is developed by her and Nader Masmoudi in their work under preparation.

[1] It is shown in [KS1, KS2] that the left-hand side of (0.4) converges to zero as $k \to \infty$ for any $f \in C_b(H)$; however, the rate of convergence is not specified.

Notation

We abbreviate a pair of random variables ξ_1, ξ_2 or points u_1, u_2 to $\xi_{1,2}$ and $u_{1,2}$, respectively. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, for any integer $k \geq 1$ we denote by Ω^k the space $\Omega \times \cdots \times \Omega$ (k times) endowed with the σ-algebra $\mathcal{F} \times \cdots \times \mathcal{F}$ and the measure $\mathbb{P} \times \cdots \times \mathbb{P}$. For a random variable ξ, we denote by $\mathcal{D}(\xi)$ its distribution.

For a Banach space H, we shall use the following spaces and sets:

$C_b(H)$ is the space of bounded continuous functions on H with the supremum norm $\|\cdot\|_\infty$.

$L(H)$ is the space of bounded Lipschitz functions on H endowed with the natural norm $\|\cdot\|_L$ (see Section 1).

$\mathcal{M}(H)$ is the space of signed Borel measures on H with bounded variation.

$\mathcal{P}(H)$ is the set of probability measures $\mu \in \mathcal{M}(H)$; this space is endowed with two different metrics described in Section 1.

$\mathcal{P}(H, \mathcal{A})$ is the set of measures $\mu \in \mathcal{P}(H)$ with support in a closed set $\mathcal{A}$.

$\mu_v(k)$ is the measure $\mathfrak{P}(k, v, \cdot)$, where $\mathfrak{P}$ is the Markov transition function for (0.2).

$B_H(R)$ is the closed ball of radius $R > 0$ centred at zero.

Acknowledgements. The authors thank Roger Tribe and Sergei Foss for fruitful discussions of the coupling approach during the Symposium "Stochastic Fluid Equations" in Warwick on January 19–20, 2001, and at seminars in Heriot-Watt University, respectively. The authors are also grateful to Jan Kristensen for useful remarks on functional analysis. This research was supported by EPSRC, grant GR/N63055/01.

1 Measures on Hilbert spaces

Let H be a separable Hilbert space with the Borel σ-algebra $\mathcal{B}(H)$ and let $\mathcal{M}(H)$ be the space of signed Borel measures with bounded variation. We denote by $\mathcal{P}(H)$ the set of probability measures $\mu \in \mathcal{M}(H)$ and by $\mathcal{P}(H, \mathcal{A})$ the subset in $\mathcal{P}(H)$ consisting of measures supported by a closed set $\mathcal{A} \subset H$. For any measure $\mu \in \mathcal{M}(H)$ and any function $f \in C_b(H)$, we write

$$(\mu, f) = \int_H f(u)\, d\mu(u) = \int_H f(u)\mu(du).$$

We shall use two different topologies on $\mathcal{P}(H)$. The first of them is given by the variation norm on $\mathcal{M}(H)$:

$$\|\mu\|_{\text{var}} = \sup_{\Gamma \in \mathcal{B}(H)} |\mu(\Gamma)|.$$

The distance defined by this norm on $\mathcal{P}(H)$ can be characterised in terms of densities. Namely, let us assume that $\mu_1, \mu_2 \in \mathcal{P}(H)$ are absolutely continuous with respect to a fixed Borel measure m, finite or infinite. (Such a measure always exists; for instance, one can take $m = (\mu_1 + \mu_2)/2$.) In this case, we have

$$\|\mu_1 - \mu_2\|_{\text{var}} = \frac{1}{2}\int_H |p_1(u) - p_2(u)|\, dm(u), \tag{1.1}$$

where $p_i(u)$, $i = 1, 2$, is the density of μ_i with respect to m. The space $\mathcal{P}(H)$ is complete with respect to $\|\cdot\|_{\text{var}}$.

To define a second topology, we denote by $L(H)$ the space of real-valued bounded Lipschitz functions on H with the norm

$$\|f\|_L := \Bigl(\sup_{u\in H}|f(u)|\Bigr) \vee \Bigl(\sup_{u\neq v}\frac{|f(u)-f(v)|}{\|u-v\|}\Bigr).$$

Let $\|\cdot\|_L^*$ be the dual norm on $\mathcal{M}(H)$:

$$\|\mu\|_L^* = \sup_{\|f\|_L\le 1}\bigl|(\mu, f)\bigr|.$$

It is clear that the norm $\|\cdot\|_L^*$ defines a metric on $\mathcal{P}(H)$.

Lemma 1.1. *The space $\mathcal{P}(H)$ is complete with respect to the metric $\|\cdot\|_L^*$.*

Proof. Suppose that $\{\mu_n\} \subset \mathcal{P}(H)$ is a sequence such that $\|\mu_n - \mu_m\|_L^* \to 0$ as $m, n \to \infty$. Let $L^*(H)$ be the space of continuous functionals on $L(H)$. Regarding μ_n as elements of $L^*(H)$, we conclude that the sequence $\{\mu_n\}$ converges (in the norm $\|\cdot\|_L^*$) to a limit $\ell \in L^*(H)$, and we have

$$\ell(f) = \lim_{n\to\infty}(\mu_n, f), \quad f \in L(H). \tag{1.2}$$

In view of the corollary[2] from Theorem 1 in [GS, Chapter VI, §1], there is a measure $\mu \in \mathcal{P}(H)$ such that $\ell(f) = (\mu, f)$. This completes the proof. □

Note that, in the case when H is finite-dimensional, the fact that the functional ℓ in (1.2) is a measure is implied by the following well-known result (for instance, see [H, Theorem 2.1.7]): any nonnegative distribution is a measure; in particular, any positive functional $\ell \in L^*(H)$ is a measure as well.

Let $\mathfrak{P}(k, u, \Gamma)$, $k \ge 0$, $u \in H$, $\Gamma \in \mathcal{B}(H)$, be a Markov transition function. A set $\mathcal{A} \subset \mathcal{B}(H)$ is said be *invariant* for $\mathfrak{P}$ if

$$\mathfrak{P}(k, u, \mathcal{A}) = 1 \quad \text{for all} \quad k \ge 0, \quad u \in \mathcal{A}.$$

Lemma 1.2. *Let $\mathcal{A} \in \mathcal{B}(H)$ be an invariant set for $\mathfrak{P}(k, u, \Gamma)$. Suppose that there is $k_0 \ge 1$ and a sequence ζ_k, $k \ge k_0$, going to zero as $k \to \infty$ such that*

$$\|\mathfrak{P}(k, u, \cdot) - \mathfrak{P}(k, v, \cdot)\|_L^* \le \zeta_k \quad \textit{for} \quad k \ge k_0, \quad u, v \in \mathcal{A}. \tag{1.3}$$

Then there is a unique measure $\mu \in \mathcal{P}(H, \mathcal{A})$ such that

$$\|\mathfrak{P}(k, u, \cdot) - \mu\|_L^* \le \zeta_k \quad \textit{for} \quad k \ge k_0, \quad u \in \mathcal{A}. \tag{1.4}$$

[2] The corollary of Theorem 1 in [GS, Chapter VI, §1] claims, in fact, that if the limit in (1.2) exists for any $f \in C_b(H)$, then the functional ℓ can be represented in the form $\ell(f) = (\mu, f)$, where $\mu \in \mathcal{P}(H)$. However, the same proof works also in the case under study.

Proof. Let $f \in L(H)$, $\|f\|_L \le 1$. Then, by (1.3) and the Chapman–Kolmogorov relation, for $l \ge k \ge k_0$ and $u, v \in \mathcal{A}$ we have

$$\begin{aligned} &\big|\big(\mathfrak{P}(l,v,\cdot) - \mathfrak{P}(k,u,\cdot), f\big)\big| \le \\ &\quad \le \left| \int_H \mathfrak{P}(l-k,v,dz) \int_H \big(\mathfrak{P}(k,z,dw)f(w) - \mathfrak{P}(k,u,dw)f(w)\big) \right| \le \\ &\qquad \le \zeta_k \int_H \mathfrak{P}(l-k,v,dz) = \zeta_k. \end{aligned} \tag{1.5}$$

By Lemma 1.1, the space $\mathcal{P}(H)$ is complete with respect to $\|\cdot\|_L^*$. Hence, there is a unique measure $\mu \in \mathcal{P}(H)$ such that $\|\mathfrak{P}(l,v,\cdot) - \mu\|_L^* \to 0$ as $l \to \infty$. It is clear that $\operatorname{supp} \mu \subset \mathcal{A}$ and therefore $\mu \in \mathcal{P}(H, \mathcal{A})$. Passing to the limit in (1.5) as $l \to \infty$, we obtain (1.4). □

We now recall that a pair of random variables (ξ_1, ξ_2) defined on the same probability space is called a *coupling* for given measures $\mu_1, \mu_2 \in \mathcal{P}(H)$ if $\mathcal{D}(\xi_j) = \mu_j$, $j = 1, 2$. For some basic results on the coupling, see [Lin, V] and the Appendix (Section 4).

Lemma 1.3. *If measures $\mu_1, \mu_2 \in \mathcal{P}(H)$ admit a coupling (ξ_1, ξ_2) such that*

$$\mathbb{P}\big\{\|\xi_1 - \xi_2\| > \varepsilon\big\} \le \theta, \tag{1.6}$$

where $\varepsilon > 0$ and $\theta > 0$ are some constants, then

$$\|\mu_1 - \mu_2\|_L^* \le 2\theta + \varepsilon. \tag{1.7}$$

Proof. Let $f \in L(H)$, $\|f\|_L \le 1$. Then $(\mu_{1,2}, f) = \mathbb{E}\, f(\xi_{1,2})$ and, therefore,

$$|(\mu_1 - \mu_2, f)| \le \big|\mathbb{E}\chi_Q(f(\xi_1) - f(\xi_2))\big| + \big|\mathbb{E}\chi_{Q^c}(f(\xi_1) - f(\xi_2))\big|, \tag{1.8}$$

where χ_Q and χ_{Q^c} are characteristic functions of the event $\|\xi_1 - \xi_2\| > \varepsilon$ and of its complement, respectively. By (1.6), the first term in the right-hand side of (1.8) is bounded by 2θ, while the second does not exceed $\varepsilon\|f\|_L \le \varepsilon$. This completes the proof of (1.7). □

2 A class of random dynamical systems

Let H be a Hilbert space with a norm $\|\cdot\|$ and an orthonormal basis $\{e_j\}$ and let $S \colon H \to H$ be an operator satisfying conditions (A) – (C) below:

(A) *For any $R > r > 0$ there exist positive constants $a = a(R, r) < 1$ and $C = C(R)$ and an integer $n_0 = n_0(R, r) \ge 1$ such that*

$$\|S(u_1) - S(u_2)\| \le C(R)\|u_1 - u_2\| \quad \textit{for all} \quad u_1, u_2 \in B_H(R), \tag{2.1}$$

$$\|S^n(u)\| \le \max\{a\|u\|, r\} \quad \textit{for} \quad u \in B_H(R), \quad n \ge n_0. \tag{2.2}$$

Let η_k, $k \geq 1$, be a sequence of i.i.d. H-valued random variables that are defined on a probability space $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and have the form (0.1), where $b_j \geq 0$ are some constants such that

$$\sum_{j=1}^{\infty} b_j^2 < \infty, \tag{2.3}$$

and $\{\xi_{jk}\}$ is a family of independent real-valued random variables such that $|\xi_{jk}| \leq 1$ for all j, k, and $\omega_1 \in \Omega_1$. We consider the following RDS in H:

$$u^k = S(u^{k-1}) + \eta_k =: F^{\omega_1}(u^{k-1}), \quad k \geq 1. \tag{2.4}$$

It follows from (0.1) and (2.3) that the distribution of η_k is supported by the Hilbert cube K,

$$K = \left\{ u = \sum_{j=1}^{\infty} u_j e_j : |u_j| \leq b_j \text{ for all } j \geq 1 \right\}.$$

Therefore, if the initial state u^0 of the RDS (2.4) belongs to a set $B \subset H$, then $u^k \in \mathcal{A}_k(B)$ for all $k \geq 1$ and $\omega_1 \in \Omega_1$, where $\mathcal{A}_0(B) = B$ and

$$\mathcal{A}_k(B) = S\big(\mathcal{A}_{k-1}(B)\big) + K \quad \text{for} \quad k \geq 1.$$

The next condition expresses the property of existence of a bounded absorbing set for the system in question.

(B) *There exists $\rho > 0$ such that for any bounded set $B \subset H$ there is an integer $k_0 \geq 1$ such that $\mathcal{A}_k(B) \subset B_H(\rho)$ for $k \geq k_0$.*

Clearly, inequality (2.2) and condition (B) are satisfied if $\|S(u)\| \leq \gamma \|u\|$ for all $u \in H$ and some positive constant $\gamma < 1$.

To formulate the last condition, we introduce some notations. For a subspace $E \subset H$, we denote by $E^\perp$ its orthogonal complement in H. For an integer $N \geq 1$, let H_N be the finite-dimensional subspace generated by the vectors $e_1, \dots, e_N$ and let P_N and Q_N be the orthogonal projections onto H_N and $H_N^\perp$, respectively.

(C) *For any $R > 0$ there is a decreasing sequence $\gamma_N(R) > 0$ tending to zero as $N \to \infty$ such that*

$$\big\|\mathsf{Q}_N\big(S(u_1) - S(u_2)\big)\big\| \leq \gamma_N(R) \|u_1 - u_2\| \quad \textit{for all} \quad u_1, u_2 \in B_H(R).$$

Finally, we specify the random variables $\{\xi_{jk}\}$:

(D) *For any j, the random variables ξ_{jk}, $k \geq 1$, have the same distribution $\pi_j(dr) = p_j(r)\, dr$, where the densities $p_j(r)$ are functions of bounded variation such that $\operatorname{supp} p_j \subset [-1, 1]$ and $\int_{|r| \leq \varepsilon} p_j(r)\, dr > 0$ for all $j \geq 1$ and $\varepsilon > 0$. We normalise the functions p_j to be continuous from the right.*

The RDS (2.4) defines a family of Markov chains in H with the transition function

$$\mathfrak{P}(k,v,\Gamma)=\mathbb{P}\{u^k\in\Gamma\},$$

where $(u^k,k\ge 0)$ is the solution of (2.4) such that $u^0=v$. Let $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ be the corresponding semigroups (see the Introduction for their definition). Continuity of S (see condition (A)) and the Lebesgue theorem on dominated convergence imply that the transition function satisfies the Feller condition: if $f\in C_b(H)$, then $\mathfrak{P}_k f\in C_b(H)$ for all $k\ge 1$.

Let $\rho>0$ be the constant in condition (B). We introduce the set

$$\mathcal{A}=\overline{\bigcup_{k\ge 1}\mathcal{A}_k\bigl(B_H(\rho)\bigr)}. \tag{2.5}$$

It is clear that $\mathcal{A}$ is an invariant set for the RDS (2.4): if $u^0\in\mathcal{A}$, then $u^k\in\mathcal{A}$ for all $k\ge 1$ and $\omega_1\in\Omega_1$. Moreover, it follows from condition (C) that the set $\mathcal{A}$ is compact in H. (Note that the union in (2.5) is taken over $k\ge 1$ and therefore $B_H(\rho)$ is not a subset of $\mathcal{A}$.)

Our goal is to prove the following result:

Theorem 2.1. *There is an integer $N\ge 1$ such that if* (0.3) *holds, then the RDS* (2.4) *has a unique stationary measure $\mu\in\mathcal{P}(H,\mathcal{A})$. Moreover, for any $R>0$ there is $C_R>0$ such that*

$$\bigl|\mathfrak{P}_k f(u)-(\mu,f)\bigr|\le C_R e^{-c\sqrt{k}}\|f\|_L \quad \textit{for} \quad k\ge 0, \quad \|u\|\le R,$$

where $f\in L(H)$ is an arbitrary function and $c>0$ is a constant not depending on f, u, R, and k.

Condition (B) and the definition of $\mathcal{A}$ imply that for any $R>0$ there is an integer $l\ge 1$ depending on R such that $\mathfrak{P}(l,u,\mathcal{A})=1$ for any $u\in B_H(R)$. Hence, we can restrict our consideration to the invariant set $\mathcal{A}$. In view of Lemma 1.2, Theorem 2.1 will be established if we show that there are positive constants C and c and an integer $k_0\ge 1$ such that

$$\|\mathfrak{P}(k,u,\cdot)-\mathfrak{P}(k,v,\cdot)\|_L^*\le C\,e^{-c\sqrt{k}} \quad \text{for} \quad k\ge k_0, \quad u,v\in\mathcal{A}. \tag{2.6}$$

3 Proof of the main result

We first establish some auxiliary assertions and then use them to prove inequality (2.6), which implies the required result.

3.1 Auxiliary assertions

We begin with a simple observation. Let $R>0$ be so large that $B_H(R)\supset\mathcal{A}$. To simplify notation, we denote $B=B_H(R)$.

Lemma 3.1. *For any $d > 0$ there is an integer $l = l(d) \geq 0$ and a constant $\varkappa = \varkappa(d) > 0$ such that*

$$\mathbb{P}\big\{\|u^l(v)\| \leq d/2 \text{ for all } v \in B\big\} \geq \varkappa. \tag{3.1}$$

Proof. Let a and n_0 be the constants in condition (A) that correspond to the parameters R (the radius of B) and $r = d/4$ and let $l = n_0 m$, where m is the smallest integer such that $a^m R \leq d/4$. If $\eta_k = 0$ in (2.4) for $1 \leq k \leq l$, then, in view of (2.2), we have

$$\|u^l(v)\| \leq \max\{a^m R, d/4\} = d/4 \quad \text{for all} \quad v \in B.$$

By continuity, there is $\gamma > 0$ such that if

$$\|\eta_k\| \leq \gamma \quad \text{for} \quad 1 \leq k \leq l, \tag{3.2}$$

then

$$\|u^l(v)\| \leq d/2. \tag{3.3}$$

It follows from (2.3) and condition (D) that the event (3.2) has a positive probability $\varkappa$. Inequality (3.1) follows now from (3.3). $\square$

To simplify notation, for any $v \in H$ we denote by $\mu_v(k)$ the measure $\mathfrak{P}(k, v, \cdot) \in \mathcal{P}(H)$. For any measurable space $(X, \mathcal{B}(X))$ and any integer $k \geq 1$, we denote by X^k the direct product $X \times \cdots \times X$ endowed with the product σ-algebra $\mathcal{B}^k(X) = \mathcal{B}(X) \times \cdots \times \mathcal{B}(X)$.

Lemma 3.2. *There is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an integer $N \geq 1$, and a constant $C > 0$ such that if (0.3) holds, then for any $u_1, u_2 \in B$ the measures $\mu_{u_{1,2}}(1)$ admit a coupling $V_{1,2} = V_{1,2}(u_1, u_2; \omega)$ that possesses the following properties:*

(i) *The maps $V_{1,2}$ are measurable with respect to the σ-algebra $\mathcal{B}^2(H) \times \mathcal{F}$ as functions of $(u_1, u_2, \omega) \in B^2 \times \Omega$.*

(ii) *Let $d = \|u_1 - u_2\|$. Then*

$$\mathbb{P}\big\{\|V_1 - V_2\| \geq d/2\big\} \leq Cd. \tag{3.4}$$

Let us note that inequality (3.4) is nontrivial only in the case $Cd < 1$.

Proof. Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ be the probability space on which the random variables $\{\eta_k\}$ are defined and let $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be the probability space constructed in Theorem 4.2 for the measures $\nu_{1,2}$ specified below. We shall show that the set $\Omega = \Omega_1 \times \Omega_2$ endowed with the natural σ-algebra and probability of direct product is the required probability space.

The random variables $V_{1,2}$ are sought in the form

$$V_1 = S(u_1) + \xi_1, \quad V_2 = S(u_2) + \xi_2,$$

where $\xi_{1,2}$ are some random variables on Ω such that $\mathcal{D}(\xi_1) = \mathcal{D}(\xi_2) = \mathcal{D}(\eta_1)$. It is clear that $\mathcal{D}(V_{1,2}) = \mu_{u_{1,2}}(1)$ and that (i) holds. To define the random variables $\xi_{1,2}$, we specify their projections $\mathsf{P}_N\xi_{1,2}$ and $\mathsf{Q}_N\xi_{1,2}$, where $N \ge 1$ is a sufficiently large integer which is chosen below.

We set

$$\mathsf{Q}_N\xi_1 = \mathsf{Q}_N\xi_2 = \mathsf{Q}_N\tilde{\eta}_1,$$

where $\tilde{\eta}_1$ is the natural extension of η_1 to Ω, i.e., $\tilde{\eta}_1(\omega) = \eta_1(\omega_1)$ for $\omega = (\omega_1, \omega_2) \in \Omega$. To define $\mathsf{P}_N\xi_{1,2}$, let us write $\nu_{1,2} := \mathsf{P}_N\mu_{u_{1,2}}(1)$ and assume that we have proved the inequality

$$\|\nu_1 - \nu_2\|_{\mathrm{var}} \le Cd, \tag{3.5}$$

where $C > 0$ is a constant not depending on $u_{1,2} \in B$. In view of Theorem 4.2, there is a maximal coupling $\Xi_{1,2}(u_1, u_2; \omega_2)$ for the measures $\nu_{1,2}$ that is measurable with respect to $(u_1, u_2, \omega_2) \in B^2 \times \Omega_2$:

$$\mathbb{P}\{\Xi_1 \ne \Xi_2\} = \|\nu_1 - \nu_2\|_{\mathrm{var}} \le Cd. \tag{3.6}$$

Retaining the same notation for the natural extensions of Ξ_1 and Ξ_2 to Ω, we now set

$$\mathsf{P}_N\xi_{1,2} = \Xi_{1,2} - \mathsf{P}_N S(u_{1,2})$$

and note that $\mathsf{P}_N V_1 \ne \mathsf{P}_N V_2$ if and only if $\Xi_1 \ne \Xi_2$. Let $N \ge 1$ be so large that $\gamma_N(R) \le 1/2$ (see condition (C)). In this case, if $\mathsf{P}_N V_1 = \mathsf{P}_N V_2$, then

$$\|V_1 - V_2\| = \big\|\mathsf{Q}_N(V_1 - V_2)\big\| = \big\|\mathsf{Q}_N(S(u_1) - S(u_2))\big\| \le \|u_1 - u_2\|/2 \le d/2.$$

Inequality (3.4) follows now from (3.6).

Thus, it remains to establish (3.5). To this end, we set $v_{1,2} = \mathsf{P}_N S(u_{1,2})$ and note that, in view of (2.1),

$$\|v_1 - v_2\| \le C(R)d. \tag{3.7}$$

Since $b_j \ne 0$ for $1 \le j \le N$, condition (D) implies that $\mathcal{D}(\mathsf{P}_N\eta_1) = p(x)\,dx$, where dx is the Lebesgue measure on the finite-dimensional space H_N and

$$p(x) = \prod_{j=1}^{N} q_j(x_j), \quad q_j(x_j) = b_j^{-1} p_j(x_j/b_j), \quad x = (x_1, \dots, x_N) \in H_N,$$

is a bounded function with support in the set $\mathsf{P}_N K$. It follows that

$$\nu_{1,2} = \mathcal{D}(v_{1,2} + \mathsf{P}_N\eta_1) = p(x - v_{1,2})\,dx.$$

Therefore, by (1.1),

$$\|v_1 - v_2\|_{\mathrm{var}} = \frac{1}{2}\int_{H_N} |p(x - v_1) - p(x - v_2)|\,dx.$$

We claim that

$$\int_{H_N} |p(x-v_1)-p(x-v_2)|\,dx \le |v_1-v_2| \sum_{j=1}^N b_j^{-1}\operatorname{Var}(p_j), \tag{3.8}$$

where $\operatorname{Var}(p_j)$ stands for the total variation of p_j. The required inequality (3.5) follows immediately from (3.7) and (3.8).

To prove (3.8), we first assume that p_j are C^1-smooth functions. In this case, we have

$$\begin{aligned}
&\int_{H_N} |p(x-v_1)-p(x-v_2)|\,dx \\
&\le |v_1-v_2| \int_{H_N}\int_0^1 \big|(\nabla p)(x-\theta v_1-(1-\theta)v_2)\big|\,d\theta dx \\
&= |v_1-v_2| \int_{H_N} \big|(\nabla p)(x)\big|\,dx \le |v_1-v_2| \sum_{j=1}^N \int_{\mathbb{R}} \big|\partial_{x_j} q_j(x_j)\big|\,dx_j \\
&= |v_1-v_2| \sum_{j=1}^N \operatorname{Var}(q_j).
\end{aligned}$$

It remains to note that $\operatorname{Var}(q_j) = b_j^{-1}\operatorname{Var}(p_j)$.

Inequality (3.8) in the general case can be easily derived by a standard approximation procedure; we omit the corresponding arguments. □

We now combine Lemmas 3.1 and 3.2 to obtain a coupling $U_{1,2}^k(u_1,u_2)$ for the measures $\mu_{u_{1,2}}(k)$, $k \ge 1$. Let $l = l(d)$ and $C > 0$ be the constants in Lemmas 3.1 and 3.2 and let $d_0 > 0$ be so small that

$$Cd_0 \le 1/4.$$

We set $d_r = 2^{-r} d_0$, $r \ge 1$.

For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we shall denote by $(\Omega^k, \mathcal{F}^k, \mathbb{P}^k)$ the direct product of its k independent copies. Points of the latter will be denoted by $\boldsymbol{\omega}^k = (\omega_1, \dots, \omega_k)$.

Lemma 3.3. *Suppose that the conditions of Lemma 3.2 are satisfied. Let $u_1, u_2 \in \mathcal{A}$ and $d = \|u_1 - u_2\|$. Then for any $k \ge 1$ the measures $\mu_{u_{1,2}}(k)$ admit a coupling $U_{1,2}^k = U_{1,2}^k(u_1,u_2;\boldsymbol{\omega}^k)$, $\boldsymbol{\omega}^k \in \Omega^k$, such that the following assertions hold:*

(i) *The maps $U_{1,2}^k(u_1,u_2;\boldsymbol{\omega}^k)$ are measurable with respect to $(u_1,u_2,\boldsymbol{\omega}^k) \in \mathcal{A}^2 \times \Omega^k$.*

(ii) *There is a constant $\theta > 0$ not depending on u_1, u_2, and k such that*

$$\mathbb{P}^k\big\{\|U_1^k - U_2^k\| \le d_r\big\} \ge \theta \quad \textit{for all} \quad k \ge r + l(d_0), \quad u_1, u_2 \in \mathcal{A}. \tag{3.9}$$

(iii) *If* $\|u_1 - u_2\| \le d_r$, *then*

$$\mathbb{P}^k\big\{\|U_1^k - U_2^k\| \le d_{k+r}\big\} \ge 1 - 2^{-r-1} \quad \textit{for all} \quad k \ge 1, \quad r \ge 0. \tag{3.10}$$

Proof. Let us recall that for any $(u_1, u_2) \in B \times B$ a coupling $V_{1,2}(u_1, u_2; \omega)$ was constructed in Lemma 3.2. We set

$$U_j(u_1, u_2; \omega) = \begin{cases} V_j(u_1, u_2; \omega) & \text{if} \quad \|u_1 - u_2\| \le d_0, \\ F^\omega(u_j) & \text{if} \quad \|u_1 - u_2\| > d_0, \end{cases}$$

where $j = 1, 2$ and $F^\omega(u)$ is given by (2.4). We define random variables $U_{1,2}^k$ on $(\Omega^k, \mathcal{F}^k)$ by the following rule: if $\|u_1 - u_2\| > d_0$, then

$$U_j^k(u_1, u_2; \boldsymbol{\omega}^k) = F^{\omega_k} \circ \cdots \circ F^{\omega_1}(u_j)$$

for $k \le l(d_0)$ and

$$U_j^k(u_1, u_2; \boldsymbol{\omega}^k) = U_j\big(U_1^{k-1}(u_1, u_2; \boldsymbol{\omega}^{k-1}), U_2^{k-1}(u_1, u_2; \boldsymbol{\omega}^{k-1}); \omega_k\big) \tag{3.11}$$

for $k > l(d_0)$, where $\boldsymbol{\omega}^k = (\boldsymbol{\omega}^{k-1}, \omega_k) = (\omega_1, \dots \omega_k)$ and $U_j^0(u_1, u_2) = u_j$. If $\|u_1 - u_2\| \le d_0$, then $U_{1,2}^0(u_1, u_2) = u_{1,2}$ and for $k \ge 1$ the random variables $U_j^k(u_1, u_2; \boldsymbol{\omega}^k)$ are inductively defined by (3.11).

We claim that $U_{1,2}^k$ satisfy assertions (i) – (iii) of the lemma. Indeed, the measurability of the maps $U_{1,2}^k$ is obvious since they are compositions of measurable maps. To prove (3.9), we first note that it is sufficient to consider the case $k = l + r$, $l = l(d_0)$. We introduce the following events in Ω^{l+r}:

$$\begin{aligned} Q^+ &= \big\{\|U_1^l - U_2^l\| \le d_0\big\}, \\ Q^- &= \big\{\|U_1^l - U_2^l\| > d_0\big\}, \\ Q &= \big\{\|U_1^{l+r} - U_2^{l+r}\| \le d_r\big\}. \end{aligned}$$

By Lemma 3.1, we have

$$\mathbb{P}^k(Q) = \mathbb{P}^k(Q|Q^+)\mathbb{P}(Q^+) + \mathbb{P}^k(Q|Q^-)\mathbb{P}(Q^-) \ge \varkappa\,\mathbb{P}^k(Q|Q^+). \tag{3.12}$$

If we assume that (3.10) is proved for $r = 0$, then (3.12) will imply the required estimate (3.9) with $\theta = \varkappa/2$. Thus, it remains to established (iii).

For a fixed $r \ge 0$, we set

$$Q_k^+ = \big\{\|U_1^k - U_2^k\| \le d_{k+r}\big\}, \quad Q_k^- = \big\{\|U_1^k - U_2^k\| > d_{k+r}\big\}$$

and denote by p_k^+ and p_k^- the probabilities of Q_k^+ and Q_k^-, respectively. Using (3.4) with $d = d_{k+r-1}$, we derive

$$p_k^+ = p_{k-1}^+\mathbb{P}^k(Q_k^+|Q_{k-1}^+) + p_{k-1}^-\mathbb{P}^k(Q_k^+|Q_{k-1}^-) \ge (1 - Cd_{k+r-1})p_{k-1}^+.$$

Since $p_0^+ = 1$, iteration of this estimate results in

$$p_k^+ \ge \lambda := \prod_{j=0}^{k-1}(1 - Cd_{j+r}). \tag{3.13}$$

Since $d_m = 2^{-m} d_0$ and $Cd_0 \leq 1/4$, we have

$$\begin{aligned}\log \lambda = \sum_{j=0}^{k-1} \log(1 - Cd_{j+r}) &\geq -C \sum_{j=0}^{k-1} d_{j+r} \\ &\geq -Cd_0 \sum_{j=0}^{\infty} 2^{-(j+r)} = -2^{1-r} Cd_0 \geq -2^{-r-1}.\end{aligned}$$

Therefore, $\lambda \geq 1 - 2^{-r-1}$. □

3.2 Proof of Theorem 2.1

As was mentioned at the end of Section 2, it is sufficient to establish inequality (2.6). In what follows, to simplify notation, we shall write $\mathbb{P}$ instead of $\mathbb{P}^k$.

1) Let us fix arbitrary $u_1, u_2 \in \mathcal{A}$ and set $T_0 = 0$ and $T_r = T_{r-1} + r + l$ for $r \geq 1$, i.e.,

$$T_r = r(r+1)/2 + rl.$$

We claim that for any integer $r \geq 0$ there is a coupling $y_{1,2}(T_r)$ on Ω^{T_r} for the measures $\mu_{u_{1,2}}(T_r)$ such that

$$\mathbb{P}\big\{\|y_1(T_r) - y_2(T_r)\| > d_r\big\} \leq C_1 \gamma^r, \tag{3.14}$$

where C_1 and $\gamma < 1$ are some positive constants.

The construction of $y_{1,2}(T_r) = y_{1,2}(T_r, u_1, u_2; \boldsymbol{\omega}^{T_r})$ and the proof of (3.14) are by induction. For $r = 0$, we set $y_j(0) = u_j$, and inequality (3.14) with $C_1 \geq 1$ is trivial in this case. Assuming that $y_{1,2}(T_i)$ are constructed for $0 \leq i \leq r$, we set

$$y_j(T_{r+1}, u_1, u_2; \boldsymbol{\omega}^{T_{r+1}}) = U_j^{r+l+1}\big(y_1(T_r, u_{1,2}; \boldsymbol{\omega}^{T_r}), y_2(T_r, u_{1,2}; \boldsymbol{\omega}^{T_r}); \boldsymbol{\omega}^{r+l+1}\big), \tag{3.15}$$

where $U_{1,2}^k(u_1, u_2; \boldsymbol{\omega}^k)$ are defined in Lemma 3.3 and $\boldsymbol{\omega}^{T_{r+1}} = (\boldsymbol{\omega}^{T_r}, \boldsymbol{\omega}^{r+l+1})$. Let us introduce the events

$$Q_r^+ = \big\{\|y_1(T_r) - y_2(T_r)\| \leq d_r\big\}, \quad Q_r^- = \big\{\|y_1(T_r) - y_2(T_r)\| > d_r\big\}$$

and denote by p_r^+ and p_r^- their probabilities. Then, in view of (3.9) and (3.10) with $k = r + l$, we have (cf. (3.12))

$$\begin{aligned}p_{r+1}^- &= \mathbb{P}(Q_{r+1}^- | Q_r^+)\mathbb{P}(Q_r^+) + \mathbb{P}(Q_{r+1}^- | Q_r^-)\mathbb{P}(Q_r^-) \\ &\leq 2^{-r-1} p_r^+ + (1-\theta) p_r^- \leq 2^{-r-1} + \gamma p_r^-,\end{aligned} \tag{3.16}$$

where $\gamma = 1 - \theta$. Without loss of generality, we can assume that $0 < \theta < 1/2$, and therefore $1 < 2\gamma < 2$. Iterating (3.16), we obtain

$$p_{r+1}^- \leq 2^{-r-1} \sum_{j=0}^{r} (2\gamma)^j + \gamma^{r+1} p_0^- \leq 2^{-r-1} \frac{(2\gamma)^{r+1} - 1}{2\gamma - 1} + \gamma^{r+1} \leq C_1 \gamma^{r+1}.$$

This completes the induction.

2) We can now prove (2.6). Let us fix arbitrary positive integers r and $m \le r+l$ and set $k = T_r + m$, so that $T_r + 1 \le k < T_{r+1}$. We define a coupling $y_{1,2}(k) = y_{1,2}(k, u_1, u_2)$ for the measures $\mu_{u_{1,2}}(k)$ by the formula (cf. (3.15))

$$y_j(k, u_1, u_2; \boldsymbol{\omega}^k) = U_j^m\bigl(y_1(T_r, u_1, u_2; \boldsymbol{\omega}^{T_r}), y_2(T_r, u_1, u_2; \boldsymbol{\omega}^{T_r}); \boldsymbol{\omega}^m\bigr).$$

In view of (3.10) and (3.14), we have (cf. (3.16))

$$\mathbb{P}\bigl\{\|y_1(k) - y_2(k)\| > d_{r+1}\bigr\} \le \mathbb{P}(Q_r^-) + 2^{-r-1}\mathbb{P}(Q_r^+) \le C_2\gamma^r, \tag{3.17}$$

where $C_2 > 0$ is a constant. Now note that $r^2/2 \le T_r \le (l+1)r^2$ for any $r \ge 0$ and therefore there are positive constants C and c such that

$$d_{r+1} \le C\, e^{-c\sqrt{k}}, \quad C_2\gamma^r \le C\, e^{-c\sqrt{k}} \quad \text{for} \quad T_r \le k < T_{r+1}.$$

Combining this with (3.17), we derive

$$\mathbb{P}\bigl\{\|y_1(k, u_1, u_2) - y_2(k, u_1, u_2)\| \ge C\, e^{-c\sqrt{k}}\bigr\} \le C\, e^{-c\sqrt{k}}. \tag{3.18}$$

By Lemma 1.3, inequality (3.18) implies that

$$\bigl\|\mu_{u_1}(k) - \mu_{u_2}(k)\bigr\|_L^* \le 3C\, e^{-c\sqrt{k}}$$

which completes the proof of (2.6) with $k_0 = T_1$. Theorem 2.1 is proved.

4 Appendix: coupling

In this appendix, we present some results on the coupling in finite-dimensional spaces in the form which we learned from S. Foss. These results are well known (e.g., see [Lin, V] for Lemma 4.1 and [BF] for Lemma 4.3).

Let $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^N)$ be two measures absolutely continuous with respect to the Lebesgue measure dx:

$$\nu_{1,2}(dx) = p_{1,2}(x)\, dx.$$

We set

$$\rho := \|\nu_1 - \nu_2\|_{\text{var}} = \frac{1}{2}\int_{\mathbb{R}^N} |p_1(x) - p_2(x)|\, dx \tag{4.1}$$

and assume first that $0 < \rho < 1$. Let

$$p := (1-\rho)^{-1} p_1 \wedge p_2, \quad \hat{p}_{1,2} := \rho^{-1}(p_{1,2} - p). \tag{4.2}$$

For $\rho = 1$ or 0, we define $p(x)$ and $p_{1,2}(x)$ as follows:

$$p(x) \equiv 0, \qquad \hat{p}_{1,2}(x) \equiv p_{1,2}(x) \qquad \text{if} \qquad \rho = 1, \tag{4.3}$$

$$p(x) \equiv p_1(x), \qquad \hat{p}_{1,2}(x) \equiv 0 \qquad \text{if} \qquad \rho = 0. \tag{4.4}$$

It is clear that

$$p_{1,2}(x) = (1-\rho)p(x) + \rho\hat{p}_{1,2}(x) \quad \text{almost everywhere.}$$

If (ξ_1, ξ_2) is a coupling for the measures (ν_1, ν_2), then for any $\Gamma \in \mathcal{B}(\mathbb{R}^N)$ we have

$$\begin{aligned}\nu_1(\Gamma) - \nu_2(\Gamma) &= \mathbb{E}\big\{\chi_\Gamma(\xi_1) - \chi_\Gamma(\xi_2)\big\} \\ &= \mathbb{E}\big\{\chi_{\{\xi_1\neq\xi_2\}}\big(\chi_\Gamma(\xi_1) - \chi_\Gamma(\xi_2)\big)\big\} \\ &\le \mathbb{P}\{\xi_1 \neq \xi_2\}.\end{aligned}$$

Therefore,

$$\mathbb{P}\{\xi_1 \neq \xi_2\} \ge \rho \equiv \|\nu_1 - \nu_2\|_{\text{var}}.$$

A coupling (ξ_1, ξ_2) for (ν_1, ν_2) is said to be *maximal* if

$$\mathbb{P}\{\xi_1 \neq \xi_2\} = \rho \equiv \|\nu_1 - \nu_2\|_{\text{var}}.$$

Lemma 4.1. *Let $\xi_{1,2}$, ξ, and α be independent random variables such that*

$$\mathbb{P}\{\alpha = 1\} = 1-\rho, \quad \mathbb{P}\{\alpha = 0\} = \rho, \quad \mathcal{D}(\xi) = p(x)\,dx, \quad \mathcal{D}(\xi_{1,2}) = \hat{p}_{1,2}(x)\,dx. \tag{4.5}$$

Then the random variables

$$\Xi_{1,2} = \alpha\xi + (1-\alpha)\xi_{1,2} \tag{4.6}$$

form a maximal coupling for $\nu_{1,2}$.

Proof. Since ξ_1 and ξ_2 are independent and their distributions possess densities with respect to the Lebesgue measure, we have $\mathbb{P}\{\xi_1 = \xi_2\} = 0$. Taking into account the relation $\alpha(1-\alpha) \equiv 0$, we get

$$\mathcal{D}(\Xi_{1,2}) = p_{1,2}(x)\,dx = \nu_{1,2}, \quad \mathbb{P}\{\Xi_1 \neq \Xi_2\} = \mathbb{P}\{\alpha = 0\} = \rho,$$

which completes the proof. □

Let us now assume that φ is a random variable in $\mathbb{R}^N$ with the distribution $\mathcal{D}(\varphi) = q(x)\,dx$, where $q \in L^1(\mathbb{R}^N)$. Consider the following family of measures depending on a parameter $v \in \mathbb{R}^N$:

$$\nu_v(dx) = \mathcal{D}(v + \varphi) = q(x-v)\,dx.$$

Let $\rho(v_1, v_2)$ be the variation distance between ν_{v_1} and ν_{v_2}. It is clear from (4.1) that $\rho(v_1, v_2)$ is measurable with respect to $v_1, v_2 \in \mathbb{R}^{2N}$. In the construction above, let us take $\nu_{1,2} = \nu_{v_{1,2}}$. Then

$$p(x) = p(x; v_1, v_2), \quad \hat{p}_{1,2}(x) = \hat{p}_{1,2}(x; v_1, v_2).$$

Clearly, the functions $p(x; v_1, v_2)$ and $\hat{p}_{1,2}(x; v_1, v_2)$ are measurable with respect to (x, v_1, v_2). Using the above observations, we construct a coupling for (ν_{v_1}, ν_{v_2}) that is measurable with respect to (v_1, v_2, ω). Namely, we have the following result:

Theorem 4.2. *There is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that for any pair $(v_1, v_2) \in \mathbb{R}^{2N}$ there are random variables $\Xi_{1,2} = \Xi_{1,2}(v_1, v_2; \omega)$ satisfying the following properties:*

(i) *The pair (Ξ_1, Ξ_2) is a maximal coupling for (ν_{v_1}, ν_{v_2}).*

(ii) *The map $\Xi(v_1, v_2; \omega) : \mathbb{R}^{2N} \times \Omega \to \mathbb{R}^N$ is measurable with respect to the σ-algebra $\mathcal{B}(\mathbb{R}^{2N}) \times \mathcal{F}$.*

To prove the theorem, we shall need the lemma below:

Lemma 4.3. *Let $\mu_z \in \mathcal{P}(\mathbb{R}^N)$, $z \in \mathbb{R}^d$, be a family of probability measures such that*

$$\mu_z(dx) = p_z(x)\, dx,$$

where $p_z \in L^1(\mathbb{R}^N_x)$ for each $z \in \mathbb{R}^d$ and $p_z(x)$ is measurable as a function of $(x, z) \in \mathbb{R}^N \times \mathbb{R}^d$. Then there is a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and a family of random variables $\zeta_z \colon \Omega \to \mathbb{R}^N$ such that $\mathcal{D}(\zeta_z) = \mu_z$ for all $z \in \mathbb{R}^d$ and $\zeta_z(x)$ is measurable with respect to (z, x).

Proof. If $N = 1$, then we take $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}, dt)$, where $\mathcal{B}$ is the Borel σ-algebra and dt is the Lebesgue measure. Denoting by $F_z(\lambda)$ the distribution function of the measure μ_z, $F_z(\lambda) = \mu_z((-\infty, \lambda])$, we set

$$\zeta_z(t) = \min\{\lambda : F_z(\lambda) \geq t\}.$$

The map $(t, z) \mapsto \zeta_z(t)$ from $[0, 1] \times \mathbb{R}^d$ to $\mathbb{R}$ is measurable, and the distribution function of $\mathcal{D}(\zeta_z)$ is equal to F_z. Thus, for $N = 1$ the lemma is proved.

We now assume that the required assertion is established for $N = L$ and prove it for $N = L + 1$. Let us write $x \in \mathbb{R}^{L+1}$ as $x = (x', y)$, where $x' \in \mathbb{R}^L$ and $y \in \mathbb{R}$. Decomposing μ_z in terms of the conditional density (see [GS]), we write

$$\mu_z(dx) = p_z(x)\, dx = p'_z(x' | \, y)\, dx' q_z(y)\, dy. \tag{4.7}$$

Here

$$q_z(y) = \int_{\mathbb{R}^L} p_z(x', y)\, dx', \quad p'_z(x' | \, y) = \frac{p_z(x', y)}{q_z(y)},$$

where we set $0/0 = \infty/\infty = 0$. Applying the induction hypothesis with z replaced by (z, y), we find a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ and a measurable map

$$\zeta'_z(\omega', y) \colon \Omega' \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^L$$

such that $\mathcal{D}\big(\zeta'_z(\cdot, y)\big) = p'_z(x' | \, y)\, dx'$ for each $(z, y) \in \mathbb{R}^d \times \mathbb{R}$. Applying the first step of the proof, we construct a measurable map $\xi_z(t) : [0, 1] \times \mathbb{R}^d \to \mathbb{R}$ such that $\mathcal{D}(\xi_z) = q_z(\lambda)\, d\lambda$. We now set $\Omega = \Omega' \times [0, 1]$ and

$$\zeta_z(\omega', t) = \big(\zeta'_z(\omega', \xi_z(t)), \xi_z(t)\big) \in \mathbb{R}^{L+1}.$$

We have constructed a measurable map $\Omega \times \mathbb{R}^d \to \mathbb{R}^{L+1}$ such that, for any fixed $z \in \mathbb{R}^d$, its distribution is given by the right-hand side of (4.7). □

Proof of Theorem 4.2. Applying Lemma 4.2 to measures in $\mathbb{R}^N$ given by the densities p and $\hat{p}_{1,2}$, we construct probability spaces $(\Sigma_j, \mathcal{S}_j, \mathbb{P}_j)$, $j = 0, 1, 2$, and random variables $\xi^j_{(v_1,v_2)}$ on Σ_j such that

$$\mathcal{D}(\xi^0_{(v_1,v_2)}) = p(x; v_1, v_2)\,dx, \quad \mathcal{D}(\xi^j_{(v_1,v_2)}) = \hat{p}_j(x; v_1, v_2)\,dx, \quad j = 1, 2. \tag{4.8}$$

We also define a random variable $\alpha_\rho : [0,1] \to \{0,1\}$, $\rho = \rho(v_1, v_2)$, by the formula

$$\alpha_\rho(t) = \chi_{[0,1-\rho]}(t),$$

where $[0,1]$ is endowed with the Borel σ-algebra and the Lebesgue measure, and $\chi_{[0,r]}$ is the characteristic function of the interval $[0,r]$.

We now define the required probability space as the set

$$\Omega = \Sigma_0 \times \Sigma_1 \times \Sigma_2 \times [0,1]$$

with the σ-algebra and the probability of direct product. The natural extensions[3] of α_ρ and $\xi^j_{(v_1,v_2)}$, $j = 0, 1, 2$, to Ω (for which we retain the same notations) form a quadruple of independent random variables satisfying (4.8) and also the relations

$$\mathbb{P}\{\alpha_\rho = 1\} = 1 - \rho(v_1, v_2), \quad \mathbb{P}\{\alpha_\rho = 0\} = \rho(v_1, v_2).$$

A maximal coupling (Ξ_1, Ξ_2) for the measures (ν_{v_1}, ν_{v_2}) that satisfies assertion (ii) of the theorem can now be defined by formula (4.6), in which $\alpha = \alpha_\rho$, $\xi = \xi^0_{(v_1,v_2)}$, and $\xi_j = \xi^j_{(v_1,v_2)}$, $j = 1, 2$. $\square$

References

[BF] A. A. Borovkov, S. G. Foss, *Stochastically recursive sequences and their generalizations*, Siberian Advances in Mathematics **2** (1992), no. 1, 16–81.

[BL] X. Bressaud, C. Liverani, *Anosov diffeomorphism and coupling*, to appear in Ergodic Theory Dynam. Systems.

[BKL] J. Bricmont, A. Kupiainen, R. Lefevere, *Exponential mixing for the 2D stochastic Navier–Stokes dynamics*, preprint.

[EMS] W. E, J. C. Mattingly, Ya. G. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation*, preprint.

[GS] I. I. Gihman, A. V. Skorohod, *The Theory of Stochastic Processes I*, Springer-Verlag, Berlin–Heidelberg–New York, 1980.

[H] L. Hörmander, *The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis*, Springer-Verlag, Berlin, 1983.

[3]For instance, the extension of α_ρ is given by $\alpha_\rho(\omega) = \alpha_\rho(t)$, where $\omega = (\omega_0, \omega_1, \omega_2, t) \in \Omega$.

[KS1] S. Kuksin, A. Shirikyan, *Stochastic dissipative PDE's and Gibbs measures*, Comm. Math. Phys. **213** (2000), 291–330.

[KS2] S. Kuksin, A. Shirikyan, *On dissipative systems perturbed by bounded random kick-forces*, Submitted to Ergodic Theory Dynam. Systems. (www.ma.hw.ac.uk/~kuksin)

[KS3] S. Kuksin, A. Shirikyan, *Ergodicity for the randomly forced 2D Navier–Stokes equations*, preprint. (www.ma.hw.ac.uk/~kuksin)

[Lin] T. Lindvall, *Lectures on the Coupling Method*, John Wiley & Sons, New York, 1992.

[V] A. Yu. Veretennikov, *Parametric and non-parametric estimation of Markov chains*, Moscow State University Press, Moscow, 2000. (in Russian)

[Y] L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.

Sergei Kuksin
Department of Mathematics
Heriot-Watt University
Edinburgh EH14 4AS
Scotland, UK
and
Steklov Institute of Mathematics
8 Gubkina St., 117966 Moscow
Russia
E-mail: S.B.Kuksin@ma.hw.ac.uk

Armen Shirikyan
Department of Mathematics
Heriot-Watt University
Edinburgh EH14 4AS
Scotland, UK
and
Institute of Mechanics of MSU
1 Michurinskii Av., 119899 Moscow
Russia
E-mail: A.Shirikyan@ma.hw.ac.uk