



Couplings via Comparison Principle and Exponential Ergodicity of SPDEs in the Hypocoelliptic Setting

Oleg Butkovsky^{1,2} , Michael Scheutzow¹

¹ Institut für Mathematik, Technische Universität Berlin, MA 7-5, Fakultät II, Strasse des 17. Juni 136, 10623 Berlin, Germany. E-mail: oleg.butkovskiy@gmail.com, ms@math.tu-berlin.de

² Weierstrass Institute, Mohrenstraße 39, 10117 Berlin, Germany

Received: 8 July 2019 / Accepted: 4 June 2020

Published online: 29 September 2020 – © The Author(s) 2020

Abstract: We develop a general framework for studying ergodicity of order-preserving Markov semigroups. We establish natural and in a certain sense optimal conditions for existence and uniqueness of the invariant measure and exponential convergence of transition probabilities of an order-preserving Markov process. As an application, we show exponential ergodicity and exponentially fast synchronization-by-noise of the stochastic reaction–diffusion equation in the hypoelliptic setting. This refines and complements corresponding results of Hairer and Mattingly (Electron J Probab 16:658–738, 2011).

1. Introduction

The goal of this article is to build a framework for analyzing ergodic properties of order-preserving Markov processes and to provide simple, verifiable, yet general enough sufficient conditions for exponential ergodicity. This framework turns out to be especially powerful for investigating ergodicity of order-preserving stochastic PDEs with highly degenerate additive forcing. Our main example is the stochastic reaction–diffusion equation in the hypoelliptic setting. We show that even if noise comes to the system only from one Brownian motion, then (under certain conditions) this SPDE has a unique invariant measure and its transition probabilities converges to it exponentially in the Wasserstein metric. We also establish exponentially fast synchronization-by-noise of the solutions to this equation. This refines [18, Remark 8.22] and complements [18, Theorem 8.21].

In the mathematical physics literature there is a growing interest in ergodic behavior of nonlinear PDEs forced by smooth in space noise acting only on a few Fourier modes, see, e.g., [5, 16, 23, 29, 30]. Since the noise is smooth in space, it is usually relatively easy to show that these SPDEs have a unique solution and that this solution is a Markov process. On the other hand, since the solution at any fixed time is an infinite-dimensional

Oleg Butkovsky: Supported in part by DFG Research Unit FOR 2402 and European Research Council Grant 683164.

Michael Scheutzow: Supported in part by DFG Research Unit FOR 2402.

random variable and the noise acts only onto finitely many degrees of freedom, these processes do not get “enough” noise and hence they are typically only Feller but not strong Feller, see also the discussion in [15, Section 9]. This makes analyzing their ergodic behavior much more challenging.

Indeed, recall that ergodicity of strong Feller processes can be established using the standard classical approach, which combines a local mixing condition on a certain set (the small set condition) and a recurrence condition, see, e.g., [31]. Unfortunately, this method is usually not applicable for Markov processes which are only Feller and not strong Feller because they do not have good mixing properties, see also a detailed discussion in [19, Section 1]). To study ergodicity of these processes three alternative strategies have been suggested recently.

The first approach was developed in [16–18], see also [15, Section 11]. It introduces the asymptotic strong Feller (ASF) property, which serves as a replacement for the strong Feller property. It is shown there that if a Markov process satisfies ASF as well as certain recurrence and topological irreducibility conditions then the process is exponentially ergodic. Note that for many Markov processes verifying ASF might be rather challenging. In particular, while this method works quite well for stochastic Navier–Stokes (SNS) equations on a torus in the hypoelliptic setting, it is not so clear how to check ASF for the SNS equation on a bounded domain, see the discussion in [11, Section 1].

Another approach establishes exponential ergodicity using generalized couplings [2, 11, 12, 19, 24, 30]. Recall that a coupling is a pair of stochastic processes with given marginal distributions. By contrast, a generalized coupling is a pair of stochastic processes, whose marginals are not necessarily equal to a prescribed pair of probability distributions, but are in a sense close to this pair. Clearly, constructing a generalized coupling is much easier than constructing a coupling. Furthermore, it is shown in the papers mentioned above that existence of a generalized coupling with certain nice properties yields exponential ergodicity. This approach works quite well for a large class of SPDEs in the effectively elliptic setting (that is, when noise acts in a finite but large enough number of directions), but is less useful for studying SPDEs in the hypoelliptic setting.

Finally, the third main approach was introduced in [19]. It utilizes the notion of a d -small set (a generalization of a small set), which is particularly well adapted to the study of Markov processes with bad mixing properties. This approach provides another set of sufficient conditions for exponential ergodicity, and it works quite well with stochastic delay equations and nonlinear autoregressions. Unfortunately, verifying this set of conditions for SPDEs is rather difficult.

Our new approach developed in this paper is somehow orthogonal to all of the strategies mentioned above. It is specifically targeted at order preserving Markov semigroups, that is, the semigroups which map increasing bounded functions to increasing bounded functions. On the one hand, this significantly reduces the applicability of this approach; for example, it cannot be used to study stochastic Navier–Stokes equations. On the other hand, for order-preserving Markov processes (e.g., stochastic reaction–diffusion equations) it allows to obtain exponential ergodicity under very weak assumptions; this is rather difficult (or maybe impossible) to achieve with other methods.

The main result (Theorems 2.3 and 2.4) is quite general. It shows that if an order-preserving Markov semigroup satisfies additionally a swap condition (i.e., two Markov processes started with initial conditions x, y with $x \preceq y$ can change their order by time 1 with a small but positive probability), then under a standard Lyapunov-type condition

as well as a certain technical assumption the process is exponentially ergodic. We also show that this swap condition cannot be omitted.

We apply the obtained theorems to establish exponential ergodicity of stochastic reaction–diffusion equations on a d -dimensional torus \mathbb{T}^d , $d \in \mathbb{N}$

$$du(t, \xi) = [\Delta u(t, \xi) + f(u(t, \xi), \xi)]dt + \sum_{k=1}^m \sigma_k(\xi) dW^k(t), \quad \xi \in \mathbb{T}^d, t \geq 0, \quad (1.1)$$

where (W^1, W^2, \dots, W^m) are independent standard Brownian motions; f, σ_k are continuous functions acting from $\mathbb{R} \times \mathbb{T}^d \rightarrow \mathbb{R}$ and $\mathbb{T}^d \rightarrow \mathbb{R}$, respectively and satisfying certain conditions. It is clear that if $m = 0$ (no noise), then this equation might have multiple invariant measures. On the other hand, if $m = \infty$ (noise acts in every direction), then, under certain additional assumptions on σ , the process u is strong Feller and it can be shown by the classical methods that it has a unique invariant measure [6, Sections 7 and 11]. Thus, it is natural to ask what the smallest number of directions m that have to be perturbed by noise is, so that Eq. (1.1) still has a unique invariant measure.

Using the ASF method described above, it was shown in [18, Remark 8.22] that if f is a polynomial and $m = 3$, then Eq. (1.1) has a unique invariant measure and is exponentially ergodic. We extend this result and show exponential ergodicity of u already if $m = 1$ (that is, when noise acts only in one direction), see Theorem 4.5 and Remark 4.6; we also do not rely on a specific form of f . Note that the convergence to the invariant measure is established in the Wasserstein metric; Theorem 4.8 shows that in the case $m < \infty$ this is optimal. Namely, if no additional assumptions are imposed, then, contrary to the case $m = \infty$, the transition probabilities of Markov process (1.1) might **not** converge to the invariant measure in the total variation metric. Finally, in Theorem 4.9, we show that any two solutions to (1.1) launched with the same noise from different initial conditions converge to each other exponentially fast (synchronization by noise).

The idea that order-preservation helps to obtain better convergence rates of a Markov process is not new; it was used to study interacting particle systems since 1970s, see [25]. However, it is not clear at all how to extend the techniques used in this book beyond the framework of interacting particle models. Order-preserving Markov processes on a general state spaces were considered in [28, 32]. However, the methods developed there rely on the small set condition. Since in the current paper we study processes on a general state space with bad mixing properties, where this condition might not hold, unfortunately the ideas of [25, 28, 32] cannot be applied in our case.

It is interesting to compare our results with [8]. In that paper the authors consider an order-preserving random dynamical system (RDS) with two additional properties: it has a unique invariant measure and it weakly converges to this measure. It is shown there that this implies that any two trajectories of the RDS converge to each other in probability. By contrast, in the current paper we start with an order-preserving Markov process and prove uniqueness of an invariant measure and convergence of transition probabilities.

Our main tool is a new version of the coupling method specifically tailored for order-preserving Markov processes, see the proof of Theorem 2.3. This is combined with an analysis of the relations between stochastic domination and expected distance of random variables, see Sect. 3. There we continue the study initiated in [4, Proposition 1] and [8, Proposition 2.4]. Note however, that the methods introduced in [4] and [8] cannot be used to get a quantitative bound even in the case where the Markov process has state space \mathbb{R} , see Example 3.2. Therefore we apply a new technique.

While we study in detail only the stochastic reaction–diffusion equations on the torus, the strategy developed in this paper should also work in a very similar way for other order-preserving SPDEs including stochastic reaction–diffusion equation on a bounded domain and stochastic porous medium equations. We would like to mention also that after the preprint of our paper became available, our technique was extended and adapted to study regularization by noise for singular SPDEs [10].

The rest of the paper is organized as follows. We present our main results in Sect. 2. In Sect. 3 we investigate the relations between stochastic domination and average distance of random variables. Section 4 is devoted to a detailed study of ergodicity of stochastic reaction–diffusion equations. The proofs of the main results are placed in Sect. 5.

Convention on constants. Throughout the paper C denotes a positive constant whose value may change from line to line.

2. A General Framework for Ergodicity for Order-Preserving Markov Processes

We begin by introducing the basic notation. Let (E, ρ) be a Polish space with partial order \preceq such that the set

$$\Gamma := \{(x, y) \in E \times E : x \preceq y\} \tag{2.1}$$

is closed. Let $\mathcal{E} = \mathcal{B}(E)$ be the Borel σ -field. For sets $A, B \in \mathcal{E}$ we will write $A \preceq B$ if for any $a \in A, b \in B$ we have $a \preceq b$. Denote by $\mathcal{P}(E)$ the set of all probability measures on (E, \mathcal{E}) . Let $\{P_t\}_{t \geq 0} = \{P_t(x, A), x \in E, A \in \mathcal{E}\}_{t \geq 0}$ be a Markov transition function over E and denote by $\{\mathbf{P}_x, x \in E\}$ the corresponding Markov family; that is \mathbf{P}_x is the law of the Markov process $\{X_t, t \geq 0\}$ with the given transition function and initial condition $X_0 = x$. The law of X will be understood in the sense of finite-dimensional distributions; that is, we will not rely on the trajectory-wise properties of X .

For a measurable function $r : E \times E \rightarrow [0, 1]$, we consider the corresponding coupling distance $W_r : \mathcal{P}(E) \times \mathcal{P}(E) \rightarrow \mathbb{R}_+$ given by

$$W_r(\mu, \nu) := \inf_{\lambda \in \mathcal{C}(\mu, \nu)} \int_{E \times E} r(x, y) \lambda(dx, dy), \quad \mu, \nu \in \mathcal{P}(E),$$

where $\mathcal{C}(\mu, \nu)$ is the set of all couplings between μ and ν , i.e., probability measures on $(E \times E, \mathcal{E} \otimes \mathcal{E})$ with marginals μ and ν . If r is a lower semicontinuous metric on E , then W_r is the usual Kantorovich–Wasserstein distance. If r is the discrete metric, i.e., $r(x, y) = \mathbb{1}(x \neq y)$, then W_r is the total variation distance, which will be denoted further by d_{TV} .

Let us now recall the standard definitions related to the partial order \preceq ; we refer to, e.g., [26, Chapter IV] for a detailed discussion.

Definition 2.1. (i) A function $f : E \rightarrow \mathbb{R}$ is called *increasing* if for any $x, y \in E$ with $x \preceq y$ we have $f(x) \leq f(y)$.

(ii) Let $\mu, \nu \in \mathcal{P}(E)$ be two probability measures. We say that ν *stochastically dominates* μ and denote it by $\mu \preceq_{st} \nu$ if for any bounded measurable increasing function $f : E \rightarrow \mathbb{R}$ we have

$$\int_E f d\mu \leq \int_E f d\nu.$$

(iii) We say that a Markov transition function $\{P_t\}_{t \geq 0}$ is *order preserving* if for any $t > 0$ and $x, y \in E$ such that $x \leq y$ we have

$$P_t(x, \cdot) \leq_{st} P_t(y, \cdot).$$

In other words, a Markov transition function is order preserving if it maps bounded increasing functions to bounded increasing functions. Examples of Markov processes with an order preserving transition function include stochastic reaction-diffusion equations, stochastic porous media equations and others, see, e.g., [8].

Remark 2.2. Strassen’s theorem (see, e.g., [26, Theorem IV.2.4]) provides the following coupling definition of stochastic domination, which is equivalent to the one stated above. We have $\mu \leq_{st} \nu$ if and only if there exist random elements $X, Y : \Omega \rightarrow E$ such that $\text{Law}(X) = \mu, \text{Law}(Y) = \nu$ and $X \leq Y$.

Now we are ready to present our main results.

Theorem 2.3. *Suppose that there exist measurable functions $V : E \rightarrow [0, \infty), \varphi : E \rightarrow \mathbb{R}, d : E \times E \rightarrow [0, \infty)$, such that the following conditions hold:*

- (1) *the Markov transition function $(P_t)_{t \geq 0}$ is order-preserving;*
- (2) *the function V is a Lyapunov function, that is, there exist constants $\gamma, K > 0$ such that*

$$P_t V(x) \leq V(x) - \gamma \int_0^t P_s V(x) ds + Kt, \quad t \geq 0, x \in E; \tag{2.2}$$

- (3) *if $x, y \in E$ and $x \leq y$, then $0 \leq d(x, y) \leq \varphi(y) - \varphi(x)$;*
- (4) *for any $x \in E$ we have $M(x) := \sup_{t \geq 0} P_t \varphi^2(x) < \infty$;*
- (5) *there exist sets $A, B \in \mathcal{E}$ and $\varepsilon > 0$ such that $A \leq B$ and for any $x \in \{V \leq 4K/\gamma\}$ we have*

$$P_1(x, A) > \varepsilon \text{ and } P_1(x, B) > \varepsilon. \tag{2.3}$$

Then for any $\theta > 0$ there exist constants $C, \lambda > 0$ such that for any $x, y \in E$

$$W_{d \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) \leq C(1 + V(x) + V(y))^\theta (1 + M(x))^\theta \exp(-\lambda t), \quad t \geq 0. \tag{2.4}$$

Theorem 2.4. *Suppose that all conditions of Theorem 2.3 are satisfied. Assume further that*

- (6) *there exists $\delta \in (0, 1]$ such that $\rho \leq d^\delta$, where ρ is the metric on the Polish space E .*
- (7) *there exists $K > 0, \kappa > 0$ such that $M(x)^\kappa \leq K + KV(x)$ for all $x \in E$.*

Then the Markov semigroup has a unique invariant measure π . Further, for any $\theta > 0$ there exist constants $C, \lambda > 0$ such that for any $x \in E$

$$W_{\rho \wedge 1}(P_t(x, \cdot), \pi) \leq C(1 + V(x))^\theta \exp(-\lambda t), \quad t \geq 0. \tag{2.5}$$

Proof. (Sketch of the proof of Theorems 2.3 and 2.4) Here, for the convenience of the reader, we provide just a very brief roadmap of the proof; a complete proof is given in Sect. 5.

Fix $x, y \in E, t > 0$. The proof splits into two independent parts. First, we use a new version of the coupling method and utilize conditions (1), (2), (5) to construct random elements Z^x, Z^y, \tilde{Z}^x taking values in E with the following properties:

$$\begin{aligned} \text{Law}(Z^x) &= \text{Law}(\tilde{Z}^x) = P_t(x, \cdot), \quad \text{Law}(Z^y) = P_t(y, \cdot); \\ \mathbb{P}(Z^x \leq Z^y \leq \tilde{Z}^x) &\geq 1 - C_1(1 + V(x) + V(y))e^{-C_2 t}, \end{aligned} \tag{2.6}$$

for some universal constants $C_1, C_2 > 0$. Second, using the ideas developed in Sect. 3, we transform the bound (2.6) into the following inequality:

$$\mathbb{E}[d(Z^x, Z^y) \wedge 1] \leq 1 - C_3(1 + V(x) + V(y))(1 + M(x))e^{-C_4t}, \tag{2.7}$$

where $C_3, C_4 > 0$ are again some universal constants. It is at this step, where we are using conditions (3) and (4). Since $\text{Law}(Z^x) = P_t(x, \cdot)$ and $\text{Law}(Z^y) = P_t(y, \cdot)$, inequality (2.7) yields (2.4) and (2.5). \square

Theorems 2.3 and 2.4 provide general sufficient conditions for an order-preserving Markov process to be ergodic. Condition (3) of Theorem 2.3 is actually a condition on the space (E, \preceq) and d rather than on the Markov semigroup. As shown below it is satisfied in many natural situations, including $E = \mathbb{R}$, $E = L_p$, $p \in [1, \infty)$, $E = \mathcal{B}_{p,\infty}^\alpha$, $\alpha < 0$, $p \in [1, \infty)$; the latter stands for the Besov space of regularity α and integrability p , see, e.g., [1, Section 2.7 and Proposition 2.93]. Thus, the only additional assumption for exponential ergodicity (apart from the standard Lyapunov and moment-type conditions) is the swap condition (5). It tells that the state space E contains two sets, one preceding the other, and locally uniformly in the initial condition the Markov process has a small chance to be in either of these sets. The following simple example explains why this condition cannot be dropped.

Example 2.5. Introduce the following trivial order on E : for $x, y \in E$ we have $x \preceq y$ if and only if $x = y$. It is clear, that for this order any Markov semigroup is order-preserving. We also see that the set Γ defined in (2.1) is closed; furthermore, conditions (3) and (4) of Theorem 2.3 trivially hold with $\varphi \equiv 0$. Thus, any Markov process on E , that has a Lyapunov function satisfies conditions (5)–(4) of Theorem 2.3. It is well-known that this is not enough for uniqueness of the invariant measure. Thus, the swap condition (5) cannot be omitted.

We also would like to emphasize that the swap condition (5) is very different in nature from the small set condition or other minorization-type conditions, which are imposed within the classical framework, see e.g., [33]. Indeed, a minorization condition guarantees good mixing properties of transition kernels and, in particular, bounds on the total variation distance between the kernels. On the other hand, the swap condition does not yield such bounds since nothing was assumed about mixing on the sets A and B . Lemma 5.9 shows how the swap condition can be verified for the stochastic reaction-diffusion equation.

Now let us provide natural examples of spaces E for which condition (3) of Theorem 2.3 holds.

Example 2.6. Let S be a countable set. Put $E = \{0, 1\}^S$ equipped with the distance $d(x, y) := \sum_{i=1}^\infty 2^{-i}|x_i - y_i|$, where $x, y \in \{0, 1\}^S$ and $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$. This space often appears within the context of interacting particle systems. Consider the following partial order: if $x, y \in \{0, 1\}^S$, then

$$x \preceq_{inc} y \text{ if and only if } x_i \leq y_i \text{ for all } i \in \mathbb{N}.$$

Then condition (3) holds for the function $\varphi(x) := \sum_{i=1}^\infty 2^{-i}x_i$, $x \in \{0, 1\}^S$.

Example 2.7. Put $E = \mathbb{R}$ equipped with the standard distance, $d(x, y) := |x - y|$, $x, y \in \mathbb{R}$, and consider the standard order \leq . Then condition (3) holds for the function $\varphi(x) := x$, $x \in \mathbb{R}$.

Example 2.8. Put $E := L_p(D, \mathbb{R})$, where $p \geq 1$ and an arbitrary domain $D \subset \mathbb{R}^n$, $n \in \mathbb{N}$. Let $\|\cdot\|_{L_p}$ be the standard L_p norm in this space. Consider the following partial order

$$x \preceq_{pos} y \text{ if and only if } x(\xi) \leq y(\xi) \text{ for almost all } \xi \in D \tag{2.8}$$

and let $d(x, y) := \|x - y\|_{L_p(D)}^p$, $x, y \in L_p(D, \mathbb{R})$. Then there exists a function φ such that the partial order \preceq_{pos} and function d introduced above satisfy condition (3) of Theorem 2.3. Furthermore, the set Γ defined in (2.1) is closed. We postpone the proof of this statement to Sect. 5.

Example 2.9. This example is due to [10]. Put $E := \mathcal{B}_{p,\infty}^\alpha(D, \mathbb{R})$, where $p \geq 1$, $\alpha < 0$ and an arbitrary domain $D \subset \mathbb{R}^n$, $n \in \mathbb{N}$. Let $\|\cdot\|_{\mathcal{B}_{p,\infty}^\alpha}$ be the standard Besov norm in this space, see, e.g., [1, Definition 2.68]. Consider the following partial order

$$x \preceq_{Besov} y \text{ if and only if } \langle x, \varphi \rangle \preceq_{pos} \langle y, \varphi \rangle \text{ for any non-negative } \varphi \in \mathcal{C}^\infty$$

and let $d(x, y) := \|x - y\|_{\mathcal{B}_{p,\infty}^\alpha}^p$, $x, y \in \mathcal{B}_{p,\infty}^\alpha(D, \mathbb{R})$. Then, as shown in [10, Lemma A.1], there exists a function φ such that the partial order \preceq_{Besov} and function d satisfy condition (3) of Theorem 2.3.

3. Stochastic Domination and Distance Between Random Variables

In this section we explore the connections between stochastic domination and expected distance of random variables, thus continuing the analysis initiated in [8]. Recall that we are given a Polish space E with metric ρ . The following statement played a key role in establishing synchronization-by-noise results in [8].

Proposition 3.1 ([8, Proposition 2.4]). *Let X_t, Y_t be two stochastic processes taking values in E such that $X_t(\omega) \preceq Y_t(\omega)$ for all $t \geq 0, \omega \in \Omega$. Suppose that X_t and Y_t converge weakly as $t \rightarrow \infty$ to a random element with the law μ . Then,*

$$\mathbf{E}[\rho(X_t, Y_t) \wedge 1] \rightarrow 0 \text{ as } t \rightarrow \infty. \tag{3.1}$$

It is natural to ask whether the above statement can be quantified. More precisely, assume additionally that

$$d_{TV}(\text{Law}(X_t), \mu) \leq r(t), \quad d_{TV}(\text{Law}(Y_t), \mu) \leq r(t),$$

for some rate function r going to 0 as $t \rightarrow \infty$. One can ask whether these bounds guarantee a quantitative estimate on $\mathbf{E}[\rho(X_t, Y_t) \wedge 1]$. Quite surprisingly, the answer to this question is negative: without any additional assumptions $\mathbf{E}[\rho(X_t, Y_t) \wedge 1]$ might tend to 0 very slowly even when $r(t)$ converges exponentially fast. This is illustrated by the following example.

Example 3.2. Let $E = \mathbb{R}$ be equipped with the standard order \leq . Let $(p_i)_{i \in \mathbb{Z}_+}$ be a sequence of positive numbers summing up to 1. Let X be a random variable which with probability p_i is uniformly distributed on $[2^i, 2^{i+1}]$, $i \in \mathbb{Z}_+$. Define for $n \in \mathbb{Z}_+$

$$Y_n := X \mathbb{1}(X \notin [2^n, 2^{n+1}]) + (X + 1) \mathbb{1}(X \in [2^n, 2^{n+1}]); \quad X_n = X.$$

Then for each n we clearly have $X_n \leq Y_n$ and it is immediate to see that for $\mu := \text{Law}(X)$ one has

$$d_{TV}(\text{Law}(X_n), \mu) = 0; \quad d_{TV}(\text{Law}(Y_n), \mu) \leq p_n 2^{-n}.$$

On the other hand,

$$\mathbf{E}[|X_n - Y_n| \wedge 1] = p_n,$$

which is much slower than $p_n 2^{-n}$.

Thus, to quantify bounds in (3.1) we need to impose extra assumptions. This is where condition (3) of Theorem 2.3 pops up.

Lemma 3.3. *Assume that condition (3) of Theorem 2.3 is satisfied for measurable functions $\varphi: E \rightarrow \mathbb{R}$, $d: E \times E \rightarrow [0, 1]$. Suppose further that there exist a function $\psi: E \rightarrow \mathbb{R}_+$ and $k \in (0, 1)$ such that*

$$|\varphi(x) - \varphi(y)| \leq d(x, y)^k (\psi(x) + \psi(y)), \quad x, y \in E. \tag{3.2}$$

Let X, Y be random elements $\Omega \rightarrow E$ such that $X \preceq Y$ a.s., $\mathbf{E}|\varphi(X)| < \infty$, $\mathbf{E}|\varphi(Y)| < \infty$ and $W_d(\text{Law}(X), \text{Law}(Y)) \leq \varepsilon$ for some $\varepsilon > 0$. Then we have

$$\mathbf{E}d(X, Y) \leq \varepsilon^k ((\mathbf{E}[\psi(X)^{1/(1-k)}])^{1-k} + (\mathbf{E}[\psi(Y)^{1/(1-k)}])^{1-k}). \tag{3.3}$$

Proof. We begin by observing that since $X \preceq Y$ a.s., condition (3) of Theorem 2.3 yields

$$\mathbf{E}d(X, Y) \leq \mathbf{E}\varphi(Y) - \mathbf{E}\varphi(X). \tag{3.4}$$

Fix $K > 1$ and let \tilde{X}, \tilde{Y} be random elements such that $\text{Law}(\tilde{X}) = \text{Law}(X)$, $\text{Law}(\tilde{Y}) = \text{Law}(Y)$ and $\mathbf{E}d(\tilde{X}, \tilde{Y}) \leq K\varepsilon$ (they exist since $W_d(\text{Law}(X), \text{Law}(Y)) \leq \varepsilon$). Then we continue (3.4) as follows, using the fact that d is bounded by 1:

$$\begin{aligned} \mathbf{E}d(X, Y) &\leq \mathbf{E}\varphi(\tilde{Y}) - \mathbf{E}\varphi(\tilde{X}) \\ &\leq \mathbf{E}[d(\tilde{X}, \tilde{Y})^k (\psi(\tilde{X}) + \psi(\tilde{Y}))] \\ &\leq K^k \varepsilon^k ((\mathbf{E}[\psi(\tilde{X})^{1/(1-k)}])^{1-k} + (\mathbf{E}[\psi(\tilde{Y})^{1/(1-k)}])^{1-k}) \\ &= K^k \varepsilon^k ((\mathbf{E}[\psi(X)^{1/(1-k)}])^{1-k} + (\mathbf{E}[\psi(Y)^{1/(1-k)}])^{1-k}). \end{aligned}$$

Since $K > 1$ was arbitrary, this yields the statement of the lemma. □

The lemma above will be very useful for obtaining exponential bounds on the synchronization-by-noise. On the other hand, to complete the second part of the proof of Theorem 2.3 (see its sketch above in Sect. 2), we need to solve the opposite problem. More precisely, Lemma 3.3 considers the case when one random element is less than the other everywhere but their laws are different. The following lemma studies the situation where the laws of the random elements coincide, but one is less than the other with probability smaller than 1.

Lemma 3.4. *Assume that condition (3) of Theorem 2.3 is satisfied for measurable functions $\varphi: E \rightarrow \mathbb{R}$, $d: E \times E \rightarrow [0, 1]$. Let X, Y, \tilde{X} be random elements $\Omega \rightarrow E$ such that $\text{Law}(X) = \text{Law}(\tilde{X})$ and*

$$\mathbf{P}(X \leq Y \leq \tilde{X}) \geq 1 - \varepsilon$$

for some $\varepsilon > 0$. Then for any $p, q > 1$ such that $1/p + 1/q = 1$ we have

$$\mathbf{E}d(X, Y) \leq 2\varepsilon^{1/p} (\mathbf{E}|\varphi(X)|^q)^{1/q} + \varepsilon.$$

Again, we see that the statement of the lemma is satisfied only under some extra conditions (one needs condition (3) of Theorem 2.3 and a reasonable bound on $E|\varphi(X)|^q$). The following two examples show that these extra conditions cannot be dropped.

Example 3.5. Let $E = \mathcal{C}([0, 1], \mathbb{R})$ be equipped with the sup norm $\|\cdot\|$ and order \preceq_{pos} (defined as in (2.8)). For $x, y \in E$ put $d(x, y) := \|x - y\| \wedge 1$. Let $n \in \mathbb{N}$. For $k \in \{1, 2, \dots, n\}$ let f_k be an element of $\mathcal{C}([0, 1], \mathbb{R})$ taking values in $[0, 1]$ such that

$$f_k(\xi) = 1 \text{ for } \xi \in [0, k/n]; \quad f_k(\xi) = 0 \text{ for } \xi \in [(k + 1)/n, 1].$$

Then it is easy to see that $f_k \preceq_{pos} f_l$ whenever $k \leq l \leq n$. Let η be uniformly distributed on the set $\{1, 2, \dots, n\}$. Put $X := f_\eta$ and $Y = \tilde{X} := f_{\eta+1}$, where the summation is taken mod n . Then X and \tilde{X} are bounded, $\text{Law}(X) = \text{Law}(\tilde{X})$, $\mathbf{P}(X \preceq Y \preceq \tilde{X}) = 1 - 1/n$, but $E[\|X - Y\| \wedge 1] = 1$.

It is easy to see that in this example condition (3) of Theorem 2.3 does not hold. Indeed, if a function φ satisfies this condition, then for any $n \in \mathbb{N}, k \in \{1, 2, \dots, n - 1\}$ we have $\varphi(f_{k+1}) \geq \varphi(f_k) + d(f_k, f_{k+1}) = \varphi(f_k) + 1$. Therefore

$$\varphi(\check{1}) = \varphi(f_n) \geq \varphi(f_1) + n - 1 \geq \varphi(\check{0}) + n, \tag{3.5}$$

where $\check{1}$ and $\check{0}$ denote elements of E identically equal to 1 and 0, respectively. Since n was arbitrary, we see that (3.5) implies that $\varphi(\check{1})$ cannot be finite, which is impossible.

Example 3.6. Let $E = \mathbb{R}$ be equipped with the standard order \leq . For $x, y \in E$ put $d(x, y) := |x - y| \wedge 1$. Let $n \in \mathbb{N}$ and let X be uniformly distributed on the set $\{1, 2, \dots, n\}$. Let $Y = \tilde{X} = (X + 1) \mathbb{1}(X < n) + \mathbb{1}(X = n)$. Then condition (3) of Theorem 2.3 holds with $\varphi(x) = x$, $\text{Law}(X) = \text{Law}(\tilde{X})$, $\mathbf{P}(X \leq Y \leq \tilde{X}) = 1 - 1/n$, but $E[\|X - Y\| \wedge 1] = 1$.

Thus, we see that in this example condition (3) of Theorem 2.3 holds, but $E|\varphi(X)| = EX = n/2 + 1/2$ can be arbitrarily large for large n .

Proof of Lemma 3.4. We begin by observing that the function φ is increasing, that is $x \leq y$ implies $\varphi(x) \leq \varphi(y)$ for any $x, y \in E$. Fix $p, q > 1$ such that $1/p + 1/q = 1$. Without loss of generality, we can also assume that $E|\varphi(X)|^q < \infty$ (otherwise the statement of the lemma is trivial).

Let $\Lambda := \{\omega : X(\omega) \leq Y(\omega) \leq \tilde{X}(\omega)\}$. Then, by assumption, $\mathbf{P}(\Lambda) \geq 1 - \varepsilon$ and we get (recall that $d \leq 1$)

$$\begin{aligned} Ed(X, Y) &= E[d(X, Y) \mathbb{1}(\Lambda)] + E[d(X, Y) \mathbb{1}(\Omega \setminus \Lambda)] \\ &\leq E[(\varphi(Y) - \varphi(X)) \mathbb{1}(\Lambda)] + \mathbf{P}(\Omega \setminus \Lambda) \\ &\leq E[(\varphi(\tilde{X}) - \varphi(X)) \mathbb{1}(\Lambda)] + \varepsilon \\ &= E(\varphi(\tilde{X}) - \varphi(X)) - E[(\varphi(\tilde{X}) - \varphi(X)) \mathbb{1}(\Omega \setminus \Lambda)] + \varepsilon. \end{aligned} \tag{3.6}$$

Since $\text{Law}(X) = \text{Law}(\tilde{X})$, we see that $E(\varphi(\tilde{X}) - \varphi(X)) = 0$. Applying the Cauchy-Schwarz inequality, we finally obtain from (3.6)

$$Ed(X, Y) \leq 2\varepsilon^{1/p} (E|\varphi(X)|^q)^{1/q} + \varepsilon.$$

□

4. Exponential Ergodicity of the Stochastic Reaction–Diffusion Equation in the Hypocoelliptic Setting

Fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We consider the stochastic reaction–diffusion equation (1.1) evolving on a d -dimensional torus \mathbb{T}^d , $d \in \mathbb{N}$. Note that we have equipped this equation with periodic boundary conditions just to simplify the exposition and to emphasize the main ideas. Everything should work in a similar way for Eq. (1.1) on a bounded domain equipped with Dirichlet boundary conditions.

We are interested in analytically and probabilistically strong solutions to this equation. Recall the following notions.

Definition 4.1 (see, e.g., [9, Definition 1.7]). Let $u_0 \in L_2(\mathbb{T}^d)$. An $L_2(\mathbb{T}^d)$ -valued, continuous in time process $u: \Omega \times [0, \infty) \rightarrow L_2(\mathbb{T}^d)$ is called an *analytically generalized strong solution* to (1.1) with initial condition u_0 if $u(0) = u_0$ and for any $\varepsilon > 0$, $t > \varepsilon$ the following holds:

$$\int_{\varepsilon}^t \|\Delta u(s)\|_{L_2(\mathbb{T}^d)} + \|f(u(s, \cdot), \cdot)\|_{L_2(\mathbb{T}^d)} ds < \infty, \quad \mathbb{P}\text{-a.s.}; \tag{4.1}$$

$$u(t) = u(\varepsilon) + \int_{\varepsilon}^t (\Delta u(s) + f(u(s, \cdot), \cdot)) ds + \sum_{k=1}^m \sigma_k (W^k(t) - W^k(\varepsilon)), \quad \mathbb{P}\text{-a.s.} \tag{4.2}$$

Definition 4.2 (see, e.g., [7, Section 6.1], [9, Definition 1.3]). Let $u_0 \in L_2(\mathbb{T}^d)$. An $L_2(\mathbb{T}^d)$ -valued, continuous in time process $u: \Omega \times [0, \infty) \rightarrow L_2(\mathbb{T}^d)$ is called an *analytically strong solution* to (1.1) with initial condition u_0 if conditions (4.1) and (4.2) hold for $\varepsilon = 0$ and any $t > 0$.

Definition 4.3 (see, e.g., [14, Remark 5.2]). A solution to (1.1) (strong or generalized strong) is called a *probabilistically strong solution* if it is adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$.

In case of ambiguity, a solution to (1.1) with initial condition $x \in L_2(\mathbb{T}^d)$ will be denoted by u^x . We refer to [7, 9, 14] for a detailed discussion of relations between different notions of a solution to SPDEs.

We will make the following assumption on f and σ_i .

Assumption A. *The function f is jointly continuous and locally Lipschitz in the first variable. Moreover, the following condition holds (dissipativity outside of a compact set): there exist $K_1, K_2, K_3 > 0$ such that for any $x, y \in \mathbb{R}, \xi \in \mathbb{T}^d$*

$$xf(x, \xi) \leq K_1 - K_2x^2; \tag{4.3}$$

$$(x - y)(f(x, \xi) - f(y, \xi)) \leq K_3(x - y)^2. \tag{4.4}$$

Suppose that

$$\sigma_k \in \mathcal{C}^2(\mathbb{T}^d), \quad \text{for each } k = 1, \dots, m. \tag{4.5}$$

As shown in Theorem 4.4 below, the condition on σ guarantees local existence of strong solution and condition (4.3) prevents blow-up. Conditions (4.3) and (4.4) are satisfied, for example, for $f(x, \xi) := Kx - x^3$, $K \in \mathbb{R}$.

Recall the definition of the order \leq_{pos} in (2.8).

Theorem 4.4. *Suppose that Assumption A holds. Then*

- (i) *for any $u_0 \in C^2(\mathbb{T}^d)$ Eq. (1.1) has a unique analytically and probabilistically strong solution u with initial condition u_0 ; furthermore u has a version which is continuous on $[0, \infty) \times \mathbb{T}^d$;*
- (ii) *for any $u_0 \in L_2(\mathbb{T}^d)$ Eq. (1.1) has a unique analytically generalized strong and probabilistically strong solution u with initial condition u_0 ; furthermore u has a version which is continuous on $(0, \infty) \times \mathbb{T}^d$;*
- (iii) *this solution of Eq. (1.1) is a homogeneous Markov process with state space $L_2(\mathbb{T}^d)$;*
- (iv) *there exists a set $\Omega' \subset \Omega$ of full measure such that if $x, y \in L_2(\mathbb{T}^d)$ and $x \preceq_{pos} y$, then*

$$u^x(t, \omega) \preceq_{pos} u^y(t, \omega), \quad t \geq 0, \omega \in \Omega';$$

- (v) *the corresponding Markov transition function is order preserving with respect to the order \preceq_{pos} ;*
- (vi) *for any $x \in L_2(\mathbb{T}^d)$, $t \geq 0$ we have $\mathbb{E} \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2 < \infty$; moreover there exists a constant $C > 0$ such that the following energy estimates hold for any $x \in L_2(\mathbb{T}^d)$, $0 \leq s \leq t$*

$$\begin{aligned} \mathbb{E} \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2 &\leq \mathbb{E} \|u^x(s)\|_{L_2(\mathbb{T}^d)}^2 - K_2 \int_s^t \mathbb{E} \|u^x(r)\|_{L_2(\mathbb{T}^d)}^2 dr \\ &\quad + (K_1 + \|\sigma\|_{L_2(\mathbb{T}^d)}^2)(t - s), \end{aligned} \tag{4.6}$$

$$\mathbb{E} \|u^x(t)\|_{L_2(\mathbb{T}^d)}^4 \leq \|x\|_{L_2}^4 \exp(-K_2 t) + C, \tag{4.7}$$

where the constants K_1 and K_2 were defined in (4.3).

The proof of the theorem is given in Sect. 5.2. Let us make here a couple of remarks about the theorem. Usually in the literature it is assumed additionally that f is a polynomial or f is globally Lipschitz or that at least some growth bounds on $|f(x, \xi)|$ hold, see, e.g., [7, Section 7], [9, Section 4.2], [18, Section 8.3]. Here we established existence and uniqueness of the solutions to (1.1) without these additional restrictions. The main challenge is that the corresponding Nemytskii operator

$$F(u)(\xi) := f(u(\xi), \xi)$$

is not $L_2(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ (or at least $L_p(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ for some large p). Therefore it is difficult to apply here any fixed point principle. Furthermore if u^n is a sequence of solutions to (1.1) with smooth initial conditions u_0^n , and u^n converges to some u in L_2 , then it is not clear at all that this u is a solution to (1.1). To overcome these obstacles, inspired by some ideas from [9], we have extended the corresponding PDE result [27, Propositions 7.3.1] to L_2 initial conditions.

The fact that for irregular initial data ($u_0 \in L_2(\mathbb{T}^d)$) Eq. (1.1) might not have an analytically strong solution is not surprising. Indeed, even for the standard heat equation $\partial_t u = \Delta u$ on \mathbb{T}^1 we have $\|\Delta u(t)\|_{L_2(\mathbb{T}^d)} \leq Ct^{-1} \|u(0)\|_{L_2(\mathbb{T}^d)}$ and thus $\int_0^t \|\Delta u(s)\|_{L_2(\mathbb{T}^d)} ds$ might be infinite.

Now let us present the main result of this section and establish ergodicity of the stochastic reaction–diffusion equation in the hypoelliptic setting. By Theorem 4.4(iii) the solution to (1.1) is a Markov process with the state space $L_2(\mathbb{T}^d)$. Let $(P_t)_{t \geq 0}$ be the family of Markov transition probabilities associated with this process. We introduce the following condition.

Assumption B. *There exist $\varepsilon > 0$ and $\lambda_k \in \mathbb{R}$, $k = 1, \dots, m$ such that*

$$\sum_{i=1}^m \lambda_k \sigma_k(\xi) > \varepsilon \text{ for all } \xi \in \mathbb{T}^d.$$

Theorem 4.5. *Suppose that Assumptions A and B hold. Then the stochastic reaction–diffusion equation (1.1) has a unique invariant measure π . Furthermore, there exist $C > 0$, $\lambda > 0$ such that for all $x \in L_2(\mathbb{T}^d)$ we have*

$$W_{\|\cdot\|_{L_2} \wedge 1}(P_t(x, \cdot), \pi) \leq C(1 + \|x\|_{L_2}^2) \exp(-\lambda t), \quad t \geq 0. \tag{4.8}$$

Remark 4.6. Assumption B is satisfied for example if $m \geq 1$ and $\sigma_1(\xi) > \varepsilon$ for all $\xi \in \mathbb{T}^d$.

Remark 4.7. Assumption B is imposed only to guarantee the swap condition, see Lemma 5.9. If the solution to (1.1) satisfies (5.43), then only Assumptions A is needed in Theorem 4.5.

As mentioned in the introduction, it is clear that if $m = 0$, then Eq. (1.1) might have multiple invariant measures. On the other hand, if $m = \infty$, Eq. (1.1) has a unique invariant measure (of course, under certain assumption on the rate of decay of $\|\sigma_k\|_{L_2}$ as $k \rightarrow \infty$) [6, Sections 7 and 11]. The first result showing uniqueness of the invariant measure for finite m (that is when noise acts not in all directions) is due to Hairer [12], who showed it for m large enough. This was later improved by Hairer and Mattingly [18, Remark 8.22], where uniqueness of the invariant measure and exponential ergodicity is proven if $d = 1$ and $m = 3$. Remark 4.6 refines this result and shows that if the noise acts only in one direction and the corresponding σ_1 is positive, then the solution to Eq. (1.1) is exponentially ergodic. Note that contrary to [18, Remark 8.22] we also do not rely here on the specific polynomial form of the drift f and do not assume that the space dimension $d \leq 3$ as in [18, Assumption RD.2].

The next result shows that, in general, the convergence of transition probabilities to the invariant measure in (4.8) might not take place in the total variation metric, and thus the $W_{\|\cdot\|_{L_2} \wedge 1}$ metric can not be replaced in (4.8) by the d_{TV} metric. As discussed in the introduction, this happens due to the fact that the stochastic system do not get “enough” noise. A related result for the case when the drift f is linear can be found in [15, Proposition 9.7].

Theorem 4.8. *Suppose that Assumption A holds for a function f which does not depend on space (so $f = f(x)$, $x \in \mathbb{R}$). Then there exists a function σ satisfying additionally Assumption B and an initial condition $x \in L_2(\mathbb{T}^d)$, such that*

$$d_{TV}(P_t(x, \cdot), \pi) = 1,$$

for any $t \geq 0$.

The proof of Theorem 4.8 is given in Sect. 5.2.

Proof of Theorem 4.5. Let us verify that all conditions of Theorems 2.3 and 2.4 are satisfied. Let $E = L_2(\mathbb{T}^d)$ with partial order \leq_{pos} (recall its definition in (2.8)). Set

$$\begin{aligned} V(x) &= \|x\|_{L_2}^2, \quad d(x, y) := \|x - y\|_{L_2}^2 \quad x, y \in L_2(\mathbb{T}^d), \\ \varphi(x) &:= 2 \int_{\mathbb{T}^d} x(z)^2 \operatorname{sign}(x(z)) \, dz, \quad x \in L_2(\mathbb{T}^d). \end{aligned} \tag{4.9}$$

The Markov semigroup corresponding to (1.1) is order-preserving with respect to \preceq_{pos} by Theorem 4.4(v) and thus condition (1) of Theorem 2.3 holds. Condition (2) of Theorem 2.3 is satisfied thanks to energy bound (4.6). It follows from Example 2.8 and the definition of φ that condition (3) of Theorem 2.3 is met. Condition (4) follows from (4.6) with $M(x) := 2\|x\|_{L_2(\mathbb{T}^d)}^4 + C$, $x \in L_2(\mathbb{T}^d)$. By Lemma 5.9 condition (5) holds. Conditions (6) and (7) hold trivially with $\delta = 1/2$ and $\kappa = 1/2$, respectively. Thus, all conditions of Theorems 2.3 and 2.4 are satisfied. Therefore the process u has a unique invariant measure and (4.8) holds. \square

Finally, we establish exponentially fast synchronization by noise for solutions to (1.1): that is, we will prove that any two solutions with initial conditions $x, y \in L_2(\mathbb{T}^d)$ launched with the same noise will converge to each other in probability exponentially fast.

Theorem 4.9. *Suppose that Assumptions A and B hold. Then there exist $C > 0$, $\lambda > 0$ such that for any $x, y \in L_2(\mathbb{T}^d)$ we have*

$$\mathbf{E}[\|u^x(t) - u^y(t)\|_{L_2(\mathbb{T}^d)} \wedge 1] \leq C(1 + \|x\|_{L_2(\mathbb{T}^d)}^2 + \|y\|_{L_2(\mathbb{T}^d)}^2) \exp(-\lambda t), \quad t \geq 0. \quad (4.10)$$

Note that the fact that $\mathbf{E}[\|u^x(t) - u^y(t)\|_{L_2(\mathbb{T}^d)} \wedge 1] \rightarrow 0$ as $t \rightarrow \infty$ follows immediately from Theorem 4.5 and [8, Proposition 2.4]. Unfortunately, as discussed in Sect. 3, the techniques developed in [8] do not provide a quantitative bound on convergence rate. Therefore to prove this theorem we use the toolkit developed in Sect. 3.

Proof of Theorem 4.9. Fix arbitrary $x_0, y_0 \in L_2(\mathbb{T}^d)$, $t \geq 0$. Define

$$z_-(\xi) := x_0(\xi) \wedge y_0(\xi), \quad z_+(\xi) := x_0(\xi) \vee y_0(\xi), \quad \xi \in \mathbb{T}^d.$$

Put $X := u^{z_-}(t)$, $Y := u^{z_+}(t)$. Then by the order-preserving property (Theorem 4.4(iv)) we have

$$X \preceq_{pos} u^{x_0}(t) \preceq_{pos} Y, \quad X \preceq_{pos} u^{y_0}(t) \preceq_{pos} Y, \quad \text{a.s.}$$

Therefore

$$[\|u^{x_0}(t) - u^{y_0}(t)\|_{L_2(\mathbb{T}^d)} \wedge 1] \leq [\|X - Y\|_{L_2(\mathbb{T}^d)} \wedge 1]. \quad (4.11)$$

Now let us check that all the conditions of Lemma 3.3 hold. Let $E = L_2(\mathbb{T}^d)$ with partial order \preceq_{pos} . Set $d(x, y) := \|x - y\|_{L_2}^2 \wedge 1$, $x, y \in L_2(\mathbb{T}^d)$, and define φ as in (4.9). Then it follows from Example 2.8 that condition (3) of Theorem 2.3 is satisfied. Furthermore, if $d(x, y) = 1$, then we have

$$|\varphi(x) - \varphi(y)| \leq (|\varphi(x)| + |\varphi(y)|) \leq 2d(x, y)^{1/2}(\|x\|_{L_2}^2 + \|y\|_{L_2}^2). \quad (4.12)$$

If $d(x, y) < 1$, then one has

$$\begin{aligned} |\varphi(x) - \varphi(y)| &\leq 2 \int_{\mathbb{T}^d} |x(z)^2 \mathbb{1}(x(z) > 0) - y(z)^2 \mathbb{1}(y(z) > 0)| dz \\ &\quad + 2 \int_{\mathbb{T}^d} |x(z)^2 \mathbb{1}(x(z) < 0) - y(z)^2 \mathbb{1}(y(z) < 0)| dz \\ &\leq 4 \int_{\mathbb{T}^d} |x(z) - y(z)| (|x(z)| + |y(z)|) dz \\ &\leq 4\sqrt{2}\|x - y\|_{L_2} (\|x\|_{L_2} + \|y\|_{L_2}) \\ &\leq 4\sqrt{2}d(x, y)^{1/2} (2 + \|x\|_{L_2}^2 + \|y\|_{L_2}^2), \end{aligned} \quad (4.13)$$

where in the second inequality we used the following bound

$$|a^2 \mathbb{1}(a > 0) - b^2 \mathbb{1}(b > 0)| \leq |a - b|(|a| + |b|), \quad a, b \in \mathbb{R}.$$

Combining (4.12) and (4.13), we get

$$|\varphi(x) - \varphi(y)| \leq d(x, y)^{1/2}(\psi(x) + \psi(y)),$$

where we put $\psi(x) := 4\sqrt{2}(1 + \|x\|_{L_2}^2)$, $x \in L_2(\mathbb{T}^d)$. Thus, bound (3.2) holds.

It is also clear that by definition we have $X \preceq_{pos} Y$. Further, by Theorem 4.4(vi) we have

$$\mathbb{E}|\varphi(X)| \leq 2\mathbb{E}\|X\|_{L_2}^2 = 2\mathbb{E}\|u^{z^-}(t)\|_{L_2}^2 < \infty$$

and, similarly, $\mathbb{E}|\varphi(Y)| < \infty$. Finally, by Theorem 4.5 there exist $C, \lambda > 0$ such that

$$\begin{aligned} W_d(\text{Law}(X), \text{Law}(Y)) &\leq W_{\|\cdot\|_{L_2} \wedge 1}(\text{Law}(X), \text{Law}(Y)) \\ &\leq C(1 + \|z_-\|_{L_2}^2 + \|z_+\|_{L_2}^2) \exp(-\lambda t) \\ &= C(1 + \|x_0\|_{L_2}^2 + \|y_0\|_{L_2}^2) \exp(-\lambda t). \end{aligned}$$

Thus, all the conditions of Lemma 3.3 are met. Therefore inequality (3.3) and energy bound (4.7) yield

$$\begin{aligned} \mathbb{E}d(X, Y) &\leq C(1 + \|x_0\|_{L_2}^2 + \|y_0\|_{L_2}^2) \exp(-\lambda t/2)(1 + \sqrt{\mathbb{E}\|X\|_{L_2}^4} + \sqrt{\mathbb{E}\|Y\|_{L_2}^4}) \\ &\leq C(1 + \|x_0\|_{L_2}^2 + \|y_0\|_{L_2}^2) \exp(-\lambda t/2)(1 + \|z_-\|_{L_2}^2 + \|z_+\|_{L_2}^2) \\ &= C(1 + \|x_0\|_{L_2}^2 + \|y_0\|_{L_2}^2)^2 \exp(-\lambda t/2). \end{aligned}$$

This bound allows us to conclude.

$$\mathbb{E}\sqrt{d(X, Y)} \leq \sqrt{\mathbb{E}d(X, Y)} \leq C(1 + \|x_0\|_{L_2}^2 + \|y_0\|_{L_2}^2) \exp(-\lambda t/4).$$

Taking into account (4.11), we obtain (4.10). □

5. Proofs

5.1. Proofs of the results of Sect. 2. To prove Theorem 2.3 we need a couple of lemmas. The first lemma is quite standard and is in the spirit of [31, Section 15.2]. The lemma deals with the connection between the finiteness of the exponential moment of the first return time of a Markov process to a set and the existence of a Lyapunov function. However we were not able to find in the literature precisely this statement (we found only a number of related ones). Thus, for the convenience of the reader and for the sake of completeness we provide the full proof of the lemma.

Let us consider a measurable space $(\tilde{E}, \tilde{\mathcal{E}})$ and a Markov kernel Q on it. Let $Z = (Z_n)_{n \in \mathbb{Z}_+}$ be a Markov process with transition kernel Q . For a set $A \in \tilde{\mathcal{E}}$ introduce the first return time to A

$$\tau_A := \inf\{n \geq 1 : Z_n \in A\}.$$

Lemma 5.1. *Suppose there exists a measurable function $\mathcal{V}: \tilde{E} \rightarrow [1, \infty)$ such that for some $\lambda \in (0, 1)$, $K > 0$*

$$Q\mathcal{V} \leq \lambda\mathcal{V} + K. \quad (5.1)$$

Fix

$$M > (K/(1 - \lambda)) \vee 1 \quad (5.2)$$

and put $r := (\lambda + K/M)^{-1}$.

(i) We have

$$\mathbf{E}_x r^{\tau(\mathcal{V} \leq M)} \leq \mathcal{V}(x), \quad x \in \tilde{E} \setminus \{\mathcal{V} \leq M\}; \quad (5.3)$$

$$\mathbf{E}_x r^{\tau(\mathcal{V} \leq M)} \leq rQ\mathcal{V}(x), \quad x \in \{\mathcal{V} \leq M\}. \quad (5.4)$$

(ii) Further, suppose that for a set $A \in \tilde{\mathcal{E}}$ we have for some $\varepsilon > 0$

$$\inf_{x \in \{\mathcal{V} \leq M\}} Q(x, A) \geq \varepsilon. \quad (5.5)$$

Then, for any $1 < l < r^{\frac{|\log(1-\varepsilon)| \wedge \log M}{2 \log M}}$ there exists a constant $C = C(\lambda, l, K, M, \varepsilon) > 0$ such that for any $x \in \tilde{E}$

$$\mathbf{E}_x l^{\tau_A} \leq C\mathcal{V}(x). \quad (5.6)$$

Proof. (i) The proof uses the standard argument. First note, that it follows from the Lyapunov condition (5.1) and the definition of r that

$$Q\mathcal{V}(x) \leq r^{-1}\mathcal{V}(x), \quad x \in \tilde{E} \setminus \{\mathcal{V} \leq M\}. \quad (5.7)$$

Put now for $n \in \mathbb{Z}_+$

$$\tau^{(n)} := \tau_{\{\mathcal{V} \leq M\}} \wedge n \wedge \inf\{k \in \mathbb{Z}_+ : r^k \mathcal{V}(Z_k) \geq n\}.$$

Then by Dynkin's formula for discrete time Markov chains (see, e.g., [31, Theorem 11.3.1]) we have for any $x \in \tilde{E}$

$$\mathbf{E}_x r^{\tau^{(n)}} \mathcal{V}(Z_{\tau^{(n)}}) = \mathcal{V}(x) + \mathbf{E}_x \left[\sum_{i=1}^n r^{i-1} \mathbb{1}(i-1 < \tau^{(n)}) (\mathbf{E}[r\mathcal{V}(Z_i) | Z_{i-1}] - \mathcal{V}(Z_{i-1})) \right]. \quad (5.8)$$

If $x \in \tilde{E} \setminus \{\mathcal{V} \leq M\}$, then by definition for any $i \in [1, \tau^{(n)}]$ we have $Z_{i-1} \in \tilde{E} \setminus \{\mathcal{V} \leq M\}$. Therefore, (5.7) implies for any $i \in [1, n]$

$$r^{i-1} \mathbb{1}(i-1 < \tau^{(n)}) (\mathbf{E}[r\mathcal{V}(Z_i) | Z_{i-1}] - \mathcal{V}(Z_{i-1})) \leq 0.$$

Combining this with (5.8), we get

$$\mathbf{E}_x r^{\tau^{(n)}} \leq \mathbf{E}_x r^{\tau^{(n)}} \mathcal{V}(Z_{\tau^{(n)}}) \leq \mathcal{V}(x), \quad x \in \tilde{E} \setminus \{\mathcal{V} \leq M\},$$

which by Fatou's lemma yields (5.3).

If $x \in \{\mathcal{V} \leq M\}$, then by the Markov property and bound (5.3),

$$\begin{aligned} \mathbf{E}_x r^{\tau(\mathcal{V} \leq M)} &= r\mathbf{P}_x(Z_1 \in \{\mathcal{V} \leq M\}) + r\mathbf{E}_x[\mathcal{V}(Z_1) \mathbb{1}(Z_1 \notin \{\mathcal{V} \leq M\})] \leq r\mathbf{E}_x \mathcal{V}(Z_1) \\ &= rQ\mathcal{V}(x), \end{aligned}$$

which is (5.4).

(ii) Fix $l \in (1, r^{\frac{|\log(1-\varepsilon)| \wedge \log M}{2 \log M}})$. For $n \in \mathbb{Z}_+$ denote by $\tau^M(n)$ the n th return time to the set $\{\mathcal{V} \leq M\}$:

$$\begin{aligned} \tau^M(1) &:= \inf\{n \geq 0: Z_n \in \{\mathcal{V} \leq M\}\}; \\ \tau^M(n) &:= \inf\{n \geq \tau^M(n-1) + 1: Z_n \in \{\mathcal{V} \leq M\}\}, \quad n \geq 2. \end{aligned}$$

Introduce random variables

$$I_n := \mathbb{1}(Z_{\tau^M(n)+1} \in A), \quad n \in \mathbb{Z}_+.$$

We see that the event $\{I_n = 1\}$ corresponds to the reaching of the desired set A after n th visit to the set $\{\mathcal{V} \leq M\}$. We have for any $x \in \tilde{E}$

$$\begin{aligned} \mathbb{E}_x l^{\tau^M} &\leq \sum_{k=1}^{\infty} \mathbb{E}_x [l^{\tau^M(k)+1} \mathbb{1}(I_1 = \dots = I_{k-1} = 0, I_k = 1)] \\ &\leq \sum_{k=1}^{\infty} \mathbb{E}_x [l^{\tau^M(k)+1} \mathbb{1}(I_1 = \dots = I_{k-1} = 0)]. \end{aligned} \tag{5.9}$$

Define

$$D(M, l) := \sup_{x \in \{\mathcal{V} \leq M\}} \mathbb{E}_x [l^{\tau(\mathcal{V} \leq M)} \mathbb{1}(Z_1 \notin A)].$$

Note that by the strong Markov property and the definition above, for any $x \in \tilde{E}$

$$\begin{aligned} &\mathbb{E}_x [l^{\tau^M(k)+1} \mathbb{1}(I_1 = \dots = I_{k-1} = 0)] \\ &= \mathbb{E}_x [l^{\tau^M(k-1)+1} \mathbb{1}(I_1 = \dots = I_{k-2} = 0) \mathbb{E}_{Z_{\tau^M(k-1)}} (l^{\tau(\mathcal{V} \leq M)} \mathbb{1}(Z_1 \notin A))] \\ &\leq D(M, l) \mathbb{E}_x [l^{\tau^M(k-1)+1} \mathbb{1}(I_1 = \dots = I_{k-2} = 0)] \\ &\leq D(M, l)^{k-1} \mathbb{E}_x [l^{\tau^M(1)+1}] \\ &\leq l D(M, l)^{k-1} (\mathcal{V}(x) \vee M), \end{aligned} \tag{5.10}$$

where in the last inequality we used (5.3) and the fact that $l < r$. Applying the Cauchy–Schwarz inequality and (5.4), we deduce

$$D(M, l) \leq \sup_{x \in \{\mathcal{V} \leq M\}} (\mathbb{E}_x [l^{2\tau(\mathcal{V} \leq M)}])^{1/2} (1 - \varepsilon)^{1/2} \leq M^{\log l / \log r} (1 - \varepsilon)^{1/2} < 1, \tag{5.11}$$

where the last inequality follows from the definitions of l and r . Thus, (5.11) together with (5.9), (5.10), and the obvious bound $(\mathcal{V}(x) \vee M) \leq M \mathcal{V}(x)$ yields

$$\mathbb{E}_x l^{\tau^M} \leq l (\mathcal{V}(x) \vee M) \sum_{k=1}^{\infty} D(M, l)^{k-1} \leq C(\lambda, l, K, M, \varepsilon) \mathcal{V}(x),$$

which implies (5.6). □

The next lemma explains why conditions (3) and (4) are imposed in Theorem 2.3.

Lemma 5.2. *Assume that condition (3) of Theorem 2.3 is satisfied for measurable functions $\varphi: E \rightarrow \mathbb{R}$, $d: E \times E \rightarrow [0, 1]$. Let $X, Y, \tilde{X}, \tilde{Y}$ be random elements $\Omega \rightarrow E$ such that $\text{Law}(X) = \text{Law}(\tilde{X})$, $\text{Law}(Y) = \text{Law}(\tilde{Y})$ and*

$$\mathbf{P}(X \preceq Y) \geq 1 - \varepsilon_1, \quad \mathbf{P}(\tilde{Y} \preceq \tilde{X}) \geq 1 - \varepsilon_2$$

for some $\varepsilon_1, \varepsilon_2 > 0$. Then

$$\mathbf{E}d(X, Y) \leq 2\sqrt{\varepsilon_1 + \varepsilon_2}(\mathbf{E}\varphi(X)^2)^{1/2} + \varepsilon_1 + \varepsilon_2.$$

Proof. It follows from the gluing lemma (see, e.g., [38, p. 23]) that there exist random elements Z_1, Z_2, Z_3 such that $\text{Law}(Z_1, Z_2) = \text{Law}(X, Y)$ and $\text{Law}(Z_2, Z_3) = \text{Law}(\tilde{Y}, \tilde{X})$. Clearly, we also have $\text{Law}(Z_1) = \text{Law}(Z_3)$.

Let Λ be the set $\{\omega : Z_1(\omega) \preceq Z_2(\omega) \preceq Z_3(\omega)\}$. Then by above

$$\mathbf{P}(\Lambda) \geq \mathbf{P}(Z_1 \preceq Z_2) + \mathbf{P}(Z_2 \preceq Z_3) - 1 = \mathbf{P}(X \preceq Y) + \mathbf{P}(\tilde{Y} \preceq \tilde{X}) - 1 \geq 1 - \varepsilon_1 - \varepsilon_2.$$

Thus, we immediately get from Lemma 3.4 that

$$\mathbf{E}d(X, Y) = \mathbf{E}d(Z_1, Z_2) \leq 2\sqrt{\varepsilon_1 + \varepsilon_2}(\mathbf{E}\varphi(X)^2)^{1/2} + \varepsilon_1 + \varepsilon_2.$$

□

Proof of Theorem 2.3. The proof of the theorem consists of several steps.

Step 1. In this step we fix $x, y \in E$ and $t > 0$. Let $\{X^x(s), s \geq 0\}$ and $\{X^y(s), s \geq 0\}$ be independent Markov processes with the laws \mathbf{P}_x and \mathbf{P}_y , correspondingly. Let $\mathcal{F}_s := \sigma(X^x(r), X^y(r), r \leq s)$, $s \leq t$, be the natural filtration. Introduce stopping times

$$\begin{aligned} \tau_{x \preceq y} &:= \inf\{n \in \mathbb{Z}_+ : X^x(n) \preceq X^y(n)\}, \\ \tau_{y \preceq x} &:= \inf\{n \in \mathbb{Z}_+ : X^y(n) \preceq X^x(n)\}. \end{aligned}$$

Note that by definition these stopping times $\tau_{x \preceq y}$ and $\tau_{y \preceq x}$ take only countably many values.

Let us now extend the state space E and add an additional element denoted by \diamond . We assume that $\diamond \preceq \diamond$ and do not impose any further partial order relations between \diamond and other elements of E .

Introduce the following random elements:

$$\begin{aligned} \eta^x &:= X^x(t) \mathbb{1}(\tau_{x \preceq y} \leq t) + \diamond \mathbb{1}(\tau_{x \preceq y} > t), \\ \eta^y &:= X^y(t) \mathbb{1}(\tau_{x \preceq y} \leq t) + \diamond \mathbb{1}(\tau_{x \preceq y} > t). \end{aligned}$$

We claim now that $\text{Law}(\eta^x) \preceq_{st} \text{Law}(\eta^y)$. Indeed, let $f: E \cup \{\diamond\} \rightarrow \mathbb{R}$ be an arbitrary bounded measurable increasing function. Then

$$\mathbf{E}f(\eta^x) = f(\diamond)\mathbf{P}(\tau_{x \preceq y} > t) + \sum_{i=0}^{\lfloor t \rfloor} \mathbf{E}[f(X^x(t)) \mathbb{1}(\tau_{x \preceq y} = i)]. \tag{5.12}$$

Note that for any $i = 0, \dots, \lfloor t \rfloor$ we have

$$\begin{aligned} \mathbf{E}[f(X^x(t)) \mathbb{1}(\tau_{x \preceq y} = i)] &= \mathbf{E}[\mathbb{1}(\tau_{x \preceq y} = i) \mathbf{E}(f(X^x(t)) | \mathcal{F}_i)] \\ &= \mathbf{E}[\mathbb{1}(\tau_{x \preceq y} = i) \int_E f(z) P_{t-i}(X^x(i), dz)]. \end{aligned} \tag{5.13}$$

Recall that the kernel P_{t-i} is order preserving and thus for any $z_1, z_2 \in E$ such that $z_1 \leq z_2$ we have

$$\int_E f(z)P_{t-i}(z_1, dz) \leq \int_E f(z)P_{t-i}(z_2, dz).$$

Since on the set $\{\tau_{x \leq y} = i\}$ we have $X^x(i) \leq X^y(i)$, we can continue (5.13) in the following way:

$$\begin{aligned} \mathbf{E}[f(X^x(t)) \mathbb{1}(\tau_{x \leq y} = i)] &\leq \mathbf{E}\left[\mathbb{1}(\tau_{x \leq y} = i) \int_E f(z)P_{t-i}(X^y(i), dz)\right] \\ &= \mathbf{E}[f(X^y(t)) \mathbb{1}(\tau_{x \leq y} = i)]. \end{aligned}$$

Combining this with (5.12), we finally deduce

$$\mathbf{E}f(\eta^x) \leq f(\diamond)\mathbf{P}(\tau_{x \leq y} > t) + \sum_{i=0}^{\lfloor t \rfloor} \mathbf{E}[f(X^y(t)) \mathbb{1}(\tau_{x \leq y} = i)] = \mathbf{E}f(\eta^y).$$

Since f was an arbitrary bounded measurable increasing function, we see that indeed $\text{Law}(\eta^x) \leq_{st} \text{Law}(\eta^y)$.

Step 2. We have shown that $\text{Law}(\eta^x) \leq_{st} \text{Law}(\eta^y)$. Therefore, by the Strassen theorem (see, e.g., [26, Theorem IV.2.4]) there exist random variables ξ^x and ξ^y such that $\text{Law}(\xi^x) = \text{Law}(\eta^x)$, $\text{Law}(\xi^y) = \text{Law}(\eta^y)$ and $\xi^x \leq \xi^y$ a.s.

It follows from the construction, that $\mathbf{P}(\eta^x \neq X^x(t)) \leq \mathbf{P}(\tau_{x \leq y} > t)$ and $\mathbf{P}(\eta^y \neq X^y(t)) \leq \mathbf{P}(\tau_{x \leq y} > t)$. Now we will apply twice the gluing lemma (see, e.g., [38, p. 23]). First, we apply it to the pairs $(X^x(t), \eta^x)$ and (ξ^x, ξ^y) and construct two random variables Y^x and Y^y such that $\text{Law}(Y^x) = \text{Law}(X^x(t))$, $\text{Law}(Y^y) = \text{Law}(\xi^y) = \text{Law}(\eta^y)$ and $\mathbf{P}(Y^x \leq Y^y) \geq 1 - \mathbf{P}(\tau_{x \leq y} > t)$. Then, we apply the gluing lemma to the pairs (Y^x, Y^y) and (η^y, X^y) . We deduce that there exist random variables Z^x and Z^y such that

$$\begin{aligned} \text{Law}(Z^x) &= \text{Law}(X^x(t)), \quad \text{Law}(Z^y) = \text{Law}(X^y(t)); \\ \mathbf{P}(Z^x \leq Z^y) &\geq 1 - 2\mathbf{P}(\tau_{x \leq y} > t). \end{aligned}$$

In a similar way, we construct another pair of random variables \tilde{Z}^x and \tilde{Z}^y with the following properties

$$\begin{aligned} \text{Law}(\tilde{Z}^x) &= \text{Law}(X^x(t)), \quad \text{Law}(\tilde{Z}^y) = \text{Law}(X^y(t)); \\ \mathbf{P}(\tilde{Z}^y \leq \tilde{Z}^x) &\geq 1 - 2\mathbf{P}(\tau_{y \leq x} > t). \end{aligned}$$

Step 3. Now it follows from Step 2 and Lemma 5.2 that

$$\begin{aligned} W_{d \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) &\leq \mathbf{E}(d(Z^x, Z^y) \wedge 1) \\ &\leq 2\sqrt{2}(\mathbf{P}(t \leq \tau_{y \leq x}) + \mathbf{P}(t \leq \tau_{x \leq y}))^{1/2}(1 + M(x)^{1/2}). \end{aligned} \tag{5.14}$$

Thus everything boils down to bounding the probabilities $\mathbf{P}(t \leq \tau_{y \leq x})$ and $\mathbf{P}(t \leq \tau_{x \leq y})$. We will do it using Lemma 5.1.

Step 4. First we note that, clearly, $\{(x_1, x_2) \in E \times E : x_1 \leq x_2\} \supset A \times B$ (recall the definitions of the sets A and B in condition (5) of the theorem). Therefore,

$$\tau_{x \leq y} \leq \inf\{n \in \mathbb{N} : (X^x(n), X^y(n)) \in A \times B\}. \tag{5.15}$$

Now we apply Lemma 5.1(ii) to the state space $\tilde{E} := E \times E$, the kernel Q on it defined by

$$Q((x_1, x_2), A_1 \times A_2) := P_1(x_1, A_1)P_1(x_2, A_2), \quad x_1, x_2 \in E, A_1, A_2 \in \mathcal{E},$$

the set $A \times B$, and the Lyapunov function $\mathcal{V}(x_1, x_2) := 1 + V(x_1) + V(x_2)$, $x_1, x_2 \in E$. It follows from (2.2) and the Gronwall inequality that

$$Q\mathcal{V}(x_1, x_2) \leq e^{-\gamma}\mathcal{V}(x_1, x_2) + (2K/\gamma + 1)(1 - e^{-\gamma}).$$

Therefore, condition (5.1) is met. Take $M = 4K/\gamma + 1$. It is clear that this choice of M satisfies (5.2). Thanks to assumption (2.3) and the definitions of \mathcal{V} and the sets A, B , we see that condition (5.5) holds for the set $A \times B$ in place of A and the positive constant M chosen above. Therefore all conditions of Lemma 5.1(ii) are satisfied and, thus there exist constants $C > 0, \lambda > 0$ that do not depend on x, y such that

$$\mathbf{E}_{x,y}e^{\lambda\tau_{A \times B}} \leq C(1 + V(x) + V(y)).$$

This, combined with (5.15) and the Chebyshev inequality implies

$$\mathbf{P}(t \leq \tau_{y \leq x}) \leq C(1 + V(x) + V(y))e^{-\lambda t}.$$

Using exactly the same argument, we also get that

$$\mathbf{P}(t \leq \tau_{x \leq y}) \leq C(1 + V(x) + V(y))e^{-\lambda t}.$$

Substituting these bounds into (5.14), we finally deduce

$$W_{d \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) \leq C(1 + V(x) + V(y))(1 + M(x))e^{-\lambda t/2}.$$

Taking into account that obviously $W_{d \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) \leq 1$, we obtain (2.4). \square

Proof of Theorem 2.4. We begin by observing that by Jensen’s inequality and condition (6) of the theorem we have for any measures $\mu, \nu \in \mathcal{P}(E)$

$$W_{\rho \wedge 1}(\mu, \nu) \leq W_{d^\delta \wedge 1}(\mu, \nu) \leq (W_{d \wedge 1}(\mu, \nu))^\delta. \tag{5.16}$$

We apply now Theorem 2.3 with $\theta := \frac{\kappa}{\delta(1+\kappa)}$. Taking into account (5.16), we obtain that there exist $C > 0, \lambda > 0$ such that for any $x, y \in E$ we have

$$W_{\rho \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) \leq C(1 + V(x) + V(y))e^{-\lambda t}, \quad t \geq 0 \tag{5.17}$$

where we have also used condition (7) of the theorem.

From here the proof follows the standard line of argument. Note that by [21, Theorem 17.24] for any fixed $t \geq 0$, the mapping $x \mapsto P_t(x, \cdot)$ is measurable. Therefore by [38, Theorem 4.8 and Corollary 5.22] the mapping $(x, y) \mapsto W_{\rho \wedge 1}(P_t(x, \cdot), P_t(y, \cdot))$ is measurable and the function $W_{\rho \wedge 1}$ is convex in both arguments. Therefore we immediately get from (5.17) that for any $\mu, \nu \in \mathcal{P}(E)$

$$W_{\rho \wedge 1}(P_t\mu, P_t\nu) \leq C(1 + \mu(V) + \nu(V))e^{-\lambda t}, \quad t \geq 0. \tag{5.18}$$

Now we are ready to prove the existence of an invariant measure for $\{P_t, t \geq 0\}$. Fix any $x \in E$. We make use of (5.18) to obtain that for any $t, s > 0$

$$\begin{aligned} W_{\rho \wedge 1}(P_t(x, \cdot), P_{t+s}(x, \cdot)) &= W_{\rho \wedge 1}(P_t \delta_x, P_t(P_s \delta_x)) \\ &\leq C(1 + V(x) + P_s V(x))e^{-\lambda t} \\ &\leq C(1 + 2V(x) + K/\gamma)e^{-\lambda t}. \end{aligned}$$

Thus the sequence of measures $\{P_t(x, \cdot)\}$ is Cauchy in the space $(\mathcal{P}(E), W_{\rho \wedge 1})$. Since this space is complete (see, e.g., [38, Theorem 6.18]), there exists a measure $\pi \in \mathcal{P}(E)$ such that $W_{\rho \wedge 1}(P_t(x, \cdot), \pi) \rightarrow 0$ as $t \rightarrow \infty$. Hence for any $y \in E$ we have

$$W_{\rho \wedge 1}(P_t(y, \cdot), \pi) \leq W_{\rho \wedge 1}(P_t(x, \cdot), \pi) + W_{\rho \wedge 1}(P_t(x, \cdot), P_t(y, \cdot)) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

Therefore for any $t \geq 0$ by the monotone convergence theorem we deduce

$$W_{\rho \wedge 1}(P_t \pi, \pi) \leq \int_E W_{\rho \wedge 1}(P_t(y, \cdot), \pi) \pi(dy) \rightarrow 0 \text{ as } t \rightarrow \infty,$$

where we have used again that the function $W_{\rho \wedge 1}$ is convex. Therefore the measure π is invariant for $\{P_t, t \geq 0\}$.

Now let us show the uniqueness of the invariant measure. First, note that if π is an arbitrary invariant measure for $\{P_t, t \geq 0\}$, then $\pi(V) < \infty$, see [13, Proposition 4.24]. Thus if π_1 and π_2 are two invariant measures, then by (5.18) we have for any $t \geq 0$

$$W_{\rho \wedge 1}(\pi_1, \pi_2) = W_{\rho \wedge 1}(P_t \pi_1, P_t \pi_2) \leq C(1 + \pi_1(V) + \pi_2(V))e^{-\lambda t}.$$

Taking the limit in the right-hand side of the above inequality we get $\pi_1 = \pi_2$.

Finally, bound (2.5) follows directly from (5.18) and the fact that $\pi(V) < \infty$. □

Proof of Example 2.8. First let us show that the set Γ defined in (2.1) is closed. Let $(x_n)_{n \in \mathbb{Z}_+}, (y_n)_{n \in \mathbb{Z}_+}$ be two sequences of elements of $L_p(D, \mathbb{R})$ such that $x_n \preceq_{pos} y_n$ and $\|x_n - x\|_{L_p(D)} \rightarrow 0, \|y_n - y\|_{L_p(D)}$ as $n \rightarrow \infty$. We claim now that $x \preceq_{pos} y$.

Indeed, by passing to an appropriate subsequence (n_k) we see that $x_{n_k} \rightarrow x$ almost everywhere (a.e.) and $y_{n_k} \rightarrow y$ a.e. as $n \rightarrow \infty$. Since for each k we have $x_{n_k} \leq y_{n_k}$ a.e., it is now easily seen that $x \leq y$ a.e. and thus $x \preceq_{pos} y$. Therefore, the set Γ is closed.

Now let us introduce the function φ . Put

$$\varphi(x) := 2^{p-1} \int_D |x(z)|^p \text{sign}(x(z)) dz, \quad x \in L_p(D, \mathbb{R}).$$

Let us verify that the partial order \preceq_{pos} and functions φ and d satisfy condition (3) of Theorem 2.3. Our original proof of this was a bit long; the following simple proof was suggested to us by the referee, to whom we are very grateful for this.

Fix any $x, y \in L_p(D, \mathbb{R})$ such that $x \preceq_{pos} y$. Note that for any real numbers $a < b, p \geq 1$ one has

$$|a - b|^p \leq 2^{p-1} (|b|^p \text{sign}(b) - |a|^p \text{sign}(a)).$$

Using this inequality, we get

$$\begin{aligned} 0 \leq d(x, y) &= \|x - y\|_{L_p(D)}^p \\ &\leq 2^{p-1} \int_D (|y(z)|^p \text{sign}(y(z)) - |x(z)|^p \text{sign}(x(z))) dz \\ &= \varphi(y) - \varphi(x), \end{aligned}$$

and thus condition (3) of Theorem 2.3 holds. □

5.2. *Proofs of the results of Sect. 4.* We begin with the following auxiliary statement which provides Gronwall-type bounds that will be very useful in the sequel.

Lemma 5.3. *Let C_1, C_2, C_3, C_4 be positive constants. Let X be a continuous \mathcal{F}_t -adapted process taking values in $[0, \infty)$ such that $X(0) = x$ and*

$$X(t) \leq X(s) - C_1 \int_s^t X(r)dr + M(t) - M(s) + C_2(t - s), \quad 0 \leq s \leq t, \quad (5.19)$$

where M is a continuous local \mathcal{F}_t -martingale with $M(0) = 0$ and

$$d\langle M \rangle_t \leq (C_3X(t) + C_4)dt. \quad (5.20)$$

Then for any $t \geq 0$ we have $\mathbf{E}X(t) < +\infty$ and the following bounds hold:

$$\mathbf{E}X(t) \leq \mathbf{E}X(s) - C_1 \int_s^t \mathbf{E}X(r)dr + C_2(t - s), \quad 0 \leq s \leq t, \quad (5.21)$$

$$\mathbf{E}X(t)^2 \leq x^2 \exp(-C_1t) + C, \quad t \geq 0, \quad (5.22)$$

where $C = C(C_1, C_2, C_3, C_4) > 0$ is some constant independent of t and x .

Proof. We begin with the proof of (5.21). For $n \in \mathbb{N}$ introduce the stopping time $\tau_n := \inf\{t \geq 0 : |M(t)| \geq n\}$. Then it follows from (5.19) that for any $n \in \mathbb{N}, t \geq 0$

$$\mathbf{E}X(t \wedge \tau_n) \leq x + C_2(t \wedge \tau_n).$$

Using Fatou's lemma, we deduce that for any $t \geq 0$ we have $\mathbf{E}X(t) \leq x + C_2t$ and thus $\mathbf{E}X(t) < \infty$. This and (5.20) imply that for any $t \geq 0$ we have $\mathbf{E}\langle M \rangle_t \leq C_4t + C_3xt + C_2C_3t^2/2 < \infty$. Therefore M is a martingale and (5.21) follows immediately from (5.19).

To establish (5.22), we introduce a process Y , which is a solution to the following equation

$$Y(t) = x - C_1 \int_0^t Y(r)dr + M(t) + C_2t, \quad t \geq 0. \quad (5.23)$$

We claim that for any $t \geq 0$ we have $X(t) \leq Y(t)$. Indeed, assume the contrary and suppose that for some $T_0 > 0$ we have $X(T_0) > Y(T_0)$. Then, arguing as in [20, Proof of Proposition 9.2], we introduce

$$S_0 := \sup\{t \in [0, T_0] : X(t) \leq Y(t)\}.$$

Since the processes X and Y are continuous we have $S_0 < T_0$ and $X(t) > Y(t)$ for $t \in (S_0, T_0)$. Then (5.19) and (5.23) imply

$$X(T_0) - Y(T_0) \leq X(S_0) - Y(S_0) - C_1 \int_{S_0}^{T_0} (X(t) - Y(t)) dt < 0$$

which contradicts the fact that $X(T_0) > Y(T_0)$. Therefore such T_0 does not exist and $X(t) \leq Y(t)$ for any $t \geq 0$.

Note now that by Ito's formula and (5.20)

$$\begin{aligned} dY(t)^2 &= (-2C_1Y(t)^2 + 2C_2Y(t))dt + d\langle M \rangle_t + dN(t) \\ &\leq (-2C_1Y(t)^2 + (2C_2 + C_3)Y(t) + C_4)dt + dN(t) \\ &\leq (-C_1Y(t)^2 + C_5)dt + dN(t), \end{aligned} \quad (5.24)$$

where $C_5 := (2C_2 + C_3)^2/(4C_1) + C_4$ and N is a continuous local martingale.

Consider a stopping time $\sigma_n := \inf\{t \geq 0 : |N(t)| \geq n\}$. Then, using (5.24), we get for any $0 \leq s \leq t$

$$\mathbb{E}Y(t \wedge \sigma_n)^2 \leq \mathbb{E}Y(s \wedge \sigma_n)^2 - C_1 \int_s^t \mathbb{E}Y(r \wedge \sigma_n)^2 dr + C_5(t - s).$$

Therefore, the Gronwall inequality yields

$$\mathbb{E}X(t \wedge \sigma_n)^2 \leq \mathbb{E}Y(t \wedge \sigma_n)^2 \leq x^2 \exp(-C_1 t) + C_5/C_1, \quad t \geq 0.$$

The proof of (5.22) is completed by passing to the limit in the above inequality using Fatou’s lemma. \square

In all the remaining theorems and lemmas of this section we will assume that Assumption A is satisfied. To establish Theorem 4.4 we split the process u into the following two parts. The first part, denoted by w , is the stochastic convolution, that is a unique analytically and probabilistically strong solution of the following equation

$$dw(t, \xi) = \Delta w(t, \xi)dt + \sum_{k=1}^m \sigma_k(\xi)dW^k(t), \quad \xi \in \mathbb{T}^d, t \geq 0, \tag{5.25}$$

with the initial condition $w(0) = 0$.

Lemma 5.4. *The function w is well-defined. Furthermore, there exists a set $\Omega' \subset \Omega$ such that $\mathbb{P}(\Omega') = 1$ and for any $\varepsilon > 0, T > 0, \omega \in \Omega'$ there exists $C = C(\varepsilon, T, \omega)$ such that for any $\xi_1, \xi_2 \in \mathbb{T}^d, t_1, t_2 \in [0, T]$ one has*

$$|w(t_1, \xi_1, \omega) - w(t_2, \xi_2, \omega)| \leq C(|t_1 - t_2|^{1/2-\varepsilon} + |\xi_1 - \xi_2|^{1-\varepsilon}). \tag{5.26}$$

Proof. Existence and uniqueness of a strong solution of (5.25) follows from [7, Example 5.39] and assumption (4.5). Bound (5.26) follows from [7, Theorem 5.15(ii)]. \square

For $T > 0$ denote now

$$M_T = M_T(\omega) := \sup_{\substack{t \in [0, T] \\ \xi \in \mathbb{T}^d}} |w(t, \xi, \omega)|. \tag{5.27}$$

By Lemma 5.4, for any $\omega \in \Omega', T > 0$ we have $M_T(\omega) < \infty$.

The second ingredient of decomposition of u is much more tricky. We fix now arbitrary $\omega \in \Omega', T > 0$ and consider the (deterministic) PDE

$$dv(t, \xi) = \Delta v(t, \xi)dt + f(v(t, \xi) + w(t, \xi, \omega), \xi)dt, \quad \xi \in \mathbb{T}^d, t \in [0, T] \tag{5.28}$$

with the initial condition $v(0) = v_0 \in \mathcal{C}^2(\mathbb{T}^d)$. We begin with the case when the initial condition is smooth enough.

Lemma 5.5. *For any $v_0 \in \mathcal{C}^2(\mathbb{T}^d)$ Eq. (5.28) has a unique analytically strong solution v with the initial condition v_0 . Furthermore $v \in \mathcal{C}([0, T], \mathcal{C}^2(\mathbb{T}^d))$ and there exists $C > 0$ such that for any $t \in [0, T]$*

$$\|v(t)\|_{L_\infty(\mathbb{T}^d)} \leq C(\|v_0\|_{L_2(\mathbb{T}^d)} + M_t)(1 + t^{-d/4}) + C_f t, \tag{5.29}$$

where $C_f := \sup_{x \geq 0} \sup_{\eta \in \mathbb{T}^d} (\text{sign}(x)f(x, \eta)) \vee 0$.

Proof. We begin with the uniqueness part. Suppose that v and \bar{v} are two strong solutions to Eq. (5.28) with the same initial condition $v_0 \in \mathcal{C}^2(\mathbb{T}^d)$. Then, by the chain rule

$$\begin{aligned} & \frac{d}{dt} \|v(t) - \bar{v}(t)\|_{L_2(\mathbb{T}^d)}^2 \\ &= 2 \int_{\mathbb{T}^d} (v(t, \xi) - \bar{v}(t, \xi)) (f(v(t, \xi) + w(t, \xi, \omega), \xi) - f(\bar{v}(t, \xi) + w(t, \xi, \omega), \xi)) d\xi \\ & \quad - 2 \|\nabla(v(t) - \bar{v}(t))\|_{L_2(\mathbb{T}^d)}^2 \\ & \leq C \|v(t) - \bar{v}(t)\|_{L_2(\mathbb{T}^d)}^2, \end{aligned} \tag{5.30}$$

where the last inequality follows from assumption (4.4). Since $v(0) = \bar{v}(0)$, an application of the Gronwall lemma immediately implies that $v(t) = \bar{v}(t)$ for any $t \in [0, T]$.

To show existence of a strong solution to (5.28) we fix the initial condition $v_0 \in \mathcal{C}^2(\mathbb{T}^d)$ and $T > 0$. Note that the function

$$\mathbb{R} \times [0, T] \times \mathbb{T}^d \ni (x, t, \xi) \mapsto f(x + w(t, \xi, \omega), \xi) \tag{5.31}$$

is locally Lipschitz in x , Hölder in t and continuous in ξ thanks to (5.26) and Assumption A. Therefore it satisfies [27, Inequality (7.3.2)]. Thus, by [27, Propositions 7.3.1.i, 7.3.1.ii, 7.1.10.iii] equation (5.28) has a local solution on some interval $[0, \delta]$, $\delta > 0$ and this solution is in $\mathcal{C}([0, \delta], \mathcal{C}^2(\mathbb{T}^d))$. Let us now show that Eq. (5.28) has a global solution.

It follows from (4.3) that there exist $x_- \leq x_+ \in \mathbb{R}$ such that $f(x, \xi) \geq 0$ for any $x \leq x_-$, $\xi \in \mathbb{T}^d$ and $f(x, \xi) \leq 0$ for any $x \geq x_+$, $\xi \in \mathbb{T}^d$. Consider the following PDE

$$d\psi(t, \xi) = \Delta\psi(t, \xi)dt + (f(\psi(t, \xi) + w(t, \xi, \omega), \xi) \vee 0)dt, \quad \xi \in \mathbb{T}^d, t \geq 0, \tag{5.32}$$

with the initial condition $\psi(0) = (x_+ + M_T) \vee \sup_{\eta} v_0(\eta)$ (recall the definition of M_T in (5.27)). By above, the constant function

$$\psi(t, \xi) := (x_+ + M_T) \vee \sup_{\eta} v_0(\eta)$$

solves this equation on $[0, T]$. On the other hand, since $v(0, \xi) \leq \psi(0, \xi)$ and $f(x + w(t, \xi, \omega), \xi) \leq f(x + w(t, \xi, \omega), \xi) \vee 0$, the comparison theorem for reaction–diffusion PDEs [36, Theorem 10.1] implies that for every interval $[0, \delta]$ where Eq. (5.28) has a local solution we have

$$v(t, \xi) \leq (x_+ + M_T) \vee \sup_{\eta} v_0(\eta), \quad t \in [0, \delta], \xi \in \mathbb{T}^d.$$

By a similar argument,

$$v(t, \xi) \geq (x_- - M_T) \wedge \inf_{\eta} v_0(\eta), \quad t \in [0, \delta], \xi \in \mathbb{T}^d.$$

Thus, for any interval $[0, \delta]$ where Eq. (5.28) has a local solution, this solution is uniformly bounded by a constant which does not depend on δ . Therefore, by [27, Proposition 7.3.1.v] Eq. (5.28) has a global solution on $[0, T]$. By [27, Proposition 7.1.10.iii] this solution is in $\mathcal{C}([0, T], \mathcal{C}^2(\mathbb{T}^d))$.

Finally, to obtain the desired bound on $\|v(t)\|_{L_\infty(\mathbb{T}^d)}$ in terms of $\|v_0\|_{L_2(\mathbb{T}^d)}$ rather than $\|v_0\|_{L_\infty(\mathbb{T}^d)}$ we fix $t \in [0, T]$ and consider again PDE (5.32) with the initial condition $\psi(0, \xi) := (v_0(\xi) \vee 0) + M_t$. By the comparison principle, we have

$$v(t, \xi) \leq \psi(t, \xi), \quad \xi \in \mathbb{T}^d. \tag{5.33}$$

Since $\psi(0, \xi) \geq M_t$ for all $\xi \in \mathbb{T}^d$ and the drift is nonnegative, it is immediate to see that $\psi(s, \xi) \geq M_t$ for all $s \in [0, t]$, $\xi \in \mathbb{T}^d$. Therefore, taking into account (5.27), we get that for all $s \in [0, t]$, $\xi \in \mathbb{T}^d$

$$f(\psi(s, \xi) + w(s, \xi, \omega), \xi) \vee 0 \leq 0 \vee \sup_{\substack{x \geq 0 \\ \eta \in \mathbb{T}^d}} f(x, \eta) =: C_+ < \infty. \tag{5.34}$$

Introduce now ψ_+ , which is the strong solution of the following PDE

$$d\psi_+(s, \xi) = \Delta\psi_+(s, \xi)ds + C_+ds, \quad \xi \in \mathbb{T}^d, s \in [0, t],$$

with the initial condition $\psi_+(0) := \psi(0)$. Using (5.34), we see that the comparison principle implies that

$$\psi(t, \xi) \leq \psi_+(t, \xi), \quad \xi \in \mathbb{T}^d. \tag{5.35}$$

Let p_s be the heat kernel on the torus \mathbb{T}^d . Then it is straightforward to see that for any $\xi \in \mathbb{T}^d$

$$\begin{aligned} |\psi_+(t, \xi)| &= |p_t * \psi_0(\xi) + C_+t| \leq C_+t + \|\psi_0\|_{L_2(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} p_t(\xi)^2 d\xi \right)^{1/2} \\ &\leq C_+t + C\|\psi_0\|_{L_2(\mathbb{T}^d)}(1 + t^{-d/4}). \end{aligned}$$

Combining this with (5.33) and (5.35), we deduce

$$\begin{aligned} v(t, \xi) &\leq C_+t + C\|\psi_0\|_{L_2(\mathbb{T}^d)}(1 + t^{-d/4}) \\ &\leq t \left(\sup_{\substack{x \geq 0 \\ \eta \in \mathbb{T}^d}} f(x, \eta) \vee 0 \right) + C(\|v_0\|_{L_2(\mathbb{T}^d)} + M_t)(1 + t^{-d/4}). \end{aligned}$$

By a similar argument, we get

$$v(t, \xi) \geq t \left(\inf_{\substack{x \leq 0 \\ \eta \in \mathbb{T}^d}} f(x, \eta) \wedge 0 \right) - C(\|v_0\|_{L_2(\mathbb{T}^d)} + M_t)(1 + t^{-d/4}).$$

This yields (5.29). □

Now let us move on to less regular initial data.

Lemma 5.6. *For any $v_0 \in L_\infty(\mathbb{T}^d)$ Eq. (5.28) has a unique analytically generalized strong solution v with initial condition v_0 . Furthermore, we have $v \in C((0, T], C^2(\mathbb{T}^d))$.*

Proof. We begin with the uniqueness part. Suppose that v and \bar{v} are two analytically generalized strong solutions to Eq. (1.1) with the same initial condition $v_0 \in L_\infty(\mathbb{T}^d)$. Fix arbitrary $t > 0$. Arguing as in the proof of Lemma 5.5, we see that by the Gronwall lemma there exists $C = C(t) > 0$ such that for any $\delta \in (0, t)$ we have

$$\|v(t) - \bar{v}(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C \|v(\delta) - \bar{v}(\delta)\|_{L_2(\mathbb{T}^d)}^2. \tag{5.36}$$

By definition, $v(\delta) \rightarrow v(0)$ and $\bar{v}(\delta) \rightarrow \bar{v}(0)$ in $L_2(\mathbb{T}^d)$ as $\delta \rightarrow 0$. Since $v(0) = \bar{v}(0)$, by passing to the limit as $\delta \rightarrow 0$ in (5.36), we deduce that $v(t) = \bar{v}(t)$.

Note again that the function defined in (5.31) is locally Lipschitz in x , Hölder in t and continuous in ξ . Therefore, by [27, Proposition 7.3.1.i] there exists $\delta > 0$ such that on the interval $[0, \delta]$ Eq. (5.28) has an analytically generalized strong solution v with initial condition v_0 and $v \in \mathcal{C}((0, \delta], \mathcal{C}^2(\mathbb{T}^d))$. Since $v(\delta) \in \mathcal{C}^2(\mathbb{T}^d)$, by Lemma 5.5 we can construct an analytically strong solution v on $[\delta, T]$ with the initial condition $v(\delta)$ and $v \in \mathcal{C}([\delta, T], \mathcal{C}^2(\mathbb{T}^d))$. By gluing these two solutions together we get an analytically generalized strong solution $v \in \mathcal{C}((0, T], \mathcal{C}^2(\mathbb{T}^d))$. \square

To consider even less regular initial data (recall that we are interested in the initial conditions from $L_2(\mathbb{T}^d)$) we need the following lemma about approximations of solutions to (5.28).

Lemma 5.7. *Let $(v_0^n)_{n \in \mathbb{Z}_+}$ be a sequence of $\mathcal{C}^2(\mathbb{T}^d)$ functions converging in L_2 to $v_0 \in L_2(\mathbb{T}^d)$. Let v^n be the analytically strong solution of (5.28) with the initial condition v_0^n . Then*

- (i) v^n converges in $\mathcal{C}([0, T], L_2(\mathbb{T}^d))$ to some $v \in \mathcal{C}([0, T], L_2(\mathbb{T}^d))$;
- (ii) for each $t \in (0, T]$ we have $v(t) \in L_\infty(\mathbb{T}^d)$;
- (iii) if \bar{v} is an analytically generalized strong solution of (5.28) with the initial condition v_0 , then for each $t \in [0, T]$ we have $\bar{v}(t, \xi) = v(t, \xi)$ for almost all $\xi \in \mathbb{T}^d$.

Proof. (i) Let $n, m \in \mathbb{Z}_+$. Then arguing as in the derivation of (5.30) we get by the chain rule and assumption (4.4)

$$\frac{d}{dt} \|v^n(t) - v^m(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C \|v^n(t) - v^m(t)\|_{L_2(\mathbb{T}^d)}^2.$$

By the Gronwall lemma, this implies that

$$\sup_{t \in [0, T]} \|v^n(t) - v^m(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C(T) \|v_0^n - v_0^m\|_{L_2(\mathbb{T}^d)}^2.$$

Now the statement of the lemma follows immediately from the completeness of the space $\mathcal{C}([0, T], L_2(\mathbb{T}^d))$ ([21, Theorem I.4.19]).

(ii) Fix $t \in (0, T]$. By Lemma 5.5, we have the following uniform over $n \geq N$ bound, where N is large enough:

$$\|v^n(t)\|_{L_\infty(\mathbb{T}^d)} \leq C(t) (\|v_0^n\|_{L_2(\mathbb{T}^d)} + M_t) + C_f t \leq 2C(t) (\|v_0\|_{L_2(\mathbb{T}^d)} + M_t) + C_f t. \tag{5.37}$$

Since $v^n(t)$ converges to $v(t)$ in L_2 , bound (5.37) implies $v(t) \in L_\infty(\mathbb{T}^d)$.

(iii) Let \bar{v} be an analytically generalized strong solution of (5.28) with the initial condition v_0 . Then arguing as in part (i) of the lemma, we get for any $\delta > 0$

$$\sup_{t \in [\delta, T]} \|\bar{v}(t) - v^n(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C(T) \|\bar{v}(\delta) - v^n(\delta)\|_{L_2(\mathbb{T}^d)}^2.$$

By passing to the limit as $\delta \rightarrow 0$ and using the fact that by definition $\bar{v}(\delta) \rightarrow v_0$ in L_2 , we get

$$\sup_{t \in [0, T]} \|\bar{v}(t) - v^n(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C(T) \|v_0 - v_0^n\|_{L_2(\mathbb{T}^d)}^2.$$

However, by part (i) of the lemma

$$\sup_{t \in [0, T]} \|v(t) - v^n(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C(T) \|v_0 - v_0^n\|_{L_2(\mathbb{T}^d)}^2.$$

This implies that $\bar{v}(t) = v(t)$ as elements of $L_2(\mathbb{T}^d)$ for any $t \in [0, T]$. □

The next lemma establishes existence of solutions to Eq. (5.28) with $L_2(\mathbb{T}^d)$ initial data. It relies on Lemma 5.7 and extends [27, Propositions 7.3.1].

Lemma 5.8. *For any $v_0 \in L_2(\mathbb{T}^d)$ Eq. (5.28) has a unique analytically generalized strong solution v with the initial condition v_0 . Furthermore, we have $v \in \mathcal{C}((0, T], \mathcal{C}^2(\mathbb{T}^d))$.*

Proof. The proof of the uniqueness part is the same as in Lemma 5.6.

Let us show existence of a solution to (5.28). Let $(v_0^n)_{n \in \mathbb{Z}_+}$ be a sequence of $\mathcal{C}^2(\mathbb{T}^d)$ functions converging in L_2 to $v_0 \in L_2(\mathbb{T}^d)$. Let v^n be the analytically strong solution of (5.28) with the initial condition v_0^n (it exists thanks to Lemma 5.5). By Lemma 5.7(i,ii) there exists $v \in \mathcal{C}([0, T], L_2(\mathbb{T}^d))$ such that

$$v(t) \in L_\infty(\mathbb{T}^d), \quad t \in (0, T] \quad \text{and} \quad \sup_{t \in [0, T]} \|v^n(t) - v(t)\|_{L_2} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (5.38)$$

We claim now that v is an analytically generalized strong solution to (5.28) with the initial condition v_0 .

Indeed, we have by construction $v(0) = v_0$ and $v \in \mathcal{C}([0, T], L_2(\mathbb{T}^d))$ as a limit of the functions from the space $\mathcal{C}([0, T], L_2(\mathbb{T}^d))$. Fix any $\varepsilon \in (0, T]$ and let us verify that identity (4.2) holds.

By (5.38), we have $v(\varepsilon/2) \in L_\infty$. Therefore, by Lemma 5.8 there exists a process $\bar{v} \in \mathcal{C}((\varepsilon/2, T], \mathcal{C}^2(\mathbb{T}^d))$, which is an analytically generalized strong solution of (5.28) on interval $[\varepsilon/2, T]$ with the initial condition $v(\varepsilon/2)$. Therefore, by Lemma 5.7(iii) we have $\bar{v}(t) = v(t)$ for any $t > \varepsilon/2$. Thus, identity (4.2) holds and for any $t \geq \varepsilon$ the function $v(t)$ has a version which is in $\mathcal{C}^2(\mathbb{T}^d)$. Since ε was arbitrary, this implies the statement of the lemma. □

Now we are ready to present the proof of Theorem 4.4.

Proof of Theorem 4.4. (i) The proof of the uniqueness part is the same as in Lemma 5.5. To show existence of a strong solution to (1.1) we fix an initial condition $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$, $T > 0$ and first show existence on the time interval $[0, T]$. Let v be an analytically strong solution to (5.28) with the initial condition $v(0) = u_0$. Recall the definition of Ω' from Lemma 5.4 and put now

$$u(t, \xi, \omega) := v(t, \xi, \omega) + w(t, \xi, \omega), \quad t \in [0, T], \xi \in \mathbb{T}^d, \omega \in \Omega'$$

and $u(t, \xi, \omega) = 0$ for $\omega \in \Omega \setminus \Omega'$. It follows from Lemmas 5.4 and 5.5 that the function u is an analytically strong solution to (1.1) with the initial condition u_0 . To show the

adaptiveness of u , we introduce a function u^n , $n \in \mathbb{Z}_+$, which is an analytically and probabilistically strong solution on $[0, T]$ to the following equation

$$u^n(t, \xi) = u_0(\xi) + \int_0^t [\Delta u^n(s, \xi) + f^n(u^n(s, \xi), \xi)] dt + \sum_{k=1}^m \sigma_k(\xi) W^k(t), \quad \xi \in \mathbb{T}^d, t \geq 0, \tag{5.39}$$

where for $x \in \mathbb{R}$, $\xi \in \mathbb{T}^d$ we put $f^n(x, \xi) := f((x \wedge n) \vee (-n), \xi)$. Since f^n is uniformly bounded, it follows from [22, Chapter II, Theorem 2.1] that u^n is well-defined; thus, identity (5.39) holds on some set Ω_n of full measure. On the other hand, by uniqueness, we have

$$\begin{aligned} A_n &:= \bigcap_{k=1}^\infty \Omega_k \cap \Omega' \cap \left\{ \sup_{t \in [0, T], \xi \in \mathbb{T}^d} |u(t, \xi)| \leq n \right\} \\ &= \bigcap_{k=1}^\infty \Omega_k \cap \Omega' \cap \left\{ \sup_{t \in [0, T], \xi \in \mathbb{T}^d} |u^n(t, \xi)| \leq n \right\} \end{aligned}$$

and $u^n = u$ on A_n . Since for each $\omega \in \Omega'$ the function u is bounded, we see that $\mathbb{P}(A_n) \rightarrow 1$ as $n \rightarrow \infty$ and $A_n \subset A_{n+1}$. This implies that u^n converges to u a.s. as $n \rightarrow \infty$. Since each u^n is (\mathcal{F}_t) -adapted, their limit u is also (\mathcal{F}_t) -adapted. Therefore u is a probabilistically strong solution of (1.1) on $[0, T]$. Since T was arbitrary, this and uniqueness imply that u is an analytically and probabilistically strong solution of (1.1) on $[0, \infty)$. The continuity of u follows from continuity of v and w (Lemmas 5.4 and 5.5).

(ii) The proof of the uniqueness part is the same as in Lemma 5.6. To show existence fix $T > 0$ and put again now

$$u(t, \xi, \omega) := v(t, \xi, \omega) + w(t, \xi, \omega), \quad t \in [0, T], \xi \in \mathbb{T}^d, \omega \in \Omega'$$

and $u(t, \xi, \omega) = 0$ for $\omega \in \Omega \setminus \Omega'$. Here v is an analytically generalized strong solution to (5.28) with the initial condition $v(0) = u_0$. It follows from Lemmas 5.4 and 5.6 that the function u is an analytically generalized strong solution to (1.1) with the initial condition u_0 . To show the adaptiveness we consider $(u_0^n)_{n \in \mathbb{Z}_+}$, a sequence of $C^2(\mathbb{T}^d)$ functions converging in L_2 to $u_0 \in L_2(\mathbb{T}^d)$. Let u^n be the analytically and probabilistically strong solution of (1.1) with the initial condition u_0^n (it exists thanks to part (i) of the theorem). By Lemma 5.7(iii), we have that for each $t > 0$ the function $u^n(t)$ converges to $u(t)$ in $L_2(\mathbb{T}^d)$ \mathbb{P} -a.s. Since $u^n(t)$ is \mathcal{F}_t -adapted, we see that $u(t)$ is also (\mathcal{F}_t) -adapted. Therefore u is a probabilistically strong solution of (1.1) on $[0, T]$. The proof ends in the same way as in the part (i) of the theorem.

(iii) We begin with the following two observations. First, we note that part (ii) of the theorem implies that the generalized strong solution of equation (1.1) is unique.

Second, for $x \in L_2(\mathbb{T}^d)$ let us denote by u^x the generalized strong solution of (1.1) with the initial condition x . Then, by the chain rule and Gronwall's inequality we get that for each $t > 0$ there exists $C(t) > 0$ such that for any $x, y \in L_2(\mathbb{T}^d)$ we have

$$\|u^x(t) - u^y(t)\|_{L_2(\mathbb{T}^d)}^2 \leq C(t) \|x - y\|_{L_2(\mathbb{T}^d)}^2,$$

where we also took into account assumption (4.4).

Now the Markov property of u follows from these two properties by repeating literally the argument from [3, Proposition 4.1].

(iv) Take the set Ω' defined in Lemma 5.4. Fix $t > 0, \omega \in \Omega', x, y \in L_2(\mathbb{T}^d)$ such that $x \preceq_{pos} y$. Let u^x, u^y be the generalized strong solutions of (1.1) with the initial conditions x and y respectively. Let us show now that $u^x(t, \omega) \preceq_{pos} u^y(t, \omega)$ a.s.

We begin with the case when $x, y \in \mathcal{C}^2(\mathbb{T}^d)$. Let w be the stochastic convolution (recall its definition in Lemma 5.4). Let v^x, v^y be analytically strong solutions of (5.28) with the initial conditions x and y respectively. Then by Lemma 5.5 v^x and v^y are continuous on $[0, t] \times \mathbb{T}^d$. Therefore, by comparison principle [36, Theorem 10.1] we have $v^x(t, \omega) \preceq_{pos} v^y(t, \omega)$ and hence $u^x(t, \omega) = v^x(t, \omega) + w(t, \omega) \preceq_{pos} v^y(t, \omega) + w(t, \omega) = u^y(t, \omega)$.

In the general case, we consider $(x^n)_{n \in \mathbb{Z}_+}$ and $(y^n)_{n \in \mathbb{Z}_+}$, two sequences of $\mathcal{C}^2(\mathbb{T}^d)$ functions converging in L_2 to x and y , respectively. By the above, for each $n \in \mathbb{Z}_+$ we have

$$u^{x^n}(t, \omega) \preceq_{pos} u^{y^n}(t, \omega). \tag{5.40}$$

On the other hand, by part (ii) of the theorem and Lemma 5.7 we have $u^{x^n}(t, \omega) \rightarrow u^x(t, \omega), u^{y^n}(t, \omega) \rightarrow u^y(t, \omega)$ in $L_2(\mathbb{T}^d)$ as $n \rightarrow \infty$. This together with (5.40) yields $u^x(t) \preceq_{pos} u^y(t)$.

(v) Follows immediately from part (iv) of the theorem and the Strassen’s theorem.

(vi) Fix $x \in L_2(\mathbb{T}^d)$. Then u^x is an analytically generalized strong solution of (1.1). Therefore by Ito’s lemma, taking into account (4.3), we deduce for any $0 < s \leq t$

$$\begin{aligned} \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2 &\leq \|u^x(s)\|_{L_2(\mathbb{T}^d)}^2 - 2K_2 \int_s^t \|u^x(r)\|_{L_2(\mathbb{T}^d)}^2 dr + (2K_1 + \|\sigma\|_{L_2(\mathbb{T}^d)}^2)(t - s) \\ &\quad + M(t) - M(s), \end{aligned} \tag{5.41}$$

where the constants K_1, K_2 are defined in (4.3) and M is a continuous local martingale with $M(0) = 0$ and

$$d\langle M \rangle_t = 4 \sum_{k=1}^m \int_{\mathbb{T}^d} (u^x(t, \xi))^2 \sigma_k(\xi)^2 d\xi dt \leq 4 \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2 \sum_{k=1}^m \|\sigma_k\|_{L_\infty(\mathbb{T}^d)}^2 dt.$$

Using the fact that u^x is continuous in $L_2(\mathbb{T}^d)$, we can pass to the limit in (5.41) as $s \rightarrow 0$ to deduce that (5.41) is also valid for $s = 0$.

Therefore all the conditions of Lemma 5.3 are satisfied for the process $X(t) := \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2$. Thus, $\mathbb{E} \|u^x(t)\|_{L_2(\mathbb{T}^d)}^2 < \infty$ for any $t > 0$, inequality (4.6) follows from (5.21) and inequality (4.7) follows from (5.22). \square

Before we formulate and prove the final lemma we recall that under Assumption A the solution u with initial condition $u_0 \in L_2(\mathbb{T}^d)$ satisfies the following mild form of the SPDE (1.1) [7, Theorem 5.4]

$$\begin{aligned} u(t, \xi) &= \int_{\mathbb{T}^d} p_t(\xi - \zeta) u(\varepsilon, \zeta) d\zeta + \sum_{k=1}^m \int_\varepsilon^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) \sigma_k(\zeta) d\zeta dW^k(s) \\ &\quad + \int_\varepsilon^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) f(u(s, \zeta), \zeta) d\zeta ds, \end{aligned}$$

for each $t > 0$, $\varepsilon \in (0, t]$ and $\xi \in \mathbb{T}^d$, where p denotes the fundamental solution of the heat equation on \mathbb{T}^d . Letting $\varepsilon \downarrow 0$, we see that almost surely

$$\begin{aligned}
 u(t, \xi) &= \int_{\mathbb{T}^d} p_t(\xi - \zeta) u_0(\zeta) d\zeta + \sum_{k=1}^m \int_0^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) \sigma_k(\zeta) d\zeta dW^k(s) \\
 &+ \lim_{\varepsilon \downarrow 0} \int_\varepsilon^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) f(u(s, \zeta), \zeta) d\zeta ds,
 \end{aligned}
 \tag{5.42}$$

where we used the Cauchy-Schwarz inequality to obtain convergence of the first term.

For an arbitrary element $x \in L_2(\mathbb{T}^d)$ let S_x and L_x be the sets of elements from $L_2(\mathbb{T}^d)$ that are smaller (respectively larger) than x , that is

$$S_x := \{f \in L_2(\mathbb{T}^d) : f \preceq x\}, \quad L_x := \{f \in L_2(\mathbb{T}^d) : x \preceq f\}.$$

It is easy to see that the sets S_x and L_x are closed. For $a \in \mathbb{R}$ put

$$\check{a}(\xi) := a, \quad \xi \in \mathbb{T}^d.$$

Lemma 5.9. *Suppose that Assumptions A and B hold. Let (P_t) be the semigroup associated with Eq. (1.1). Then for any $M > 0$ there exists $\varepsilon > 0$ such that*

$$\inf_{\substack{x \in L_2(\mathbb{T}^d) \\ \|x\|_{L_2} \leq M}} P_1(x, S_\check{0}) > \varepsilon, \quad \inf_{\substack{x \in L_2(\mathbb{T}^d) \\ \|x\|_{L_2} \leq M}} P_1(x, L_\check{0}) > \varepsilon.
 \tag{5.43}$$

Proof. By symmetry it suffices to prove the first claim. We show it in two steps.

Step 1. At this step we will use only Assumption A. We claim that for every $M > 0$ and $T \in (0, 1]$ there exist $\Gamma \in \mathbb{R}$ and $t_0 \in (0, T]$ such that

$$\inf_{\substack{x \in L_2(\mathbb{T}^d) \\ \|x\|_{L_2} \leq M}} P_{t_0}(x, S_\check{\Gamma}) > 0.$$

First we note, that it is sufficient to prove the claim only for large enough M . Observe now that by (4.3) there exists $\gamma \in \mathbb{R}$ such that $f(z, \xi) \leq 0$ for all $z \geq \gamma$ and all $\xi \in \mathbb{T}^d$. Fix arbitrary $\beta > \gamma$, and assume that M is large enough so that $\|\check{\beta}\|_{L_2} \leq M$. By order preservation (Theorem 4.4 (iv)) it suffices to show that for every $M > 0$, $T \in (0, 1]$ there exist $\Gamma \in \mathbb{R}$ and $t_0 \in (0, T]$ such that

$$\inf_{\substack{x \in L_2(\mathbb{T}^d), \check{\beta} \preceq_{pos} x \\ \|x\|_{L_2} \leq M}} P_{t_0}(x, S_\check{\Gamma}) > 0.$$

By continuity of solutions (Theorem 4.4) there exists $t_0 \in (0, T]$ such that

$$\kappa := \mathbf{P}(\check{\gamma} \preceq_{pos} u^{\check{\beta}}(s) \text{ for all } s \in [0, t_0]) > 0.$$

For $x \in L_2(\mathbb{T}^d)$ we define

$$\Omega(x) := \{\omega : \check{\gamma} \preceq_{pos} u^x(s) \text{ for all } s \in [0, t_0]\}.$$

Fix now arbitrary $x \in L_2(\mathbb{T}^d)$ satisfying $\check{\beta} \preceq_{pos} x$ and $\|x\|_{L_2} \leq M$. Then, using order preservation again, we see that

$$\mathbf{P}(\Omega(x)) \geq \kappa.$$

Choose $\tilde{\Gamma}$ so large that

$$\begin{aligned} P(\tilde{\Omega}) &:= P\left(\sum_{k=1}^m \int_0^r \int_{\mathbb{T}^d} p_{r-s}(\xi - \zeta) \sigma_k(\zeta) d\zeta dW^k(s) \leq \tilde{\Gamma} \text{ for all } 0 \leq r \leq t_0, \xi \in \mathbb{T}^d\right) \\ &\geq 1 - \frac{\kappa}{2}. \end{aligned}$$

Using (5.42), we see that on the set $\tilde{\Omega} \cap \Omega(x)$ (which has measure at least $\kappa/2$) the solution u^x satisfies for all $\xi \in \mathbb{T}^d$

$$u^x(t_0, \xi) \leq \left(\int_{\mathbb{T}^d} p_{t_0}^2(\xi - \zeta) d\zeta\right)^{1/2} \|x\|_{L_2} + \tilde{\Gamma} \leq \left(\int_{\mathbb{T}^d} p_{t_0}^2(\xi - \zeta) d\zeta\right)^{1/2} M + \tilde{\Gamma} =: \Gamma,$$

where we used the Cauchy–Schwarz inequality in the first inequality and the fact that $f(z, \xi) \leq 0$ for $z \geq \gamma$ and $\xi \in \mathbb{T}^d$. The proof of the claim of Step 1 is complete.

Step 2. This is the only part of the proof of Theorem 4.5, where Assumption B is used. We claim that for every $\Gamma \in \mathbb{R}$ and $\tau > 0$ we have

$$P_\tau(\check{\Gamma}, S_\check{\gamma}) > 0.$$

To show this claim, first of all, we assume, without loss of generality, that $\Gamma > 0$. It suffices to show that for some $\tau_0 = \tau_0(\Gamma) > 0$ the statement holds for every $\tau \in (0, \tau_0]$. Fix numbers $\lambda_k, k = 1, \dots, m$ as in Assumption B with $\varepsilon = 1$ and define $\Lambda(\xi) := \sum_{k=1}^m \lambda_k \sigma_k(\xi) \geq 1$ and $\hat{\Lambda} := \max_{\xi \in \mathbb{T}^d} \Lambda(\xi)$. Choose $Q > \hat{\Lambda}(\Gamma + 2) + 2$ and fix $\tau_0 \in (0, 1]$ such that $\tau_0 |f(z, \xi)| \leq 1$ for all $\xi \in \mathbb{T}^d$ and $z \in [\Gamma - Q, \Gamma + 2]$ (such τ_0 exists since f is continuous). Fix $\tau \in (0, \tau_0]$ and $\Theta := \frac{\Gamma+2}{\tau}$. Then, for

$$\tilde{W}^k(s) := \Theta \lambda_k s + W^k(s), \quad s \geq 0, \quad k \in \{1, \dots, m\}$$

and $t > 0$, the solution $u^{\check{\Gamma}}$ satisfies

$$\begin{aligned} u^{\check{\Gamma}}(t, \xi) &= \Gamma - \int_0^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) \Lambda(\zeta) \Theta d\zeta ds \\ &\quad + \sum_{k=1}^m \int_0^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) \sigma_k(\zeta) d\zeta d\tilde{W}^k(s) \\ &\quad + \int_0^t \int_{\mathbb{T}^d} p_{t-s}(\xi - \zeta) f(u^{\check{\Gamma}}(s, \zeta), \zeta) d\zeta ds \\ &=: \Gamma - I_1(t, \xi) + I_2(t, \xi, \omega) + I_3(t, \xi, \omega). \end{aligned}$$

Note that for each $t \geq 0$,

$$I_1(t, \xi) \in [\Theta t, \Theta \hat{\Lambda} t].$$

Let

$$\hat{\Omega} := \{\omega : |I_2(s, \xi, \omega)| \leq 1 \text{ for all } s \in [0, 1], \xi \in \mathbb{T}^d\}.$$

Then $P(\hat{\Omega}) > 0$. Let $\psi := \inf\{s \geq 0 : u^{\check{\Gamma}}(s) \notin [\Gamma - Q, \Gamma + 2]\}$. If $\psi < \tau$ and $u(\psi, \xi) = \Gamma + 2$ for some $\xi \in \mathbb{T}^d$, then, on the set $\hat{\Omega} \cap \{\psi < \tau\}$ we have

$$\Gamma + 2 \leq \Gamma - \Theta \psi + 1 + 1,$$

so $\psi = 0$ which is impossible (since u is continuous). If, on the other hand, $\psi < \tau$ and $u(\psi, \xi) = \Gamma - Q$ for some $\xi \in \mathbb{T}^d$, then, on the set $\tilde{\Omega} \cap \{\psi < \tau\}$, we have

$$\Gamma - Q \geq \Gamma - \Theta \hat{\Lambda} \psi - 2,$$

which is also impossible by the definition of Q and Θ .

Thus, $\psi \geq \tau$ almost surely on $\hat{\Omega}$. Therefore, on $\hat{\Omega}$, we have

$$u^{\tilde{\Gamma}}(\tau, \xi) \leq \Gamma - \Theta \tau + 2 \leq 0$$

for every $\xi \in \mathbb{T}^d$, so the proof of Step 2 is complete.

Using order preservation once more, we see that Step 1 and Step 2 imply the assertion in the lemma. \square

Finally, before we present the proof of Theorem 4.8, we need to introduce some additional notation. If for some $a \in \mathbb{R}$, $x \in L_2(\mathbb{T}^d)$ we have $x(\xi) = a$ for all $\xi \in \mathbb{T}^d$, then we will write

$$\Pi x = a; \quad \Pi^{-1} a = x.$$

Introduce also the set of all constants in the space $L_2(\mathbb{T}^d)$, that is,

$$A_c := \{\Pi^{-1} a \mid a \in \mathbb{R}\} \subset L_2(\mathbb{T}^d).$$

Let $(P_t)_{t \geq 0}$ be the transition function associated with the solution to (1.1).

Lemma 5.10. *Consider Eq. (1.1) with $m = 1$ and $\sigma_1 \equiv 1$. Suppose that the drift f does not depend on space. This equation has the following properties:*

- (i) if $u(0) \in A_c$, then $u(t) \in A_c$ a.s. for any $t \geq 0$;
- (ii) if $u(0) \notin A_c$, then $u(t) \notin A_c$ a.s. for any $t \geq 0$;
- (iii) this equation has a unique invariant measure π and $\pi(A_c) = 1$.

Proof. (i) Suppose that $u(0) \in A_c$. Let $a = \Pi[u(0)]$. Consider a stochastic differential equation

$$X_t^a = a + \int_0^t f(X_s^a) ds + W_t. \tag{5.44}$$

Since the function f is locally Lipschitz and satisfies the one-sided growth condition, Eq. (5.44) has a unique strong solution [34, Proposition 2.1(b)]. It is immediate to see that the function $u(t, \xi) := X^a(t)$ is an analytically and probabilistically strong solution of (1.1) with the initial condition $u(0)$. By Theorem 4.4 this is the unique solution to (1.1). Thus $u(t) \in A_c$ for any $t \geq 0$.

(ii) Fix the initial condition $x \notin A_c$. First, consider the case $x \in \mathcal{C}^2(\mathbb{T}^d)$. Let a_1, a_2 be real numbers such that $a_1 \leq x(\xi) \leq a_2$ for all $\xi \in \mathbb{T}^d$. Let $\Omega'' \subset \Omega$ be the set of full measure where the trajectories of the Brownian motion W are continuous and identity (4.2) holds for $\varepsilon = 0$ for the initial conditions $\Pi^{-1} a_1, \Pi^{-1} a_2$, and x . Set $\Omega_0 := \Omega'' \cap \Omega'$, where Ω' is defined in Theorem 4.4(iv). It follows that $\mathbb{P}(\Omega_0) = 1$.

Assume now the contrary. That is suppose that for some $T > 0$, $\omega \in \Omega_0$ we have $u^x(0) = x$ and $u^x(T, \omega) \in A_c$. Denote $b := \Pi[u^x(T, \omega)]$. By Theorem 4.4(iv) and part (i) of the theorem, we have for all $\xi \in \mathbb{T}^d$

$$X_T^{a_1}(\omega) = u^{\Pi^{-1} a_1}(T, \xi, \omega) \leq u^x(T, \xi, \omega) \leq u^{\Pi^{-1} a_2}(T, \xi, \omega) = X_T^{a_2}(\omega).$$

It follows immediately from the Gronwall lemma, the fact that f is locally Lipschitz, and the comparison principle that solutions to (5.44) are continuous with respect to the initial condition. Therefore there exists some initial condition $a \in (a_1, a_2)$ such that $X^a(0, \omega) = a$ and $X^a(T, \omega) = b$. Denote now $v(t, \xi) := u(T - t, \xi, \omega) - X^a(T - t, \omega)$, $t \in [0, T]$, $\xi \in \mathbb{T}^d$. Then we have $v(0, \xi) = 0$ for any $\xi \in \mathbb{T}^d$. On the other hand, for any $t \in [0, T]$, $\xi \in \mathbb{T}^d$ we have

$$\begin{aligned} & |\Delta v(t, \xi) + \partial_t v(t, \xi)| \\ &= |\Delta u(T - t, \xi, \omega) - f(X^a(T - t, \omega)) - \Delta u(T - t, \xi, \omega) - f(u(T - t, \xi, \omega))| \\ &\leq C|X^a(T - t, \omega) - (u(T - t, \xi, \omega))| \\ &= C|v(t, \xi)|, \end{aligned} \tag{5.45}$$

where we also used that the function f is locally Lipschitz and the processes $u(\omega)$ and $X^a(\omega)$ are bounded. However, inequality (5.45) together with the fact that $v(0) = 0$ implies that $v(t) = 0$ for all $t \in [0, T]$ [37, Theorem 3.0.4]. This yields that

$$0 = v(T, \xi) = u(0, \xi) - a, \quad \xi \in \mathbb{T}^d.$$

However it was assumed that $u(0) \notin A_c$. This contradiction finishes the proof for the case $x \in C^2(\mathbb{T}^d)$.

In the general case, $x \in L_2(\mathbb{T}^d)$, we will use the Markov property of the solutions to (1.1), that is, Theorem 4.4(iii). We have for any $0 < s < t$

$$\begin{aligned} P_t(x, A_c) &= \int_{L_2(\mathbb{T}^d)} P_s(x, dy) P_{t-s}(y, A_c) = \int_{C^2(\mathbb{T}^d)} P_s(x, dy) P_{t-s}(y, A_c) \\ &= \mathbf{P}(u^x(s) \in A_c), \end{aligned}$$

where the first identity is the Markov property, the second identity follows from the fact that $u^x(s) \in C^2(\mathbb{T}^d)$ a.s. by Lemma 5.8, and the third identity follows from the fact that $P_{t-s}(y, A_c) = 0$ if $y \in C^2(\mathbb{T}^d) \setminus A_c$ established above. Since s was arbitrary, we can pass to the limit as $s \rightarrow 0$. Using the fact that the process $u^x(\omega)$ is by definition continuous in $L_2(\mathbb{T}^d)$ and $x \notin A_c$, we get

$$\mathbf{P}(u^x(s) \in A_c) \rightarrow 0, \quad \text{as } s \rightarrow 0,$$

which completes the proof.

(iii) It follows from Theorem 4.5 that SPDE (1.1) has a unique invariant measure π and that for any $x \in L_2(\mathbb{T}^d)$ one has $P_t(x, \cdot) \rightarrow \pi$ weakly as $t \rightarrow \infty$. Fix $x_0 \in A_c$. Then, by part (i) of the proof and the Portmanteau theorem (see, e.g., [35, Theorem III.1.1(II)]), one has

$$1 = \limsup_{t \rightarrow \infty} P_t(x_0, A_c) \leq \pi(A_c),$$

where we also used the fact that the set A_c is closed in $L_2(\mathbb{T}^d)$. This proves the statement of the lemma. □

Proof of Theorem 4.8. Take $m = 1$ and $\sigma_1 \equiv 1$. Fix any $x \in L_2(\mathbb{T}^d) \setminus A_c$. Then for any $t > 0$ by Lemma 5.10(ii) one has $P_t(x, A_c) = 0$. On the other hand, by Lemma 5.10(iii), $\pi(A_c) = 1$. This implies the statement of the theorem. □

Acknowledgements. The authors are deeply indebted to Benjamin Gess and Konstantinos Dareiotis for their help, patience and detailed explanations of some parts of the theory of parabolic PDEs. We also would like to thank Máté Gerencsér, Alessandra Lunardi, Lenya Ryzhik, Vladimír Šverák, and Pavlos Tsatsoulis for useful comments and helpful discussions. We are also grateful to two anonymous reviewers for providing helpful comments on an earlier draft of the manuscript. Part of the work on the project has been done during the visit of OB to the Institute of Science and Technology—Austria (IST) and Max-Planck-Institut für Mathematik in den Naturwissenschaften (Leipzig, Germany). OB is very grateful to these institutions for providing excellent working conditions. OB has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No. 683164).

Funding Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Bahouri, H., Chemin, J.-Y., Danchin, R.: *Fourier Analysis and Nonlinear Partial Differential Equations*. Springer, Heidelberg (2011)
2. Butkovsky, O., Kulik, A., Scheutzow, M.: Generalized couplings and ergodic rates for SPDEs and other Markov models. *Ann. Appl. Probab.* **30**, 1–39 (2020)
3. Butkovsky, O., Scheutzow, M.: Invariant measures for stochastic functional differential equations. *Electr. J. Probab.* **22**, 1–23 (2017)
4. Chueshov, I., Scheutzow, M.: On the structure of attractors and invariant measures for a class of monotone random systems. *Dyn. Syst.* **19**, 127–144 (2004)
5. Constantin, P., Glatt-Holtz, N.E., Vicol, V.: Unique ergodicity for fractionally dissipated, stochastically forced 2D Euler equations. *Commun. Math. Phys.* **330**, 819–857 (2014)
6. Da Prato, G., Zabczyk, J.: *Ergodicity for Infinite Dimensional Systems*. Cambridge University Press, Cambridge (1996)
7. Da Prato, G., Zabczyk, J.: *Stochastic Equations in Infinite Dimensions*, 2nd edn. Cambridge University Press, Cambridge (2014)
8. Flandoli, F., Gess, B., Scheutzow, M.: Synchronization by noise for order-preserving random dynamical systems. *Ann. Probab.* **45**, 1325–1350 (2017)
9. Gess, B.: Strong solutions for stochastic partial differential equations of gradient type. *J. Funct. Anal.* **263**, 2355–2383 (2012)
10. Gess, B., Tsatsoulis, P.: Synchronisation by noise for the stochastic quantisation equation in dimensions 2 and 3. *arXiv preprint arXiv:1910.07769* (2019)
11. Glatt-Holtz, N.E., Mattingly, J.C., Richards, G.: On unique ergodicity in nonlinear stochastic partial differential equations. *J. Stat. Phys.* **166**, 618–649 (2017)
12. Hairer, M.: Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probab. Theory Relat. Fields* **124**, 345–380 (2002)
13. Hairer, M.: Ergodic properties of Markov processes. Lecture notes, Univ. Warwick. <http://www.hairer.org/notes/Markov.pdf> (2006)
14. Hairer, M.: An Introduction to Stochastic PDEs. Lecture notes, Univ. Warwick. <http://www.hairer.org/notes/SPDEs.pdf> (2009)
15. Hairer, M.: Ergodic theory for Stochastic PDEs. Lecture notes, Univ. Warwick. <http://www.hairer.org/notes/Imperial.pdf> (2008)
16. Hairer, M., Mattingly, J.C.: Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Ann. Math.* **164**, 993–1032 (2006)
17. Hairer, M., Mattingly, J.C.: Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. *Ann. Probab.* **36**, 2050–2091 (2008)

18. Hairer, M., Mattingly, J.C.: A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electron. J. Probab.* **16**, 658–738 (2011)
19. Hairer, M., Mattingly, J.C., Scheutzow, M.: Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Relat. Fields.* **149**, 223–259 (2011)
20. Herzog, D., Mattingly, J.C.: Noise-induced stabilization of planar flows I. *Electron. J. Probab.* **20**, 1–43 (2015)
21. Kechris, A.: *Classical Descriptive Set Theory*. Springer, Berlin (2012)
22. Krylov, N.V., Rozovskii, B.L.: Stochastic evolution equations. *J. Math. Sci.* **16**, 1233–1277 (1981)
23. Kuksin, S., Shirikyan, A.: *Mathematics of Two-Dimensional Turbulence*. Cambridge University Press, Cambridge (2012)
24. Kulik, A., Scheutzow, M.: Generalized couplings and convergence of transition probabilities. *Probab. Theory Relat. Fields* **171**, 333–376 (2018)
25. Liggett, T.M.: *Interacting Particle Systems*. Springer, Berlin (1985)
26. Lindvall, T.: *Lectures on the Coupling Method*. Dover, New York (1992)
27. Lunardi, A.: *Analytic semigroups and optimal regularity in parabolic problems*. Springer, Berlin (2012)
28. Lund, R.B., Tweedie, R.L.: Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.* **21**, 182–194 (1996)
29. Mattingly, J.C.: Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity. *Commun. Math. Phys.* **206**, 273–288 (1999)
30. Mattingly, J.C.: Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics. *Commun. Math. Phys.* **230**, 421–462 (2002)
31. Meyn, S., Tweedie, R.: *Markov Chains and Stochastic Stability*. Springer, London (2012)
32. Roberts, G., Tweedie, R.: Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.* **37**, 359–373 (2000)
33. Rosenthal, J.: Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Am. Stat. Assoc.* **90**, 558–566 (1995)
34. Scheutzow, M., Schulze, S.: Strong completeness and semi-flows for stochastic differential equations with monotone drift. *J. Math. Anal. Appl.* **446**, 1555–1570 (2017)
35. Shiryaev, A.N.: Probability. In: *Graduate Texts in Mathematics*, vol. 95, 2nd edn, pp. xvi+623. Springer, New York. <https://doi.org/10.1007/978-1-4757-2539-1> (Translated from the first (1980) Russian edition by R. P. Boas) (1996)
36. Smoller, J.: *Shock Waves and Reaction–Diffusion Equations*. Springer, Berlin (2012)
37. Vessella, S.: Unique continuation properties and quantitative estimates of unique continuation for parabolic equations. *Handb. Diff. Equ. Evol. Equ.* **5**, 421–500 (2009)
38. Villani, C.: *Optimal Transport: Old and New*. Springer, Berlin (2008)

Communicated by M. Hairer