Exponential Mixing Properties of Stochastic PDEs Through Asymptotic Coupling

September 14, 2001

M. Hairer

Département de Physique Théorique
Université de Genève, 1211 Genève 4, Switzerland
E-mail: Martin.Hairer@physics.unige.ch

Abstract

We consider parabolic stochastic partial differential equations driven by white noise in time. We prove exponential convergence of the transition probabilities towards a unique invariant measure under suitable conditions. These conditions amount essentially to the fact that the equation transmits the noise to all its determining modes. Several examples are investigated, including some where the noise does *not* act on every determining mode directly.

Contents

1 Introduction

We are interested in the study of long-time asymptotics for parabolic stochastic partial differential equations. More precisely, the existence, uniqueness, and speed of convergence towards the invariant measure for such systems is investigated. The general setting is that of a stochastic PDE of the form

$$dx = Ax\,dt + F(x)\,dt + Q\,d\omega(t)\;, \qquad x(0) = x_0\;, \tag{1.1}$$

where x belongs to some Hilbert space $\mathcal{H}$, A is the generator of a C_0-semigroup on $\mathcal{H}$, $F : \mathcal{H} \to \mathcal{H}$ is some nonlinearity, ω is the cylindrical Wiener process on some other Hilbert space $\mathcal{W}$, and $Q : \mathcal{W} \to \mathcal{H}$ is a bounded operator. If the nonlinearity F is sufficiently "nice", there exists a unique solution $x(t)$ to (1.1) (see *e.g.* [DPZ92b]). In this paper, we investigate the asymptotic stability of (1.1). We say that the solutions of (1.1) are asymptotically stable if there exists a *unique* probability measure μ_* on $\mathcal{H}$ such that the laws of $x(t)$ converge to μ_*, independently of the initial condition x_0. We are interested in the situation where the asymptotic stability is a consequence of the noise (*i.e.* the deterministic equation $\dot{x} = Ax + F(x)$ is not asymptotically stable in the above sense), although the noise is weak, in the sense that the range of Q in $\mathcal{H}$ is "small".

The investigation of asymptotic behaviour for solutions of (1.1) goes back to the early eighties (see for example [MS99] for an excellent review article or the monograph [DPZ96] for a detailed exposition). Until recently, two approaches dominated the literature on this subject. For the first approach, sometimes called the "dissipativity method", one considers two solutions $x(t)$ and $y(t)$ of (1.1), corresponding to the same realization of the Wiener process ω, but with different initial conditions x_0 and y_0. If A and F are sufficiently dissipative, $\|x(t) - y(t)\|$ converges to 0 for large times in some suitable sense. If this convergence is sufficiently fast and uniform, it yields asymptotic stability results (see for example [DPZ92a]). Closely related to this approach are the Lyapunov function techniques, developed for semilinear equations in [Ich84]. The dissipativity method, as well as the Lyapunov function techniques, are limited by the requirement that the deterministic equation $\dot{x} = Ax + F(x)$ already shows stable behaviour.

The (linearly) unstable situations are covered by the second approach, to which we refer as the "overlap method". It consists in showing that the Markov transition semigroup associated to (1.1) has the strong Feller property and is topologically irreducible. Then, provided that the equation (1.1) shows some dissipativity, arguments as developed in the monograph [MT94], allow to bound the overlap between transition probabilities starting at two different initial points. This in turn yields strong asymptotic stability properties. The main technical difficulty of this approach is to show that the strong Feller property holds. This difficulty is usually mastered either by studying the infinite-dimensional backward Kolmogorov equation associated to (1.1) [DPZ91], or by showing that the Markov transition semigroup has good smoothing properties [DPEZ95, Cer99]. This technique is limited by the requirement that the noise be sufficiently non-degenerate. A typical requirement is that the range of Q *contains* the domain of some positive power of $-A$. To our knowledge, only one work [EH01, Hai01] shows the strong Feller property for a stochastic PDE in a situation where the range of Q is not dense in $\mathcal{H}$ (but still of finite codimension).

Very recently, a third approach, to which we refer as the "coupling method", emerged in a series of papers on the 2D Navier-Stokes equation. (See [KS01, Mat01, MY01] and the references in Section 6.) The main idea of these papers is to make a splitting $\mathcal{H} = \mathcal{H}_L \oplus \mathcal{H}_H$ of the dynamics into a finite-dimensional, linearly unstable, low-frequency part $\mathcal{H}_L$ and a remaining infinite-dimensional stable part $\mathcal{H}_H$. An important assumption on Q is then that the range of Q contains $\mathcal{H}_L$. The space $\mathcal{H}_L$ is chosen in such a way that the long-time asymptotics of the dynamics is completely dominated by the behaviour of the low-frequency part. More precisely, for any given realization $x_L(t)$ of the low-frequency part, the dynamics of the high-frequency part $x_H(t)$ will loose memory of its initial condition exponentially fast. On the low-frequency part, in turn, the noise acts in a non-degenerate way. A clever coupling argument allows to combine these two facts in order to obtain asymptotic stability results. The argument consists in coupling two realizations of (1.1) in such a way that if the low-frequency parts meet at some time τ, they remain equal for all times $t > \tau$. (Of course, one has to show that τ is finite with probability 1.) In fact, this coupling method is very close to the Gibbsian approach developed in [KS00, BKL00b, EMS01], which consisted in transforming the infinite-dimensional Markovian system on $\mathcal{H}$ to a finite-dimensional non-Markovian system on $\mathcal{H}_L$. This finite-dimensional system was shown to have exponentially decaying memory and thus techniques from statistical mechanics can be applied.

Loosely speaking, the coupling method combines the arguments of both the dissipativity method (on $\mathcal{H}_H$) and the overlap method (on $\mathcal{H}_L$). The coupling method thus yields a very powerful approach to the problem of asymptotic stability of (1.1). The conditions of applicability of this coupling method have been successively weakened in the aforementioned papers, but the existing results always require, as we already mentioned, that the noise acts directly and independently on *every* determining mode of the equation. In this paper, we extend the coupling method to problems which do not satisfy this condition. Our overall approach is similar to the one exposed by Mattingly in [Mat01], and consequently some of our proofs are closely related to the arguments exposed there. Our main new idea is to construct a coupling for which the low-frequency parts of the dynamics do not actually meet at some finite time, but converge exponentially fast towards each other. This "asymptotic coupling" is achieved through a binding construction exposed in Section 2.3, which seems to be new and can in some cases be implemented even in very degenerate situations.

In the following section, we illustrate the method of asymptotic coupling for a simple finite dimensional problem.

1.1 A toy model

Consider the following system of stochastic differential equations in $\mathbf{R}^2$:

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\,dt + d\omega(t)\;, \\ dx_2 &= (2x_2 + x_1 - x_2^3)\,dt\;. \end{aligned} \tag{1.2}$$

This equation should be interpreted in the integral sense, with $\omega \in \Omega$ a Brownian motion. Applying Hörmander's condition [Hör85, Nor86], it is easy to see that the transition probabilities of (1.2) are smooth with respect to the Lebesgue measure on $\mathbf{R}^2$. Furthermore, an easy controllability argument shows that they have support everywhere and therefore are all mutually equivalent. Since (1.2) also exhibits a strong drift towards the center of the phase space at large amplitudes, it follows by standard arguments that (1.2) possesses a unique invariant measure μ_* and that every initial condition is exponentially (in variation norm) attracted by μ_*.

The problem with this argument is that it heavily relies on the existence of some reference measure (in this case the Lebesgue measure) which is equivalent to the transition probabilities. In the infinite-dimensional setting, such a reference measure will usually not exist when the noise is sufficiently degenerate. (For an account of some cases where such a reference measure does exist in the infinite-dimensional case, see [MS99, EH01].) Furthermore, the fact that both directions in (1.2) are linearly unstable prevents one from applying the coupling method as it is presented in the previous section.

We will show that the invariant measure for (1.2) is unique, using a coupling construction which pushes solutions together at an exponential rate. This construction is asymptotic, compared to more conventional coupling constructions, which look for hitting times at which the coupled dynamics actually meets.

Before we proceed, let us explain briefly what we mean by "coupling". A coupling for (1.2) is a process $(x(t), y(t)) \in \mathbf{R}^2 \times \mathbf{R}^2$, whose marginals $x(t)$ and $y(t)$ taken separately are both solutions of (1.2) (but with different initial conditions). In general, one takes a measure P on $\Omega \times \Omega$, whose marginals are both equal to the Wiener measure W. Then a coupling for (1.2) can be constructed by drawing a pair $(\omega, \tilde{\omega}) \in \Omega \times \Omega$ distributed according to P and solving the equations

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\, dt + d\omega(t) , & dy_1 &= (2y_1 + y_2 - y_1^3)\, dt + d\tilde{\omega}(t) , \\ dx_2 &= (2x_2 + x_1 - x_2^3)\, dt , & dy_2 &= (2y_2 + y_1 - y_2^3)\, dt . \end{aligned} \tag{1.3}$$

We will carefully choose the measure P in such a way that the quantity $\|x - y\|$ converges exponentially to 0 for large times. This then yields the uniqueness of the invariant measure for (1.2).

Our main idea leading to the construction of P is to consider the following system in $\mathbf{R}^4$:

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\, dt + d\omega(t) , \\ dx_2 &= (2x_2 + x_1 - x_2^3)\, dt , \\ dy_1 &= (2y_1 + y_2 - y_1^3)\, dt + d\omega(t) + G(x_1, x_2, y_1, y_2)\, dt , \\ dy_2 &= (2y_2 + y_1 - y_2^3)\, dt , \end{aligned} \tag{1.4}$$

where $d\omega$ denotes twice the same realization of the Wiener process. We see that this equation is the same as (1.3) with $\tilde{\omega}$ defined by

$$\tilde{\omega}(t) = \omega(t) + \int_0^t G(x_1(s), x_2(s), y_1(s), y_2(s))\, ds . \tag{1.5}$$

The noise $\tilde{\omega} \in \Omega$ is distributed according to some measure $\tilde{\mathsf{W}}$ which is in general not equal to the Wiener measure W. Therefore, (1.4) does not yet define a coupling for (1.2). If G is small in the sense that the quantity

$$\int_0^\infty \|G(x_1(s), x_2(s), y_1(s), y_2(s))\|^2\, ds \tag{1.6}$$

is bounded with sufficiently high probability, then the measures $\tilde{\mathsf{W}}$ and W are equivalent. In this case, it is possible to construct a measure P on $\Omega \times \Omega$ whose marginals are W, with the important property that there exists a random time τ with $\mathbf{P}(\tau < \infty) = 1$ such that the solutions of the coupled system satisfy (1.4) for times $t \geq \tau$.

In view of the above, we have reduced the problem to finding a function G such that the solutions of (1.4) satisfy $\|y(t) - x(t)\| \to 0$ for $t \to \infty$ and (1.6) is bounded. We introduce the difference process $\varrho = y - x$, and we write

$$\dot{\varrho}_1 = 2\varrho_1 + \varrho_2 - \varrho_1(x_1^2 + x_1 y_1 + y_1^2) + G(x, y) , \tag{1.7a}$$

$$\dot{\varrho}_2 = 2\varrho_2 + \varrho_1 - \varrho_2(x_2^2 + x_2 y_2 + y_2^2) . \tag{1.7b}$$

It is easy to find a function G such that $\varrho_1 \to 0$, but this does not yet mean that ϱ_2 will go to zero. A closer look at (1.7b) shows that if we could force ϱ_1 to be very close to $-3\varrho_2$, (1.7b) could be written as

$$\dot{\varrho}_2 = -\varrho_2 + \varepsilon - \varrho_2(x_2^2 + x_2 y_2 + y_2^2) \;,$$

which is asymptotically stable. Introduce the function $\zeta = \varrho_1 + 3\varrho_2$. We then have

$$\dot{\zeta} = (\ldots) + G(x_1, x_2, y_1, y_2) \;,$$

with $(\ldots)$ an expression of the order $\|\varrho\|(1 + \|x\|^2 + \|y\|^2)$. Now we can of course choose $G = -(\ldots) - 2\zeta$. This way, the equation for ζ becomes $\dot{\zeta} = -2\zeta$ and we have the solution $\zeta(t) = \zeta(0)e^{-2t}$. Plugging this into (1.7b), we get

$$\dot{\varrho}_2 = -\varrho_2 + \zeta(0)e^{-2t} - \varrho_2(x_2^2 + x_2 y_2 + y_2^2) \;.$$

We thus have the estimate

$$|\varrho_2(t)| \le |\varrho_2(0)|e^{-t} + |\zeta(0)|e^{-2t} \;.$$

Finally, ϱ_1 is estimated by using the definition of ζ and we get

$$|\varrho_1(t)| \le |\varrho_2(0)|e^{-t} + 4|\zeta(0)|e^{-2t} \;.$$

This shows that, with G chosen this way, there exists a constant C such that

$$\|x(t) - y(t)\| \le C\|x(0) - y(0)\|e^{-t} \;,$$

for almost every realization of the noise. Since typical realizations of $x(t)$ do not grow faster than linearly, G is also of the order e^{-t}, with at most a polynomial factor in t multiplying the exponential. The main result of this paper, Theorem 4.1, shows that the above construction implies the existence and uniqueness of an invariant probability measure μ_* for the problem at hand. Furthermore, it shows that the transition probabilities converge exponentially fast towards μ_* in the Wasserstein norm (the dual norm to the Lipschitz norm on functions).

This concludes our presentation of the toy model. For a more precise statement, the reader is encouraged to actually check that the above construction allows to verify the assumptions stated in Section 5.

The remainder of this paper is organized as follows. In Section 2, we give the precise definitions for the type of coupling we will consider. In Section 3, we state the properties of the coupling that are required for our purpose. In Section 4, we prove the abstract formulation of our main ergodic theorem. In Section 5, this abstract theorem is then specialized to the case of stochastic differential equations. In Section 6 finally, we present several examples where our construction applies, although the noise does not act directly on every determining mode of the equation.

Acknowledgements

The author thanks Jean-Pierre Eckmann, Guillaume van Baalen, Emmanuel Zabey, and Jacques Rougemont for several helpful discussions and comments. This work was partially supported by the Fonds National Suisse.

2 The Coupling Construction

In this section, we explain our coupling construction. Before we start with the actual definitions of the various objects appearing in the construction, we fix our notations.

2.1 Notations

If μ is a measure on a measurable space X (in the sequel, we will always consider Polish[1] spaces) and $f : \mathsf{X} \to \mathsf{Y}$ is a measurable map, we denote by $f^*\mu$ the measure on Y defined by $(f^*\mu)(A) \equiv \mu(f^{-1}(A))$. For example, if Π is a projection on one component of a product space, $\Pi^*\mu$ denotes the marginal of μ on this component. If a natural reference measure is given on the current space, we denote by $\mathcal{D}\mu$ the density of μ with respect to the reference measure.

We define for any two measures μ and ν the measures $\mu \wedge \nu$ and $\mu \setminus \nu$. If a common reference measure is given, these operations act on densities like

$$\begin{aligned}(\mathcal{D}(\mu \wedge \nu))(x) &= \min\{\mathcal{D}\mu(x)\,, \mathcal{D}\nu(x)\}\;,\\ (\mathcal{D}(\mu \setminus \nu))(x) &= \max\{\mathcal{D}\mu(x) - \mathcal{D}\nu(x), 0\}\;.\end{aligned}$$

It immediately follows that $\mu = (\mu \wedge \nu) + (\mu \setminus \nu)$ for any two measures μ and ν. We will use the equivalent notations $\mu \le \nu$ and $\nu \ge \mu$ to say that $\mu \wedge \nu = \mu$ holds. One can check the following relations:

$$\begin{aligned}f^*(\mu \wedge \nu) &\le f^*\mu \wedge f^*\nu\;,\\ f^*(\mu \setminus \nu) &\ge f^*\mu \setminus f^*\nu\;.\end{aligned}$$

Equalities hold if f is injective.

For a given topological space X, we denote by $\mathscr{M}(\mathsf{X})$ the space of all finite signed Borel measures on X. We denote by $\mathscr{M}_1(\mathsf{X})$ the set of all probability measures on X. For $\mu \in \mathscr{M}(\mathsf{X})$, we denote by $\|\mu\|$ its total variation norm (which is simply its mass if μ has a sign).

---

[1] *i.e.* complete, separable, and metric

2.2 Definition of coupling

In this section, and until the end of the paper, we will often consider families $\mathbf{Q}_y$ of measures indexed by elements $y \in \mathsf{Y}$, with Y some Polish space. One should think of y as the initial condition of a Markov chain on Y and of $\mathbf{Q}_y$ either as its transition probabilities, or as the measure on pathspace obtained by starting from y. We will always assume that the functions $y \mapsto \mathbf{Q}_y(A)$ are measurable for every Borel set A. If $\mathbf{Q}_y$ is a family of measures on Y^n and $\mathbf{R}_y$ is a family of measures on Y^m, a family of measures $(\mathbf{RQ})_y$ on $\mathsf{Y}^{n+m} = \mathsf{Y}^n \times \mathsf{Y}^m$ can be defined on cylindrical sets in a natural way by

$$(\mathbf{RQ})_y(A \times B) = \int_A \mathbf{R}_{z_n}(B)\,\mathbf{Q}_y(dz)\;, \tag{2.1}$$

where $A \subset \mathsf{Y}^n$, $B \subset \mathsf{Y}^m$, and z_n denotes the nth component of z.

We consider a discrete-time Markovian random dynamical system (RDS) Φ on a Polish space X with the following structure. There exists a "one-step" probability space $(\Omega, \mathscr{F}, \mathsf{P})$ and Φ is considered as a jointly measurable map $\Phi : (\mathsf{X}, \Omega) \to \mathsf{X}$. The iterated maps $\Phi^n : (\mathsf{X}, \Omega^n) \to \mathsf{X}$ with $n \in \mathbf{N}$ are constructed recursively by

$$\Phi^n(x, \omega_1, \ldots, \omega_n) = \Phi(\Phi^{n-1}(x, \omega_1, \ldots, \omega_{n-1}), \omega_n)\;,$$

This construction gives rise to a Markov chain on X (also denoted by Φ) with one-step transition probabilities

$$\mathsf{P}_x \equiv \Phi(x, \cdot)^* \mathsf{P}\;.$$

The n-step transition probabilities will be denoted by P_x^n. Our main object of study will be the family of measures on pathspace generated by Φ. Take a sequence $\{\omega_i\}_{i=0}^\infty$ and an initial condition $x \in \mathsf{X}$. We then define $x_0 = x$ and $x_{i+1} = \Phi(x_i, \omega_i)$. We will denote by $\mathbf{P}_x^n$ with $n \in \mathbf{N} \cup \{\infty\}$ the measure on X^n obtained by transporting P^n with the map $\{\omega_i\} \mapsto \{x_i\}$. It is also natural to view $\mathbf{P}_x^n$ as a measure on $\mathsf{X}^n \times \Omega^n$ by transporting P^n with the map $\{\omega_i\} \mapsto \{x_i, \omega_i\}$, so we will use both interpretations.

Remark 2.1 The above setup is designed for the study of stochastic differential equations driven by additive noise. In that case, Ω is some Wiener space and Φ maps an initial condition and a realization of the Wiener process on the solution after time 1. Nevertheless, our setup covers much more general cases.

The coupling needs two copies of the pathspace, *i.e.* we will consider elements $(x, y) \in \mathsf{X}^\infty \times \mathsf{X}^\infty$. It will be convenient to use several projectors from $\mathsf{X}^N \times \mathsf{X}^N$ to its components. We define therefore (for $n \le N$):

$$\Pi_1 : (x, y) \mapsto x\;, \quad \Pi_2 : (x, y) \mapsto y\;, \quad \pi_n : (x, y) \mapsto (x_n, y_n)\;.$$

We also define $\pi_{i,n} \equiv \Pi_i \circ \pi_n$ for $i \in \{1,2\}$.

Definition 2.2 *Let Φ be a Markov chain on a Polish space X and let $\mathbf{P}_x^\infty$ be the associated family of measures on the pathspace X^∞. A* coupling *for Φ is a family $\mathbf{C}_{x,y}^\infty$ of probability measures on $\mathsf{X}^\infty \times \mathsf{X}^\infty$ satisfying*

$$\Pi_1^* \mathbf{C}_{x,y}^\infty = \mathbf{P}_x^\infty \qquad \textit{and} \qquad \Pi_2^* \mathbf{C}_{x,y}^\infty = \mathbf{P}_y^\infty \;,$$

where Π_1 and Π_2 are defined as above.

A trivial example of coupling is given by $\mathbf{C}_{x,y}^\infty = \mathbf{P}_x^\infty \times \mathbf{P}_y^\infty$. The interest of constructing a non-trivial coupling comes from the following observation. Take some suitable set of test functions $\mathcal{G}$ on X and define a norm on $\mathscr{M}(\mathsf{X})$ by

$$\|\mu\|_{\mathcal{G}} = \sup_{g\in\mathcal{G}} \langle g, \mu\rangle \;.$$

Once the existence of an invariant measure for the Markov chain Φ is established, one usually wishes to show its uniqueness by proving that Φ forgets about its past sufficiently fast, *i.e.*

$$\lim_{n\to\infty} \|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}} = 0 \;, \quad \text{for all} \quad (x,y) \in \mathsf{X}^2 \;,$$

with suitable bounds on the convergence rate as a function of the initial conditions. Now take a coupling $\mathbf{C}_{x,y}^\infty$ for Φ. It is straightforward to see that by definition the equality

$$\langle \mathsf{P}_x^n, g\rangle = \int_{\mathsf{X}\times\mathsf{X}} g(z)\, (\pi_{1,n}^* \mathbf{C}_{x,y}^\infty)(dz)$$

holds, as well as the same equality where $\pi_{1,n}$ is replaced by $\pi_{2,n}$ and P_x^n is replaced by P_y^n. Therefore, one can write

$$\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}} = \sup_{g\in\mathcal{G}} \int_{\mathsf{X}\times\mathsf{X}} (g(\Pi_1 z) - g(\Pi_2 z))\, (\pi_n^* \mathbf{C}_{x,y}^\infty)(dz) \;. \tag{2.2}$$

This equation is interesting, because it is in many cases possible to construct a coupling $\mathbf{C}_{x,y}^\infty$ such that for n large, the measure $\pi_n^* \mathbf{C}_{x,y}^\infty$ is concentrated near the diagonal $\Pi_1 z = \Pi_2 z$, thus providing through (2.2) an estimate for the term $\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}}$. This is precisely what was shown in our toy model of Section 1.1, where we constructed f in such a way that $\|x(t) - y(t)\| \to 0$ for $t \to \infty$.

2.3 The binding construction

In this subsection, we describe a specific type of coupling for a given RDS Φ. Only couplings of that type will be taken under consideration in the sequel.

Let Φ and the associated probability space $(\Omega, \mathscr{F}, \mathsf{P})$ be as above. We consider a family $\psi_{x\to y} : \Omega \to \Omega$ (the pair (x, y) belongs to X^2) of measurable functions that also have measurable inverses. We will call these functions *binding functions* for Φ. The reason for this terminology is that, given a realization $\{\omega_n\}_{n=0}^{\infty}$ of the noise and a pair of initial conditions $(x_0, y_0) \in \mathsf{X}^2$, the binding functions allow us to construct two paths $\{x_n\}$ and $\{y_n\}$ by setting

$$\tilde{\omega}_n = \psi_{x_n\to y_n}(\omega_n) , \qquad x_{n+1} = \Phi(x_n, \omega_n) , \qquad y_{n+1} = \Phi(y_n, \tilde{\omega}_n) . \tag{2.3}$$

Our aim is to find a family $\psi_{x\to y}$ such that y_n converges towards x_n in a suitable sense for large values of n. Thus, the binding functions play the role of a spring between x and y. We will say that (2.3) is a *binding construction* for Φ. We denote the inverse of $\psi_{x\to y}$ by $\psi_{x\leftarrow y}$. The reason behind this notation should be clear from the diagram below.

$$\begin{tikzcd}
x_n \arrow[d, no head] \arrow[rr, dashed] & & \psi_{x_n\to y_n} & & y_n \arrow[ll, dashed] \arrow[d, no head] \\
\Phi(\,\cdot\,, \omega_n) \arrow[d] & \omega_n \arrow[l, dashed] \arrow[rr, shift left] & & \tilde{\omega}_n \arrow[ll, shift left, "\psi_{x_n\leftarrow y_n}"] \arrow[r, dashed] & \Phi(\,\cdot\,, \tilde{\omega}_n) \arrow[d] \\
x_{n+1} & & & & y_{n+1}
\end{tikzcd} \tag{2.4}$$

The solid arrows denote the various maps and the dashed arrows denote the influences of the appearing quantities on those maps. It shows that it is also possible to achieve the binding construction by first choosing a sequence $\{\tilde{\omega}_n\}_{n=0}^{\infty}$ and then using $\psi_{x_n\leftarrow y_n}$ to construct the ω_n, thus obtaining the same set of possible realizations for (x_n, y_n). This symmetry between $\psi_{x\to y}$ and $\psi_{x\leftarrow y}$ is also explicit in (2.6) below.

Guided by the above construction, we use the binding maps to construct a coupling Markov chain Ψ on $\mathsf{X} \times \mathsf{X}$ with transition probabilities $\mathbf{C}_{x,y}$ in the following way. Define the maps

$$\begin{aligned} \Psi_{x\to y} : \Omega &\to \Omega \times \Omega & \qquad \Psi_{x\leftarrow y} : \Omega &\to \Omega \times \Omega \\ \omega &\mapsto (\omega, \psi_{x\to y}(\omega)) , & \omega &\mapsto (\psi_{x\leftarrow y}(\omega), \omega) . \end{aligned} \tag{2.5}$$

Notice that, up to some null set, the image of both maps is the set $\{(\omega, \tilde{\omega}) \mid \tilde{\omega} = \psi_{x\to y}(\omega)\}$. Then we define a family of measures $\mathsf{P}_{x,y}$ on $\Omega \times \Omega$ by

$$\mathsf{P}_{x,y} = (\Psi^*_{x\to y}\mathsf{P}) \wedge (\Psi^*_{x\leftarrow y}\mathsf{P}) = \Psi^*_{x\leftarrow y}(\mathsf{P} \wedge \psi^*_{x\to y}\mathsf{P}) . \tag{2.6}$$

According to (2.4), the measure P_{x_n,y_n} is precisely the common part between the measure obtained for $(\omega_n, \tilde{\omega}_n)$ by distributing ω_n according to P and the one obtained by distributing $\tilde{\omega}_n$ according to P. Thus both marginals of the measure $\mathsf{P}_{x,y}$ are smaller (in the sense of Section 2.1) than P. In order to have a non-trivial construction, we impose that the measures P and $\psi^*_{x\to y}\mathsf{P}$ are equivalent. The density of $\psi^*_{x\to y}\mathsf{P}$ relative to P will be denoted by $\mathcal{D}_{x,y}(\omega)$.

Considering again (2.4), the family of measures $\mathsf{P}_{x,y}$ is transported on $\mathsf{X} \times \mathsf{X}$ by defining

$$\begin{aligned} \Phi_{x,y} : \Omega \times \Omega &\to \mathsf{X} \times \mathsf{X} \\ (\omega, \tilde{\omega}) &\mapsto (\Phi(x,\omega), \Phi(y,\tilde{\omega})) \;, \end{aligned}$$

and setting

$$\mathbf{Q}_{x,y} \equiv \Phi^*_{x,y}\mathsf{P}_{x,y} \;. \tag{2.7}$$

But this does not give a transition probability function yet, since the measures $\mathsf{P}_{x,y}$ are not normalized to 1. We therefore define the family of measures $\mathbf{P}_{x,y}$ by

$$\mathbf{P}_{x,y} = \mathsf{P}_{x,y} + c_{x,y}\,(\mathsf{P} \setminus \Pi_1^*\mathsf{P}_{x,y}) \times (\mathsf{P} \setminus \Pi_2^*\mathsf{P}_{x,y}) \;,$$

where the number $c_{x,y}$ is chosen in such a way that the resulting measure is a probability measure. By a slight abuse of notation, we used here the symbol Π_i to denote the projection on the ith component of $\Omega \times \Omega$. As a matter of fact, $(\mathsf{P} \setminus \Pi_1^*\mathsf{P}_{x,y})$ and $(\mathsf{P} \setminus \Pi_2^*\mathsf{P}_{x,y})$ have the same mass, which is equal to $1 - \|\mathsf{P}_{x,y}\|$, so

$$c_{x,y} = \frac{1}{\|\mathsf{P} \setminus \Pi_2^*\mathsf{P}_{x,y}\|} \;,$$

for example. (Recall that the symbol $\|\cdot\|$ stands for the total variation norm, which is simply equal to its mass for a positive measure.) It is straightforward to show that the following holds:

Lemma 2.3 *The measures* $\mathbf{P}_{x,y}$ *satisfy* $\Pi_i^*\mathbf{P}_{x,y} = \mathsf{P}$ *for* $i = 1, 2$.

Proof. It is clear by (2.6) that $\Pi_i^*\mathsf{P}_{x,y} \le \mathsf{P}$. Thus

$$\begin{aligned} \Pi_1^*\mathbf{P}_{x,y} &= \Pi_1^*\mathsf{P}_{x,y} + c_{x,y}\|\mathsf{P} \setminus \Pi_2^*\mathsf{P}_{x,y}\|(\mathsf{P} \setminus \Pi_1^*\mathsf{P}_{x,y}) \\ &= (\mathsf{P} \wedge \Pi_1^*\mathsf{P}_{x,y}) + (\mathsf{P} \setminus \Pi_1^*\mathsf{P}_{x,y}) = \mathsf{P} \;, \end{aligned} \tag{2.8}$$

and similarly for $\Pi_2^*\mathbf{P}_{x,y}$. □

This finally allows us to define the transition probabilities for Ψ by

$$\mathbf{C}_{x,y} = \Phi^*_{x,y}\mathbf{P}_{x,y} \equiv \mathbf{Q}_{x,y} + \mathbf{R}_{x,y} \;. \tag{2.9}$$

In this expression, the only feature of $\mathbf{R}_{x,y}$ we will use is that it is a positive measure. We define $\mathbf{C}^\infty_{x,y}$ as the measure on the pathspace $\mathsf{X}^\infty \times \mathsf{X}^\infty$ obtained by iterating (2.1). Since $\Pi_1 \circ \Phi_{x,y} = \Phi(x,\cdot) \circ \Pi_1$ and similarly for Π_2, it is straightforward to verify, using Lemma 2.3, that the measure $\mathbf{C}^\infty_{x,y}$ constructed this way is indeed a coupling for Φ.

For a given step of Ψ, we say that the trajectories do couple if the step is drawn according to $\mathbf{Q}_{x,y}$ and that they don't couple otherwise.

Remark 2.4 Since $\mathbf{P}_{x,y}$ is a family of measures on $\Omega \times \Omega$, it is also possible to interpret $\mathbf{C}^n_{x,y}$ as a family of probability measures on $\mathsf{X}^n \times \mathsf{X}^n \times \Omega^n \times \Omega^n$. We will sometimes use this viewpoint in the following section. It is especially useful when the RDS Φ is obtained by sampling a continuous-time process.

Remark 2.5 It will sometimes be useful to have an explicit way of telling whether a step of Ψ is taken according to $\mathbf{Q}^\infty_{x,y}$ or according to $\mathbf{R}^\infty_{x,y}$ (*i.e.* whether the trajectories couple or not). To this end, we introduce a Markov chain $\hat\Psi$ on the augmented phase space $\mathsf{X} \times \mathsf{X} \times \{0,1\}$ with transition probabilities

$$\mathbf{P}_{x,y} = \mathbf{Q}_{x,y} \times \delta_1 + \mathbf{R}_{x,y} \times \delta_0 \;.$$

The marginal of $\hat\Psi$ on $\mathsf{X} \times \mathsf{X}$ is of course equal to Ψ. By a slight abuse of notation, we will also write $\mathbf{C}^\infty_{x,y}$ for the probability measure on pathspace induced by $\hat\Psi$.

It will be useful in the sequel to have a map that "transports" the family of maps $\psi_{x\to y}$ on Ω^n via the RDS Φ. More precisely, fix a pair $(x,y) \in \mathsf{X} \times \mathsf{X}$ of starting points and a sequence $(\omega_0,\ldots,\omega_n)$ of realizations of the noise. We then define $x_0 = x$, $y_0 = y$, and, recursively for $i = 0,\ldots,n$

$$x_{i+1} = \Phi(x_i,\omega_i) \;, \qquad y_{i+1} = \Phi(y_i, \psi_{x_i\to y_i}(\omega_i)) \;.$$

This allows us to define the family of maps $\Xi^n_{x,y} : \Omega^n \to \Omega^n$ by

$$\Xi^{n+1}_{x,y}(\omega_0,\ldots,\omega_n) \mapsto (\psi_{x_0\to y_0}(\omega_0),\ldots,\psi_{x_n\to y_n}(\omega_n)) \;. \tag{2.10}$$

Since $\psi^*_{x\to y}\mathsf{P}$ is equivalent to P, we see that $(\Xi^n_{x,y})^*\mathsf{P}^n$ is equivalent to P^n and we denote its density by $\mathcal{D}^n_{x,y}$. We also notice that the family of measures $\mathbf{Q}^n_{x,y}$ is obtained by transporting $(\Xi^n_{x,y})^*\mathsf{P}^n \wedge \mathsf{P}^n$ onto $\mathsf{X}^n \times \mathsf{X}^n$ with the maps $\Phi_{x_i,y_i} \circ \Psi_{x_i\to y_i}$. In particular, one has the equality

$$\|\mathbf{Q}^n_{x,y}\| = \|(\Xi^n_{x,y})^*\mathsf{P}^n \wedge \mathsf{P}^n\| = \int_{\Omega^n} \bigl(1 \wedge \mathcal{D}^n_{x,y}(\omega)\bigr)\, \mathsf{P}^n(d\omega) \;. \tag{2.11}$$

3 Assumptions on the Coupling

In this section, we investigate the properties of the coupling $\mathbf{C}_{x,y}^{\infty}$ constructed in the previous section. We give a set of assumptions on the binding functions $\psi_{x \to y}$ that ensure the existence and uniqueness of the invariant measure for Φ.

In order to achieve this, we want the map $\psi_{x \to y}$ to modify the noise in such a way that trajectories drawn according to $\mathbf{Q}_{x,y}$ tend to come closer together. This will be the content of Assumption **A3**. Furthermore, we want to know that this actually happens, so the noise should not be modified too much. This will be the content of assumptions **A4** and **A5**. All these nice properties usually hold only in a "good" region of the phase space. Assumptions **A1** and **A2** will ensure that visits to this good region happen sufficiently often.

3.1 Lyapunov structure

Since we are interested in obtaining exponential mixing, we need assumptions of exponential nature. Our first assumption concerns the global aspects of the dynamics. It postulates that Φ is attracted exponentially fast towards a "good" region of its state space. We achieve this by assuming the existence of a Lyapunov function for Φ.

Definition 3.1 *Let Φ by a RDS with state space X as before. A* Lyapunov function *for Φ is a function $V : \mathsf{X} \to [0, \infty]$ for which there exist constants $a \in (0,1)$ and $b > 0$, such that*

$$\int_{\Omega} V(\Phi(x,\omega))\,\mathsf{P}(d\omega) \le aV(x) + b\;, \tag{3.1}$$

for every $x \in \mathsf{X}$ with $V(x) < \infty$.

Our first assumption then reads

A1 *There exist a Lyapunov function V for Φ. Furthermore, V is such that*

$$\mathsf{P}\{\omega \mid V(\Phi(x,\omega)) < \infty\} = 1\;,$$

for every $x \in \mathsf{X}$.

For convenience, we also introduce the function $\tilde V : \mathsf{X} \times \mathsf{X} \to [0, \infty]$ defined by

$$\tilde V(x,y) = V(x) + V(y)\;.$$

Notice that $\tilde V$ is a Lyapunov function for Ψ by construction.

In some cases, when the control over the densities $\mathcal{D}_{x,y}$ is uniform enough or when the phase space is already compact (or bounded), a Lyapunov function is not needed. In such a situation, one can simply choose $V \equiv 1$.

In our case of interest, the RDS Φ is obtained by sampling a continuous-time process Φ_t at discrete times. In that setting, it is useful to have means to control excursions to large amplitudes that take place between two successive sampling times. To this end, we introduce a function $W : \mathsf{X} \times \Omega \to [0, \infty]$ given by

$$W(x, \omega) = \sup_{t \in [0,1]} V(\Phi_t(x, \omega))$$

in the continuous-time setting and by

$$W(x, \omega) = V(x)$$

in the discrete-time setting. In fact, any other choice of W is all right, as long as it satisfies the properties that are summarized in Assumption **A2** below.

Before stating these properties, we define two other functions that act on pairs of initial conditions that couple by

$$\begin{aligned} W_{x \to y}(\omega) &= W(x, \omega) + W(y, \psi_{x \to y}(\omega)) \;, \\ W_{x \leftarrow y}(\omega) &= W(x, \psi_{x \leftarrow y}(\omega)) + W(y, \omega) \;. \end{aligned} \tag{3.2}$$

We will assume that W and the binding functions are such that W, $W_{x \to y}$ and $W_{x \leftarrow y}$ do not behave much worse than V. More precisely, we will assume that:

A2 *There exists a function* $W : \mathsf{X} \times \Omega \to [0, \infty]$ *such that*

$$\operatorname*{ess\,inf}_{\omega \in \Omega} W(x, \omega) = V(x) \,, \tag{3.3a}$$

$$\int_\Omega W(x, \omega) \, \mathsf{P}(d\omega) \le c\, V(x) \,, \tag{3.3b}$$

for some constant $c > 0$. *Furthermore, there exist constants* $C > 0$ *and* $\delta \ge 1$ *such that the estimates*

$$\begin{aligned} W_{x \to y}(\omega) &\le C(1 + V(y) + W(x, \omega))^\delta \;, \\ W_{x \leftarrow y}(\omega) &\le C(1 + V(x) + W(y, \omega))^\delta \;, \end{aligned} \tag{3.4}$$

hold for the functions defined in (3.2).

The Lyapunov structure given by assumptions **A1** and **A2** ensures that W (and thus also V) does not increase too fast along a typical trajectory. In order to make this statement precise, we define for a given initial condition $x \in \mathsf{X}$ the sets $A_{x,k} \subset \Omega^\infty$ by

$$A_{x,k} = \{\omega \in \Omega^\infty \mid W(\Phi^n(x, \omega), \omega_n) \le kV(x) + kn^2 \quad \forall n > 0\} \;, \tag{3.5}$$

where k is some positive constant. The sets $A_{x,k}$ contain almost every typical realization of the noise:

Lemma 3.2 *Let Φ be a RDS satisfying assumptions* **A1** *and* **A2**. *Then, there exists a constant $C > 0$ such that*

$$\mathsf{P}^\infty(A_{x,k}) \geq 1 - \frac{C}{k}\;,$$

for every $x \in \mathsf{X}$ and every $k > 0$.

Proof. For $\omega \in \Omega^\infty$, we define $x_n = \Phi^n(x,\omega)$. Notice that by (3.3b) and the Lyapunov structure, one has the estimate

$$\mathbf{E}(W(x_n,\omega_{n+1})) \leq ca^n V(x) + \frac{bc}{1-a}\;, \tag{3.6}$$

where $\mathbf{E}$ denotes expectations with respect to P^∞. We also notice that $A_{x,k} = \bigcap_{n>0} A_{x,k}^{(n)}$ with

$$A_{x,k}^{(n)} = \{\omega \mid W(x_n,\omega_{n+1}) \leq kV(x) + kn^2\}\;.$$

Combining this with (3.6), we see that

$$\mathsf{P}^\infty(A_{x,k}^{(n)}) \geq 1 - \frac{c}{k}\frac{a^n V(x) + b(1-a)^{-1}}{V(x)+n^2}\;.$$

Therefore, the worst possible estimate for $\mathsf{P}^\infty(A_{x,k})$ is

$$\mathsf{P}^\infty(A_{x,k}) \geq 1 - \frac{c}{k}\sum_{n=1}^\infty \frac{a^n V(x) + b(1-a)^{-1}}{V(x)+n^2}\;,$$

which proves the claim. □

3.2 Binding property

The crucial property of the coupling is to bring trajectories closer together. In order to make this statement more precise, we introduce the Lipschitz norm $\|\cdot\|_{\mathrm{L}}$ defined on functions $g : \mathsf{X} \to \mathbf{R}$ by

$$\|g\|_{\mathrm{L}} = \sup_{x\in\mathsf{X}} |g(x)| + \sup_{x,y\in\mathsf{X}} \frac{|g(x)-g(y)|}{d(x,y)}\;,$$

where $d(\cdot,\cdot)$ denotes the distance in X. The dual norm on $\mathscr{M}(\mathsf{X})$ is then given by

$$\|\mu\|_{\mathrm{L}} = \sup_{\|g\|_{\mathrm{L}}=1} \int_{\mathsf{X}} g(x)\,\mu(dx)\;.$$

With this definition at hand, we make the following assumption on the coupling part $\mathbf{Q}_{x,y}^\infty$.

A3 *There exist a positive constant γ_1 and a family of constants $K \mapsto C_K$ such that, for every $K > 0$,*

$$\|\pi^*_{1,n}\mathbf{Q}^\infty_{x,y} - \pi^*_{2,n}\mathbf{Q}^\infty_{x,y}\|_{\mathrm{L}} \le C_K e^{-\gamma_1 n} , \tag{3.7}$$

holds when $\tilde V(x,y) \le K$.

Remark 3.3 The sub-probability kernels $\mathbf{Q}_{x,y}$ are smaller than the transition probabilities for the binding construction (2.3). Thus, (3.7) is implied by an inequality of the type

$$\mathbf{E}(d(x_n,y_n)) \le C\tilde V(x_0,y_0)e^{-\gamma_1 n} ,$$

where d denotes the distance in X and $\mathbf{E}$ denotes the expectation with respect to the construction (2.3).

Notice that this assumption is non-trivial only if our coupling is such that $\|\mathbf{Q}^\infty_{x,y}\| > 0$ for sufficiently many starting points. This will be ensured by the next assumption.

A4 *Let $\mathcal{D}^n_{x,y}$ be defined as in Section 2.3. We assume that for every $K > 0$, there exists a family of sets $\Gamma^K_{x,y} \subset \Omega^\infty$ and constants $c_1, c_2 > 0$ such that the estimates*

$$\mathsf{P}^\infty(\Gamma^K_{x,y}) > c_1 , \qquad \int_{\Gamma^K_{x,y}} (\mathcal{D}^n_{x,y}(\omega))^{-2}\, \mathsf{P}^n(d\omega) < c_2 , \tag{3.8}$$

hold for every $n \ge 0$, whenever $\tilde V(x,y) \le K$. The integral over $\Gamma^K_{x,y}$ in (3.8) should be interpreted as the integral over the projection of $\Gamma^K_{x,y}$ onto its n first components.

A typical choice for $\Gamma^K_{x,y}$ is $\Gamma^K_{x,y} = A_{y,k}$ or $\Gamma^K_{x,y} = A_{x,k} \cap A_{y,k}$ with k sufficiently large as a function of K. In this case, Lemma 3.2 ensures that the conditions required on $\Gamma^K_{x,y}$ are satisfied. As a consequence of Assumption **A4**, we have

Proposition 3.4 *Let $\mathbf{Q}^\infty_{x,y}$ be defined as above and suppose that assumptions* **A1** *and* **A4** *hold. Then there exists for every K a constant C_K such that $\|\mathbf{Q}^\infty_{x,y}\| \ge C_K$, whenever $\tilde V(x,y) \le K$.*

Proof. Notice first that if μ_1 and μ_2 are two equivalent probability measures with $\mu_2(dx) = \mathcal{D}(x)\,\mu_1(dx)$, then the condition

$$\int_A (\mathcal{D}(x))^{-2}\,\mu_1(dx) < c$$

implies that

$$(\mu_1 \wedge \mu_2)(A) \geq \frac{\mu_1(A)^2}{4c} ,$$

see, *e.g.* [Mat01]. Recalling (2.11), we use Lemma 3.2 and the above estimate with $\mu_1 = \mathsf{P}^n$, $\mathcal{D} = \mathcal{D}^n_{x,y}$, and $A = \Gamma^K_{x,y}$. Taking the limit $n \to \infty$ and using the assumption on $\Gamma^K_{x,y}$ proves the claim. □

Our last assumption will ensure that trajectories that have already coupled for some time have a very strong tendency to couple for all times.

In order to formulate our assumption, we introduce a family of sets $Q^n_K(x,y)$, which are the possible final states of a "coupled" trajectory of length n, starting from (x,y), and never leaving the set $\{(a,b) \mid V(a) + V(b) \leq K\}$. For a given pair of initial conditions $(x,y) \in \mathsf{X}^2$ with $\tilde{V}(x,y) \leq K$, we define the family of sets $Q^n_K(x,y) \subset \mathsf{X} \times \mathsf{X}$ recursively in the following way:

$$\begin{aligned} Q^0_K(x,y) &= \{(x,y)\} , \\ Q^{n+1}_K(x,y) &= \bigcup_{(a,b) \in Q^n_K(x,y)} \{(\Phi_{a,b} \circ \Psi_{a \to b})(\omega) \mid \omega \in \Omega \text{ and } W_{a \to b}(\omega) \leq K\} . \end{aligned}$$

Notice that we would have obtained the same sets by reversing the directions of the arrows in the definition.

We also denote by $\mathcal{D}_{x,y}(\omega)$ the density of $\psi^*_{x \to y}\mathsf{P}$ relative to P.

A5 *There exist positive constants C_2, γ_2 and ζ, such that for every $K > 0$, every $(x_0, y_0) \in \mathsf{X}^2$ with $\tilde{V}(x_0,y_0) \leq K$, and every $(x,y) \in Q^n_K(x_0,y_0)$, the estimate*

$$\int_{W_{x \leftarrow y}(\omega) \leq K} (1 - \mathcal{D}_{x,y}(\omega))^2 \, \mathsf{P}(d\omega) \leq C_2 e^{-\gamma_2 n} (1 + K)^\zeta , \tag{3.9}$$

holds for $n > \zeta \ln(1 + K)/\gamma_2$.

This assumption means that if the process couples for a time n, the density $\mathcal{D}_{x,y}$ is close to 1 on an increasingly large set, and therefore the probability of coupling for a longer time becomes increasingly large. This assumption is sufficient for the family of measures $(\mathbf{RQ}^n)_{x,y}$ to have an exponential tail at large values of n. More precisely, we have

Proposition 3.5 *Let assumptions* **A1**, **A2** *and* **A5** *hold. Then, there exists a positive constant γ_3 and, for every $K > 0$, a constant C_K such that*

$$\|(\mathbf{RQ}^n)_{x,y}\| \leq C_K e^{-\gamma_3 n} , \tag{3.10}$$

holds for every $n > 0$, whenever $\tilde{V}(x,y) \leq K$.

We first show the following elementary estimate (it is not optimal, but sufficient for our needs):

Lemma 3.6 *Let $\mu_1, \mu_2 \in \mathscr{M}_1(\mathsf{X})$ be two equivalent probability measures with $\mu_2(dx) = \mathcal{D}(x)\,\mu_1(dx)$. Then the conditions*

$$\mu_1(A) \geq 1 - \varepsilon_1 \quad \text{and} \qquad \int_A (1 - \mathcal{D}(x))^2\,\mu_1(dx) \leq \varepsilon_2\;,$$

for some measurable set A imply that

$$(\mu_1 \wedge \mu_2)(A) \geq 1 - \varepsilon_1 - \varepsilon_2^{1/2}\;.$$

Proof. Define the set $E \subset \mathsf{X}$ by

$$E = A \cap \{x \in \mathsf{X} \mid \mathcal{D}(x) \geq 1\}\;.$$

We then have

$$\begin{aligned}(\mu_1 \wedge \mu_2)(A) &= \mu_1(E) + \int_{A\setminus E} \mathcal{D}(x)\,\mu_1(dx)\\ &= \mu_1(A) - \int_{A\setminus E} (1 - \mathcal{D}(x))\,\mu_1(dx)\\ &\geq \mu_1(A) - \int_{A\setminus E} |1 - \mathcal{D}(x)|\,\mu_1(dx)\\ &\geq 1 - \varepsilon_1 - \sqrt{\int_{A\setminus E} (1 - \mathcal{D}(x))^2\,\mu_1(dx)}\;.\end{aligned}$$

This shows the claim. □

Proof of Proposition 3.5. Fix the value n and the pair (x, y). For every $c_n \geq \tilde{V}(x, y)$ (we will fix it later), we have the estimate

$$\begin{aligned}\|(\mathbf{R}\mathbf{Q}^n)_{x,y}\| &= \int_{\mathsf{X}^2} (1 - \|\psi^*_{x_n \to y_n}\mathsf{P} \wedge \mathsf{P}\|)\,(\pi_n^* \mathbf{Q}^n_{x,y})(dx_n, dy_n) \qquad (3.11)\\ &\leq (\pi_n^* \mathbf{Q}^n_{x,y})(\mathsf{X}^2 \setminus Q^n_{c_n}(x, y))\\ &\quad + \int_{Q^n_{c_n}(x,y)} (1 - \|\psi^*_{x_n \to y_n}\mathsf{P} \wedge \mathsf{P}\|)\,(\pi_n^* \mathbf{Q}^n_{x,y})(dx_n, dy_n)\;.\end{aligned}$$

Now choose another value w_n to be fixed later and consider for every (x_n, y_n) the set

$$B_n = \{\omega \in \Omega \mid W_{x_n \leftarrow y_n}(\omega) \leq w_n\}\;.$$

By the definition of $Q^n_{c_n}(x,y)$, its elements (x_n,y_n) satisfy in particular the inequality $\tilde V(x_n,y_n) \le c_n$. By Assumption **A2** and the Lyapunov structure, we have for every $(x_n,y_n) \in Q^n_{c_n}(x,y)$ the estimate

$$\mathsf{P}(B_n) \ge 1 - C\frac{c_n}{w_n^{1/\delta}} \;.$$

Combining this and Assumption **A5** with Lemma 3.6 yields

$$\|\psi^*_{x_n\to y_n}\mathsf{P}\wedge\mathsf{P}\| \ge 1 - C\frac{c_n}{w_n^{1/\delta}} - Ce^{-\gamma_2 n/2}(1+w_n)^{\zeta/2} \;,$$

as long as w_n is such that

$$w_n \ge c_n \quad\text{and}\quad n \ge \zeta\ln(1+w_n)/\gamma_2 \;. \tag{3.12}$$

It remains to give an upper bound for $(\pi^*_n\mathbf{Q}^n_{x,y})(\mathsf{X}^2 \setminus Q^n_{c_n}(x,y))$ to complete our argument. Define the sets $A^n(K) \subset \mathsf{X}^n \times \mathsf{X}^n \times \Omega^n \times \Omega^n$ by

$$A^n(K) = \{(x_i,y_i,\omega_i,\eta_i)_{i=1}^n \mid W(x_i,\omega_i) + W(y_i,\eta_i) \le K\} \;.$$

It is clear by the definition of $Q^n_{c_n}(x,y)$ that we have the equality

$$(\pi^*_n\mathbf{Q}^n_{x,y})(\mathsf{X}^2 \setminus Q^n_{c_n}(x,y)) = \mathbf{Q}^n_{x,y}(\mathsf{X}^n\times\mathsf{X}^n\times\Omega^n\times\Omega^n \setminus A^n(c_n)) \;,$$

where $\mathbf{Q}^n_{x,y}$ is considered as a measure on $\mathsf{X}^n\times\mathsf{X}^n\times\Omega^n\times\Omega^n$, following Remark 2.4. Since $\mathbf{Q}^n_{x,y} \le \mathbf{C}^n_{x,y}$, we have

$$(\pi^*_n\mathbf{Q}^n_{x,y})(\mathsf{X}^2 \setminus Q^n_{c_n}(x,y)) \le 1 - \mathbf{C}^n_{x,y}(A^n(c_n)) \le C\frac{n(\tilde V(x,y)+1)}{c_n} \;,$$

for some constant C. This last estimate is obtained in a straightforward way, following the lines of the proof of Lemma 3.2. Plugging these estimates back into (3.11) yields

$$\|(\mathbf{R}\mathbf{Q}^n)_{x,y}\| \le C\frac{n(\tilde V(x,y)+1)}{c_n} + C\frac{c_n}{w_n^{1/\delta}} + Ce^{-\gamma_2 n/2}(1+w_n)^{\zeta/2} \;.$$

At this point, we make use of our freedom to choose c_n and w_n. We set

$$c_n = \tilde V(x,y) + e^{\gamma_c n} \qquad\text{and}\qquad w_n = \tilde V(x,y) + e^{\gamma_w n} \;,$$

with γ_c and γ_w given by

$$\gamma_c = \frac{1}{2+2\delta\zeta}\gamma_2 \qquad\text{and}\qquad \gamma_w = \frac{\delta}{1+\delta\zeta}\gamma_2 \;.$$

As a consequence, there exist for any $\gamma < \gamma_c$ some constants C and c such that

$$\|(\mathbf{RQ}^n)_{x,y}\| \le C(1 + \tilde{V}(x,y))^c e^{-\gamma n} \;,$$

as long as $n \ge \zeta \ln(1 + w_n)/\gamma_2$. (Such a value of n can always be found, because the exponent γ_w is always smaller than γ_2/ζ.) In order to complete the argument, we notice that (3.10) is trivially satisfied for small values of n because $\|(\mathbf{RQ}^n)_{x,y}\|$ is always smaller than 1 by definition: it suffices to choose C_K sufficiently big. The proof of Proposition 3.5 is complete. □

4 An Exponential Mixing Result

This section is devoted to the proof of the main theorem of this paper.

Theorem 4.1 *Let Φ be a RDS with state space* X *satisfying assumptions* **A1**–**A5**. *Then, there exists a constant $\gamma > 0$ such that*

$$\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathrm{L}} \le C(1 + \tilde{V}(x,y))\, e^{-\gamma n} \;,$$

for every $(x,y) \in \mathsf{X}^2$ and every $n > 0$.

Remark 4.2 The proof of Theorem 4.1 does not rely on assumptions **A4** and **A5** directly, but on the conclusions of Propositions 3.4 and 3.5. Nevertheless, in the setting of stochastic differential equations, it seems to be easier to verify the assumptions rather than to show the conclusions of the propositions by other means.

Corollary 4.3 *If Φ satisfies assumptions* **A1**–**A5**, *it possesses a unique invariant measure μ_* and*

$$\|\mathsf{P}_x^n - \mu_*\|_{\mathrm{L}} \le C(1 + V(x))\, e^{-\gamma n} \;.$$

Proof of the corollary. To show the existence of the invariant measure μ_*, we show that for any given initial condition x with $V(x) < \infty$, the sequence of measures P_x^n is a Cauchy sequence in the norm $\|\cdot\|_{\mathrm{L}}$. We have indeed

$$\begin{aligned}
\|\mathsf{P}_x^n - \mathsf{P}_x^{n+k}\|_{\mathrm{L}} &= \sup_{\|g\|_{\mathrm{L}} \le 1} \int_{\mathsf{X}} g(z)(\mathsf{P}_x^n - \mathsf{P}_x^{n+k})(dz) \\
&= \sup_{\|g\|_{\mathrm{L}} \le 1} \int_{\mathsf{X}} \int_{\mathsf{X}} g(z)(\mathsf{P}_x^n - \mathsf{P}_y^n)(dz)\, \mathsf{P}_x^k(dy) \\
&\le \int_{\mathsf{X}} \|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathrm{L}}\, \mathsf{P}_x^k(dy) \le C e^{-\gamma n} \int_{\mathsf{X}} (1 + \tilde{V}(x,y))\, \mathsf{P}_x^k(dy) \\
&\le C e^{-\gamma n}(1 + V(x)) \;,
\end{aligned}$$

where we used the Lyapunov structure to get the last inequality.

The claim now follows immediately from the theorem, noticing that if μ_* is an invariant measure for Φ, then

$$\int_{\mathsf{X}} V(x)\,\mu_*(dx) \le \frac{b}{1-a}\;,$$

due to the Lyapunov structure and the fact that the dynamics immediately leaves the set $V^{-1}(\infty)$. □

Before we turn to the proof of Theorem 4.1, we introduce some notations and make a few remarks. By iterating (2.9), one sees that

$$\mathbf{C}^{\infty}_{x,y} = \mathbf{Q}^{\infty}_{x,y} + \sum_{n=0}^{\infty} \bigl(\mathbf{C}^{\infty}\mathbf{R}\mathbf{Q}^{n}\bigr)_{x,y}\;, \tag{4.1}$$

where the symbol $\bigl(\mathbf{C}^{\infty}\mathbf{R}\mathbf{Q}^{n}\bigr)_{x,y}$ is to be interpreted in the sense of (2.1). This expression is the equivalent, in our setting, of Lemma 2.1 in [Mat01]. Using (4.1), the Markov chain Ψ can be described by means of another Markov chain Υ on $\mathsf{Y} = (\mathsf{X}^2 \times \mathbf{N}) \cup \{\star\}$, where $\star$ corresponds to "coupling for all times" in the sense of Section 2.3. First, we define

$$K_0 = \frac{4b}{1-a}\;, \qquad \tilde{K}_0 = \{(x,y) \mid \tilde{V}(x,y) \le K_0\}\;, \tag{4.2}$$

where a and b are the constants appearing in the Lyapunov condition. This set is chosen in such a way that

$$\int_{\mathsf{X}\times\mathsf{X}} \tilde{V}(x,y)\,\mathbf{C}_{x_0,y_0}(dx,dy) \le \frac{1+a}{2}\tilde{V}(x_0,y_0)\;, \qquad \forall\,(x_0,y_0) \notin \tilde{K}_0\;. \tag{4.3}$$

At time 0, Υ is located at $(x,y,0)$. If it is located at (x,y,n) and $(x,y) \notin \tilde{K}_0$, then it makes one step according to $\mathbf{C}_{x,y}$ and n is incremented by one:

$$\mathbf{P}_{(x,y,n)} = \mathbf{C}_{x,y} \times \delta_{n+1}\;.$$

If Υ is located at (x,y,n) and $(x,y) \in \tilde{K}_0$, then it has a probability $\|\mathbf{Q}^{\infty}_{x,y}\|$ of jumping to $\star$ and a probability $\|(\mathbf{R}\mathbf{Q}^{m})_{x,y}\|$ of making m steps according to $(\mathbf{R}\mathbf{Q}^{m})_{x,y}$:

$$\mathbf{P}_{(x,y,n)} = \|\mathbf{Q}^{\infty}_{x,y}\|\delta_\star + \sum_{m=0}^{\infty} \pi^*_m(\mathbf{R}\mathbf{Q}^{m})_{x,y} \times \delta_{n+m+1}\;.$$

If Υ is located at $\star$, it remains there:

$$\mathbf{P}_\star = \delta_\star\;.$$

The Markov chain Υ induces a family $\hat{\mathbf{P}}^\infty_{x,y}$ of probability measures on Y^∞. Let $\hat\tau : \mathsf{Y}^\infty \to \mathbf{N} \cup \{\infty\}$ be the function that associates to a sequence of elements in Y the largest value of n that is reached by the sequence ($\hat\tau = 0$ if the sequence is equal to $\star$ repeated). We also define $\hat\kappa : \mathsf{Y}^\infty \to \mathbf{N} \cup \{\infty\}$ as the value of n attained at the first non-vanishing time when the sequence hits the set $\tilde K_0 \times \mathbf{N}$ ($\hat\kappa = \infty$ if this set is never reached). The construction of Υ is very close to the coupling construction of [Mat01].

The crucial observation for the proof of Theorem 4.1 is

Lemma 4.4 *Let Φ be a RDS with state space X satisfying assumptions* **A1** *and* **A3**, *and let Υ be defined as above. Then, there exists a constant C such that*

$$\|\mathsf{P}^n_x - \mathsf{P}^n_y\|_{\mathsf{L}} \le \hat{\mathbf{P}}^\infty_{x,y}(\{\hat\tau \ge n/2\}) + Ce^{-\gamma_1 n/2}\;,$$

for every $(x,y) \in \mathsf{X}^2$ and every $n > 0$.

Proof. Recall the Markov chain $\hat\Psi$ defined in Remark 2.5. We define a function τ_1 on its pathspace by

$$\begin{aligned}\tau_1 : \mathsf{X}^\infty \times \mathsf{X}^\infty \times \{0,1\}^\infty &\to \mathbf{N} \cup \{\infty\}\\ \{(x_i, y_i, b_i)\}_{i=1}^\infty &\mapsto \inf\{n \,|\, (x_n, y_n) \in \tilde K_0 \text{ and } b_i = 1\ \forall i \ge n\}\;.\end{aligned}$$

Combining (2.2) with Assumption **A5** and the definition of τ_1, one sees that

$$\|\mathsf{P}^n_x - \mathsf{P}^n_y\|_{\mathsf{L}} \le \mathbf{C}^\infty_{x,y}(\{\tau_1 \ge n/2\}) + Ce^{-\gamma_1 n/2}\;.$$

From the construction of Υ and the definition of $\hat\tau$, we see furthermore that the probability distributions of τ_1 under $\mathbf{C}^\infty_{x,y}$ and of $\hat\tau$ under $\hat{\mathbf{P}}^\infty_{x,y}$ are the same. □

Proof of Theorem 4.1. It remains to show that $\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\tau \ge n/2\})$ has an exponential tail. The key observation is the following. Let $x_n \in \mathbf{N} \cup \{-\infty\}$ with $n \ge 0$ be a Markov chain defined by

$$x_0 = 0\;,\quad x_{n+1} = \begin{cases} -\infty & \text{with probability } p_\star,\\ x_n + m & \text{with probability } p_m,\end{cases}$$

where $m \ge 1$ and, of course, $p_\star + \sum_{m=1}^\infty p_m = 1$.

Lemma 4.5 *If the p_m have an exponential tail and we define $\tau = \max_n x_n$, then the probability distribution of τ also has an exponential tail.*

Proof. The claim is an easy consequence of Kendall's theorem, but for the sake of completeness, and because the proof is quite elegant, we outline it here. Define the analytic function $p(\zeta) = \sum_{m=1}^{\infty} p_m \zeta^m$ and define q_n as the probability of τ being equal to n. Notice that, because of the exponential tail, p is analytic in a disk of radius strictly bigger than 1 around the origin. A straightforward computation shows that $q_0 = p_\star$ and, for $n > 0$,

$$q_n = p_* \star Bigl(p_n + \sum_{k_1+k_2=n} p_{k_1}p_{k_2} + \sum_{k_1+k_2+k_3=n} p_{k_1}p_{k_2}p_{k_3} + \dots\Bigr) \;,$$

which is equal to the nth Taylor coefficient of the function

$$q(\zeta) = \frac{p_\star}{1 - p(\zeta)} \;.$$

Since $p(1) = 1 - p_\star < 1$, there exists an $\varepsilon > 0$ such that $p(1+\varepsilon) < 1$. Furthermore, since the p_n are all positive, one has the estimate $|p(\zeta)| \le p(|\zeta|)$. Using Cauchy's formula on a circle of radius $1 + \varepsilon$, one gets

$$|q_n| \le \frac{p_\star}{1 - p(1+\varepsilon)} \frac{1}{(1+\varepsilon)^n} \;,$$

which shows the claim. □

Before we prove Theorem 4.1 in full generality, we restrict ourselves to the case when $(x, y) \in \tilde K_0$. It follows from the construction that $\hat\tau$ (seen as a random variable under the distribution induced by $\hat{\mathbf{P}}^\infty_{x,y}$) is dominated by the process x_n constructed above with the tail distribution of the p_m being equal to

$$\tilde p_m = \sup_{(x,y)\in \tilde K_0} \hat{\mathbf{P}}^\infty_{x,y}(\{\hat\kappa = m\}) \;.$$

This means that we define m_* as

$$m_* = \inf\Bigl\{m \Bigm| \sum_{n=m}^{\infty} \tilde p_n \le 1\Bigr\} \;,$$

and then set $p_m = \tilde p_m$ for $m \ge m_*$, $p_m = 0$ for $m < m_* - 1$, and p_{m_*-1} in such a way that the p_n sum up to 1.

Because of Lemma 4.5, it suffices to show that the tail distribution of the $\tilde p_m$ decays exponentially. We thus estimate the quantity $\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\kappa \ge n\})$. To this end, we introduce the function $\tau_\Psi : \mathsf{X}^\infty \times \mathsf{X}^\infty \to \mathbf{N} \cup \{\infty\}$ defined by

$$\tau_\Psi(x, y) = \inf\{n > 0 \mid (x_n, y_n) \in \tilde K_0\} \;.$$

Notice that, in order to have $\hat{\kappa} \geq n$, there are two possibilities. Either the first step of Υ is taken according to $(\mathbf{RQ}^m)_{x,y}$ with some $m \geq n/2$, or the corresponding realization of Ψ stays outside of $\tilde{K}_0$ for a time longer than $n/2$. This yields the estimate

$$\hat{\mathbf{P}}^\infty_{x,y}(\{\hat{\kappa} \geq n\}) \leq \sum_{m=n/2}^\infty \|(\mathbf{RQ}^m)_{x,y}\| + \frac{n}{2} \sup_{(x_0,y_0)\in\tilde{K}_0} \mathbf{C}^\infty_{x_0,y_0}(\{\tau_\Psi \geq n/2\}) ,$$

holding for $(x,y) \in \tilde{K}_0$. The first term has an exponential tail by Proposition 3.5. The second term has also an exponential tail by (4.3) and standard Lyapunov techniques (see *e.g.* [MT94, Thm 15.2.5]). This concludes the proof of Theorem 4.1 for the case $(x,y) \in \tilde{K}_0$.

In order to conclude the proof for the case $(x,y) \notin \tilde{K}_0$, notice that

$$\begin{aligned}\hat{\mathbf{P}}^\infty_{x,y}(\{\hat{\tau} \geq n\}) &\leq \sum_{m=1}^\infty \mathbf{C}^\infty_{x,y}(\{\tau_\Psi = m\}) \sup_{(x_0,y_0)\in\tilde{K}_0} \hat{\mathbf{P}}^\infty_{x_0,y_0}(\{\hat{\tau} \geq n-m\}) \\ &\leq \frac{n}{2} \sup_{(x_0,y_0)\in\tilde{K}_0} \hat{\mathbf{P}}^\infty_{x_0,y_0}(\{\hat{\tau} \geq n/2\}) + \sum_{m=n/2}^\infty \mathbf{C}^\infty_{x,y}(\{\tau_\Psi = m\}) .\end{aligned}$$

The first term is bounded by the construction above. The Lyapunov structure implies that there exists a constant $\gamma > 0$ such that the first hitting time τ_Ψ satisfies $\mathbf{E}_{(x,y)}e^{\gamma\tau_\Psi} = \mathcal{O}(\tilde{V}(x,y))$ for every $(x,y) \in \mathsf{X}^2$ (see again [MT94, Thm. 15.2.5]). This allows to bound the second term and concludes the proof of Theorem 4.1. □

5 Application to Stochastic Differential Equations

In this section, we will see how to apply Theorem 4.1 to the case when the RDS Φ is constructed by sampling the solution of a (possibly infinite-dimensional) stochastic differential equation. We will restrict ourselves to the case where the equation is driven by additive white noise. The case of multiplicative noise requires further estimates, but can also be described by the formalism exposed here.

Consider the equation described by

$$dx(t) = Ax\,dt + F(x)\,dt + Q\,d\omega(t) , \qquad x(0) = x_0 , \tag{5.1}$$

where x belongs to some separable Hilbert space $\mathcal{H}$, ω is the cylindrical Wiener process on some separable Hilbert space $\mathcal{W}$, and A, F and Q satisfy the following assumptions:

B1 *a. The linear operator $A : \mathscr{D}(A) \to \mathcal{H}$ is the generator of a strongly continuous semigroup on $\mathcal{H}$.*

b. *The operator* $e^{At}Q : \mathcal{W} \to \mathcal{H}$ *is Hilbert-Schmidt for every* $t > 0$ *and one has the estimate*

$$\int_0^1 \|e^{At}Q\|_{\mathrm{HS}}^2 \, dt < \infty \; . \tag{5.2}$$

c. *The nonlinear operator* $F : \mathscr{D}(F) \to \mathcal{H}$ *is such that, for every* $x_0 \in \mathcal{H}$*, there exists a unique, continuous stochastic process* $x(t)$ *such that* $x(s) \in \mathscr{D}(F)$ *for* $s > 0$ *and*

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}F(x(s))\, ds + \int_0^t e^{A(t-s)}Q \, d\omega(s) \; , \tag{5.3}$$

for every $t > 0$.

Remark 5.1 This assumptions simply states that there exists a unique weak solution to (5.1). Notice that we do *not* make any assumptions on the tightness of the transition probabilities for (5.1). As a consequence, existence and uniqueness results for invariant measures can in principle be deduced from Theorem 5.5 below even in cases where the semigroup e^{At} is not compact.

In order to recover the formalism used in Section 2, we follow [DPZ92b] and introduce an auxiliary Hilbert space $\hat{\mathcal{W}}$ such that there exists a continuous embedding $\iota : \mathcal{W} \hookrightarrow \hat{\mathcal{W}}$, which is Hilbert-Schmidt. We can now set $\Omega = \mathcal{C}_0([0,1], \hat{\mathcal{W}})$, the space of continuous $\hat{\mathcal{W}}$-valued functions that vanish at 0, and define P as the Wiener measure on $\hat{\mathcal{W}}$ with covariance operator $\iota\iota^*$.

We define $\Phi : \mathcal{H} \times \Omega \to \mathcal{H}$ as the map that solves (5.3) up to time 1 given an initial condition and a realization of the noise. This map is defined P-almost everywhere on Ω. We also denote by $\Phi_t : \mathcal{H} \times \Omega^\infty \to \mathcal{H}$ the map that maps an initial condition and a realization of the noise onto the solution of (5.3) after a time t.

Our next assumption is the existence of an appropriate Lyapunov function V:

B2 *There exists a measurable function* $V : \mathcal{H} \to [0, \infty]$ *and constants* $a < 1$ *and* $b, c, d > 0$ *such that*

$$\begin{aligned} \mathbf{E}_\omega V(\Phi(x,\omega)) &\le aV(x) + b \; , \\ \mathbf{E}_\omega \Big(\sup_{0 \le t \le 1} V(\Phi_t(x,\omega)) \Big) &\le cV(x) + d \; , \\ \mathsf{P}(\{\omega \mid V(\Phi(x,\omega)) = \infty\}) &= 0 \; , \end{aligned} \tag{5.4}$$

for every $x \in \mathcal{H}$*. Furthermore,* V *dominates the norm in* $\mathcal{H}$ *in the sense that* $\|x\| \le C(1 + V(x))$ *for some constant* C*.*

As is Section 3, we define $\tilde{V}(x,y) = V(x) + V(y)$.

Remark 5.2 Take $\mathcal{H}$ equal to $\mathrm{L}^2(\mathcal{O})$ for some regular bounded domain $\mathcal{O} \subset \mathbf{R}^d$, A a second-order elliptic differential operator on $\mathcal{O}$ with sufficiently smooth coefficients, and F any polynomial non-linearity of odd degree having the correct sign. The assumptions **B1** and **B2** are satisfied with $V(x) = \|x\|_\star^p$ for every power $p \geq 1$ and every "reasonable" norm $\|\cdot\|_\star$, as long as Q is "small" enough. (One can for example take for $\|\cdot\|_\star$ the L^∞ norm or a Sobolev norm.)

We now turn to the binding construction for the problem (5.1). Take a function $G : \mathcal{H}^2 \to \mathcal{W}$ and consider the $\mathcal{H}^2$-valued process (x,y) solving

$$dx(t) = Ax\,dt + F(x)\,dt + Q\,d\omega(t)\;, \tag{5.5a}$$

$$dy(t) = Ay\,dt + F(y)\,dt + Q\,G(x,y)\,dt + Q\,d\omega(t)\;. \tag{5.5b}$$

Notice that the realization of ω is the same for both components. The process (5.5) yields our binding construction for (5.1). In order to give sense to (5.5b), we introduce the $\mathcal{H}$-valued process $\varrho(t) = y(t) - x(t)$ and we define it pathwise as the solution of the ordinary differential equation

$$\dot{\varrho} = A\varrho + F(x+\varrho) - F(x) + Q\,G(x, x+\varrho)\;. \tag{5.6}$$

We assume that G is sufficiently regular to ensure the existence and uniqueness of global weak solutions to (5.6) for almost every (with respect to the measure on pathspace induced by Φ_t) continuous function $x : [0,\infty) \to \mathcal{H}$. This allows us to *define* the stochastic process $y(t)$ by $y(t) = x(t) + \varrho(t)$. We will denote by $\vec{\Phi}_t : \mathsf{X} \times \mathsf{X} \times \Omega \to \mathsf{X}$ the map that solves (5.5b) up to time t, given an initial condition for x and y, and a realization of the noise.

The above construction is invertible in the following sense. Consider the $\mathcal{H}^2$-valued process

$$d\tilde{x}(t) = A\tilde{x}\,dt + F(\tilde{x})\,dt - Q\,G(\tilde{x},\tilde{y})\,dt + Q\,d\tilde{\omega}(t)\;, \tag{5.7a}$$

$$d\tilde{y}(t) = A\tilde{y}\,dt + F(\tilde{y})\,dt + Q\,d\tilde{\omega}(t)\;, \tag{5.7b}$$

where we give sense to the equation for $\tilde{x}$ as above by setting $\tilde{\varrho} = \tilde{y} - \tilde{x}$ and solving

$$\dot{\tilde{\varrho}} = A\tilde{\varrho} + F(\tilde{y}) - F(\tilde{y} - \tilde{\varrho}) + Q\,G(\tilde{y} - \tilde{\varrho}, \tilde{y})\;.$$

We denote by $\overleftarrow{\Phi}_t : \mathsf{X} \times \mathsf{X} \times \Omega \to \mathsf{X}$ the map that solves (5.7a) up to time t, given an initial condition for $\tilde{x}$ and $\tilde{y}$, and a realization of the noise $\tilde{\omega} \in \Omega$. We see that (5.7) can be obtained from (5.5) by the substitution $d\tilde{\omega} = d\omega + G(x,y)\,dt$ and a renaming of the variables. This observation yields the invertibility of the maps $\psi_{x\to y}$ defined in Eq. (5.12) below.

We will state two more assumptions to make sure that the conclusions of Theorem 4.1 hold. First, we want G to become small as x and y become close.

B3 *There exists a constant $C > 0$ and exponents $\alpha, \beta > 0$ such that*

$$\|G(x,y)\|^2 \le C\|x-y\|^\alpha (1+\tilde V(x,y))^\beta\;, \tag{5.8}$$

for every $x,y \in \mathcal{H}$.

The last assumption ensures that the process $y(t)$ converges towards $x(t)$ for large times.

B4 *There exist positive constants C and γ such that the solutions of (5.5) and (5.7) satisfy*

$$\|\Phi_t(x,\omega) - \vec{\Phi}_t(x,y,\omega)\| \le Ce^{-\gamma t}\Bigl(1 + V(y) + \sup_{s\le t} V(\Phi_s(x,\omega))\Bigr)\;, \tag{5.9a}$$

$$\|\overleftarrow{\Phi}_t(x,y,\omega) - \Phi_t(y,\omega)\| \le Ce^{-\gamma t}\Bigl(1 + V(x) + \sup_{s\le t} V(\Phi_s(y,\omega))\Bigr)\;, \tag{5.9b}$$

for P^∞-almost every $\omega \in \Omega^\infty$. Furthermore, there exists $\delta > 0$ such that one has the estimate

$$V(\vec{\Phi}_t(x,y,\omega)) \le C\Bigl(1 + V(y) + \sup_{s\le t} V(\Phi_s(x,\omega))\Bigr)^\delta\;, \tag{5.10a}$$

$$V(\overleftarrow{\Phi}_t(x,y,\omega)) \le C\Bigl(1 + V(x) + \sup_{s\le t} V(\Phi_s(y,\omega))\Bigr)^\delta\;, \tag{5.10b}$$

for P^∞-almost every $\omega \in \Omega^\infty$ and every $t \ge 0$.

Remark 5.3 One important particular case is the choice $V(x) = \|x\|^p$, where the power p is chosen in such a way that (5.9) is satisfied. Notice that in this case, the estimates (5.10) are a straightforward consequence of (5.9).

The function G is then only required to satisfy a bound of the type

$$\|G(x,y)\|^2 \le C\|x-y\|^\alpha (1 + \|x\| + \|y\|)^q\;,$$

with α and q some arbitrary positive exponents.

It is also possible to choose $V(x) = \|x\|_\star^p$, with $\|\cdot\|_\star$ the norm of some Banach space $\mathcal{B} \subset \mathcal{H}$. In this case, (5.9) with the $\mathcal{B}$-norm replacing the $\mathcal{H}$-norm in the left-hand side implies (5.10).

Remark 5.4 An equivalent way of writing (5.9b) is

$$\|\Phi_t(x,\omega) - \vec{\Phi}_t(x,y,\omega)\| \le Ce^{-\gamma t}\Big(1 + V(x) + \sup_{s\le t} V(\vec{\Phi}_s(x,y,\omega))\Big)\;. \tag{5.11}$$

The equation (5.11) will be more natural in our examples, but (5.9) is more symmetric and more convenient for the proof of Theorem 5.5 below.

All these assumptions together ensure that exponential mixing takes place:

Theorem 5.5 *Let A, F and Q be such that assumptions* **B1** *and* **B2** *are satisfied. If there exists a function $G : \mathcal{H}^2 \to \mathcal{W}$ such that assumptions* **B3** *and* **B4** *hold, then the solution of (5.1) possesses a unique invariant measure μ_* and there exist constants $C, \gamma > 0$ such that*

$$\|\mathbf{P}^n_x - \mu_*\|_{\mathrm{L}} \le Ce^{-\gamma n}(1 + V(x))\;.$$

Proof. It suffices to show that assumptions **A1**–**A5** hold. Assumption **A1** follows immediately from Assumption **B2**. In order to check the other assumptions, we define the various objects appearing in the previous sections. We have already seen that $\mathsf{X} = \mathcal{H}$, $\Omega = \mathcal{C}_0([0,1], \hat{\mathcal{W}})$, and Φ is the solution of (5.1) at time 1.

We define the function $W : \mathcal{H} \times \Omega \to [1, \infty]$ by

$$W(x,\omega) = \sup_{t\in[0,1]} V(\Phi_t(x,\omega))\;.$$

The estimate (5.4) and the definition ensure that W satisfies (3.3a) and (3.3b). The bound (5.10) ensures that Assumption **A2** is also satisfied.

It remains to define the binding functions $\psi_{x\to y}$ and to compute the densities $\mathcal{D}^n_{x,y}$. According to the constructions (5.5) and (5.7), we define for $(x,y) \in \mathcal{H}^2$ the binding functions

$$(\psi_{x\to y}(\omega))(t) = \omega(t) + \int_0^t G(\Phi_s(x,\omega), \vec{\Phi}_s(x,y,\omega))\,ds\;, \tag{5.12a}$$

$$(\psi_{x\leftarrow y}(\omega))(t) = \omega(t) - \int_0^t G(\overleftarrow{\Phi}_s(x,y,\omega), \Phi_s(y,\omega))\,ds\;, \tag{5.12b}$$

with $t \in [0,1]$. It follows from the construction that these functions are each other's inverse. Furthermore, if we identify Ω^n with $\mathcal{C}_0([0,n], \hat{\mathcal{W}})$ in a natural way, we see that the maps $\Xi^n_{x,y}$ introduced in (2.10) are obtained from (5.12) by simply letting t take values in $[0,n]$. These observations allow us to compute the densities $\mathcal{D}^n_{x,y}(\omega)$ by Girsanov's theorem:

Lemma 5.6 *The family of densities $\mathcal{D}^n_{x,y}(\omega)$ is given by*

$$\mathcal{D}^n_{x,y}(\omega) = \exp\Big(\int_0^n G(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\, d\omega(t) - \frac{1}{2}\int_0^n \|G(\ldots)\|^2\, dt\Big) ,$$

where the arguments of G in the second term are the same as in the first term.

Proof. If we can show that Girsanov's theorem applies to our situation, then it precisely states that

$$\psi^*_{x\leftarrow y} \hat{\mathsf{P}}^n = \mathsf{P}^n ,$$

with $\hat{\mathsf{P}}^n(d\omega) = \mathcal{D}^n_{x,y}(\omega)\, \mathsf{P}^n(d\omega)$, and $\mathcal{D}^n_{x,y}(\omega)$ defined as above. Applying $\psi^*_{x\to y}$ to both sides of the equality shows the result.

We now show that Girsanov's theorem can indeed be applied. By [DPZ92b, Thm 10.14], it suffices to verify that

$$\int_{\Omega^n} \mathcal{D}^n_{x,y}(\omega)\, \mathsf{P}^n(d\omega) = 1 . \tag{5.13}$$

This can be achieved by a suitable cut-off procedure. Define for $N > 0$ the function

$$G_N(x,y) = \begin{cases} G(x,y) & \text{if } \|G(x,y)\| \le N, \\ 0 & \text{otherwise,} \end{cases}$$

and define

$$\mathcal{D}^{n,N}_{x,y}(\omega) = \exp\Big(\int_0^n G_N(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\, d\omega(t) - \frac{1}{2}\int_0^n \|G_N(\ldots)\|^2\, dt\Big) .$$

It is immediate that (5.13) holds for $\mathcal{D}^{n,N}_{x,y}$. Furthermore, it follows from Assumption **B4** that there exists a constant C_N such that $\mathcal{D}^{n,N}_{x,y}(\omega) = \mathcal{D}^n_{x,y}(\omega)$ on the set

$$\Gamma_N = \{\omega \in \mathcal{P}^n \mid \tilde{V}(\overleftarrow{\Phi}_s(x,y,\omega), \Phi_s(y,\omega)) < C_N \quad \forall s \in [0,n]\} .$$

The sets Γ_N satisfy $\lim_{N\to\infty} \mathsf{P}^n(\Gamma_N) = 1$ by (5.4) and (5.10b). This shows that (5.13) holds. Notice that the *a-priori* bounds of Assumption **B4** were crucial in this step in order to apply Girsanov's theorem. The bound (5.8) alone could lead to exploding solutions for which Girsanov's theorem does not apply. □

It is immediate that Assumption **A3** follows from Assumption **B4** and the definition of the norm $\|\cdot\|_{\mathrm{L}}$.

We now turn to the verification of Assumption **A4**. Recalling the definition (3.5), we see that in our case

$$A_{y,k} \subset B_{y,k} \equiv \{\omega \in \Omega^\infty \mid V(\Phi_s(y,\omega)) \le k(V(y) + s^2) \quad \forall s \ge 0\} .$$

As we see from the definition of $B_{y,k}$, a natural definition for a truncation $G_{y,k}$ of G (this time the truncation additionally depends on time) is

$$G_{y,k}(\tilde{x},\tilde{y},t) = \begin{cases} G(\tilde{x},\tilde{y}) & \text{if } V(\tilde{y}) \le k(V(y)+t^2), \\ 0 & \text{otherwise.} \end{cases}$$

As above, we define

$$\mathcal{D}^{n,k}_{x,y}(\omega) = \exp\Bigl(\int_0^n G_{y,k}(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega),t)\,d\omega(t) - \frac{1}{2}\int_0^n \|G_{y,k}(\ldots)\|^2\,dt\Bigr)\;.$$

By definition, $\mathcal{D}^{n,k}_{x,y}(\omega) = \mathcal{D}^{n}_{x,y}(\omega)$ for every $\omega \in B_{y,k}$. Setting $\xi = \delta(\alpha+\beta)$, we thus have the estimate

$$\begin{aligned}
\int_{A_{y,k}} &\bigl(\mathcal{D}^{n}_{x,y}(\omega)\bigr)^{-2}\,\mathsf{P}^n(d\omega) \le \int_{B_{y,k}} \bigl(\mathcal{D}^{n,k}_{x,y}(\omega)\bigr)^{-2}\,\mathsf{P}^n(d\omega) \\
&\le \Bigl(\int_{B_{y,k}} \exp\Bigl(10\int_0^n \|G_{y,k}(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega),t)\|^2\,dt\Bigr)\,\mathsf{P}^n(d\omega)\Bigr)^{1/2} \\
&\le \Bigl(\int_{B_{y,k}} \exp\Bigl(10\int_0^n Ce^{-\gamma t}\Bigl(1+V(x)+\sup_{s\le t} V(\Phi_s(y,\omega))\Bigr)^{\xi}\,dt\Bigr)\,\mathsf{P}^n(d\omega)\Bigr)^{1/2} \\
&\le \exp\Bigl(C\int_0^n e^{-\gamma t}(1+k\tilde{V}(x,y)+kt^2)^{\xi}\,dt\Bigr)\;.
\end{aligned}$$

In this expression, we used the Cauchy-Schwarz inequality to go from the first to the second line, and we used assumptions **B3** and **B4** to go from the second to the third line. Since the integral converges for $n \to \infty$, the bound is uniform in n and Assumption **A4** is verified.

The verification of Assumption **A5** is quite similar. Fix some positive constant $K > 0$ and use again the cutoff function

$$G_N(\tilde{x},\tilde{y}) = \begin{cases} G(\tilde{x},\tilde{y}) & \text{if } \|G(\tilde{x},\tilde{y})\|^2 \le N, \\ 0 & \text{otherwise.} \end{cases}$$

The precise value of N (as a function of K) will be fixed later. We also fix a pair $(x_0,y_0) \in \mathcal{H}^2$ with $\tilde{V}(x_0,y_0) \le K$, a value $n > 0$, and initial conditions $(x,y) \in Q^n_K(x_0,y_0)$. By the definition of $Q^n_K(x_0,y_0)$, there exists an element $\tilde{\omega} \in \Omega^n$ such that

$$(x,y) = (\overleftarrow{\Phi}_n(x_0,y_0,\tilde{\omega}),\Phi_n(y_0,\tilde{\omega}))\;, \tag{5.14}$$

and such that

$$\sup_{s\in[0,n]} \tilde{V}(\overleftarrow{\Phi}_s(x_0,y_0,\tilde{\omega}),\Phi_s(y_0,\tilde{\omega})) \le K\;. \tag{5.15}$$

Following the statement of Assumption **A5**, we define the set

$$B_{x,y}^K = \Bigl\{\omega \in \Omega \Bigm| \sup_{t\in[0,1]} \tilde V(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega)) \le K\Bigr\} ,$$

which is equal in our setup to the set over which integration is performed in (3.9). Being now accustomed to these truncation procedures, we define again

$$\mathcal{D}_{x,y}^{(K)}(\omega) = \exp\Bigl(\int_0^1 G_N(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\, d\omega(t) - \frac{1}{2}\int_0^1 \|G_N(\ldots)\|^2\, dt\Bigr) .$$

By (5.14) and the cocycle property, we can write the integral in the above expression as

$$\int_0^1 G_N(\overleftarrow{\Phi}_{n+t}(x_0,y_0,\tilde\omega\omega), \Phi_{n+t}(y_0,\tilde\omega\omega))\, d\omega(t) ,$$

where $\tilde\omega\omega$ is the realization of the noise which is equal to $\tilde\omega$ for a time n and then to ω for a time 1. Using the *a-priori* bound (5.15) as well as assumptions **B3** and **B4**, we thus see that there exists a constant C such that the choice

$$N = Ce^{-\alpha\gamma n}(1+K)^{\alpha+\beta} ,$$

ensures that $\mathcal{D}_{x,y}^{(K)}(\omega)$ is equal to $\mathcal{D}_{x,y}(\omega)$ for $\omega \in B_{x,y}^K$.

We then have the estimate

$$\begin{aligned}\int_{B_{x,y}^K} (1-\mathcal{D}_{x,y}(\omega))^2\,\mathsf{P}(d\omega) &\le \int_\Omega (1-\mathcal{D}_{x,y}^{(K)}(\omega))^2\,\mathsf{P}(d\omega)\\ &= \int_\Omega (\mathcal{D}_{x,y}^{(K)}(\omega))^2\,\mathsf{P}(d\omega) - 1\\ &\le \Bigl(\int_\Omega \exp\Bigl(6\int_0^1 \|G_N(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\|^2\,dt\Bigr)\,\mathsf{P}(d\omega)\Bigr)^{1/2} - 1\\ &\le \exp(Ce^{-\alpha\gamma n}(1+K)^{\alpha+\beta}) - 1 .\end{aligned}$$

If we take $n \ge \beta \ln(1+K)/\gamma$, the exponent is bounded by C and there exists a constant C' such that

$$\int_{B_{x,y}^K} (1-\mathcal{D}_{x,y}(\omega))^2\,\mathsf{P}(d\omega) \le C'e^{-\alpha\gamma n}(1+K)^{\alpha+\beta} ,$$

thus validating Assumption **A5** with $\gamma_2 = \alpha\gamma$ and $\zeta = \alpha+\beta$.

The proof of Theorem 5.5 is complete. □

6 Examples

Numerous recent results show that the invariant measure for the 2D Navier-Stokes equation (and also for other dissipative PDEs) is unique if a sufficient number of low-frequency modes are forced by the noise [BKL00a, BKL00b, EMS01, Mat01, EL01, KS00, KS01, MY01]. These results are not covered directly by Theorem 5.5, but some more work is needed. The reason is that the sets A_x^k defined in (3.5) are not the natural sets that allow to control the influence of the low-frequency modes onto the high-frequency modes in the 2D Navier-Stokes equation.

On the other hand, our formulation of Theorem 5.5 makes it quite easy to verify that the n-dimensional Ginzburg-Landau equation (in a bounded domain) shows exponential mixing properties, if sufficiently many low-frequency modes are forced by the noise. We verify this in the following subsection.

6.1 The Ginzburg-Landau equation

We consider the SPDE given by

$$du = (\Delta u + u - u^3)\,dt + Q\,dw(t)\;, \qquad u(0) = u_0\;, \tag{6.1}$$

where the function u belongs to the Hilbert space

$$\mathcal{H} = \mathrm{L}^2([-L,L]^n, \mathbf{R})\;,$$

and Δ denotes the Laplacean with periodic boundary conditions. The symbol $Q\,d\omega(t)$ stands as a shorthand for

$$Q\,d\omega(t) \equiv \sum_{i=1}^{N} q_i e_i\,d\omega_i(t)\;,$$

where $\{q_i\}_{i=1}^N$ is a collection of strictly positive numbers, e_i denotes the ith eigenvector of the Laplacean, and the ω_i are N independent Brownian motions (for some finite integer N). We also denote by λ_i the eigenvalue of Δ corresponding to e_i and we assume that they are ordered by $\ldots \le \lambda_2 \le \lambda_1 \le 0$. We will see that it is fairly easy to construct a binding function G for which the assumptions of the previous section hold with $V(u) = \|u\|$, where $\|\cdot\|$ denotes the norm of $\mathcal{H}$.

In [DPZ96], it is shown that (6.1) possesses a unique mild solution for initial conditions $u_0 \in \mathrm{L}^\infty([-L,L]^n)$. It is straightforward to extend this to every initial condition $u_0 \in \mathcal{H}$, by using the regularizing properties of the heat semigroup. Thus, Assumption **B1** holds and we denote by P_u^t the transition probabilities of the solution at time t starting from u. We have the following result:

Theorem 6.1 *There exist positive constants C and γ, and a unique measure $\mu_* \in \mathscr{M}_1(\mathcal{H})$ such that*

$$\|\mathsf{P}_u^t - \mu_*\|_{\mathrm{L}} \le C e^{-\gamma t}(1 + \|u\|) , \tag{6.2}$$

for every $u \in \mathcal{H}$ and every $t > 0$.

Proof. We verify that the assumptions of Theorem 5.5 hold. The bounds required for the verification of Assumption **B2** can be found in [Cer99, DPZ96], for example.

It remains to construct the forcing $G : \mathcal{H}^2 \to \mathbf{R}^N$ and to verify assumptions **B3** and **B4**. We consider two copies u_1 and u_2 of (6.1), with the noise $d\omega$ replaced by $d\omega + G\,dt$ in the second copy. We also denote by $\varrho = u_2 - u_1$ the difference process. It satisfies the differential equation

$$\dot\varrho = \Delta\varrho + \varrho - \varrho(u_1^2 + u_1u_2 + u_2^2) \; + Q\,G(u_1,u_2) . \tag{6.3}$$

We can project (6.3) onto the direction given by e_k. This yields

$$\dot\varrho_k = (\lambda_k + 1)\varrho_k - \Big(\varrho(u_1^2 + u_1u_2 + u_2^2)\Big)_k + q_k G_k(u_1,u_2) ,$$

for $k = 1, \ldots, N$ and

$$\dot\varrho_k = (\lambda_k + 1)\varrho_k - \Big(\varrho(u_1^2 + u_1u_2 + u_2^2)\Big)_k ,$$

for $k > N$. We choose G_k for $k = 1, \ldots, N$ as

$$G_k(u_1,u_2) = -\frac{2+\lambda_k}{q_k}\varrho_k .$$

Since G_k can only be defined this way if $q_k \neq 0$, we use at this point the fact that the noise acts directly and independently on *every unstable mode*. This requirement can be significantly weakened with the help of Theorem 5.5. We will focus next on more degenerate problems which illustrate the power of our technique.

This choice satisfies Assumption **B3**. With this choice, we can write down the evolution of the norm of ϱ as

$$\frac{d\|\varrho\|^2}{dt} = 2\langle\varrho, A\varrho\rangle - 2\langle\varrho, \varrho(u_1^2 + u_1u_2 + u_2^2)\rangle ,$$

with A the linear operator given by adding up the contribution of $\Delta + 1$ and the contribution of G. By the condition we imposed on N, there exists a constant $a > 0$ such that $\langle\varrho, A\varrho\rangle \le -a\|\varrho\|^2$. Furthermore, one has

$$\langle\varrho, \varrho(u_1^2 + u_1u_2 + u_2^2)\rangle \ge 0 .$$

We thus have the differential inequality

$$\frac{d\|\varrho\|^2}{dt} \le -2a\|\varrho\|^2 \;,$$

which implies that

$$\|\varrho(t)\| \le e^{-at}\|\varrho(0)\| \;.$$

This implies by Remark 5.3 that Assumption **B4** is also satisfied. The proof of Theorem 6.1 is complete. □

6.2 A reaction-diffusion system

Consider the following reaction-diffusion system:

$$\begin{aligned} du &= (\Delta u + 2u + v - u^3)\,dt + dw(t)\;,\\ dv &= (\Delta v + 2v + u - v^3)\,dt\;, \end{aligned} \tag{6.4}$$

where the pair (u, v) belongs to the Hilbert space

$$\mathcal{H} = \mathcal{H}_u \oplus \mathcal{H}_v = \mathrm{L}^2([-L, L], \mathbf{R}) \oplus \mathrm{L}^2([-L, L], \mathbf{R})\;.$$

The symbol Δ again denotes the Laplacean with periodic boundary conditions, and $d\omega$ is the cylindrical Wiener process on $\mathcal{H}_u$ (meaning that it is space-time white noise).

Notice that, because of the presence of v, this system does not satisfy the assumptions stated in the papers mentioned at the beginning of this section. In other words, even though the forcing is infinite-dimensional, not all the determining modes for (6.4) are forced.

We take as our Lyapunov function

$$V(u, v) = \|u\|_\infty + \|v\|_\infty\;,$$

with $\|\cdot\|_\infty$ the L^∞ norm. As in the previous subsection, one can show that with this choice of V, our problem satisfies assumptions **B1** and **B2**. We will now construct a binding function G which satisfies assumptions **B4** and **B3**. We consider, as in (5.5), two copies (u_1, v_1) and (u_2, v_2) of the system (6.4), but the noise is modified by G on the second copy. We also define $\varrho_u = u_2 - u_1$ and $\varrho_v = v_2 - v_1$. We then have

$$\begin{aligned} \dot\varrho_u &= \Delta\varrho_u + 2\varrho_u + \varrho_v - \varrho_u(u_1^2 + u_1u_2 + u_2^2) + G(u_1, u_2, v_1, v_2)\;,\\ \dot\varrho_v &= \Delta\varrho_v + 2\varrho_v + \varrho_u - \varrho_v(v_1^2 + v_1v_2 + v_2^2)\;. \end{aligned} \tag{6.5}$$

Our construction of G is inspired from the construction we presented in Section 1.1. We introduce the variable $\zeta = \varrho_u + 3\varrho_v$. Substituting this in (6.5), it defines the function G if we impose that the equation for ζ becomes

$$\dot{\zeta} = \Delta\zeta - \zeta\;, \tag{6.6}$$

so that $\|\zeta(t)\|^2 \le \|\zeta(0)\|^2 e^{-t}$. Notice that the function G achieving this identity satisfies a bound of the type

$$\|G\| \le C(\|\varrho_u\| + \|\varrho_v\|)(1 + \|u_1\|_\infty + \|u_2\|_\infty + \|v_1\|_\infty + \|v_2\|_\infty)^2\;,$$

thus satisfying Assumption **B3**. It remains to show that Assumption **B4** is satisfied. The equation for ϱ_v reads

$$\dot{\varrho}_v = \Delta\varrho_v - \varrho_v + \zeta - \varrho_v(v_1^2 + v_1 v_2 + v_2^2)\;.$$

Therefore, the norm of ϱ_v satisfies

$$\|\varrho_v(t)\|^2 \le \|\varrho_v(0)\|^2 e^{-t} + \frac{1 + \|\zeta(0)\|^2}{2} e^{-t}\;.$$

This in turn implies, through the definition of ζ and the bound on $\|\zeta(t)\|$, that a similar bound holds for $\|\varrho_u(t)\|$. This shows that the bound (5.9) is satisfied. Similar estimates hold with the L^∞ norm replacing the L^2 norm, and so Assumption **B4** is satisfied by Remark 5.3.

In fact, a straightforward computation, which can be found in [Cer99, Hai01, MS01] for example, shows that in this example, one can get a uniform estimate on the Lyapunov function V. More precisely, there exists a constant C such that for all initial conditions $x \in \mathcal{H}$,

$$\int_{\mathcal{H}} V(y)\,\mathsf{P}_x(dy) \le C\;. \tag{6.7}$$

Denoting by $\mathcal{P}_t^*$ the semigroup acting on measures generated by the solutions of (6.4), we thus have:

Theorem 6.2 *There exists a unique probability measure $\mu_* \in \mathscr{M}_1(\mathcal{H})$ such that $\mathcal{P}_t^*\mu_* = \mu_*$ for every $t \ge 0$. Furthermore, there exist constants C and γ such that*

$$\|\mathcal{P}_t^*\nu - \mu_*\|_{\mathrm{L}} \le Ce^{-\gamma t}\;, \tag{6.8}$$

for every $\nu \in \mathscr{M}_1(\mathcal{H})$.

Proof. Combining (6.7) with the results of Theorem 4.1 and a computation similar to what was done in the proof of Corollary 4.3, we get (6.8) for integer times. The generalization to arbitrary times is straightforward, using the fact that the growth rate of the difference process (ϱ_u, ϱ_v) (with $G \equiv 0$) can easily be controlled. □

Remark 6.3 In fact, the dependence on u in the right-hand side of (6.2) can be removed similarly by checking that an estimate of the type (6.7) is verified for the solutions of the stochastic Ginzburg-Landau equation (6.1).

6.3 A chain with nearest-neighbour interactions

In the previous example, the noise acted on infinitely many degrees of freedom in a non-degenerate way. As a consequence, one step was sufficient to transmit the noise to the entire system. We will now look at a much more degenerate system, where the noise acts on only *one* degree of freedom, although an *arbitrary* number of modes are linearly unstable.

Our model is given by

$$\begin{aligned} dx_0 &= (a^2 x_0 + x_1 - x_0^3)\,dt + d\omega\;, \\ \dot{x}_k &= (a^2 - k^2)x_k + x_{k-1} + x_{k+1} - x_k^3\;, \qquad k = 1, 2, \ldots\;, \end{aligned} \tag{6.9}$$

where $a \in \mathbf{R}$ is an arbitrary constant. One should think of the deterministic part of (6.9) as a very simple model for a dissipative PDE of the Ginzburg-Landau type. We will consider (6.9) in the (real) Hilbert space $\mathcal{H} = \ell^2$ endowed with its canonical orthonormal basis $\{e_k\}_{k=0}^{\infty}$. It is easy to verify that (6.9) possesses a unique solution. We denote again by $\mathcal{P}_t^*$ the semigroup acting on measures $\nu \in \mathscr{M}(\ell^2)$ generated by (6.9). We will show

Theorem 6.4 *For the problem (6.9), there exists a unique probability measure $\mu_* \in \mathscr{M}_1(\ell^2)$ such that $\mathcal{P}_t^* \mu_* = \mu_*$ for every $t \geq 0$. Furthermore, there exist constants C and γ such that*

$$\|\mathcal{P}_t^* \nu - \mu_*\|_{\mathrm{L}} \leq C e^{-\gamma t}\;,$$

for every $\nu \in \mathscr{M}_1(\mathcal{H})$.

Proof. We will take as our Lyapunov function $V(x) = \|x\|^p$ for some power of p to be fixed later. It is a straightforward task to verify that the dynamics generated by (6.9) does indeed satisfy assumptions **B1** and **B2** for this choice of V.

We next show that a bound of the type (6.7) holds for the solutions of (6.9), thus yielding the uniformity in the convergence towards the invariant measure μ_*.

Let us define the process $y(t) \in \ell^2$ by $y(t) = x(t) - \omega(t)e_0$. This process then satisfies the following system of differential equations:

$$\begin{aligned}
\dot y_0 &= a^2(y_0+\omega) + y_1 - (y_0+\omega)^3\;,\\
\dot y_1 &= (a^2-1)y_1 + y_0 + y_2 - y_1^3 + \omega\;,\\
\dot y_k &= (a^2-k^2)y_k + y_{k-1} + y_{k+1} - y_k^3\;, \qquad k = 2,3,\ldots\;.
\end{aligned} \tag{6.10}$$

We denote by $\|y\|_\infty$ the norm given by $\sup_k |y_k|$. It follows from [Lun95] that (6.10) possesses a strong solution for positive times. Furthermore, from (6.10) and the definition of the $\|\cdot\|_\infty$-norm, we see that there are constants $c_1, c_2 > 0$ such that

$$\frac{D_-\|y\|_\infty}{Dt} \le -c_1\|y\|_\infty^3 + c_2(1+|\omega(t)|^3)\;, \tag{6.11}$$

where D_-/Dt denotes the left-handed lower Dini derivative. A straightforward computation shows that (6.11) implies that there exists a constant C such such that

$$\|y(1/2)\|_\infty \le C \sup_{t\in[0,1/2]} (1+|\omega(t)|)\;,$$

independently of the initial condition. In order to conclude the proof of the estimate (6.7), it suffices to show that there exists a constant C such that

$$\mathbf{E}\|y(1/2)\| \le C(1+\|y(0)\|_\infty)\;.$$

This follows easily from the dissipativity of the nonlinearity in $\mathcal{H}$ and the fact that the semigroup generated by the linear part of (6.10) is bounded from ℓ^∞ into ℓ^2.

It remains to verify that the assumptions **B1**–**B4** are indeed satisfied for some binding function G. This, together with the uniform bound obtained above, shows that the conclusions of Theorem 6.4 hold. As for the toy model presented in Section 1.1, we consider a process $y \in \ell^2$ governed by the same equation as (6.9), but with $d\omega$ replaced by $d\omega + G(x,y)\,dt$. We then introduce the difference process $\varrho = y - x$, which is given by the solution of

$$\begin{aligned}
\dot\varrho_0 &= a^2\varrho_0 + \varrho_1 - \varrho_0(x_0^2 + x_0y_0 + y_0^2) + G(x,y)\;,\\
\dot\varrho_k &= (a^2-k^2)\varrho_k + \varrho_{k+1} + \varrho_{k-1} - \varrho_k(x_k^2 + x_ky_k + y_k^2)\;.
\end{aligned} \tag{6.12}$$

The aim of the game is to find a function G for which $\varrho(t) \to 0$ as $t \to \infty$. We can split (6.12) into "low modes" and "high modes" by introducing

$$k_* = \inf\{k > 0 \mid k^2 - a^2 \ge 3\}\;.$$

At the level of the Hilbert space ℓ^2, we set $\ell^2 = \mathcal{H}_L \oplus \mathcal{H}_H$, where $\mathcal{H}_L \approx \mathbf{R}^{k_*}$ is generated by $e_0, \ldots, e_{k_*-1}$ and $\mathcal{H}_H$ is its orthogonal complement. We denote by ϱ_L and ϱ_H the components of ϱ and by A_H the restriction (as a symmetric quadratic form) of the linear part of (6.9) to $\mathcal{H}_H$. It is by construction easy to see that

$$\langle \varrho_H, A_H \varrho_H \rangle \geq \|\varrho_H\|^2 \;.$$

As a consequence, we have for $\|\varrho_H\|^2$ the following estimate:

$$\|\varrho_H(t)\|^2 \leq e^{-t}\|\varrho_H(0)\|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_1(s)|^2\,ds \;, \tag{6.13}$$

where we defined $\zeta_1 = \varrho_{k_*-1}$. (The reason for renaming ϱ_{k_*-1} this way will become clear immediately.) It remains to construct G in such a way to get good estimates on $\|\varrho_L(t)\|^2$. In order to achieve this, we use again the same method as for the first toy model. The variable ζ_1 obeys the equation

$$\dot{\zeta}_1 = c_1\varrho_{k_*-1} + \varrho_{k_*} + \varrho_{k_*-2} - \varrho_{k_*-1}(x_{k_*-1}^2 + x_{k_*-1}y_{k_*-1} + y_{k_*-1}^2) \;,$$

with some constant $c_1 \in \mathbf{R}$. We thus introduce a new variable ζ_2 defined by

$$\zeta_2 = (c_1+1)\varrho_{k_*-1} + \varrho_{k_*} + \varrho_{k_*-2} - \varrho_{k_*-1}(x_{k_*-1}^2 + x_{k_*-1}y_{k_*-1} + y_{k_*-1}^2) \;.$$

It is important to notice two facts about this definition. The first is that it yields for $|\zeta_1|^2$ the estimate

$$|\zeta_1(t)|^2 \leq e^{-t}|\zeta_1(0)|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_2(s)|^2\,ds \;. \tag{6.14}$$

The second is that ζ_2 can be written in the form

$$\zeta_2 = \varrho_{k_*-2} + \mathcal{Q}_2(\varrho, x, y) \;,$$

where $\mathcal{Q}_2$ is a polynomial depending only on components ϱ_i, x_i and y_i with $i \geq k_* - 1$, and such that each of its terms contains at least one factor ϱ_i.

Now look at the equation for $\dot{\zeta}_2$. It is clear from the structure of ζ_2 and from the structure of the equations (6.9) and (6.12) that it can be written as

$$\dot{\zeta}_2 = -\zeta_2 + \zeta_3 \;,$$

where

$$\zeta_3 = \varrho_{k_*-3} + \mathcal{Q}_3(\varrho, x, y) \;.$$

This time, the polynomial $\mathcal{Q}_3$ depends only on components with an index $i \geq k_* - 2$. This procedure can be iterated, yielding a whole family of variables

$$\zeta_l = \varrho_{k_*-l} + \mathcal{Q}_l(\varrho, x, y) \;, \tag{6.15}$$

where the $\mathcal{Q}_l$ are polynomials depending only on indices $i \geq k_* - l + 1$, and containing at least one factor ϱ_i in each term. Furthermore, one gets for every ζ_l the estimate

$$|\zeta_l(t)|^2 \leq e^{-t}|\zeta_l(0)|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_{l+1}(s)|^2\,ds\ . \tag{6.16}$$

Notice that (6.16) is valid for $l < k_*$. For $l = k_*$, we have

$$\dot{\zeta}_{k_*} = \mathcal{Q}_{k_*+1}(\varrho, x, y) + G(x, y)\ . \tag{6.17}$$

It thus suffices to choose G in such a way that (6.17) becomes

$$\dot{\zeta}_{k_*} = -\zeta_{k_*}\ . \tag{6.18}$$

Denoting by ζ the vector $\zeta_1, \dots, \zeta_{k_*}$, we get from (6.16) and (6.18) the estimate

$$\|\zeta(t)\|^2 \leq Ce^{-\gamma t}\|\zeta(0)\|^2\ , \tag{6.19}$$

for any $\gamma \in (0,1)$. Plugging this into (6.13) yields for $\|\varrho_H\|$ the estimate

$$\begin{aligned}\|\varrho_H(t)\|^2 &\leq Ce^{-\gamma t}(\|\varrho_H(0)\|^2 + \|\zeta(0)\|^2)\\ &\leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\|)^p\ ,\end{aligned}$$

for some constants C, γ and p. It remains to get an estimate on $\|\varrho_L\|$. From (6.19) and the definition of ζ_1, we get immediately

$$|\varrho_{k_*-1}(t)|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\|)^p\ .$$

From the definition of ζ_2, we get

$$|\varrho_{k_*-2}(t)|^2 \leq C\Big(|\zeta_2(t)|^2 + |\mathcal{Q}_2(\varrho(t), x(t), y(t))|^2\Big)\ .$$

But we know that $\mathcal{Q}_2$ only depends on components of ϱ, x, and y with an index $i \geq k_* - 1$. These are precisely the components of ϱ on which we already have an estimate. We thus get

$$|\varrho_{k_*-2}(t)|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|x(t)\|)^p\ ,$$

for some other power p. Here we used the fact that $y(t) = x(t) + \varrho(t)$ to get rid of $\|y(t)\|$ in the estimate. The same reasoning can be applied to ϱ_{k_*-3}, and so forth down to ϱ_0. We finally get

$$\|\varrho_L(t)\|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|x(t)\|)^p\ , \tag{6.20}$$

for some (large) power of p. We thus verified (5.9a). The bound (5.9b) is obtained in the same way, by noticing that we can as well get the estimate

$$\|\varrho_L(t)\|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|y(t)\|)^p\ ,$$

instead of (6.20). The proof of Theorem 6.4 is complete. □

Remark 6.5 The whole construction is strongly reminiscent of what was done in [EPR99] to control a finite Hamiltonian chain of non-linear oscillators with nearest-neighbour coupling driven by thermal noise at its boundaries.

Remark 6.6 The linearity of the nearest-neighbour coupling is not essential for our argument. We could as well have replaced (6.9) by

$$\begin{aligned} dx_0 &= (a^2 x_0 + V_2'(x_1 - x_0) - V_1'(x_0))\,dt + d\omega\;, \\ \dot x_k &= (a^2 - k^2)x_k + V_2'(x_{k-1} - x_k) + V_2'(x_{k+1} - x_k) - V_1'(x_k)\;, \end{aligned}$$

with V_1 and V_2 two polynomial-like functions, *i.e.* smooth functions such that

$$\frac{d^n V_i(x)}{dx^n} \approx x^{\alpha_i - n} \qquad \text{for} \qquad |x| \to \infty\;,$$

for some $\alpha_i \geq 2$. Imposing the condition $V_2''(x) \geq c$ for some $c > 0$ yields an effective coupling between neighbours at every point of the phase space. This is sufficient to apply our construction.

References

[BKL00a] J. Bricmont, A. Kupiainen, and R. Lefevere, *Ergodicity of the 2D Navier-Stokes Equations with Random Forcing*, Preprint, 2000.

[BKL00b] J. Bricmont, A. Kupiainen, and R. Lefevere, *Exponential Mixing of the 2D Stochastic Navier-Stokes Dynamics*, Preprint, 2000.

[Cer99] S. Cerrai, *Smoothing Properties of Transition Semigroups Relative to SDEs with Values in Banach Spaces*, Probab. Theory Relat. Fields **113** (1999), 85–114.

[DPEZ95] G. Da Prato, K. D. Elworthy, and J. Zabczyk, *Strong Feller Property for Stochastic Semi-Linear Equations*, Stochastic Anal. Appl. **13** (1995), 35–45.

[DPZ91] G. Da Prato and J. Zabczyk, *Smoothing Properties of Transition Semigroups in Hilbert Spaces*, Stochastics Stochastics Rep. **35** (1991), 63–77.

[DPZ92a] G. Da Prato and J. Zabczyk, *Non-explosion, Boundedness and Ergodicity for Stochastic Semilinear Equations*, J. Diff. Equ. **98** (1992), 181–195.

[DPZ92b] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, University Press, Cambridge, 1992.

[DPZ96] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, London Mathematical Society Lecture Note Series, vol. 229, University Press, Cambridge, 1996.

[EH01] J.-P. Eckmann and M. Hairer, *Uniqueness of the Invariant Measure for a Stochastic PDE Driven by Degenerate Noise*, Commun. Math. Phys. **219** (2001), no. 3, 523–565.

[EL01] W. E and D. Liu, *Gibbsian Dynamics and Invariant Measures for Stochastic Dissipative PDEs*, Preprint, Princeton, 2001.

[EMS01] W. E, J. C. Mattingly, and Y. G. Sinai, *Gibbsian Dynamics and Ergodicity for the Stochastically Forced Navier-Stokes Equation*, to appear in Commun. Math. Phys., 2001.

[EPR99] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Entropy Production in Non-Linear, Thermally Driven Hamiltonian Systems*, J. Stat. Phys. **95** (1999), 305–331.

[Hai01] M. Hairer, *Exponential Mixing for a Stochastic PDE Driven by Denerate Noise*, Preprint, Geneva, 2001.

[Hör85] L. Hörmander, *The Analysis of Linear Partial Differential Operators I–IV*, Springer, New York, 1985.

[Ich84] A. Ichikawa, *Semilinear Stochastic Evolution Equations: Boundedness, Stability and Invariant Measures*, Stochastics **12** (1984), 1–39.

[KS00] S. B. Kuksin and A. Shirikyan, *Stochastic Dissipative PDE's and Gibbs Measures*, Commun. Math. Phys. **213** (2000), 291–230.

[KS01] S. B. Kuksin and A. Shirikyan, *A Coupling Approach to Randomly Forced Nonlinear PDE's. I*, Commun. Math. Phys. **221** (2001), 351–366.

[Lun95] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Birkhäuser, Basel, 1995.

[Mat01] J. C. Mattingly, *Exponential Convergence for the Stochastically Forced Navier-Stokes Equations and Other Partially Dissipative Dynamics*, Preprint, 2001.

[MS99] B. Maslowski and J. Seidler, *Invariant Measures for Nonlinear SPDE's: Uniqueness and Stability*, Archivum Math. **34** (1999), 153–172.

[MS01] B. Maslowski and J. Seidler, *Uniform Exponential Ergodicity of Stochastic Dissipative Systems*, Preprint, 2001.

[MT94] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, New York, 1994.

[MY01] N. Masmoudi and L.-S. Young, *Ergodic Theory of Infinite Dimensional Systems with Applications to Dissipative Parabolic PDEs*, Preprint, 2001.

[Nor86] J. Norris, *Simplified Malliavin Calculus*, Lecture Notes in Mathematics **1204** (1986), 101–130.