EXPONENTIAL ERGODICITY AND REGULARITY FOR EQUATIONS WITH LÉVY NOISE

ENRICO PRIOLA, ARMEN SHIRIKYAN, LIHU XU, AND JERZY ZABCZYK

Abstract. We prove exponential convergence to the invariant measure, in the total variation norm, for solutions of SDEs driven by α-stable noises in finite and in infinite dimensions. Two approaches are used. The first one is based on Liapunov's function approach by Harris, and the second on Doeblin's coupling argument [9]. Irreducibility and uniform strong Feller property play an essential role in both approaches. We concentrate on two classes of Markov processes: solutions of finite dimensional equations, introduced in [28], with Hölder continuous drift and a general, non-degenerate, symmetric α-stable noise, and infinite dimensional parabolic systems, introduced in [31], with Lipschitz drift and cylindrical α-stable noise. We show that if the nonlinearity is bounded, then the processes are exponential mixing. This improves, in particular, an earlier result established in [29], with a different method.

Keywords: stochastic PDEs, α-stable noise, Hölder continuous drift, Harris' theorem, coupling, total variation, exponential mixing, Ornstein-Uhlenbeck processes.

Mathematics Subject Classification (2000): 60H15, 47D07, 60J75, 35R60.

Contents

The first author was supported by the M.I.U.R. research project Prin 2008 "Deterministic and stochastic methods in the study of evolution problems". The third author gratefully acknowledges the support by Junior program *Stochastics* of Hausdorff Research Institute for Mathematics. His research is partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement nr. 258237. The fourth author gratefully acknowledges the support by the Polish Ministry of Science and Higher Education grant "Stochastic equations in infinite dimensional spaces" N N201 419039.

1. Introduction

This paper is concerned with ergodic properties of the stochastic equation

$$dX_t = [AX_t + F(X_t)]dt + dZ_t, \quad X_0 = x, \tag{1.1}$$

both in finite and infinite dimensional real Hilbert spaces H. Here A is a linear operator, F, a bounded mapping and Z, a symmetric α-stable process. Under suitable conditions, we establish exponential convergence of the solutions to the invariant measure in the variation norm. When H is infinite dimensional several nonlinear stochastic PDEs, including semilinear heat equations perturbed by Lévy noise, are of the form (1.1).

Irreducibility and uniform strong Feller properties play an essential role in our approach. They are established in the paper when the space H is finite dimensional, Z is a non-degenerate, symmetric α-stable process and F is η-Hölder continuous with $1 - \frac{\alpha}{2} < \eta \le 1$ and $1 < \alpha < 2$. Under stronger assumptions on the drift F and on the noise process Z, those properties were derived in [31] in infinite dimensions. The finite dimensional result is an important contribution of the paper of independent interest.

The stochastic PDEs driven by Lévy noises have been intensively studied since some time; e.g., see the papers [3, 1, 27, 25, 19, 31, 40], the book [26] and the references therein. Invariant measures and long-time asymptotics for stochastic systems driven by Lévy noises were studied in a number of papers. In particular, the linear case ($F \equiv 0$) was investigated, in finite dimensions, in [35] and [43] and, in infinite dimensions, in [5, 32, 10]. The case of nonlinear equations was studied in [33, 26, 21, 40, 41]. However, there are no many results on ergodicity and exponential mixing (cf. [41, 14, 29]). The paper [14] studied the exponential mixing of finite dimensional stochastic systems with jump noises, which include one dimensional SDEs driven by α-stable noise.

Some ergodic properties for SPDEs like (1.1) were also studied in [29]. There it was proved that if the supremum norm of F is *small*, then there exists a unique invariant measure, which is exponential mixing under the weak topology in the space of measures. Here we improve substantially this result, showing that the convergence to the invariant measure holds exponentially fast in the total variation norm without any smallness assumption on F. To prove this result, we have to impose a slightly stronger regularity condition on the noise with respect to [29]; this is really a mild assumption (see Remark 2.3 and Example 2.9).

As mentioned before, we also establish exponential mixing in the total variation norm for finite dimensional stochastic equations like (1.1) with a less regular drift term F and a more general noise Z. It seems that even in one dimension (when Z reduces to a standard symmetric rotationally invariant α-stable noise) our result on exponential mixing is new (cf. [40] and [14]).

We have two proofs for the exponential mixing results. The first one is based on the classical Harris' theorem, while the other is on the classical coupling argument, see Section 2.5 and also [18]. In both approaches, irreducibility and uniform strong Feller property play the crucial role. The Harris approach only needs to check some conditions involving Lyapunov functions, but it is not intuitive. The coupling proof is quite involved, but gives the intuition for understanding the way in which the dynamics converges to the ergodic measure.

Let us sketch our methods on proving the well-posedness and the structural properties of finite dimensional stochastic systems, since it has independent interest. To prove the existence and pathwise uniqueness of solutions, we only need to modify a little bit the method established in [28]. We stress that the condition $1-\frac{\alpha}{2}<\eta\le 1$ is needed to have existence and uniqueness of solutions (cf. [28]). The irreducibility and uniform strong Feller property will be established in the following two steps. First we prove irreducibility and (uniform) gradient estimates for finite dimensional Ornstein-Uhlenbeck processes driven by non-degenerate symmetric α-stable processes (related gradient estimates under different assumptions from ours are given in the recent paper [42]). Then we proceed as in [31] and deduce irreducibility and uniform gradient estimates for solutions to (1.1). Note that if $\eta<1$ then the deterministic equation may have many solutions as classical examples show. Currently, there is a great interest in understanding pathwise uniqueness for SDEs when F is not Lipschitz, see the references given in [6, 28].

The paper is organized as follows. In Section 2 we formulate basic structural properties of the solutions of (1.1) and our main ergodic results, namely Theorems 2.8 and 2.7. In Section 3 we concentrate on proving the new structural properties of finite dimensional systems. Section 4 contains decay L_p-estimates for solutions of (1.1), which are needed to prove exponential ergodicity; here we concentrate on the infinite dimensional case since in finite dimensions these estimates are straightforward. The two proofs for the exponential mixing of infinite dynamics are established in Sections 5 and 6 respectively, the former applying Harris' theorem and the latter using coupling argument. Section 6 is quite involved, in particular, exponential estimates for the first hitting time of balls are of independent interest. In Section 7 we show the exponential ergodicity for finite dimensional systems (Theorem 2.7) in a sketchy way. We have only shown the full details for the proof of Theorem 2.8 concerning SPDEs, since the finite dimensional result can be proved by similar and easier methods.

Acknowledgements. We would like to thank C. Odasso for patiently discussing with us his paper [23] and writing a note for us on the proof of inequality (6.6). We also would like to thank M. Hairer for pointing out to us the proof of Theorem 2.8 by the Harris approach.

2. Notations and main results

2.1. **Notations and assumptions.** Let H be a real separable Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and the corresponding norm $|\cdot|$. We denote by $\{e_k\}_{k\geq 1}$ an orthonormal basis, so that any vector $x \in H$ can be written as $x = \sum_{k\geq 1} x_k e_k$, where $\sum_k |x_k|^2 < \infty$. Denote by $B_b(H)$ the Banach space of bounded Borel-measurable functions $f : H \to \mathbb{R}$ with the supremum norm

$$\|f\|_0 := \sup_{x\in H} |f(x)|.$$

Let $\mathcal{B}(H)$ be the Borel σ-algebra on H and let $\mathcal{P}(H)$ be the set of probabilities on $(H, \mathcal{B}(H))$. Recall that the total variation distance between two measures $\mu_1, \mu_2 \in \mathcal{P}(H)$ is defined by

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} = \frac{1}{2} \sup_{\substack{f\in B_b(H)\\ \|f\|_0=1}} |\mu_1(f) - \mu_2(f)| = \sup_{\Gamma\in\mathcal{B}(H)} |\mu_1(\Gamma) - \mu_2(\Gamma)|.$$

Let $z(t)$ be a one-dimensional symmetric α-stable process with $0 < \alpha < 2$. Its infinitesimal generator $\mathcal{A}$ is given by

$$\mathcal{A}f(x) := \frac{1}{C_\alpha} \int_{\mathbb{R}} \frac{f(y+x) - f(x)}{|y|^{\alpha+1}} dy, \quad x \in \mathbb{R}, \tag{2.1}$$

where $C_\alpha = -\int_{\mathbb{R}} (\cos y - 1) \frac{dy}{|y|^{1+\alpha}}$; see [34] and [2]. It is well known that $z(t)$ has the following characteristic function:

$$\mathbb{E}[e^{i\lambda z(t)}] = e^{-t|\lambda|^\alpha},$$

$t \geq 0$, $\lambda \in \mathbb{R}$. A multidimensional generalization of $z(t)$ is obtained by considering an n-dimensional non-degenerate symmetric α-stable process $Z = (Z_t)$. This is a Lévy process with the additional property that

$$\mathbb{E}[e^{i\langle Z_t,u\rangle}]e^{-t\psi(u)}, \quad \psi(u) = -\int_{\mathbb{R}^d} \Big(e^{i\langle u,y\rangle} - 1 - \, i\langle u, y\rangle \, 1_{\{|y|\leq 1\}}\,(y) \Big) \nu(dy), \tag{2.2}$$

$u \in \mathbb{R}^n$, $t \geq 0$, where the Lévy (intensity) measure ν is of the form

$$\nu(D) = \int_S \mu(d\xi) \int_0^\infty 1_D(r\xi) \frac{dr}{r^{1+\alpha}}, \quad D \in \mathcal{B}(\mathbb{R}^n), \tag{2.3}$$

for some symmetric, non-zero finite measure μ concentrated on the unitary sphere $S = \{y \in \mathbb{R}^d \,:\, |y| = 1\}$ (see [34, Theorem 14.3]). Note that formula (2.3) implies that $\psi(u) = c_\alpha \int_S |\langle u, \xi\rangle|^\alpha \mu(d\xi),\ u \in \mathbb{R}^n$ (see also [34, Theorem 14.13]). The non-degeneracy hypothesis on Z is the assumption that there exists a positive constant C_α such that, for any $u \in \mathbb{R}^n$,

$$\psi(u) \geq C_\alpha |u|^\alpha. \tag{2.4}$$

This is equivalent to the fact that the support of μ is not contained in a proper linear subspace of $\mathbb{R}^n$ (see [28] for more details). Recall that the infinitesimal generator $\mathcal{A}$ of the process Z is given on the space of all infinitely differentiable functions with compact support $C_c^\infty(\mathbb{R}^n)$ by the formula,

$$\mathcal{A}f(x) = \int_{\mathbb{R}^d} \big(f(x+y) - f(x) - 1_{\{|y|\leq 1\}}\langle y, Df(x)\rangle\big)\,\nu(dy), \quad f \in C_c^\infty(\mathbb{R}^n),$$

see [34, Section 31]. Note that $Z_t = \sum_{1\leq j\leq n} \beta_j z_j(t) e_j$ (where $\{z_j(t)\}_{1\leq j\leq n}$ are i.i.d. one-dimensional symmetric α-stable processes) is in particular a non-degenerate symmetric α-stable process if each $\beta_j \neq 0$.

We will make two sets of assumptions on (1.1) depending on the dimension of the Hilbert space H. They are similar but more restrictive if $dim(H) = \infty$.

Assumption 2.1. $[dim(H) = n < \infty]$

- (A1) *A is an $n \times n$ matrix and $\max_{1\leq i\leq n} Re(\gamma_k) < 0$, where $\gamma_1, \dots, \gamma_n$ are the eigenvalues of A counted according to their multiplicity.*
- (A2) *$Z = (Z_t)$ is a symmetric non-degenerate n-dimensional α-stable process with $1 < \alpha < 2$.*
- (A3) *$F : H \to H$ is bounded and η-Hölder continuous with $1 - \frac{\alpha}{2} < \eta \leq 1$.*

Assumption 2.2. $[dim(H) = \infty]$

- (A1) *A is a dissipative operator defined by*

$$A = \sum_{k\geq 1} (-\gamma_k) e_k \otimes e_k$$

with $0 < \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_k \leq \ldots$ and $\gamma_k \to \infty$ as $k \to \infty$.
- (A2) *Z_t is a cylindrical α-stable process with $Z_t = \sum_{k\geq 1} \beta_k z_k(t) e_k$, where $\{z_k(t)\}_{k\geq 1}$ are i.i.d. symmetric α-stable processes with $1 < \alpha < 2$, $\beta_k > 0$ and there exists some $\varepsilon \in (0,1)$ such that $\sum_{k\geq 1} \frac{\beta_k^\alpha}{\gamma_k^{1-\alpha\varepsilon}} < \infty$.*
- (A3) *$F : H \to H$ is Lipschitz and bounded.*
- (A4) *There exist some $\theta \in (0,1)$ and $C > 0$ so that $\beta_k \geq C\gamma_k^{-\theta+1/\alpha}$.*

Remark 2.3. Let us comment on Assumption 2.2. The Lipschitz property guarantees that Eq. (1.1) has a unique solution, and (A4) that the solution is strong Feller. The condition $\sum_{k\geq 1} \frac{\beta_k^\alpha}{\gamma_k^{1-\alpha\varepsilon}} < \infty$ in (A2) implies that the solution to (1.1) evolves in linear subspace with compact embedding into H, see Section 4. Note that in [29] it is only required that (A2) holds for $\epsilon = 0$ (i.e., that $X_t^x \in H$, a.s.).

However our present assumption with $\epsilon > 0$ is really a mild assumption (compare also with Example 2.9).

2.2. **Structural properties of solutions.** In this subsection we formulate the structural properties of solutions both in finite and in infinite dimensions, i.e. Theorems 2.4 and 2.5. These structural properties shall play an important role in proving the exponential ergodicity.

The proof of the next theorem is quite involved and is postponed to Section 3.

Theorem 2.4. *Let $H = \mathbb{R}^n$. Under Assumption 2.1, there exists a unique strong solution X_t^x for* (1.1)*. The solutions $(X_t^x)_{x\in H}$ form a Markov process with transition semigroup P_t,*

$$P_t f(x) = \mathbb{E}[f(X_t^x)], \quad f \in B_b(H),$$

which is irreducible and such that there exists $C > 0$ with

$$|P_t f(x) - P_t f(y)| \le \frac{C\|f\|_0}{t^{1/\alpha} \wedge 1}|x-y|, \;\; x, y \in H, \; t > 0, \; f \in B_b(H). \tag{2.5}$$

The following infinite dimensional result is analogous to the previous one and is proved in [31]. Note that the noise Z considered here reduces in finite dimension to a particular case of the noise in Theorem 2.4.

Theorem 2.5. *Under Assumption 2.2, there exists a unique mild solution X_t^x for* (1.1)*,*

$$X_t^x = e^{At}x + \int_0^t e^{A(t-s)}F(X_s^x)ds + \int_0^t e^{A(t-s)}dZ_s. \tag{2.6}$$

The solutions $(X_t^x)_{x\in H}$ form a Markov process with the transition semigroup P_t. The process is irreducible and there exists $C > 0$ such that

$$|P_t f(x) - P_t f(y)| \le \frac{C\|f\|_0}{t^{1/\theta} \wedge 1}|x-y|, \;\; x, y \in H, \; t > 0, \tag{2.7}$$

where θ is given in (A4) of Assumption 2.2.

Remark 2.6. Note if $dim(H) = \infty$ then, in general, trajectories of (X_t^x) do not have a càdlàg modifications (see [4]).

2.3. **Ergodic results for finite-dimensional equations.** Let us denote by $(P_t)_{t\ge 0}$ the Markov semigroup associated with (1.1) and by $(P_t^*)_{t\ge 0}$ the dual semigroup acting on $\mathcal{P}(H)$.

The main result for the finite-dimensional case is as follows:

Theorem 2.7. *Under Assumption 2.1, the system* (1.1) *is ergodic and exponential mixing. More precisely, there exists $\mu \in \mathcal{P}(H)$ such that, for any $p \in (0, \alpha)$ and any measure $\nu \in \mathcal{P}(H)$ with finite p^{th} moment, we have*

$$\|P_t^*\nu - \mu\|_{\mathrm{TV}} \le Ce^{-ct}\left(1 + \int_H |x|^p \nu(dx)\right), \tag{2.8}$$

where $C = C(p, \alpha, A, \|F\|_0)$ *and* $c = c(p, \alpha, A, \|F\|_0)$.

One can easily adapt our proof to show that the previous theorem is also true when (Z_t) is Gaussian.

2.4. **Ergodic results in the infinite dimensional case.** The following theorem describing the long-time behaviour of (X_t^x) is the main result of the infinite-dimensional case.

Theorem 2.8. *Under Assumption 2.2, the system* (1.1) *is ergodic and exponential mixing. More precisely, there exists* $\mu \in \mathcal{P}(H)$ *so that for any* $p \in (0, \alpha)$ *and any measure* $\nu \in \mathcal{P}(H)$ *with finite* p^{th} *moment, we have*

$$\|P_t^* \nu - \mu\|_{\text{TV}} \leq C e^{-ct} \left(1 + \int_H |x|^p \nu(dx)\right), \tag{2.9}$$

where $C = C(p, \alpha, \theta, \beta, \gamma, \varepsilon, \|F\|_0)$ *and* $c = c(p, \alpha, \theta, \beta, \gamma, \varepsilon, \|F\|_0)$ *with* $\beta = (\beta_k)$, $\gamma = (\gamma_k)$.

We will apply the above theorem in the following example which was considered in [29].

Example 2.9. Consider the following stochastic semilinear equation on $D = [0, \pi]^d$ with $d \geq 1$ and the Dirichlet boundary condition:

$$\begin{cases} dX(t, \xi) = [\Delta X(t, \xi) + F(X(t, \xi))]dt + dZ_t(\xi), \\ X(0, \xi) = x(\xi), \\ X(t, \xi) = 0, \quad \xi \in \partial D, \end{cases} \tag{2.10}$$

where Z_t and F are both specified below. It is clear that Δ with Dirichlet boundary condition has the following eigenfunctions

$$e_k(\xi) = \left(\frac{2}{\pi}\right)^{\frac{d}{2}} \sin(k_1 \xi_1) \cdots \sin(k_d \xi_d), \quad k \in \mathbb{N}^d, \ \xi \in D.$$

It is easy to see that $\Delta e_k = -|k|^2 e_k$, i.e. $\gamma_k = |k|^2 = k_1^2 + \ldots + k_d^2$ for all $k \in \mathbb{N}^d$. We study the dynamics defined by (2.10) in the Hilbert space $H = L^2(D)$ with orthonormal basis $\{e_k\}_{k \in \mathbb{N}^d}$. $Z = (Z_t)$ is some cylindrical α-stable noises which, under the basis $\{e_k\}_k$, is defined by

$$Z_t = \sum_{k \in \mathbb{N}^d} |k|^\beta z_k(t) e_k$$

where $\{z_k(t)\}_k$ are i.i.d. symmetric α-stable processes with $\alpha \in (0, 2)$ and β a real number. Note that $\sum_{k \in \mathbb{N}^d} \frac{|k|^{\beta\alpha}}{|k|^2} < \infty$ if and only if $2 > d + \alpha\beta$.

From Theorems 2.5 and 2.6 in [29], one has

(1) If F is a bounded Lipschitz function and

$$2 > d + \alpha\beta, \ \ \frac{1}{\alpha} - \frac{\beta}{2} < 1,$$

or equivalently, $\frac{d}{\alpha} < \frac{2}{\alpha} - \beta < 2$, then the system (2.10) is strongly mixing.

(2) If in addition $\|F\|_0$ is sufficiently small then the system (2.10) is exponential mixing under the *weak* topology in the space of finite measures.

From Theorem 2.8 in the present paper, we have the following much stronger result:

(3) If F satisfies the conditions in (1), then the system (2.10) is exponential mixing under the *total variation* topology.

2.5. **Two approaches to exponential ergodicity.** We shall prove the exponential ergodicity results by two approaches. The first one is by applying classical Harris' theorem and the other is by coupling argument.

We shall use the following Harris' theorem. For a surprisingly short and nice proof we refer to Hairer's lecture notes [12].

Theorem 2.10 (Harris)**.** *Let P_t be a Markov semigroup in the Polish space X such that there exists $T_0 > 0$ and $V : X \to \mathbb{R}_+$ which satisfies:*
(i) there exists $\gamma < 1$ and $K > 0$ such that $P_{T_0}V(x) \le \gamma V(x) + K$, $x \in X$.
(ii) for every $R > 0$ there exists $\delta > 0$ such that

$$\|P^*_{T_0}\delta_x - P^*_{T_0}\delta_y\|_{TV} \le 2 - \delta,$$

for all $x, y \in X$ such that $V(x) + V(y) \le R$.

Then there exist some $T > 0$ and $\beta < 1$ such that

$$\int_X (1 + V(x))|P^*_T\mu - P^*_T\nu|(dx) \le \beta \int_X (1 + V(x))|\mu - \nu|(dx).$$

The key point for Harris' theorem approach is to guess the Liapuonov function V and to check conditions (i) and (ii).

To sketch the coupling approach, let us fix a large constant $T > 0$ and consider the restriction of the Markov process (X^x_t), $x \in H$, to the times proportional to T. We denote by (Y_k) the resulting discrete-time Markov process, by $\mathbb{P}_x$ the corresponding family of probability measures, and by $P_k(x, \Gamma)$ the transition function. The dissipativity of A, the boundedness of F, and the non-degeneracy of Z imply that (Y_k) is irreducible, and the first hitting time of any ball has a finite exponential moment. Furthermore, and this will follow from Theorems 2.4 and 2.5, if the initial points $x_1, x_2 \in H$ are such that $|x_1 - x_2| \le r$, with a sufficiently small $r > 0$, then

$$\|P_1(x_1, \cdot) - P_1(x_2, \cdot)\|_{\mathrm{TV}} \le \frac{1}{2}. \tag{2.11}$$

Now let (Y_k^1, Y_k^2) be a homogeneous discrete-time Markov process in the extended phase space $H \times H$ such that the following properties hold for the pair (Y_1^1, Y_1^2) under the law $\mathbb{P}_{(x_1,x_2)}$ corresponding to the initial point (x_1, x_2):

(a) The laws of Y_1^1 and Y_1^2 coincide with $P_1(x_1, \cdot)$ and $P_1(x_2, \cdot)$, respectively.
(b) If $\max(|x_1|, |x_2|) > r$ and $x_1 \neq x_2$, then the random variables Y_1^1 and Y_1^2 are independent.
(c) If $\max(|x_1|, |x_2|) \leq r$ and $x_1 \neq x_2$, then

$$\mathbb{P}_{(x_1,x_2)}\{Y_1^1 \neq Y_1^2\} = \|P_1(x_1, \cdot) - P_1(x_2, \cdot)\|_{\mathrm{TV}}.$$

(d) If $x_1 = x_2$, then $Y_1^1 = Y_1^2$ with probability 1.

Such a chain can be constructed with the help of maximal coupling of measures; see Section 6. Combining properties (a)–(d) with irreducibility of (Y_k) and inequality (2.11), it is possible to prove that the stopping time $\rho = \min\{k \geq 0 : Y_k^1 = Y_k^2\}$ is $\mathbb{P}_{(x_1,x_2)}$-almost surely finite and has a finite exponential moment. Moreover, it follows from (d) that $Y_k^1 = Y_k^2$ for $k \geq \rho$. We can thus write

$$|P_k(x_1, \Gamma) - P_k(x_2, \Gamma)| = |\mathbb{E}_{(x_1,x_2)}(I_\Gamma(Y_k^1) - I_\Gamma(Y_k^2))| \leq \mathbb{P}_{(x_1,x_2)}\{\rho > k\}, \tag{2.12}$$

where $\Gamma \subset H$ is an arbitrary Borel subset and I_Γ stands for its indicator function. Since ρ has a finite exponential moment, the right-hand side of (2.12) can be estimated by $\mathrm{const}\, e^{-\gamma k}$. Taking the supremum over all Borel subsets Γ, we conclude that the total variation distance between $P_k(x_1, \Gamma)$ and $P_k(x_2, \Gamma)$ goes to zero exponentially fast for any initial points $x_1, x_2 \in H$. This implies the required uniqueness and exponential mixing.

In conclusion, let us note that, in the context of randomly forced PDE's, the coupling argument can be modified to cover the case of degenerate noises. We refer the reader to [15, 20, 36] for discrete-time random perturbations, to [13, 11, 16, 37, 24] for a white noise, to [22] for a compound Poisson process, and to the book [17] for further references on this subject. We believe that a similar approach can be developed in the case of dissipative PDE's driven by Lévy noises.

3. Proof of structural properties, dim H $< \infty$

In this section, we concentrate on proving Theorem 2.4, which can be done in the following steps.

Step 1. Existence and uniqueness. Since (with $X_t = X_t^x$)

$$X_t = x + \int_0^t AX_s ds + \int_0^t F(X_s)ds + Z_t, \tag{3.1}$$

defining $v(t) = X_t - Z_t$, one can construct a càdlàg adapted solution, by working ω by ω and using a compacteness argument.

Uniqueness holds even in the limiting case $\alpha = 1$. When $A = 0$ it follows directly from [28]. In the present case of $A \neq 0$, since the drift in [28] was supposed to be bounded and $x \mapsto Ax$ is an unbounded mapping, to prove pathwise uniqueness one can proceed into two different ways. First one can adapt the computations in [28] using a standard stopping time argument. To this purpose, we only note that if X_t is one solution starting from $x \in \mathbb{R}^n$ then formula in [28, Lemma 4.2] continue to hold if t is replaced by $t \wedge \tau_R$, $R > 0$, where

$$\tau_R = \inf\{t \geq 0; |X_t| \leq R\}.$$

Another method consists in introducing the process $Y_t = e^{-At}X_t$. Clearly Y_t satisfies the following equation

$$dY_t = e^{-At}F(e^{At}Y_t)dt + e^{-At}dZ_t. \tag{3.2}$$

According to [28] with small modifications (due to the fact that now the drift is bounded but also time-dependent), (3.2) has a unique strong solution such that

$$Y_t = x + \int_0^t e^{-As}F(e^{As}Y_s)ds + \int_0^t e^{-As}Z_s,$$

and this is equivalent to (3.1).

Step 2. Markov property. This follows from the uniqueness by standard considerations.

Step 3. Uniform strong Feller estimate (2.7).

In order to adapt the method used in the proof of [31, Theorem 5.7], we need gradient estimates like

$$\|DR_tf\|_0 \leq \frac{c}{t^{1/\alpha}}\|f\|_0, \;\; t \in (0,1], \; f \in B_b(H). \tag{3.3}$$

for the OU semigroup R_t corresponding to $F = 0$ in (3.1).

Remark 3.1. Some related estimates were obtained in a recent paper [42] which however does not cover the present situation. We also mention [38] which contains a Bismut-Elworthy-Li formula for jump diffusion semigroups (even without a Gaussian part). We cannot apply [38] since our Lévy measure ν in general does not have a C^1-density with respect to the Lebesgue measure in $\mathbb{R}^n \setminus \{0\}$.

The next result seems to be of independent interest.

Theorem 3.2. *Let $H = \mathbb{R}^n$. Assume that $Z = (Z_t)$ is an n-dimensional symmetric non-degenerate α-stable process, $\alpha \in (0,2)$. Consider any real $n \times n$ matrix A. Then gradient estimates* (3.3) *holds for the OU semigroup R_t associated to*

$$dX_t = AX_tdt + dZ_t, \;\; X_0 = x.$$

Proof. Let us fix $f \in B_b(H)$ and $t \in (0,T]$. It is known (see, for instance, [30]) that

$$R_tf(x) = \int_H f(e^{tA}x + y)p_t(y)(dy),$$

$$p_t(y) = \frac{1}{(2\pi)^n} \int_H e^{-i\langle y,h\rangle} \exp\Big(- \int_0^t \psi(e^{sA^*}h)ds\Big) dh,$$

where ψ is the exponent (or symbol) of the Lévy process Z (see (2.2)). We write

$$R_t f(x) = \frac{1}{(2\pi)^n} \int_H f(z) \Big(\int_H e^{-i\langle z,h\rangle} e^{i\langle e^{tA^*}h,x\rangle} e^{-\int_0^t \psi(e^{sA^*}h)ds} dh\Big) dz.$$

(1). Recall the rescaling property

$$\psi(us) = s^\alpha \psi(u), \quad s \geq 0,$$

and $u \in H$. The non-degeneracy assumption (2.4) implies that there exists the directional derivative along any fixed direction $l \in H$, $|l| = 1$ (cf. Section 3 in [28]),

$$D_l R_t f(x) = \frac{i}{(2\pi)^n} \int_H f(z) \Big(\int_H e^{-i\langle z,h\rangle} e^{i\langle e^{tA^*}h,x\rangle} \, \langle e^{tA^*}h, l\rangle \, e^{-\int_0^t \psi(e^{sA^*}h)ds} dh\Big) dz.$$

Let $e^{tA^*}h = k$. We have

$$\begin{aligned} D_l R_t f(x) &= \frac{ie^{-t\, tr(A)}}{(2\pi)^n} \int_H f(z) \Big(\int_H e^{-i\langle z, e^{-tA^*}k\rangle} e^{i\langle k,x\rangle} \, \langle k, l\rangle \, e^{-\int_0^t \psi(e^{(s-t)A^*}k)ds} dk\Big) dz \\ &= \frac{i}{(2\pi)^n} \int_H f(e^{tA}\xi) \Big(\int_H e^{-i\langle \xi,k\rangle} e^{i\langle k,x\rangle} \, \langle k, l\rangle \, e^{-\int_0^t \psi(e^{-rA^*}k)dr} dk\Big) d\xi \\ &= \frac{i}{(2\pi)^n} \int_H f(e^{tA}\xi) \Big(\int_H e^{i\langle k,(x-\xi)\rangle} \, \langle k, l\rangle \, e^{-\int_0^t \psi(e^{-rA^*}k)dr} dk\Big) d\xi. \end{aligned}$$

Let us introduce

$$\phi_t(v) = \frac{1}{(2\pi)^n} \int_H e^{i\langle k,v\rangle} \, \langle k, l\rangle \, e^{-\int_0^t \psi(e^{-rA^*}k)dr} dk.$$

It is clear that we get

$$\|D_l R_t f\|_0 \leq \frac{C_1}{t^{1/\alpha}} \|f\|_0, \quad t \in (0, 1].$$

(and so (3.3)) if we are able to prove that

$$\|\phi_t\|_{L^1(H)} \leq \frac{C_1}{t^{1/\alpha}}, \quad t \in (0, 1], \tag{3.4}$$

where $L^1(H) = L^1(\mathbb{R}^n)$ with respect to the Lebesgue measure.

(2). Let us check (3.4). Using the rescaling property, we have

$$\begin{aligned} \phi_t(v) &= \frac{1}{(2\pi)^n} \int_H e^{i\langle k,v\rangle} \, \langle k, l\rangle \, \exp\left\{ -\frac{1}{t} \int_0^t \psi(e^{-rA^*} t^{1/\alpha} k) dr \right\} dk \\ &= \frac{1}{(2\pi)^n \, t^{n/\alpha}} \int_H \exp\left\{ i\langle \frac{h}{t^{1/\alpha}}, v\rangle \right\} \, \langle \frac{h}{t^{1/\alpha}}, l\rangle \, \exp\left\{ -\frac{1}{t} \int_0^t \psi(e^{-rA^*} h) dr \right\} dh \\ &= \frac{1}{t^{1/\alpha}} \frac{1}{(2\pi)^n \, t^{n/\alpha}} \int_H \exp\left\{ i\langle \frac{v}{t^{1/\alpha}}, h\rangle \right\} \, \langle h, l\rangle \, \exp\left\{ -\frac{1}{t} \int_0^t \psi(e^{-rA^*} h) dr \right\} dh. \end{aligned}$$

Since (with the change of variable: $v/t^{1/\alpha} = w$)

$$\int_H |\phi_t(v)| dv = \frac{1}{t^{1/\alpha}} \frac{1}{(2\pi)^n} \int_H \Big| \int_H e^{i\langle w,h\rangle} \langle h, l\rangle \exp\left\{ -\frac{1}{t}\int_0^t \psi(e^{-rA^*}h)dr \right\} dh \Big| dw,$$

in order to prove (3.4) we need to show that

$$\|\varphi_t\|_{L^1(H)} \le C_1, \quad t \in (0,1], \tag{3.5}$$

where

$$\varphi_t(w) = \frac{1}{(2\pi)^n} \int_H e^{-i\langle w,h\rangle} \langle h, l\rangle \exp\left\{ -\frac{1}{t}\int_0^t \psi(e^{-rA^*}h)dr \right\} dh.$$

(3). Let us now show (3.5). Write $\psi = \psi_1 + \psi_2$,

$$\psi_1(u) = \int_{\{|y|\le 1\}} \big(1 - \cos\langle u, y\rangle\big)\nu(dy), \qquad \psi_2 = \psi - \psi_1,$$

so that

$$\varphi_t(w) = \frac{1}{(2\pi)^n} \int_H e^{-i\langle w,h\rangle} \langle h, l\rangle \, e^{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr} e^{-\frac{1}{t}\int_0^t \psi_2(e^{-rA^*}h)dr} dh.$$

Now consider the random variable

$$Y_t = \frac{1}{t^{1/\alpha}} \int_0^t e^{-(t-s)A} dZ_s^2, \quad t \in (0,1],$$

where $Z^2 = (Z_t^2)$ is a Lévy process having exponent ψ_2. It is easy to check that its law μ_t has characteristic function $e^{-\frac{1}{t}\int_0^t \psi_2(e^{-rA^*}h)dr}$, i.e.,

$$\hat{\mu}_t(h) = \exp\left\{ -\frac{1}{t}\int_0^t \psi_2(e^{-rA^*}h)dr \right\}, \quad h \in H.$$

Now suppose that there exists $g_t \in L^1(H)$, $t \in (0,1]$, such that

$$\hat{g}_t(h) = \langle h, l\rangle \exp\left\{ -\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr \right\}. \tag{3.6}$$

Then, by well known properties of the Fourier transfom (see Proposition 2.5 in [34]) we would get

$$\hat{g}_t \cdot \hat{\mu}_t = \widehat{g_t * \mu_t}$$

and, using the Fourier inversion formula,

$$\varphi_t(w) = (g_t * \mu_t)(w)$$

so that $\|\varphi_t\|_{L^1} \le \|g_t\|_{L^1}$, $t \in (0,1]$. Thus to prove (3.5) and get the assertion, it remains to show that (3.6) holds and moreover that

$$\|g_t\|_{L^1(H)} \le C_1, \quad t \in (0,1]. \tag{3.7}$$

(4). Now we show (3.6) and (3.7). Note that

$$\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\} = \exp\left\{-\frac{1}{t}\int_0^t dr \int_{\{|y|\le 1\}} \big(1-\cos(\langle e^{-rA^*}h, y\rangle)\big)\nu(dy)\right\}$$

$$= \exp\left\{-\frac{1}{t}\int_0^t \psi(e^{-rA^*}h)dr\right\}\ \exp\left\{\frac{1}{t}\int_0^t dr \int_{\{|y|>1\}} \big(1-\cos(\langle e^{-rA^*}h, y\rangle)\big)\nu(dy)\right\}$$

$$\le \exp\left\{2\nu(\{|y|>1\})\right\}\ \exp\left\{-\frac{C_\alpha}{t}\int_0^t |e^{-rA^*}h|^\alpha dr\right\}.$$

Since $|h| \le c_2|e^{-rA^*}h|$, $h\in H$, $r\in[0,T]$, it follows that

$$\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\} \le c_1 e^{-c_3|h|^\alpha},\quad h\in H,\ t\in(0,1]. \tag{3.8}$$

We find easily that $\psi_1 \in C^\infty(H)$ and so, using also (3.8) we deduce that the mapping $h\mapsto \langle h,l\rangle\, e^{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr}$ is in the Schwartz space $\mathcal{S}(H)$, for any $t\in(0,1]$. It follows that there exists $g_t\in\mathcal{S}(H)$ such that (3.6) holds. By the inversion formula,

$$g_t(w) = \frac{1}{(2\pi)^n}\int_H e^{-i\langle w,h\rangle}\,\langle h,l\rangle\,\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\}dh,\quad w\in H.$$

Now we show (3.7), by proving that for any multiindex $\beta=(\beta_1,\dots,\beta_n)\in\mathbb{Z}_+^n$, there exists c_T such that (with $w^\beta := w_1^{\beta_1}\cdots w_n^{\beta_n}$)

$$\sup_{w\in H}|w^\beta g_t(w)| = c_1 < \infty,\quad t\in]0,1] \tag{3.9}$$

(note that the constant c_1 is independent of t). Indeed once (3.9) is proved then

$$\|g_t\|_{L^1} \le c_1'\int_H \frac{1}{1+|w|^{2n}}dw = c_1'' < \infty.$$

We will check (3.9) only for $w^\beta = w_j$, i.e. $\beta=(0,\dots,1,\dots,0)$ with 1 in the j-th position. The proof in the general case is similar.

We have, integrating by parts and using estimate (3.8),

$$w_j\,g_t(w) = \frac{1}{(2\pi)^n}\int_H w_j e^{-i\langle w,h\rangle}\,\langle h,l\rangle\,\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\}dh$$

$$= \frac{i}{(2\pi)^n}\int_H \partial_{h_j}\big(e^{-i\langle w,h\rangle}\big)\,\langle h,l\rangle\,\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\}dh$$

$$= -\frac{i}{(2\pi)^n}\int_H e^{-i\langle w,h\rangle}\,l_j\,\exp\left\{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr\right\}dh$$

$$-\frac{i}{(2\pi)^n}\int_H e^{-i\langle w,h\rangle}\,\langle h,l\rangle\,e^{-\frac{1}{t}\int_0^t \psi_1(e^{-rA^*}h)dr}\Big(-\frac{1}{t}\int_0^t \langle D\psi_1(e^{-rA^*}h), e^{-rA^*}e_j\rangle dr\Big)dh.$$

Using (3.8) and the fact the $|D\psi_1(u)| \le c_5|u|$, $u \in H$, get easily that

$$\sup_{w\in H} |w_j\, g_t(w)| = c_1 < \infty, \quad t \in]0,1].$$

The proof is complete. $\square$

Step 4. Irreducibility. A Markov semigroup P_t is called *irreducible* at $t_0 > 0$ and $x \in H$ if for all non-empty open set $\Gamma \subset H$, we have

$$P_{t_0} 1_\Gamma(x) := P(t_0; x, \Gamma) > 0,$$

where $P(t_0; ., .) : H \times \mathcal{B}(H) \to [0,1]$ is the transition probability at the time t_0. We also call the underlying process X_t^x is irreducible at t_0.

We cannot argue as in the proof of [31, Theorem 5.3] since the drift F is only Hölder continuous. Note, however, that if we prove that the Ornstein-Uhlenbeck process $Z_A = (Z_A(t))$,

$$Z_A(t) = \int_0^t e^{A(t-s)} dZ_s \tag{3.10}$$

(starting at $x = 0$), is irreducible then we can obtain irreducibility for the solution X^x using the following quite general result of independent interest.

Proposition 3.3. *Assume that for each $t > 0$ the support of $Z_A(t)$ is the whole space. Then the process (X_t^x) is irreducible, for any $x \in H$.*

Proof. Fix $t > 0$, $a > 0$ and let $r > 0$ be any positive number. Then

$$X_{t+a} = e^{Aa} X_t + \int_t^{t+a} e^{A(t+a-s)} F(X_s) ds + \int_t^{t+a} e^{A(t+a-s)} dZ_s.$$

Let z be any element in the support of the distribution of the random variable $e^{Aa}X_t$. Then, by the very definition, the event

$$B = \{|e^{Aa}X_t - z| < r/3\}$$

is of positive probability. Since $||F||_0 < \infty$, there exists $c > 0$ such that for each $t \ge 0$ and for each positive b with probability 1

$$\Big| \int_t^{t+b} e^{A(t+b-s)} F(X_s) ds \Big| \le cb,$$

In particular the above inequality holds for $b = a$. Let us fix x and y in H. Then

$$X_{t+a} - y = (e^{Aa}X_t - z) + \int_t^{t+a} e^{A(t+a-s)} F(X_s) ds + \Big(\int_t^{t+a} e^{A(t+a-s)} dZ_s - y + z \Big)$$

Define an event C

$$C = \Big\{ \Big| y - z - \int_t^{t+a} e^{A(t+a-s)} dZ_s \Big| < r/3 \Big\},$$

which, by the assumption, is of positive probability. The events B and C are independent and therefore the probability of $B \cap C$ is positive. On this event, and thus with positive probability, we have the estimate:

$$|X_{t+a} - y| \leq \frac{r}{3} + ca + \frac{r}{3}.$$

Starting from number a such that $ca < r/3$ we have with positive probability

$$|X_{t+a} - y| \leq r.$$

To finish the proof we should replace $t + a$ and t with t and $t - a$. $\square$

By the previous result we know that the proof of Step 4 is complete once the following theorem has been proved.

Theorem 3.4. *Let $H = \mathbb{R}^n$. Assume that $Z = (Z_t)$ is an n-dimensional symmetric non-degenerate α-stable process, $\alpha \in (0, 2)$. Consider any real $n \times n$ matrix A. Then, for all $t > 0$, $X(t) = Z_A(t)$ (given in (3.10) and starting at $x = 0$) is irreducible.*

Proof. By the non-degenerate assumption (2.4) there exists n points $a_1, \ldots, a_n \in S$ such that $a_k \in supp(\mu)$ for $1 \leq k \leq n$ and $span\{a_1, \ldots, a_n\} = \mathbb{R}^n$. Since μ is symmetric, $-a_1, \ldots, -a_n \in supp(\mu)$. It is clear that for any $\varepsilon > 0$, $\mu(B_s(\pm a_k, \varepsilon)) > 0$ where $B_s(a_k, \varepsilon) = \{y \in S; |y - a_k| < \varepsilon\}$.

For each k, let us now consider the affines $\mathcal{F}_{k,+} := \{ra_k, r > 1\}$ and $\mathcal{F}_{k,-} := \{-ra_k, r > 1\}$. For any point $y_k \in \{ra_k, -\infty < r < \infty\}$, there exist $y_{k,+} \in \mathcal{F}_{k,+}$ and $y_{k,-} \in \mathcal{F}_{k,-}$ such that $y_k = y_{k,+} + y_{k,-}$. Define $\mathcal{F}^+_{k,\varepsilon} := \{(x, r) : x \in B_s(a_k, \varepsilon), r > 1\}$, $\mathcal{F}^-_{k,\varepsilon} = \{(x, r) : x \in B_s(-a_k, \varepsilon), r > 1\}$, Take $\varepsilon > 0$ small enough to make $\mathcal{F}^{\pm}_{i,\varepsilon} \cap \mathcal{F}^{\pm}_{j,\varepsilon} = \emptyset$ for $i \neq j$ and $\mathcal{F}^+_{i,\varepsilon} \cap \mathcal{F}^-_{i,\varepsilon} = \emptyset$ for each i.

Decompose ν as the sum of two measures ν_1, ν_2 such that

$$\nu = \nu_1 + \nu_2,$$

and one of the measures, say $\nu_1 = \nu 1_{(\cup_{k=1}^n \mathcal{F}^+_{k,\varepsilon}) \cup (\cup_{k=1}^n \mathcal{F}^-_{k,\varepsilon})}$, is finite. We can assume that the process Z is the sum of two independent Lévy processes $Z^1$and Z^2, with the Lévy measures ν_1 and ν_2 respectively. Note that

$$X^1(t) := \int_0^t e^{A(t-s)} dZ^1_s, \quad t \geq 0,$$

is a compound Poisson process. Since $supp(\mu_1) \subset supp(\mu_1 * \mu_2)$ for any two measures μ_1 and μ_2, it is enough to prove the irreducibility of X^1.

Let us fix $t > 0$, $y \in H$ and $r > 0$. It is enough to show that

$$\mathbb{P}(|X^1(t) - y| < r) > 0.$$

Let M be a number such that for all $s \in (0, 1)$:

$$|e^{As} z| \leq M|z|, \quad |(e^{As} - I)z| \leq Ms|z|, \ z \in H.$$

Write $y = \sum_{k=1}^{n} y_k a_k$ where $y_1, \ldots, y_n \in \mathbb{R}$, for each k we have two points $y_{k,+} \in \mathcal{F}_{k,+}$ and $y_{k,-} \in \mathcal{F}_{k,-}$ and positive number $\delta < 1$ such that:

$$y_{k,+} + y_{k,-} = y_k a_k, \quad \delta M\left(|y_{k,+}| + |y_{k,-}|\right) < \frac{r}{2n}.$$

Choose $\varepsilon > 0$ sufficiently small, the probability that the process Z^1 will perform exactly $2n$ jumps $\xi_{1,-} \in \mathcal{F}^-_{1,\varepsilon}, \xi_{1,+} \in \mathcal{F}^+_{1,\varepsilon}, \ldots, \xi_{n,-} \in \mathcal{F}^-_{n,\varepsilon}, \xi_{n,+} \in \mathcal{F}^+_{n,\varepsilon}$ before t at moments $\tau_{1,-} < \tau_{1,+} < \tau_{2,-} < \tau_{2,+} < \ldots < \tau_{n,-} < \tau_{n,+} < t$ such that

$$\tau_{1,-} > t - \delta, \quad |\xi_{k,-} - y_{k,-}| < \frac{r}{4nM}, \ \ldots, \ |\xi_{k,+} - y_{k,+}| < \frac{r}{4nM}, \quad k = 1, \cdots, n,$$

is positive. Therefore, at least with the same probability, the following relations hold:

$$\begin{aligned}
&\left|\int_0^t e^{(t-s)A} dZ^1_s - y\right| \\
&= \left|\sum_{j=1}^{n} e^{A(t-\tau_{j,-})}\xi_{j,-} + e^{A(t-\tau_{j,+})}\xi_{j,+} - y\right| \\
&= \left|\sum_{j=1}^{n} e^{A(t-\tau_{j,-})}(\xi_{j,-} - y_{j,-}) + e^{A(t-\tau_{j,+})}(\xi_{j,+} - y_{j,+})\right| \\
&\quad + \left|\sum_{j=1}^{n} (e^{A(t-\tau_{j,-})} - I) y_{j,-} + (e^{A(t-\tau_{j,+})} - I) y_{j,+}\right| \\
&\leq \sum_{j=1}^{n} M\left(|\xi_{j,-} - y_{j,-}| + |\xi_{j,+} - y_{j,+}|\right) + \sum_{j=1}^{n} \delta M\left(|y_{j,-}| + |y_{j,+}|\right) < r.
\end{aligned}$$

This finishes the proof. □

The proof of Theorem 2.4 is now complete.

4. Estimates of the solution, dim H $= \infty$

This section contains some preparation for the proof of Theorem 2.8, giving some estimates for the solution (2.6). Recall that the Ornstein-Uhlenbeck process is defined by

$$Z_A(t) = \int_0^t e^{A(t-s)} dZ_s = \sum_{k \geq 1} Z_{A,k}(t) e_k \tag{4.1}$$

where

$$Z_{A,k}(t) = \int_0^t e^{-\gamma_k(t-s)} \beta_k dz_k(s).$$

For any $\varepsilon \geq 0$, define

$$H^\varepsilon = \left\{ x = \sum_{k\geq 1} x_k e_k \in H : \sum_{k\geq 1} \gamma_k^{2\varepsilon} |x_k|^2 < \infty \right\}.$$

Note that H^ε coincides with the domain of $(-A)^\varepsilon$ and that $H^0 = H$. Denote further by $|\cdot|_\varepsilon$ the norm of H^ε. For $x \in H^\varepsilon$ and $R > 0$, we denote by $B_\varepsilon(x, R)$ the closed ball in H^ε of radius R centered at x. We shall write $B_\varepsilon(R) := B_\varepsilon(0, R)$ and $B(x, R) := B_0(x, R)$.

Lemma 4.1. *The following assertions hold:*
(i) $Z_A(t) \in H^\varepsilon$ a.s. for all $t > 0$.
(ii) For any $p \in (0, \alpha)$, we have

$$\mathbb{E}|Z_A(t)|_\varepsilon^p \leq C \left(\sum_{k\geq 1} |\beta_k|^\alpha \, \frac{1 - e^{-\alpha\gamma_k t}}{\alpha \gamma_k^{1-\alpha\varepsilon}} \right)^{\frac{p}{\alpha}}, \tag{4.2}$$

where $C = C(\alpha, p) > 0$.

Proof. (i). By (4.7) in [31] we have

$$\mathbb{E}[e^{i\lambda Z_{A,k}(t)}] = e^{-|\lambda|^\alpha c_k^\alpha(t)},$$

where $c_k(t) = \beta_k \left(\frac{1-e^{-\alpha\gamma_k t}}{\alpha\gamma_k} \right)^{1/\alpha}$. Hence, $Z_{A,k}(t)$ has the same distribution as $c_k(t)\xi_k$ for all $k \geq 1$ where $\{\xi_k\}_{k\geq 1}$ are i.i.d. with $\mathbb{E}[e^{i\lambda\xi_1}] = e^{-|\lambda|^\alpha}$. We shall use Proposition 3.3 in [31], which claims that

$$(q_k \xi_k)_{k\geq 1} \in l^2 \;\; a.s. \Longleftrightarrow \sum_{k\geq 1} |q_k|^\alpha < \infty,$$

where $q_k \in \mathbb{R}$ for all k. From this it is easy to check that

$$\sum_{k\geq 1} (\gamma_k)^{2\varepsilon} \left[c_k(t)\xi_k \right]^2 < \infty \;\; a.s. \Longleftrightarrow \sum_{k\geq 1} \frac{\beta_k^\alpha}{\gamma_k^{1-\alpha\varepsilon}} < \infty.$$

Since $Z_A(t)$ has the same distribution as $(c_k(t)\xi_k)_{k\geq 1}$, (i) is clearly true.

(ii). We follow the argument in the proof of [31, Theorem 4.4]. Take a Rademacher sequence $\{r_k\}_{k\geq 1}$ in a new probability space $(\Omega', \mathcal{F}', \mathbb{P}')$, i.e. $\{r_k\}_{k\geq 1}$ are i.i.d. with $\mathbb{P}\{r_k = 1\} = \mathbb{P}\{r_k = -1\} = \frac{1}{2}$. By the following Khintchine inequality: for any $p > 0$, there exists some $C(p) > 0$ such that for arbitrary real sequence $\{h_k\}_{k\geq 1}$,

$$\left(\sum_{k\geq 1} h_k^2 \right)^{1/2} \leq C(p) \left(\mathbb{E}' \left| \sum_{k\geq 1} r_k h_k \right|^p \right)^{1/p}.$$

By this inequality, one has

$$
\begin{aligned}
\mathbb{E}|Z_A(t)|_\varepsilon^p = \mathbb{E}\left(\sum_{k\geq 1}\gamma_k^{2\varepsilon}|Z_{A,k}(t)|^2\right)^{p/2} &\leq C\mathbb{E}\mathbb{E}'\left|\sum_{k\geq 1} r_k\gamma_k^\varepsilon Z_{A,k}(t)\right|^p \\
&= C\mathbb{E}'\mathbb{E}\left|\sum_{k\geq 1} r_k\gamma_k^\varepsilon Z_{A,k}(t)\right|^p,
\end{aligned}
\tag{4.3}
$$

where $C = C^p(p)$. For any $\lambda \in \mathbb{R}$, by the fact of $|r_k| = 1$ and formula (4.7) of [31], one has

$$
\begin{aligned}
\mathbb{E}\exp\left\{i\lambda\sum_{k\geq 1} r_k\gamma_k^\varepsilon Z_{A,k}(t)\right\} &= \exp\left\{-|\lambda|^\alpha\sum_{k\geq 1}|\beta_k|^\alpha\gamma_k^{\varepsilon\alpha}\int_0^t e^{-\alpha\gamma_k(t-s)}ds\right\} \\
&= \exp\left\{-|\lambda|^\alpha\sum_{k\geq 1}\gamma_k^{\varepsilon\alpha}c_k^\alpha(t)\right\}.
\end{aligned}
$$

Now we use (3.2) in [31]: if X is a symmetric random variable satisfying $\mathbb{E}\left[e^{i\lambda X}\right] = e^{-\sigma^\alpha|\lambda|^\alpha}$ for some $\alpha \in (0,2)$ and any $\lambda \in \mathbb{R}$, then $\mathbb{E}|X|^p = C(\alpha,p)\sigma^p$ for all $p \in (0,\alpha)$. Since $\sum_{k\geq 1}\gamma_k^{\varepsilon\alpha}c_k^\alpha(t) < \infty$, it is clear to see

$$
\mathbb{E}\left|\sum_{k\geq 1} r_k\gamma_k^\varepsilon Z_{A,k}(t)\right|^p = C(\alpha,p)\left(\sum_{k\geq 1}|\beta_k|^\alpha\ \frac{1-e^{-\alpha\gamma_k t}}{\alpha\gamma_k^{1-\alpha\varepsilon}}\right)^{\frac{p}{\alpha}},
$$

from which and (4.3) we get (4.2). $\square$

Lemma 4.2. *Let (X_t^x) be the solution to Eq. (1.1) with $x \in H^\varepsilon$. For any $p \in (0,\alpha)$, there exist some constants $C_1 = C_1(p) > 0$ and $C_2 = C_2(p,\varepsilon,\gamma,\beta,\|F\|_0) > 1$ such that*

$$
\mathbb{E}|X_t^x|_\varepsilon^p \leq C_1 e^{-p\gamma_1 t}|x|_\varepsilon^p + C_2, \qquad \forall\ t > 0,
\tag{4.4}
$$

where $C_1(p) \leq 1$ for $p \in (0,1]$ and $C_1(p) = 3^{p-1}$ otherwise.

Proof. By (2.6), we have

$$
X_t = e^{At}x + \int_0^t e^{A(t-s)}F(X_s)ds + Z_A(t).
$$

It is easy to see

$$
|e^{At}x|_\varepsilon \leq e^{-\gamma_1 t}|x|_\varepsilon.
$$

By the easy inequality $|(-A)^\sigma e^{At}|_{L(H)} \leq C(\sigma)t^{-\sigma}$, $t \geq 0$, $\sigma > 0$, one has

$$
\begin{aligned}
\left|\int_0^t e^{A(t-s)}F(X_s)ds\right|_\varepsilon &\le \int_0^t |(-A)^\varepsilon e^{A(t-s)/2}|_{L(H)}|e^{A(t-s)/2}F(X_s)|ds \\
&\le C(\varepsilon)\int_0^t (t-s)^{-\varepsilon}e^{-\gamma_1(t-s)/2}ds\|F\|_0 \\
&\le C(\varepsilon,\gamma_1)\|F\|_0.
\end{aligned}
$$

for all $t>0, x\in H$ and $\omega\in\Omega$. Furthermore, from (4.2),

$$\mathbb{E}|Z_A(t)|_\varepsilon^p \le C(p,\alpha,\beta,\gamma,\varepsilon), \qquad \forall\, p\in(0,\alpha).$$

Now we use the following trivial inequality: for any $a,b,c\ge 0$,

$$
\begin{aligned}
&(a+b+c)^p \le (a^p+b^p+c^p)\,, \quad p\le 1; \\
&(a+b+c)^p \le 3^{p-1}(a^p+b^p+c^p)\,, \quad p>1.
\end{aligned}
$$

Combining the above three estimates and the inequality, we can easily see that (4.4) is true. □

Lemma 4.3. *Let (X_t^x) be the solution to Eq. (1.1). For any $p\in(0,\alpha)$, we have*

$$\mathbb{E}|X_t^x|_\varepsilon^p \le C\left(t^{-\varepsilon p}|x|^p + t^{p-\varepsilon p}\|F\|_0^p + 1\right) \tag{4.5}$$

for all $t>0$, where $C=C(p,\alpha,\beta,\gamma,\varepsilon)$.

Proof. By (2.6) and (4.2), we have

$$
\begin{aligned}
\mathbb{E}|X_t^x|_\varepsilon^p &\le C_1\left[|A^\varepsilon e^{At}x|^p + \mathbb{E}\left(\int_0^t |A^\varepsilon e^{A(t-s)}|_{L(H)}|F(X_s^x)|ds\right)^p + \mathbb{E}|Z_A(t)|_\varepsilon^p\right] \\
&\le C_2\left[t^{-\varepsilon p}|x|^p + \left(\int_0^t (t-s)^{-\varepsilon}ds\right)^p \|F\|_0^p + 1\right] \\
&\le C_3\left(t^{-\varepsilon p}|x|^p + t^{p-\varepsilon p}\|F\|_0^p + 1\right),
\end{aligned}
$$

where $C_1=C_1(p)$ and $C_i=C_i(p,\alpha,\beta,\gamma,\varepsilon)$ $(i=2,3)$. □

5. Proof of Theorem 2.8 by Harris' approach, dim H $=\infty$

Let us split the proof into the following three steps.

Step 1. The existence of an invariant measure was established in [29]. Let us prove that any invariant measure μ has finite p^{th} moment $(p<\alpha)$:

$$\mathfrak{m}_p(\mu) := \int_H |x|^p\mu(dx) < \infty \quad \text{for any } p\in(0,\alpha). \tag{5.1}$$

Indeed, by (2.6) and the trivial inequality

$$(a+b)\wedge c \le a\wedge c + b\wedge c, \quad a,b,c\in\mathbb{R}^+,$$

for all $t>0$ and $n\in\mathbb{N}$, we have

$$|X_t^x|^p\wedge n \le \left[\big(C_p e^{-p\gamma_1 t}|x|^p\big)\wedge n + C_p\left|\int_0^t e^{A(t-s)}F(X_s)ds\right|^p + C_p|Z_A(t)|^p\right].$$

Using a similar calculation as in Lemma 4.2, we obtain

$$\mathbb{E}\big(|X_t^x|^p\wedge n\big) \le \big(C_p e^{-p\gamma_1 t}|x|^p\big)\wedge n + C,$$

where $C=C(\alpha,\beta,\gamma,p,\|F\|_0)$. Integrating this inequality against $\mu(dx)$, we get

$$\mu(|x|^p\wedge n)\le \mu\left[\big(C_p e^{-p\gamma_1 t}|x|^p\big)\wedge n\right]+C.$$

Passing to the limit first as $t\to\infty$ and then as $n\uparrow\infty$, we complete the proof of (5.1).

Step 2. To prove the uniqueness of an invariant measure and inequality (2.9), it suffices to show that

$$\|P_{kT}(x_1,\cdot)-P_{kT}(x_2,\cdot)\|_{\mathrm{TV}} \le C\,(1+|x_1|^p+|x_2|^p)e^{-ckT}, \quad x_1,x_2\in H, \tag{5.2}$$

where C and c are positive constants not depending on x_1, x_2, and k. Indeed, if (5.2) is established, then for any measures $\nu_1,\nu_2\in\mathcal{P}(H)$ with finite p^{th} moment we derive

$$\|P_{kT}^*\nu_1-P_{kT}^*\nu_2\|_{\mathrm{TV}} \le C\,\big(1+\mathfrak{m}_p(\nu_1)+\mathfrak{m}_p(\nu_2)\big)e^{-ckT}, \quad k\in\mathbb{N}. \tag{5.3}$$

This implies, in particular, that an invariant measure is unique. Moreover, writing any $t\ge 0$ in the form $t=kT+s$ with $0\le s<T$ and using inequalities (5.3) and (4.4), we obtain

$$\begin{aligned}\|P_t^*\nu_1-P_t^*\nu_2\|_{\mathrm{TV}} &= \|P_{kT}^*(P_s^*\nu_1)-P_{kT}^*(P_s^*\nu_2)\|_{\mathrm{TV}}\\ &\le C\,\big(1+\mathfrak{m}_p(P_s^*\nu_1)+\mathfrak{m}_p(P_s^*\nu_2)\big)e^{-ckT}\\ &\le C_1\big(1+\mathfrak{m}_p(\nu_1)+\mathfrak{m}_p(\nu_2)\big)e^{-ct}.\end{aligned}$$

This estimate readily implies the required inequality (2.9).

Note that (5.2) holds if we are able to apply Theorem 2.10 to equation (1.1) with $V(x)=|x|^p$ and $p\in(0,\alpha)$. Indeed, once this is done, we obtain that there exists $T>0$ such that

$$\begin{aligned}\|P_{kT}(x_1,\cdot)-P_{kT}(x_2,\cdot)\|_{TV} &\le \int_H (1+V(x))|P_{kT}^*\delta_{x_1}-P_{kT}^*\delta_{x_2}|(dx)\\ &\le \beta^k\int_H(1+V(x))|\delta_{x_1}-\delta_{x_2}|(dx)\\ &\le 2\beta^k\big(1+|x_1|^p+|x_2|^p\big), \quad k\ge 1.\end{aligned}$$

This immediately implies (5.2).

Step 3. It remains to check the conditions (i) and (ii) in Theorem 2.10. Choosing $V(x) = |x|^p$ with $p \in (0, \alpha)$ and applying Lemma 4.2 with $\varepsilon = 0$ and $T_0 > \frac{\log(1+C_1)}{p\gamma_1}$, one immediately get (i).

To prove (ii), we shall use the following auxiliary lemma, which has been proved in [31].

Lemma 5.1 (Theorem 5.4, [31]). *Let (X_t^x) be the solution to Eq. (1.1). Then (X_t^x) is irreducible on H, i.e., for any $t > 0$ and $B(y, r)$ with arbitrary $y \in H$ and $r > 0$, we have*

$$\mathbb{P}\left(X_t^x \in B(y, r)\right) > 0. \tag{5.4}$$

Let x and y satisfy $|x|^p + |y|^p \le R$. By Lemma 4.3 we know that, for any fixed $T_0 > 0$,

$$\mathbb{E}[|X_{T_0}^x|_\epsilon^p] + \mathbb{E}[|X_{T_0}^y|_\epsilon^p] \le C(|x|^p + |y|^p + 1) \le C_1.$$

It follows that there exists some $R_1 > 0$ such that

$$\mathbb{P}\left(|X_{T_0}^x|_\varepsilon \le R_1\right) > 1/2, \quad \mathbb{P}\left(|X_{T_0}^y|_\varepsilon \le R_1\right) > 1/2.$$

Since $\gamma_k \to \infty$, $B_\varepsilon(M)$ is compact in H. By Lemma 5.1, for any $r > 0$ we have some $\delta(r) > 0$ such that

$$\inf_{x \in B_\varepsilon(R_1)} \mathbb{P}\left(X_{T_0}^x \in B(r)\right) \ge 2\delta. \tag{5.5}$$

By Markov property and the above three inequalities,

$$\mathbb{P}\left(X_{2T_0}^x \in B(r)\right) > \delta, \quad \mathbb{P}\left(X_{2T_0}^y \in B(r)\right) > \delta.$$

Without loss of generality, in the next computations we assume that X_t^x and X_t^y are independent (this is true if the driven noises of X_t^x and X_t^y are independent). By Markov property and Theorem 2.5,

$$\begin{aligned}
&\|P_{3T_0}^*\delta_x - P_{3T_0}^*\delta_y\|_{TV} = \frac{1}{2}\sup_{\|\phi\|_0 \le 1} |\mathbb{E}[P_{T_0}\phi(X_{2T_0}^x) - P_{T_0}\phi(X_{2T_0}^y)]| \\
&\le \mathbb{P}\{X_{2T_0}^x \notin B(r)\} + \mathbb{P}\{X_{2T_0}^y \notin B(r)\} \\
&\quad + \frac{1}{2}\mathbb{E}\left\{\sup_{\|\phi\|_0 \le 1} |P_{T_0}\phi(X_{2T_0}^x) - P_{T_0}\phi(X_{2T_0}^y)|, X_{2T_0}^x \in B(r), X_{2T_0}^y \in B(r)\right\} \\
&\le 2 - \mathbb{P}\{X_{2T_0}^x \in B(r)\} - \mathbb{P}\{X_{2T_0}^y \in B(r)\} + Cr\mathbb{P}\{X_{2T_0}^x \in B(r)\}\mathbb{P}\{X_{2T_0}^y \in B(r)\} \\
&\le 2 - \delta,
\end{aligned}$$

as $r > 0$ is sufficiently small. This finishes the proof.

6. Proof of Theorem 2.8 by coupling, dim H $= \infty$

In this section, we shall prove Theorem 2.8 by the Doeblin coupling argument, which gives much more intuitions for understanding the way that the dynamics converges to the ergodic measure.

6.1. **Construction of the coupling chain.** Let us first give some preliminary about maximal coupling.

Definition 6.1. Let μ_1, $\mu_2 \in \mathcal{P}(H)$. A pair of random variables (ξ_1, ξ_2) defined on the same probability space is called a *coupling* for (μ_1, μ_2) if $\mathcal{D}(\xi_i) = \mu_i$ for $i = 1, 2$, where $\mathcal{D}(\cdot)$ denotes the distribution of random variable. A coupling (ξ_1, ξ_2) is said to be *maximal* if

$$\mathbb{P}\{\xi_1 \neq \xi_2\} = \|\mu_1 - \mu_2\|_{\mathrm{TV}}, \tag{6.1}$$

and the random variable ξ_1 and ξ_2 conditioned on the event $N := \{\xi_1 \neq \xi_2\}$ are independent. The latter condition means that, for any $A_1, A_2 \in \mathcal{B}(H)$, one has

$$\mathbb{P}\big(\{\xi_1 \in A_1\} \cap \{\xi_2 \in A_2\} \,|\, N\big) = \mathbb{P}\big(\xi_1 \in A_1 \,|\, N\big)\, \mathbb{P}\big(\xi_2 \in A_2 \,|\, N\big).$$

In what follows, we shall the need the following lemma whose proof can be found in [39, 18, 17].

Lemma 6.2. *For any two measures $\mu_1, \mu_2 \in \mathcal{P}(H)$, there exists a maximal coupling. Moreover, if (ξ_1, ξ_2) is a maximal coupling, then we have* [1]

$$\mathbb{P}(\xi_1 \in A, \xi_2 \in A) \geq \mathbb{P}(\xi_1 \in A)\, \mathbb{P}(\xi_2 \in A), \qquad \forall\ A \in \mathcal{B}(H). \tag{6.2}$$

Now let us construct an auxiliary Markov chain in the extended phase space $H \times H$. Let $T > 0$ be some fixed real number to be chosen later. For any $x := (x_1, x_2) \in H \times H$, denote by $M(x) = (M_1(x), M_2(x))$ the maximal coupling of $(P_T)^*\delta_{x_1}$ and $(P_T)^*\delta_{x_2}$. Let us define a transition function $\tilde{P}_T(x, \cdot)$ on the space $H \times H$ such that

$$\tilde{P}_T(x; A_1 \times A_2) = \begin{cases} P_T(x_1, A_1 \cap A_2) \ \text{ if } \ x_1 = x_2, \\ \mathcal{D}(M_1(x), M_2(x))(A_1 \times A_2) \ \text{ if } x_1, x_2 \in B(r) \text{ with } x_1 \neq x_2, \\ P_T(x_1, A_1) P_T(x_2, A_2) \quad \text{otherwise}, \end{cases}$$

where $A_1, A_2 \in \mathcal{B}(H)$ are arbitrary sets, $P_T(x_i, \cdot)$ is the transition probability of $X_T^{x_i}$ for $i = 1, 2$, and $\mathcal{D}(\cdot)$ denotes the distribution of a random variable. For any $A \in \mathcal{B}(H \times H)$, $\tilde{P}_T(x, A)$ is uniquely defined by a classical approximation procedure. Now the transition function $\tilde{P}_T(x, \cdot)$ is well defined.

6.2. **Hitting times τ^ε and τ.** We denote by $(X_1(kT), X_2(kT))_{k \in \mathbb{Z}^+}$ the Markov chain whose transition function is equal to $\tilde{P}_T(x, \cdot)$; here $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$. Clearly, for each $i = 1, 2$, $(X_i(kT))$ is also a Markov chain and has the same distribution as $(X_{kT}^{x_i})$. We shall write $X(kT) = (X_1(kT), X_2(kT))$ for $k \in \mathbb{Z}^+$.

For any $r, M > 0$, define the hitting times

$$\tau^\varepsilon = \inf\{kT; |X_1(kT)|_\varepsilon + |X_2(kT)|_\varepsilon \leq M\}, \tag{6.3}$$

$$\tau = \inf\{kT; |X_1(kT)| + |X_2(kT)| \leq r\}, \tag{6.4}$$

[1] Inequality (6.2) is true for any pair of random variables that are independent conditioned on the event $\{\xi_1 \neq \xi_2\}$.

where $\varepsilon \in (0,1)$ is the constant in Assumption 2.2. Recall that the infimum over an empty set is equal to $+\infty$.

6.2.1. *Estimates of the hitting time τ^ε.* The main result of this subsection is the following theorem, which is in fact a step for estimating τ.

Theorem 6.3. *For any $p \in (0,\alpha)$ and sufficiently large $T > 0$ there is a constant $M = M(p,T,\alpha,\beta,\gamma,\varepsilon)$ such that, for any $x = (x_1,x_2) \in H \times H$,*

$$\mathbb{E}_x \left[e^{\eta\tau^\varepsilon}\right] \le C\big(1 + |x_1|^p + |x_2|^p\big) \tag{6.5}$$

where $\eta > 0$ is sufficiently small, and $C = C(p,T,\alpha,\beta,\gamma,\varepsilon,\|F\|_0,\eta)$

To prove Theorem 6.3, we first establish two auxiliary lemmas.

Lemma 6.4. *For any $p \in (0,\alpha)$, the Markov chain $(X(kT))$ satisfies the inequality*

$$\mathbb{E}_x\big(|X_1(T)|_\varepsilon^p + |X_2(T)|_\varepsilon^p\big) \le C_1 e^{-p\gamma_1 T}\big(|x_1|_\varepsilon^p + |x_2|_\varepsilon^p\big) + 2C_2,$$

where C_1 and C_2 are the same as in Lemma 4.2.

Proof. By definition of coupling and Lemma 4.2, we have

$$\mathbb{E}_x|X_i(T)|_\varepsilon^p = \mathbb{E}|X_T^{x_i}|_\varepsilon^p \le C_1(p)e^{-p\gamma_1 T}|x_i|_\varepsilon^p + C_2$$

for $i = 1,2$. From the above inequality, we complete the proof. □

Lemma 6.5. *For any $p \in (0,\alpha)$ and sufficiently large $T > 0$, there exist positive constants $q = q(p,\gamma) \in (0,1)$ and $M = M(p,T,\alpha,\beta,\gamma,\|F\|_0,\varepsilon)$ such that*

$$\mathbb{P}_x(\tau^\varepsilon > kT) \le q^k\left(1 + |x_1|_\varepsilon^p + |x_2|_\varepsilon^p\right) \quad \textit{for any } x = (x_1,x_2) \in H^\varepsilon \times H^\varepsilon. \tag{6.6}$$

Proof. The proof follows the idea in [8]. Let us take $T > 0$ so large that the coefficient in front of $|x|_\varepsilon^p$ in inequality (4.4) is smaller than 1. In this case, setting $\mathbb{P} = \mathbb{P}_x$, $\mathbb{E} = \mathbb{E}_x$, and

$$|x|_\varepsilon^p = |x_1|_\varepsilon^p + |x_2|_\varepsilon^p,$$

we can write

$$\mathbb{E}\big[|X(kT+T)|_\varepsilon^p \,\big|\mathcal{F}_{kT}\big] \le q^2\,|X(kT)|_\varepsilon^p + 2C_2 \tag{6.7}$$

where $q > 0$ is defined by the relation $q^2 = C_1 e^{-p\gamma_1 T} < 1$. By Chebyshev inequality,

$$\mathbb{P}\left(|X(kT+T)|_\varepsilon > M|\mathcal{F}_{kT}\right) \le \frac{q^2}{M^p}|X(kT)|_\varepsilon^p + \frac{2C_2}{M^p}. \tag{6.8}$$

Denote

$$B_k = \{|X(jT)|_\varepsilon > M; j = 0,\dots,k\}$$

and

$$p_k = \mathbb{P}(B_k), \quad e_k = \mathbb{E}\big(|X(kT)|_\varepsilon^p 1_{B_k}\big),$$

integrating (6.8) over B_k, one has

$$p_{k+1} \le \frac{q^2}{M^p} e_k + \frac{2C_2}{M^p} p_k. \tag{6.9}$$

Moreover, by integrating (6.7) over B_k,

$$e_{k+1} \le \mathbb{E}\big(|X(kT+T)|_\varepsilon^p 1_{B_k}\big) \le q^2 e_k + 2C_2 p_k. \tag{6.10}$$

From (6.9) and (6.10), one has

$$\begin{pmatrix} e_{k+1} \\ p_{k+1} \end{pmatrix} \le \begin{pmatrix} q^2 & 2C_2 \\ \frac{q^2}{M^p} & \frac{2C_2}{M^p} \end{pmatrix} \begin{pmatrix} e_k \\ p_k \end{pmatrix}, \tag{6.11}$$

which clearly implies

$$q^2 e_{k+1} + 2C_2 p_{k+1} \le \left(q^2 + \frac{2C_2}{M^p}\right)(q^2 e_k + 2C_2 p_k) \tag{6.12}$$

We can choose $M = M(p, T, \alpha, \beta, \gamma, \varepsilon, \|F\|_0)$ so that

$$q^2 + 2C_2/M^p \le q.$$

Thus we clearly have from (6.12)

$$q^2 e_k + 2C_2 p_k \le q^k \left(q^2 e_0 + 2C_2 p_0\right),$$

This inequality, together with the easy fact $p_k = \mathbb{P}_x(\tau^\varepsilon > kT)$, immediately implies the required estimate (6.6) since $C_2 > 1$ in inequality (4.4). □

Proof of Theorem 6.3. By the definition of coupling and (4.5), for any $p \in (0, \alpha)$ we have

$$\mathbb{E}_x\big(|X_1(T)|_\varepsilon^p + |X_2(T)|_\varepsilon^p\big) = \mathbb{E}|X_T^{x_1}|_\varepsilon^p + \mathbb{E}|X_T^{x_2}|_\varepsilon^p \le C_4\left(1 + |x_1|^p + |x_2|_\varepsilon^p\right) \tag{6.13}$$

where $C_4 = C_4(p, T, \alpha, \beta, \gamma, \varepsilon, \|F\|_0)$.

For any $x = (x_1, x_2) \in H \times H$, by Markov property, (6.6) and the above inequality, we easily have

$$\begin{aligned} \mathbb{E}_x\left[e^{\eta\tau^\varepsilon}\right] &= \mathbb{E}_x\left(e^{\eta\tau^\varepsilon} 1_{\{\tau^\varepsilon \le T\}}\right) + \mathbb{E}_x\left(e^{\eta\tau^\varepsilon} 1_{\{\tau^\varepsilon > T\}}\right) \\ &\le e^{\eta T} + \mathbb{E}_x\left\{1_{\{\tau^\varepsilon > T\}} \mathbb{E}_{X(T)}\left[e^{\eta\tau^\varepsilon}\right]\right\} \\ &\le e^{\eta T} + C_5 \mathbb{E}_x\left[1 + |X_1(T)|_\varepsilon^p + |X_2(T)|_\varepsilon^p\right] \\ &\le C_6(1 + |x_1|^p + |x_2|^p) \end{aligned} \tag{6.14}$$

where $C_i = C_i(p, \alpha, \eta, \gamma, \beta, \varepsilon, \|F\|_0, T)$ $(i = 5, 6)$.

□

6.2.2. *Estimates of the hitting time τ.*

Theorem 6.6. *For any $p \in (0, \alpha)$ and sufficiently large $T > 0$, there exist positive constants $\lambda = \lambda(T, p, \alpha, \beta, \gamma, \|F\|_0, r)$ and $C = C(p, \alpha, \beta, \gamma, \|F\|_0, r, T)$ such that*

$$\mathbb{E}_x[e^{\lambda\tau}] \le C(1 + |x_1|^p + |x_2|^p). \tag{6.15}$$

The key point of the proof is to use Theorem 6.3 and Lemma 6.7 below. The argument is quite general, for simplicity, let us give its heuristic idea by using (X_{kT}), (note the difference between X_{kT} and $X(kT)$), as follows:

(i) Since $B_\varepsilon(M)$ is compact in H, by irreducibility and uniform strong Feller property we have that $\inf_{z\in B_\varepsilon(0,M)} P_T(z, B(r)) = p > 0$. Therefore, as long as X_{kT} is in $B_\varepsilon(M)$, it has the probability at least p to jump into $B(r)$ at $(k+1)T$.

(ii) Suppose that (X_{kT}) enters $B_\varepsilon(M)$ for j times *before* it jumps into $B(r)$, by strong Markov property and (i) this event happens with some probability less than $(1-p)^j$.

(iii) If $\tau = kT$ for some large kT (i.e. the process first enters $B(r)$ at kT), j is also large. Thus $\mathbb{P}(\tau = kT) \le (1-p)^j$ is small.

Let us now make the above heuristic argument rigorous for $(X(kT))$. We first need to establish the following lemma.

Lemma 6.7. *For any compact set $\mathcal{K} \subset H \times H$ and any $R > 0$, there exists some constant $\delta = \delta(\mathcal{K}, R) > 0$ such that*

$$\inf_{x\in\mathcal{K}} \mathbb{P}_x\{X(T) \in B(R) \times B(R)\} > 0. \tag{6.16}$$

Proof. To show (6.16), we split the argument into the following three cases.

(i) As $x \notin B(r) \times B(r)$ with $x_1 \ne x_2$, $X_1(T)$ and $X_2(T)$ are independent. Therefore, by Lemma 5.1 one has

$$\begin{aligned}\mathbb{P}_x(X(T) \in B(R) \times B(R)) &= \mathbb{P}_x\left(X_1(T) \in B(R)\right)\mathbb{P}_x\left(X_2(T) \in B(R)\right)\\ &= \mathbb{P}\left(X_T^{x_1} \in B(R)\right)\mathbb{P}\left(X_T^{x_2} \in B(R)\right) > 0.\end{aligned}$$

(ii) As $x = (x_1, x_2)$ with $x_1 = x_2$, we have $X_1(T) = X_2(T)$. Hence,

$$\mathbb{P}_x(X(T) \in B(R) \times B(R)) = \mathbb{P}\left(X_T^{x_1} \in B(R)\right) > 0.$$

(iii) As $x \in B(r) \times B(r)$ with $x_1 \ne x_2$, by the maximal coupling property (6.2) one has

$$\begin{aligned}\mathbb{P}_x(X(T) \in B(R) \times B(R)) &= \mathbb{P}_x(M(x) \in B(R) \times B(R)) \\ &\geq \mathbb{P}_x(M_1(x) \in B(R))\mathbb{P}_x(M_2(x) \in B(R)) \\ &= \mathbb{P}(X_T^{x_1} \in B(R))\mathbb{P}(X_T^{x_2} \in B(R)) > 0,\end{aligned}$$

where $M(x) = (M_1(x), M_2(x))$ is the maximal coupling of $(P_T^*\delta_{x_1}, P_T^*\delta_{x_2})$.

From (i)-(iii) it is clear that

$$\mathbb{P}_x(X(T) \in B(R) \times B(R)) \geq \mathbb{P}(X_T^{x_1} \in B(R))\mathbb{P}(X_T^{x_2} \in B(R)).$$

By Feller property of P_T and Lemma 5.1, for any open subset $O \subset H$ the function $x \mapsto P_T(x, O)$ is positive and lower semi-continuous. Hence, it is separated from zero on any compact subset. Therefore, there is a constant $\delta = \delta(x, R, T) > 0$ so that

$$\inf_{x \in \mathcal{K}} \mathbb{P}(X_T^{x_1} \in B(R))\mathbb{P}(X_T^{x_2} \in B(R)) > 0. \tag{6.17}$$

From the above two inequality, we complete the proof. □

Proof of Theorem 6.6. Take $M = M(p, T, \alpha, \beta, \gamma, \varepsilon, \|F\|_0)$ defined in Theorem 6.3, and *simply write*

$$|x|^p = |x_1|^p + |x_2|^p, \qquad x = (x_1, x_2) \in H \times H.$$

Let us prove the theorem in the following four steps:

Step 1. Write $\tau_0^\varepsilon = 0$, $\tau_1^\varepsilon = \tau^\varepsilon$ and define

$$\tau_{k+1}^\varepsilon = \inf\{jT > \tau_k^\varepsilon; |X_1(jT)|_\varepsilon + |X_2(jT)|_\varepsilon \leq M\}$$

for all integer $k \geq 1$. Since $(X(kT))$ is a discrete time Markov chain, it is strong Markovian. By Theorem 6.3 and Poincare inequality $|z| \leq \frac{1}{\gamma_1^\varepsilon}|z|_\varepsilon$ for any $z \in H^\varepsilon$, we have

$$\mathbb{E}_{X(\tau_k^\varepsilon)}\left[e^{\eta(\tau_{k+1}^\varepsilon - \tau_k^\varepsilon)}\right] \leq C(1 + |X(\tau_k^\varepsilon)|^p) \leq c(1 + M^p), \tag{6.18}$$

where $c = C\,(1 + 2^p/\gamma_1^{\varepsilon p})$ and $C = C(p, \alpha, \beta, \gamma, \|F\|_0, r, T)$ is the same as in Theorem 6.3. The above inequality, together with strong Markov property, implies

$$\begin{aligned}\mathbb{E}_x[e^{\eta\tau_k^\varepsilon}] &= \mathbb{E}_x\left[e^{\eta\tau_1^\varepsilon}\mathbb{E}_{X(\tau_1^\varepsilon)}\left[e^{\eta(\tau_2^\varepsilon - \tau_1^\varepsilon)} \cdots \mathbb{E}_{X(\tau_{k-1}^\varepsilon)}\left[e^{\eta(\tau_k^\varepsilon - \tau_{k-1}^\varepsilon)}\right] \cdots\right]\right] \\ &\leq c^k(1 + M^p)^{k-1}(1 + |x|^p).\end{aligned} \tag{6.19}$$

Step 2. Since $B_\varepsilon(M) \subset\subset H$, by Lemma 6.7 we have

$$\inf_{y \in B_\varepsilon(M) \times B_\varepsilon(M)} \mathbb{P}_y\big(X(T) \in B(r) \times B(r)\big) = \sigma,$$

for all $r > 0$, where $\sigma = \sigma(\varepsilon, M, r, T) > 0$. Therefore, for some $\sigma \in (0, 1)$,

$$\inf_{|y|_\varepsilon \le M} \mathbb{P}_y\big(X(T) \in B(r) \times B(r)\big) \ge \sigma, \tag{6.20}$$

where $|y|_\varepsilon = |y_1|_\varepsilon + |y_2|_\varepsilon$.

Step 3. Given any $k \in \mathbb{N}$, define

$$\rho_k = \sup\{j;\ \tau_j^\varepsilon \le kT\}.$$

Clearly, $\tau_{\rho_k+1}^\varepsilon > kT$. For any $k \in \mathbb{N}$, one has

$$\begin{aligned}
\mathbb{P}_x(\tau = kT) &= \sum_{j=0}^{k} \mathbb{P}_x(\tau = kT, \rho_k = j) \\
&= \sum_{j=0}^{l} \mathbb{P}_x(\tau = kT, \rho_k = j) + \sum_{j=l+1}^{k} \mathbb{P}_x(\tau = kT, \rho_k = j) \\
&=: I_1 + I_2
\end{aligned} \tag{6.21}$$

where $l < k$ is some integer number to be chosen later.

Step 4. Let us estimate the above I_1 and I_2. By the definition of ρ_k, Chebyshev inequality and strong Markov property, we have

$$\begin{aligned}
\mathbb{P}_x(\tau = kT, \rho_k = j) &\le \mathbb{P}_x\left(\tau_j^\varepsilon > kT/2\right) + \mathbb{P}_x\left(\tau_j^\varepsilon \le kT/2,\ \rho_k = j\right) \\
&\le \mathbb{P}_x\left(\tau_j^\varepsilon > kT/2\right) + \mathbb{P}_x\left(\tau_j^\varepsilon \le kT/2,\ \tau_{j+1}^\varepsilon > kT\right) \\
&\le e^{-\eta kT/2}\mathbb{E}_x\left[e^{\eta\tau_j^\varepsilon}\right] + \mathbb{E}_x\left[\mathbb{P}_{X(\tau_j^\varepsilon)}\left(\tau_{j+1}^\varepsilon - \tau_j^\varepsilon > kT/2\right)\right]
\end{aligned}$$

By (6.19) and (6.18), the above inequality implies

$$\mathbb{P}_x(\tau = kT, \rho_k = j) \le c^j(1 + M^p)^{j-1}(1 + |x|^p)e^{-\eta kT/2} + c(1 + M^p)e^{-\eta kT/2}.$$

Hence,

$$\begin{aligned}
I_1 &\le \left[c^{l+1}(1 + M^p)^{l+1}(1 + |x|^p) + lc(1 + M^p)\right] e^{-\eta kT/2} \\
&\le c^{l+2}(1 + M^p)^{l+2}(1 + |x|^p)e^{-\eta kT/2}.
\end{aligned} \tag{6.22}$$

Now we estimate I_2. For $j > l$, by the definitions of τ and ρ_k, strong Markov property and (6.20), we have

$$\mathbb{P}_x\left(\tau = kT, \rho_k = j\right) \le \mathbb{P}_x\left(|X(\tau_1^\varepsilon)| > r, \ldots, |X(\tau_j^\varepsilon)| > r\right) \le (1 - \sigma)^j.$$

Hence,

$$I_2 \le \frac{1}{\sigma}(1-\sigma)^{l+1}. \tag{6.23}$$

Taking $\bar\eta = \frac{\eta}{4\log(c+cM^p)}$ and $l = [\bar\eta kT]$, we have

$$I_1 \le e^{-k\eta T/4}\big(1+|x|^p\big), \quad I_2 \le \frac{1}{\sigma}\exp\Big\{-kT\bar\eta\log\frac{1}{1-\sigma}\Big\}.$$

Combining the above estimates of I_1 and I_2, and taking $2\lambda = \frac{\eta}{4}\wedge\bar\eta\log\frac{1}{1-\sigma}$, we have

$$\mathbb{P}_x(\tau = kT) \le \left(c^2+\frac{1}{\sigma}\right)e^{-2\lambda kT}\,(1+|x|^p)$$

From the above inequality, we immediately obtain the desired estimate. □

6.3. **Final part of the coupling proof.** It is divided into two steps.

Step 1. By the same reason as in Steps 1 and 2 in Section 5, to prove the uniqueness of an invariant measure and inequality (2.9), it suffices to show that

$$\|P_{kT}(x_1,\cdot)-P_{kT}(x_2,\cdot)\|_{\mathrm{TV}} \le C\,(1+|x_1|^p+|x_2|^p)e^{-ckT}, \quad x_1,x_2\in H, \tag{6.24}$$

where C and c are positive constants not depending on x_1, x_2, and k. Let $(X_1(t),X_2(t))$, $t\in T\mathbb{Z}$, be the chain constructed in Section 6.1. Define the stopping time

$$\rho = \min\{kT : k\in\mathbb{N}, X_1(kT)=X_2(kT)\},$$

where the minimum over an empty set is equal to $+\infty$. Suppose we have proved that

$$\mathbb{P}_x\{\rho > kT\} \le Ce^{-\eta kT}(1+|x_1|^p+|x_2|^p), \tag{6.25}$$

where $x=(x_1,x_2)\in H\times H$ is arbitrary, and the positive constants η and C do not depend on x. In this case, using the fact that $X_1(kT)=X_2(kT)$ for $k\ge l$ as soon as $X_1(lT)=X_2(lT)$, we can write

$$\begin{aligned}\big|P_{kT}(x_1,\Gamma)-P_{kT}(x_2,\Gamma)\big| &= \big|\mathbb{E}_x 1_\Gamma\big(X_1(kT)\big)-\mathbb{E}_x 1_\Gamma\big(X_2(kT)\big)\big|\\ &= \mathbb{E}_x\Big(1_{\{\rho>kT\}}\big|1_\Gamma\big(X_1(kT)\big)-1_\Gamma\big(X_2(kT)\big)\big|\Big)\\ &\le \mathbb{P}_x\{\rho>kT\}.\end{aligned}$$

Using (6.25), we obtain

$$\big|P_{kT}(x_1,\Gamma)-P_{kT}(x_2,\Gamma)\big| \le Ce^{-\eta kT}\,(1+|x_1|^p+|x_2|^p).$$

Taking the supremum over all $\Gamma\in\mathcal{B}(H)$, we arrive at the required inequality (5.2).

Step 2. Thus, it remains to establish (6.25). To this end, we first note that if $r>0$ is sufficiently small, then

$$\mathbb{P}_x\{X_1(T)\ne X_2(T)\} \le 1/2 \quad \text{for any } x\in B(r)\times B(r). \tag{6.26}$$

Indeed, by Theorem 2.4, for any function $f \in B_b(H)$ with $\|f\|_0 \le 1$ we have

$$\big|(P_T(x_1,\cdot),f)-(P_T(x_2,\cdot),f)\big| = |P_Tf(x_1)-P_Tf(x_2)| \le C_1|x_1-x_2| \quad \text{for } x_1,x_2 \in H.$$

Recalling the definition of the total variation distance, we see that

$$\|P_T(x_1,\cdot)-P_T(x_2,\cdot)\|_{\mathrm{TV}} \le 1/2, \quad x_1,x_2 \in B(r),$$

where $r>0$ is sufficiently small. Since $\big(X_1(T),X_2(T)\big)$ is a maximal coupling for the pair $\big(P_T(x_1,\cdot),P_T(x_2,\cdot)\big)$, by (6.1) we arrive at (6.26).

We now introduce the iterations $\{\tau_n\}$ of the stopping time τ defined by (6.4):

$$\tau_1=\tau, \quad \tau_{n+1}=\inf\{jT>\tau_n : |X_1(jT)|+|X_2(jT)|\le r\}.$$

An argument similar to that used in Step 1 of the proof of Theorem 6.6 shows that

$$\mathbb{E}_x e^{\lambda\tau_n} \le K^n(1+|x_1|^p+|x_2|^p),$$

where $K>1$ and $\lambda>0$ do not depend on $x_1,x_2\in H$ and $n\ge 1$. By the Chebyshev inequality, it follows that

$$\mathbb{P}_x\{\tau_n>kT\} \le e^{-\lambda kT}K^n(1+|x_1|^p+|x_2|^p). \tag{6.27}$$

Let us define the events

$$\Gamma_n=\{X_1(\tau_m+T)\ne X_2(\tau_m+T) \text{ for } 1\le m\le n\}$$

and set $P_n(x)=\mathbb{P}_x(\Gamma_n)$. By (6.26) and the strong Markov property, we have

$$\mathbb{P}_x\big\{X_1(\tau_n+T)\ne X_2(\tau_n+T)\,|\,\mathcal{F}_{\tau_n}\big\} \le \mathbb{P}_{X(\tau_n)}\{X_1(T)\ne X_2(T)\}\le 1/2$$

It follows that

$$\begin{aligned}P_n(x)&=\mathbb{P}_x\big(\Gamma_{n-1}\cap\{X_1(\tau_n+T)\ne X_2(\tau_n+T)\}\big)\\&=\mathbb{E}_x\big(1_{\Gamma_{n-1}}\mathbb{P}_x\{X_1(\tau_n+T)\ne X_2(\tau_n+T)\,|\,\mathcal{F}_{\tau_n}\}\big)\le\frac{1}{2}P_{n-1}(x),\end{aligned}$$

whence, by iteration, we get $P_n(x)\le 2^{-n}$ for any $n\ge 1$. Combining this with (6.27), for any integers $n,k\ge 1$ we obtain

$$\begin{aligned}\mathbb{P}_x\{\rho>kT\}&=\mathbb{P}_x\{\rho>kT,\tau_n<kT\}+\mathbb{P}_x\{\rho>kT,\tau_n\ge kT\}\\&\le\mathbb{P}_x(\Gamma_n)+\mathbb{P}_x\{\tau_n\ge kT\}\\&\le 2^{-n}+e^{-\lambda kT}K^n(1+|x_1|^p+|x_2|^p).\end{aligned}$$

Taking $n=\varepsilon k$ with a sufficiently small $\varepsilon>0$, we arrive at the required inequality (6.25). The proof of Theorem 2.8 is complete.

7. Proofs of exponential mixing when dim H $< +\infty$

First of all, by Theorem 2.5 of [29], the system in (3.1) has at least one invariant measure. To prove Theorem 2.7, we can use the Harris method or the coupling argument.

In both approaches we need also the decay estimates for solutions given in Lemmas 4.2 and 4.3. These can be easily adapted to the strong solution X_t in (3.1) (indeed by the Gronwall lemma, starting from (3.1), we get that $\mathbb{E}|Z_A(t)|^p < \infty$ for any $p \in (0, \alpha)$).

For the first Harris approach, in order to verify the two conditions in Theorem 2.10 we can repeat the same argument given in Section 5.

For the coupling approach, the key point is irreducibility and gradient estimates of Theorem 2.4. Using a similar (but easier) argument as in Section 6, we can prove Theorem 2.7 in the following three steps:

(1) constructing the coupling and defining the stopping time τ exactly as in Section 6.1;

(2) proving the exponential estimate (6.15);

(3) using the same argument as in Section 6.3 which involves the coupling time.

Finally, let us emphasize that unlike the infinite dimensional setting, we do not need to introduce H^ε and τ^ε to get some compactness, since any finite-dimensional closed ball is automatically compact.

References

1. S. Albeverio, V. Mandrekar, and B. Rüdiger, *Existence of mild solutions for stochastic differential equations and semilinear equations with non-Gaussian Lévy noise*, Stochastic Process. Appl. **119** (2009), no. 3, 835-863.
2. S. Albeverio, B. Rüdiger, and J. L. Wu, *Invariant measures and symmetry property of Lévy type operators*, Potential Anal. **13** (2000), no. 2, 147-168.
3. S. Albeverio, J. L. Wu, and T. S. Zhang, *Parabolic SPDEs driven by Poisson white noise*, Stochastic Process. Appl. **74** (1998), no. 1, 21-36.
4. Z. Brzezniak, B. Goldys, P. Imkeller, S. Peszat, E. Priola and J. Zabczyk, *Time irregularity of generalized Ornstein-Uhlenbec processes*, C. R. Acad. Sci. Paris Ser. Math. **348** (2010), 273-276.
5. A. Chojnowska-Michalik, *On processes of Ornstein-Uhlenbeck in Hilbert spaces*, Stochastics **21** (1987), 251-286.
6. G. Da Prato and F. Flandoli, *Pathwise uniqueness for a class of SDE in Hilbert spaces and applications*, J. Funct. Anal. **259** (2010), no. 1, 243-267.
7. G. Da Prato and J. Zabczyk, *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.
8. A. Debussche, *Stochastic Navier-Stokes equations: well posedness and ergodic properties,* preprint (available on http://php.math.unifi.it/users/cime/).
9. W. Doeblin, *Éléments d'une théorie générale des chaînes simples constantes de Markoff*, Ann. Sci. École Norm. Sup. (3) **57** (1940, 61–111.
10. T. Funaki and B. Xie, *A stochastic heat equation with the distributions of Lévy processes as its invariant measures*, Stochastic Process. Appl. **119** (2009), no. 2, 307-326.

11. M. Hairer, *Exponential mixing properties of stochastic PDEs through asymptotic coupling*, Probab. Theory Related Fields **124** (2002), no. 3, 345–380. MR 1939651 (2004j:60135)
12. M. Hairer, *An introduciton to Stochastic PDEs*, http://www.hairer.org/notes/SPDEs.pdf.
13. Jonathan C. Mattingly, *Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics*, Comm. Math. Phys. **230** (2002), no. 3, 421–462. MR 1937652 (2004a:76039)
14. Alexey M. Kulik, *Exponential ergodicity of the solutions to SDE's with a jump noise*, Stochastic Process. Appl. **119** (2009), no. 2, 602–632. MR 2494006 (2010i:60176)
15. S. Kuksin and A. Shirikyan, *A coupling approach to randomly forced nonlinear PDEs. I*, Comm. Math. Phys. **221** (2001), no. 2, 351–366.
16. ______, *Coupling approach to white-forced nonlinear PDEs*, J. Math. Pures Appl. (9) **81** (2002), no. 6, 567602.
17. ______, *Mathematics of 2D Statistical Hydrodynamics*, manuscript of a book (available on www.u-cergy.fr/shirikyan/book.html)
18. T. Lindvall, *Lectures on the coupling method*, Dover Publications, Mineola, NY, 2002.
19. C. Marinelli and M. Röckner, *Well-posedness and ergodicity for stochastic reaction-diffusion equations with multiplicative Poisson noise,* Electron. J. Probab. **15** (2010) 1529-1555.
20. N. Masmoudi and L.-S. Young, *Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs*, Comm. Math. Phys. **227** (2002), no. 3, 461–481.
21. H. Masuda, *Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps.* Stochastic Process. Appl. 117 (2007), no. 1, 35-56.
22. Vahagn Nersesyan, *Polynomial mixing for the complex Ginzburg–Landau equation perturbed by a random force at random times*, J. Evol. Equ. **8** (2008), no. 1, 129.
23. Cyril Odasso, *Exponential mixing for the 3D stochastic Navier-Stokes equations*, Comm. Math. Phys. **270** (2007), no. 1, 109–139. MR MR2276442 (2008e:60190)
24. ______, *Exponential mixing for stochastic PDEs: the non-additive case*, Probab. Theory Related Fields **140** (2008), no. 1-2, 41–82. MR 2357670 (2009k:60148)
25. B. Øksendal, *Stochastic partial differential equations driven by multi-parameter white noise of Lévy processes*, Quart. Appl. Math. **66** (2008), no. 3, 521-537.
26. S. Peszat and J. Zabczyk, *Stochastic partial differential equations with Lévy noise*, Encyclopedia of Mathematics and its Applications, vol. 113, Cambridge University Press, Cambridge, 2007, An evolution equation approach.
27. S. Peszat and J. Zabczyk, *Stochastic heat and wave equations driven by an impulsive noise*, Stochastic partial differential equations and applications VII, Lect. Notes Pure Appl. Math., vol. 245, Chapman & Hall/CRC, Boca Raton, FL, 2006, pp. 229-242.
28. E. Priola, *Pathwise uniqueness for singular SDEs driven by stable processes*, to appear in Osaka Journal of Mathematics (available on http://www.newton.ac.uk/preprints/NI10062.pdf).
29. E. Priola, L. Xu and J. Zabczyk, *Exponential mixing for some SPDEs with Lévy noise*, arXiv:1010.4530, to appear in Stochastic and Dynamics.
30. E. Priola, J. Zabczyk, *Densities for Ornstein-Uhlenbeck processes with jumps,* Bulletin of the London Mathematical Society, 41 (2009), 41-50.
31. E. Priola and J. Zabczyk, *Structural properties of semilinear SPDEs driven by cylindrical stable processes*, published on line on Probab. Theory Related Fields (arXiv:0810.5063v1).
32. E. Priola and J. Zabczyk, *On linear evolution equations with cylindrical Lévy noise* to appear in Proceedings "SPDE's and Applications - VIII", Quaderni di Matematica, Seconda Università di Napoli (arXiv:0908.0356v1).
33. A. Rusinek, *Mean reversion for HJMM forward rate models*, Adv. in Appl. Probab. **42** (2010), no. 2, 371-391.

34. K.I. Sato : *Lévy processes and infinite divisible distributions,* Cambridge University Press, Cambridge, 1999.
35. K.I. Sato and M. Yamazato: *Stationary processes of Ornstein-Uhlenbeck type*, Lect.Notes in Math. **1021** (1983),541-551.
36. Armen Shirikyan, *Exponential mixing for 2D Navier-Stokes equations perturbed by an unbounded noise*, J. Math. Fluid Mech. **6** (2004), no. 2, 169–193. MR 2053582 (2005c:37160)
37. ______, *Exponential mixing for randomly forced partial differential equations: method of coupling*, Instability in models connected with fluid flows. II, Int. Math. Ser. (N. Y.), vol. 7, Springer, New York, 2008, pp. 155-188. MR 2459266 (2009i:35359).
38. A. Takeuchi, *The Bismut-Elworthy-Li type formulae for stochastic differential equations with jumps,* Preprint arXiv:1002.1384, to appear in Journal of Theoretical Probability.
39. H. Thorisson, *Coupling, stationarity, and regeneration*, Springer-Verlag, New York, 2000.
40. L. Xu and B. Zegarliński, *Ergodicity of the finite and infinite dimensional α-stable systems*, Stoch. Anal. Appl. **27** (2009), no. 4, 797-824.
41. Lihu Xu and Bogusław Zegarliński, *Existence and exponential mixing of infinite white alpha-stable systems with unbounded interactions*, Electron. J. Probab. **15** (2010), 1994–2018.
42. F.Y. Wang, *Gradient Estimate for Ornstein-Uhlenbeck Jump Processes*, to appear in Stoch. Proc. Appl., available at arXiv:1005.5023.
43. J. Zabczyk, *Stationary distributions for linear equations driven by general noise*, Bull. Pol. Acad. Sci. **31** (1983), 197-209.

Dipartimento di Matematica, Università di Torino, via Carlo Alberto 10, 10123 Torino, Italy
E-mail address: `enrico.priola@unito.it`

Department of Mathematics, University of Cergy–Pontoise, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France
E-mail address: `Armen.Shirikyan@u-cergy.fr`

Department of Mathematics, Brunel University, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom
E-mail address: `Lihu.Xu@brunel.ac.uk`

Institute of Mathematics, Polish Academy of Sciences, P-00-950 Warszawa, Poland
E-mail address: `zabczyk@impan.pl`